

Investigating the Impact of COVID-19 Vaccine on COVID-19 Pandemic

Kelvin Njuki

1 Introduction

1.1 Background Information

Early last year (2020), the World was stormed by a deadly novel human coronavirus disease that was later named as COVID-19. The virus was declared as a pandemic by the World Health Organization on March 2020. Businesses and schools were closed as one way of curbing the spread of the virus. Countries closed their boundaries to limit the number of people entering their jurisdiction. The number of people with confirmed COVID-19 cases started increasing as the virus spread from one region to the other. Scientists and scholars moved with speed to identify how the virus was spreading from one person to the other. This discovery led to introduction of some measures that would reduce the spread. These included, sanitizing your hands after touching most accessible surfaces, avoiding shaking of hands, washing your hands multiple times in a day, keeping social distance while in public places, among others. However, these measures were not enough because they could only slow the spread. Further research was conducted and especially on ways to weaken the virus in a human body. This led to start of the vaccine development in different schools and industries. Several vaccines such as Pfizer-BioNTech, Moderna, Johnson & Johnson's Janssen, Oxford-AstraZeneca, Sinopharm BIBP, Sputnik V, CoronaVac, Covaxin, Sputnik Light, among others were developed. Some of these vaccines are still not fully approved by the Food and Drug Administration (FDA), but they are being used for emergency purpose. None of these vaccines has 100% efficiency. However, the scholars and scientists believed that the vaccine would significantly help in reducing the spread of the virus. That is, reducing the number of new cases and deaths. Most of the researchers have addressed the vaccine development progress, Wu (2020), acceptance by people, Lazarus et al. (2021), vaccine efficacy, Lipsitch and Dean (2020), psychological characteristics, Madison et al. (2021), side effects, Saeed et al. (2021), among others. There is little attention on whether the vaccine has significantly reduced the COVID-19 numbers. This study focuses on investigating the effect of the vaccine on COVID-19 pandemic. The results will shed light on how the World is doing on COVID-19 numbers, that is, is the vaccine reliable and is it a permanent solution to the pandemic?

1.2 Research Problem

As alluded earlier, there is little attention on finding out if the vaccine has significantly reduced the COVID-19 numbers. This study will investigate this problem and report the findings. More specifically, the study aims at answering the following research questions:

- i) Is there a negative association between the number of deaths and the number of vaccinations received? That is, can we conclude that as the number of vaccinations increase, the number of new deaths is decreasing?
- ii) Is the number of new cases negatively associated with the number of vaccinations received? That is, can we conclude that as the number of vaccinations increase, the number of new cases is decreasing?
- iii) Has the vaccine reduced the number of positive cases per test?
- iv) Does there exist a significant difference in the number of vaccinations received among different countries?
- v) How is the Gross Domestic Product (GDP) per capita affecting the vaccinations in different continents? Have the continents with high GDP per capita vaccinated more people?

1.3 Data Description

The data used in this study was obtained from Our World in Data platform (Link: COVID-19 Dataset by Our World in Data). The data was collected from February 2019 until to date. It is being updated after every two hours. The data was gathered from various sources such as, Center for Systems and Engineering at Johns Hopkins University (confirmed cases and deaths data), Our World in Data official reports (testing and vaccinations data), other data variables were collected by various sources such as United Nations, World Bank, Global Burden of Disease, Blavatnik School of Government etc. The data set had 67 variables and 136250 observations. The missing observations (NAs) were omitted. The study used only nine variables. Among the nine, two are categorical variables, one is a date variable, and six are quantitative variables. These are,

- **location** - Geographical location
- **continent** - Continent of the geographical location
- **date** - Date of observation
- **population** - Population (latest available values). See https://github.com/owid/covid-19-data/blob/master/scripts/input/un/population_latest.csv for full list of sources

- **new_cases_smoothed_per_million** - New confirmed cases of COVID-19 (7-day smoothed) per 1,000,000 people. “For countries that don’t report new cases data on a daily basis, we assume that numbers changed equally on a daily basis over any periods in which no data was reported. This produces a complete series of daily figures, which is then averaged over a rolling 7-day window.”
- **new_deaths_smoothed_per_million** - New deaths attributed to COVID-19 (7-day smoothed) per 1,000,000 people
- **positive_rate** - The share of COVID-19 tests that are positive, given as a rolling 7-day average.
- **total_vaccinations_per_hundred** - Total number of COVID-19 vaccination doses administered per 100 people in the total population
- **gdp_per_capita** - Gross domestic product at purchasing power parity (constant 2011 international dollars), most recent year available.

To achieve comparison of the number of people vaccinated per continent, we computed the proportion of people vaccinated. This was important because the continents had different population sizes. So the proportion vaccinated was computed as follows;

$$\text{prop_vaccinated} = \{\text{Total number of people vaccinated}\} / \{\text{Total population size}\}$$

2 Methods

The study aimed at finding out if the COVID-19 vaccines had a significant impact on the pandemic. To answer this research problem, there were various statistical methods employed, that is, nonparametric methods with their equivalent parametric methods for comparison purpose. Nonparametric methods used were Spearman correlation test - which was performed to find out if there was a negative association between the new cases and the vaccinations, new deaths and the vaccinations, and positive rate and the vaccinations. Further, the method was used to find if there was any association between GDP per capita and the vaccinations. Pearson correlation test, which is an equivalent parametric method was conducted on the same variables and the results were compared. To find out if there was any significant difference in the number of vaccinations among the continents, and GDP per capita among the continents, Kruskal-Wallis test was performed and compared with its equivalent parametric method (One-Way ANOVA). The tests were conducted at 5% significance level and the analysis was done using R software.

3 Results

3.1 Exploratory Data Analysis

Exploring the data before inferential data analysis is essential because it gives an insight into the data. It helps the researcher decide on the appropriate analysis methods for the study (either parametric or nonparametric). Parametric methods are commonly used when all the underlined assumptions are met. The assumptions includes normality, homoscedasticity, large sample size among others. If the assumptions are not met, then nonparametric methods are used.

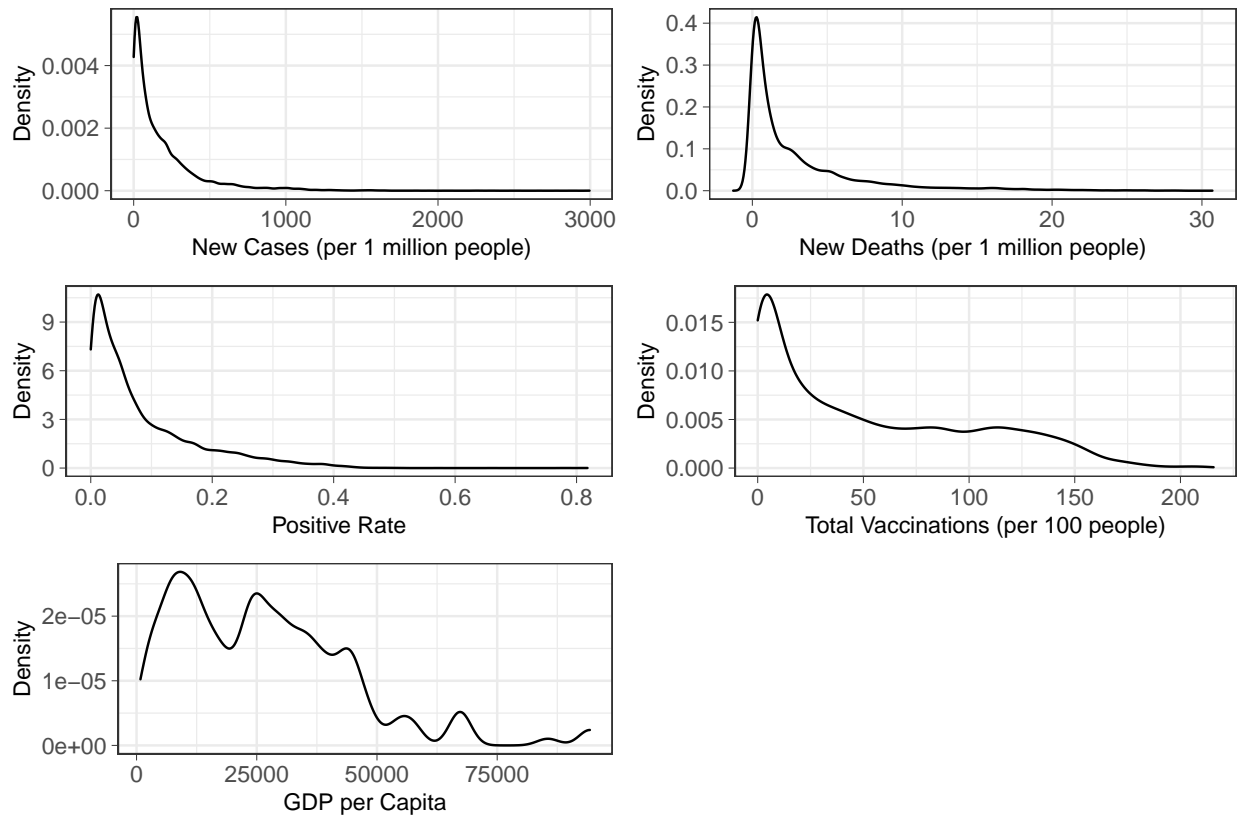


Figure 1: Density curves for new cases, deaths, positive rate and vaccinations

In this study, we generated density curves to find out if the variables were normally distributed. Figure 1 above reports that all the five quantitative variables are heavily skewed to the right. This implies that for valid results, nonparametric methods should be used. Further, we plotted scatter plots for our variables to find out if there was any linear association between them.

Figure 2 below shows some association between vaccinations and other variables, but it is quite difficult to tell the strength and direction of the association. Therefore, conducting inferential analysis such as Pearson and/or Spearman correlation would give us a better conclusion.

To find out if the number of vaccinations and GDP per capita varied among the continents, we generate a bar plots of the vaccinations and GDP per capita, each faceted by the continent. The plot (Figure 3) shows some difference in these variables among the continents. However, we are not sure if the difference is significant. Therefore, conducting inferential analysis such Kruskal-Wallis and/or One-Way ANOVA would answer this question.

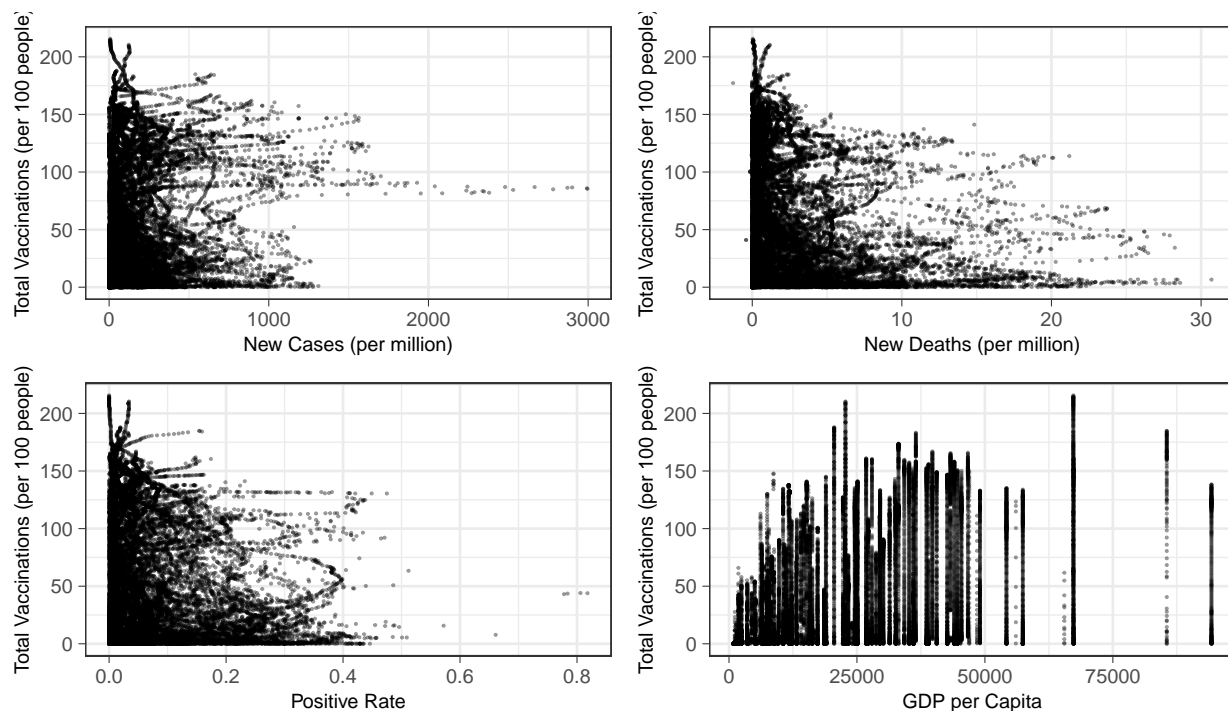


Figure 2: Scatter plots for new cases, deaths, positive rate against vaccinations

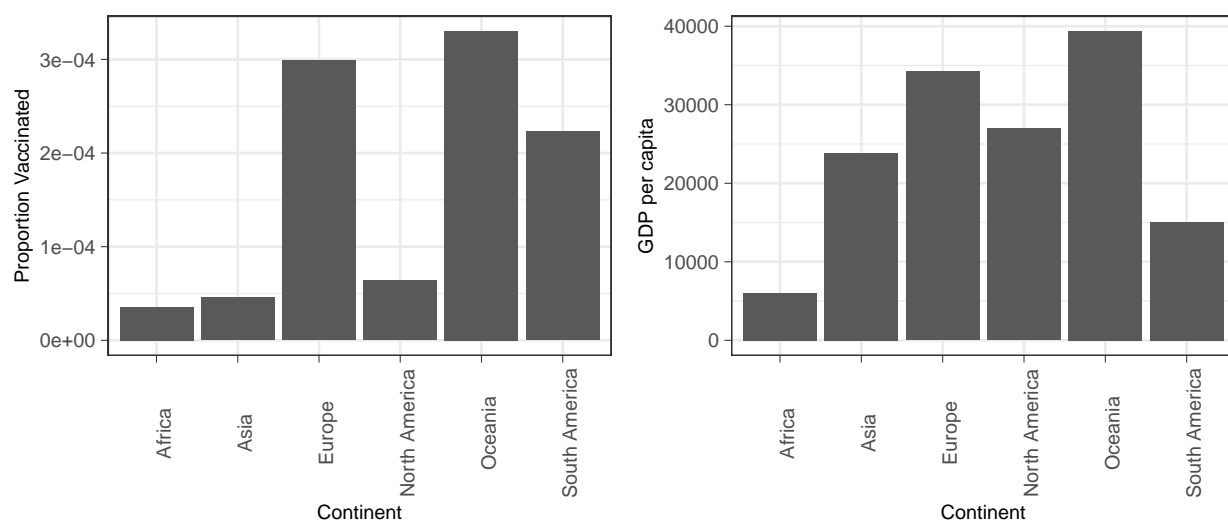


Figure 3: Bar plots for number of vaccinations and GDP per capita across the continents

3.2 Inferential Analysis

Tables 1, 2 and 3 below contains the results from the inferential analysis methods. The discussions of the results is below.

3.2.1 Negative association between new cases and vaccinations?

The results from Spearman correlation test (nonparametric) suggests that there is no negative association between the number of new cases and the vaccinations ($p\text{-value} = 1$) at 5% significance level. Instead, there is a “very” weak positive association. An equivalent parametric method (Pearson correlation test) reports a similar conclusion.

3.2.2 Negative association between new deaths and vaccinations?

Spearman correlation test (nonparametric) shows a weak negative association between vaccinations and the number of new deaths ($p\text{-value} = 0$). The Pearson correlation test also shows presence of a negative association between the two variables ($p\text{-value} = 0$).

3.2.3 Negative association between positive rate and vaccinations?

Positive rate is the share of COVID-19 tests that are positive. From the results below, both Spearman and Pearson reports a weak, negative association between vaccinations and the number of positive cases per test ($p\text{-value} = 0$).

3.2.4 Association between GDP per Capita and vaccinations

GDP per capita plays a major role when it comes to budgeting in different countries. It was worth investigating if the number of vaccinations were affected by the GDP per capita. Both Spearman and Pearson correlation tests reported a weak positive association between these two variables.

Table 1: Spearman correlation test (Nonparametric Method) results

| Variables | P-value | $\hat{\rho}$ | Strength | Direction |
|--------------------------------|--------------------|--------------|----------|-----------|
| New Cases vs Vaccinations | 1 | 0.0967 | Weak | Positive |
| New Deaths vs Vaccinations | 0 | -0.1693 | Weak | Negative |
| Positive Rate vs Vaccinations | 0 | -0.2782 | Weak | Negative |
| GDP per Capita vs Vaccinations | $< 2.2\text{e-}16$ | 0.3935 | Weak | Positive |

Table 2: Pearson correlation test (Parametric Method) results

| Variables | P-value | $\hat{\rho}$ | Strength | Direction |
|--------------------------------|-----------|--------------|----------|-----------|
| New Cases vs Vaccinations | 1 | 0.0780 | Weak | Positive |
| New Deaths vs Vaccinations | 0 | -0.2203 | Weak | Negative |
| Positive Rate vs Vaccinations | < 2.2e-16 | -0.2525 | Weak | Negative |
| GDP per Capita vs Vaccinations | < 2.2e-16 | 0.3776 | Weak | Positive |

3.2.5 Any difference in the GDP per Capita among the continents?

Kruskal-Wallis rank sum test (nonparametric) shows a significant difference in the GDP per capita among the continents. The equivalent parametric method (One-Way ANOVA) agrees with these results.

3.2.6 Any difference in the number of vaccinations among the continents?

Kruskal-Wallis rank sum test (nonparametric) shows there is a significant difference in the number of vaccinations among the continents. The equivalent parametric method (One-Way ANOVA) reports similar results. This could be explained by the GDP per capita in these continents. We have seen before that GDP is positively correlated with the number of vaccinations.

Table 3: P-values from a nonparametric (Kruskal-Wallis) and parametric (One-Way ANOVA) method

| Variables | Kruskal-Wallis test | One-Way ANOVA test |
|------------------------------|---------------------|--------------------|
| GDP per Capita vs Continents | < 2.2e-16 | < 2.2e-16 |
| Vaccinations vs Continents | < 2.2e-16 | < 2.2e-16 |

4 Discussion

COVID-19 vaccine brought a lot of hope to the World. The expectation was that the vaccine would significantly reduce the number of new cases and deaths. Although, the results have showed that the vaccine has a negative association with the number of new deaths, this association is very weak (Section 3.2.2). The number of positive cases per test is also negatively associated with the vaccinations, which implies that the number of positive cases per test is reducing as more people continue getting vaccinated (Section 3.2.3). This is a sign that the vaccine is reducing the spread of the virus. However, it is doing so at a very slow rate as we have seen the correlation coefficients were very low (weak association). The number of new cases has a positive association with the number of vaccinations (Section 3.2.1). Although, the association is very weak, it is wake up call. This suggests that the vaccination is yet to significantly reduce the number of new cases. The proportion of people vaccinated is significantly differing among the continents. This could be due to the difference in the GDP per capita.

5 Conclusion and Recommendations

Generally, the COVID-19 is reducing the number of deaths and positive cases per test, but at a very slow rate. The number of new cases are still being reported but they are not as many as when the pandemic started. This implies that if the vaccinations campaign continues and people accept to be vaccinated, the number of new cases and deaths will eventually go down significantly. I would recommend that people continue taking precautions to avoid increase on new cases and deaths. The people who are not vaccinated should consider getting vaccinated to reduce the number of deaths. All other measures as advised by the Center for Disease Control should continue as we hope the numbers will eventually decrease significantly.

6 References

- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). Nonparametric statistical methods (Vol. 751). John Wiley & Sons.
- Lazarus, Jeffrey V, Scott C Ratzan, Adam Palayew, Lawrence O Gostin, Heidi J Larson, Kenneth Rabin, Spencer Kimball, and Ayman El-Mohandes. 2021. “A Global Survey of Potential Acceptance of a COVID-19 Vaccine.” *Nature Medicine* 27 (2): 225–28.
- Lipsitch, Marc, and Natalie E Dean. 2020. “Understanding COVID-19 Vaccine Efficacy.” *Science* 370 (6518): 763–65.
- Madison, Annelise A, M Rosie Shrout, Megan E Renna, and Janice K Kiecolt-Glaser. 2021. “Psychological and Behavioral Predictors of Vaccine Efficacy: Considerations for COVID-19.” *Perspectives on Psychological Science* 16 (2): 191–203.
- Saeed, Balsam Qubais, Rula Al-Shahrabi, Shaikha Salah Alhaj, Zainab Mansour Alkokhardi, and Ahmed Omar Adrees. 2021. “Side Effects and Perceptions Following Sinopharm COVID-19 Vaccination.” *International Journal of Infectious Diseases* 111: 219–26.
- Wu, Suh-Chin. 2020. “Progress and Concept for COVID-19 Vaccine Development.” *Biotechnology Journal*.