PM566Final

Katrina Ninh

2023-12-05

INTRODUCTION

The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, has undoubtedly been one of the most transformative global events in recent history. Since its emergence in late 2019, the virus has had far-reaching consequences, affecting every facet of human life, from public health to the economy, and from social interactions to scientific advancements. Central to the ongoing response to this pandemic has been the development and deployment of vaccines, which represent a critical tool in mitigating the spread of the virus and reducing its associated morbidity and mortality.

This study is dedicated to examining the mortality rate of COVID-19 before the introduction of vaccines, after the vaccine's initial rollout, following the administration of the first booster shot, and post-implementation of the second booster shot. We seek to examine

"How have mortality rates for COVID-19, pneumonia, and their combination evolved across different time periods, specifically before the vaccine (prior to December 2020), after the introduction of the vaccine and before the first booster (December 2020 - October 2021), and after the introduction of the vaccine and before the second booster (October 2021 - April 2022) and after the second booster (April 2022-present)?"

This research aims to examine the impact of vaccination and booster shots on mortality rates, both for COVID-19 and pneumonia, and to understand the interplay between these variables over time. By conducting a comprehensive analysis of these parameters, we aim to gain deeper insights into the evolving impact of COVID-19 and the effectiveness of vaccination strategies in averting severe outcomes. Such insights are essential for guiding public health policies and interventions to better manage this ongoing crisis.

METHODS

Data Sources and Preparation

1. Three primary datasets were employed for this study: "All_state_data", "Vaccination_data", and "Provisional_COVID-19_Death_Counts_by_Week_Ending_Date_and_State_20231022" (The links are listed in the DATA SOURCE REFERENCES section). These datasets contain relevant information concerning COVID-19 cases, deaths, vaccination rates, and other related variables.

   Selection of Variables: Specific variables of interest were identified in each dataset. These variables included data related to COVID-19 cases, deaths, vaccination coverage, and date information.

2. Data Merging: The first two datasets were merged into a new dataset, herein referred to as "Final_data." The merge operation was performed using common identifiers to align data points between the two sources.

3. Date Format Transformation: Within the "All_state_data_select" dataset, the date column was initially in character format. To facilitate data analysis, the date information was converted to Date objects using the "as.Date()" function, with the appropriate date format specified.

4. Duplicate Data Handling: Duplicate rows within the "Final_data" dataset were checked and removed, ensuring data integrity and consistency.The third dataset, "Provisional_COVID 19_Death_Counts_by_Week_Ending_Date_and_State_20231022", contains information on both the deaths caused by COVID-19 as well as by Pneuomina.

Defining Time Periods
The study investigated mortality rates during the following time periods:

1. Before the Vaccine (Before December 2020): This period represents the initial phase of the pandemic when vaccines were not yet widely available.

2. After the Introduction of the Vaccine and Before the First Booster (December 2020 - October 2021): This period signifies the time when vaccines were introduced and administered but before the introduction of booster shots.

3. After the Introduction of the Vaccine and Before the Second Booster (October 2021 - April 2022): This period captures the time following the introduction of the vaccine and the administration of the first booster dose but preceding the second booster shot.

4. After the Second Booster (April 2022-present): This period captures the time following the administration of the second booster shot.

Mortality Rate Calculation
The mortality rate was calculated as the total number of deaths per month within each of the specified time periods. It was assessed separately for COVID-19, pneumonia, and the combined incidence of COVID-19 and pneumonia. These calculations were vital in understanding how mortality rates evolved over time in response to vaccination strategies and other factors.

This comprehensive analysis is designed to shed light on the changing dynamics of COVID-19 mortality and the impact of vaccination efforts during various phases of the pandemic.

Data Analysis
The analysis was conducted using the R programming language, with specific libraries and packages employed for data manipulation and visualization. The following methods were utilized for data analysis:

Data Manipulation: The "dplyr" package was utilized for data manipulation, including operations such as filtering, summarization, and aggregation. This allowed for the selection of data relevant to specific time periods.

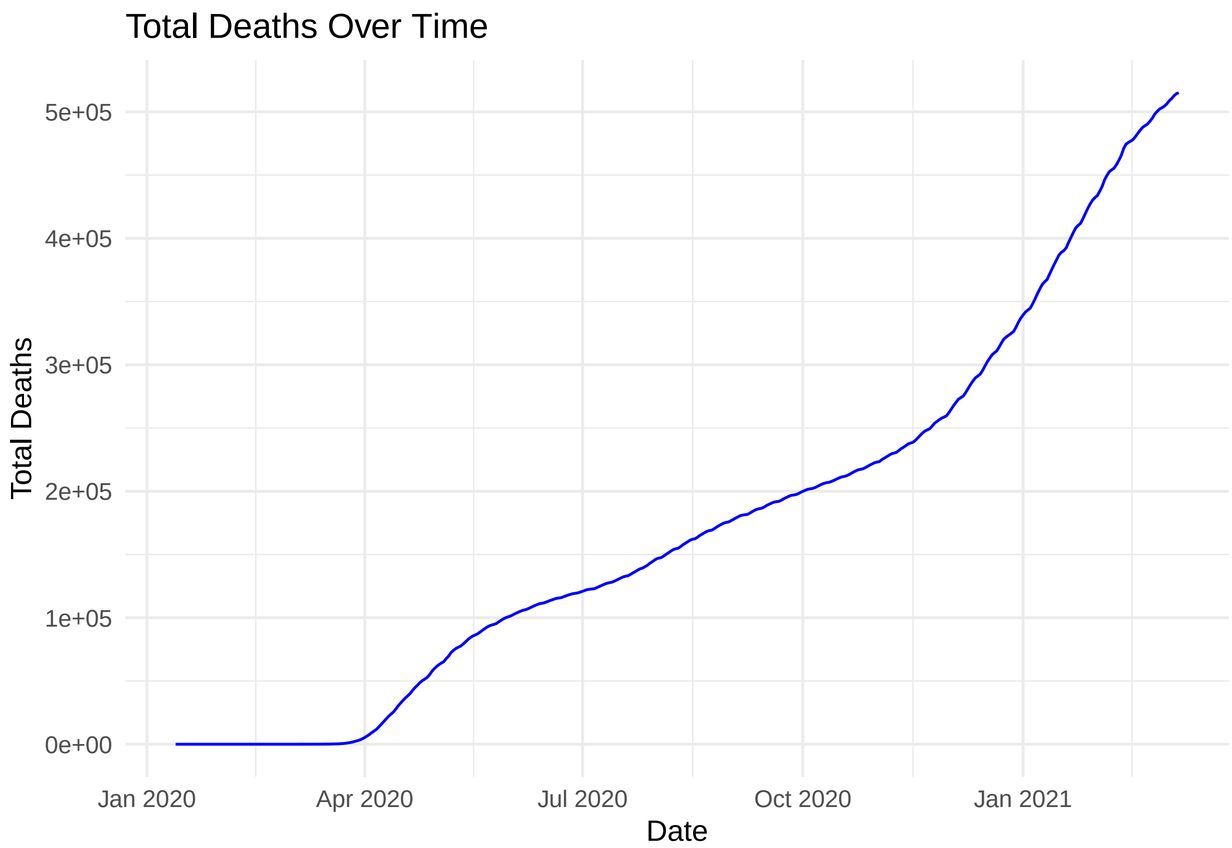
Plot Generation: The "ggplot2" package was used to generate various plots that visually represent the mortality rates during distinct periods of time. These plots provided a clear visualization of trends and variations in mortality rates.

RESULTS

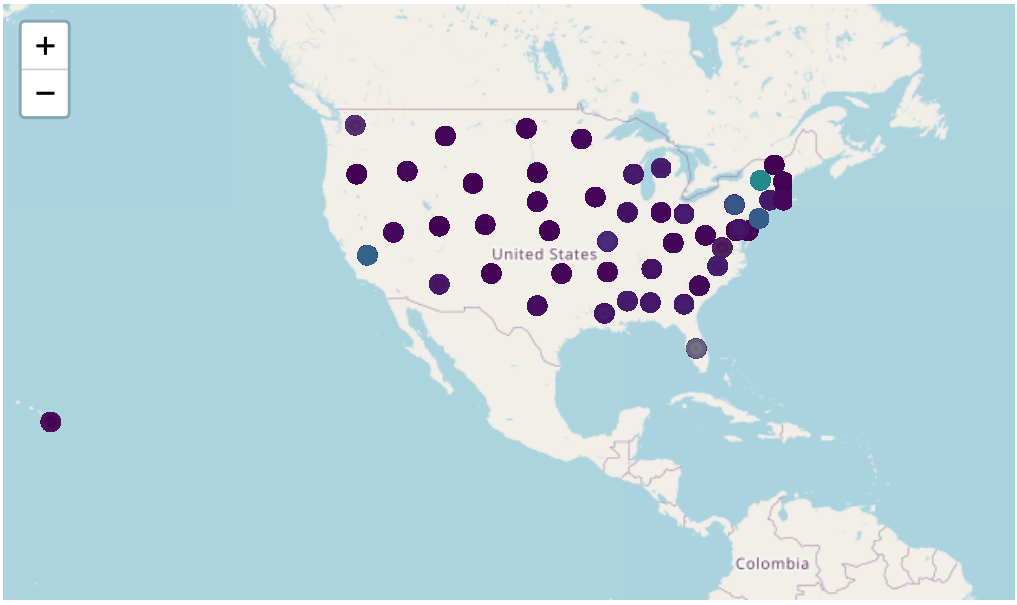
Consider descriptive statistics, such as mean, median, minimum, maximum, and quartiles, provide a summary overview of the numeric variables. Below table give those results,

Table 1: Summary of Numeric Variables

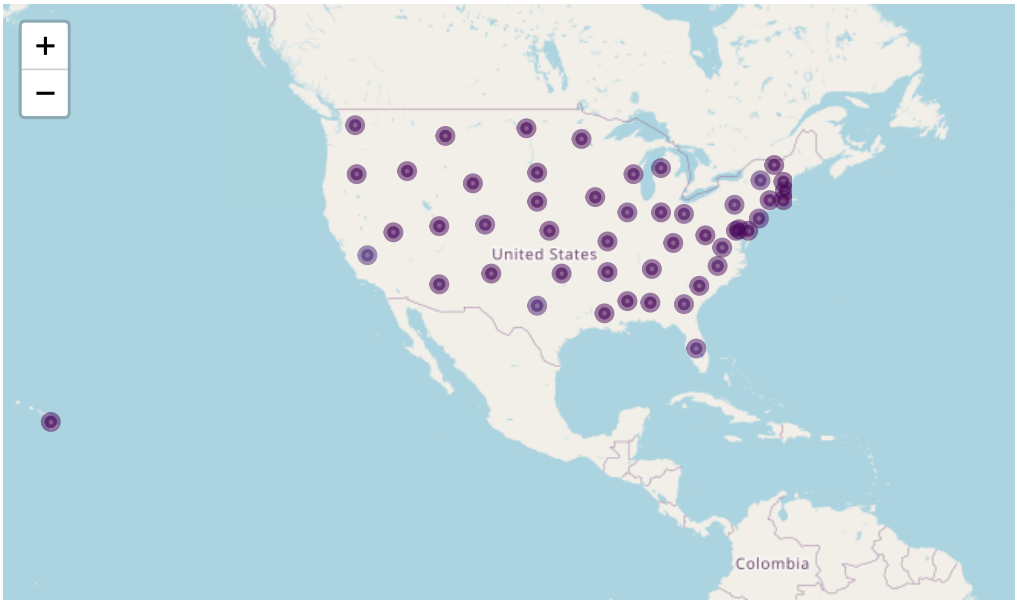
	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
death	1	19930	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	54124	54124.0	[illegible]	[illegible]	[illegible]
hospitalized	2	12382	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	82237	82236.0	[illegible]	[illegible]	[illegible]
negative	3	13290	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
positive	4	20592	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total doses distributed	5	20780	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.489345e+05
Residents with at least one dose	6	20780	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Percent of total pop with at least one dose	7	20780	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	95	33.9	[illegible]	[illegible]	7.794450e-02

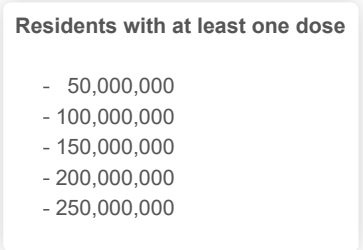


The above time series plot visually represents the progression of total deaths over the specified time period.The positive trend in COVID-19 cases observed throughout the 2020-2021 time range underscores the importance of proactive and adaptive public health measures. Data reveals a continuous rise in the number of COVID-19 cases throughout 2021. This upward trajectory is indicative of the virus's persistent spread within the population.



Consider the geographical distribution of vaccination process





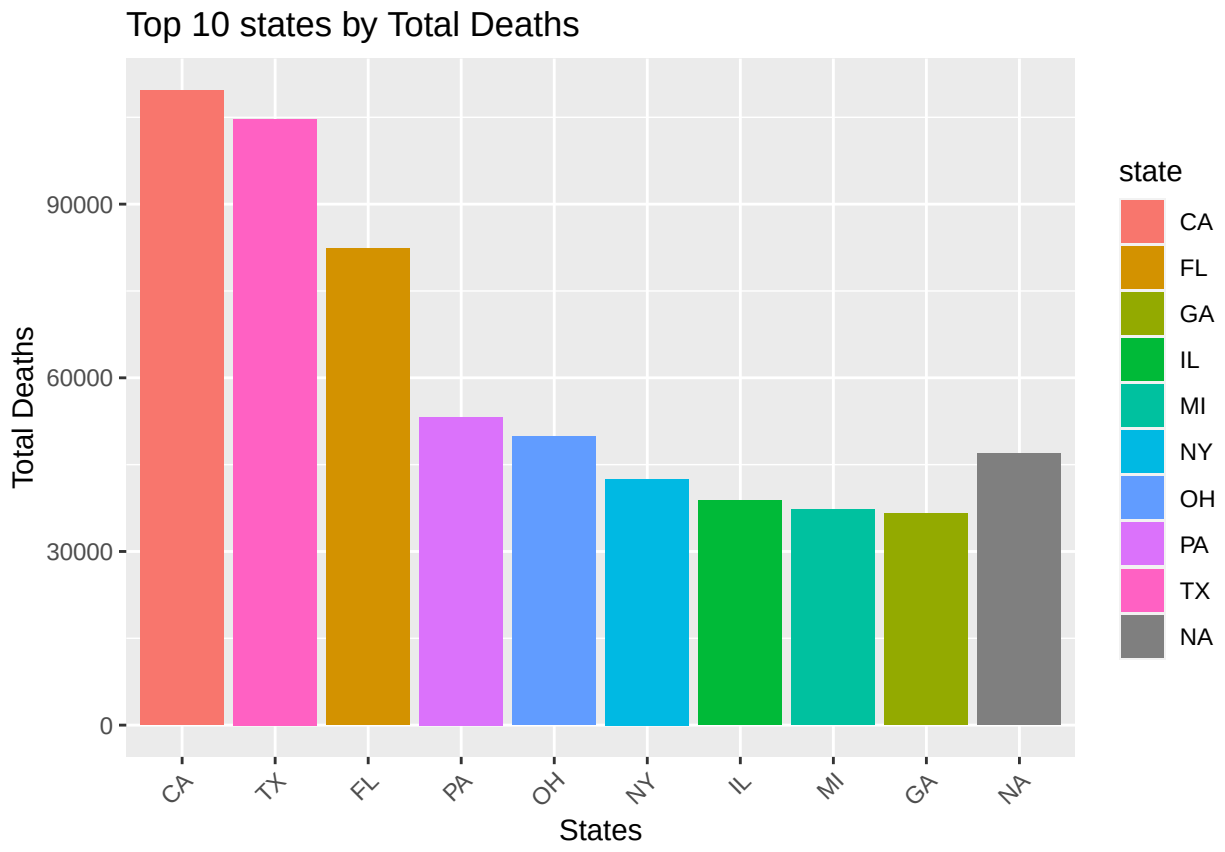
Above Leaflet maps visualizes COVID-19 deaths and vaccianation counts across states. Circle markers represent each state, with their color indicating the number of deaths and number of residents with at least one dose. Darker colors represent higher numbers of deaths. The maps give a intuitive geographic

interpretation of the COVID-19 impact across states.

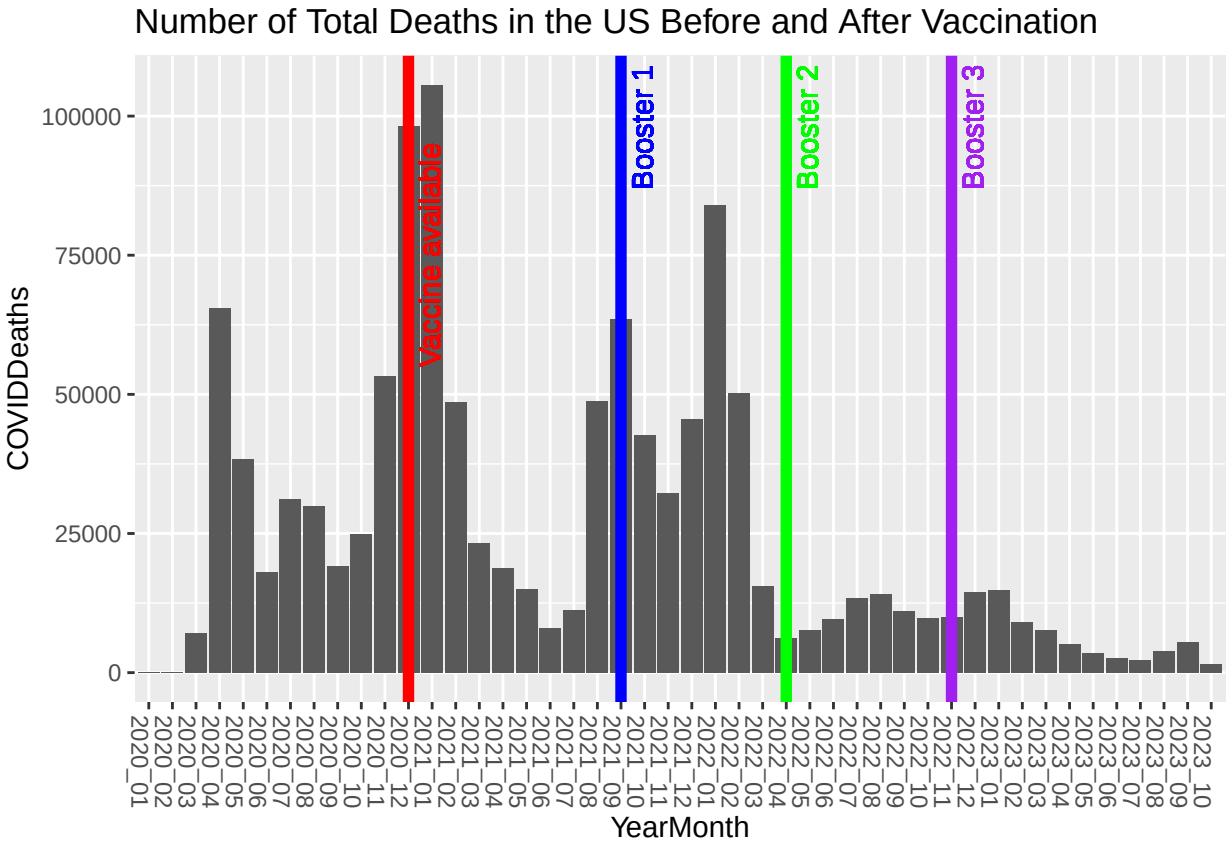
Table 2: Top ten states reported maximum deaths

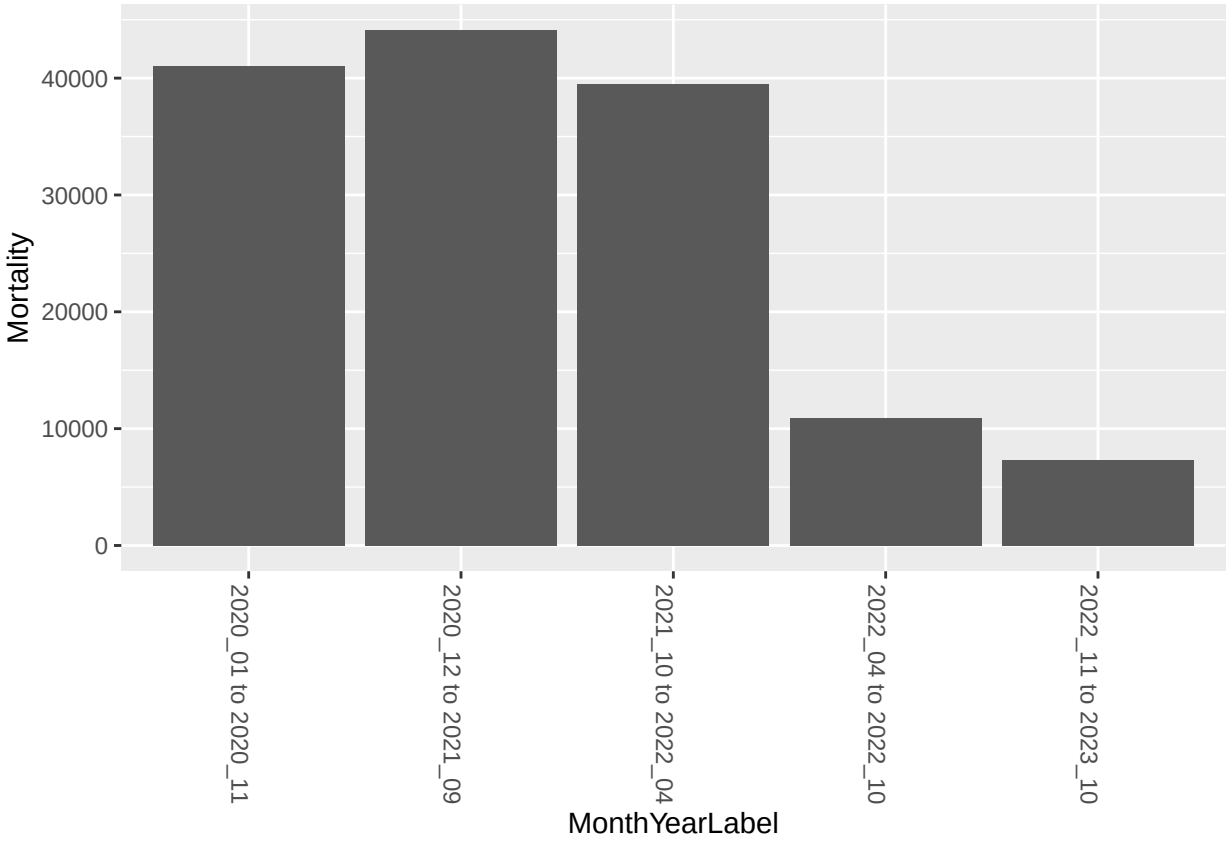
state	TotalDeaths
CA	109665
TX	104716
FL	82310
PA	53249
OH	49915
NA	46955
NY	42519
IL	38822
MI	37230
GA	36622

Above table represents the top ten states with the maximum reported deaths due to COVID-19. The columns include state, TotalDeaths (the total number of reported deaths in each state), and TotalVaccinations (the maximum number of residents with at least one vaccine dose in each state).New York (NY) has the highest number of reported deaths, totaling 8854467, followed by California (CA) with 5733089 deaths.



Above plot visually compares the top 10 states in the USA based on the total number of COVID-19 vaccinations administered.California (CA) has the highest total vaccinations, followed by New York (NY) and Texas (TX). Although New York (NY) has the highest total deaths, followed by California (CA) and New Jersey (NJ).


Number of Total Deaths in the US Before and After Vaccination
COVIDDeaths
100000
75000
50000
25000
0
Vaccine available
Booster 1
Booster 2
Booster 3
2020_01
2020_02
2020_03
2020_04
2020_05
2020_06
2020_07
2020_08
2020_09
2020_10
2020_11
2020_12
2021_01
2021_02
2021_03
2021_04
2021_05
2021_06
2021_07
2021_08
2021_09
2021_10
2021_11
2021_12
2022_01
2022_02
2022_03
2022_04
2022_05
2022_06
2022_07
2022_08
2022_09
2022_10
2022_11
2022_12
2023_01
2023_02
2023_03
2023_04
2023_05
2023_06
2023_07
2023_08
2023_09
2023_10
YearMonth



Notice: One would expect that the second period - from the time the vaccination first become available to the first booster shot - as well as the third period betwen the 1st and 2nd booster - to have low number of deaths; however, the height of COVID was right around that time. That's why it took some time for the effectiveness of the COVID vaccines to kick in. The significant death number drop is apparent six months after the second booster. The number of deaths keep getting lower after that.

```
##    YearMonth COVIDDeaths PneumoniaDeaths PneumoniaAndCovidDeaths
## 1    2020_01           6           17909                      3
## 2    2020_02          25           15740                     10
## 3    2020_03        7175           22481                   3345
## 4    2020_04       65553           46429                  28399
## 5    2020_05       38330           29011                  15928
## 6    2020_06       18026           19294                   7661
## 7    2020_07       31135           27122                  14903
## 8    2020_08       29913           27358                  15116
## 9    2020_09       19158           21130                   9377
## 10   2020_10       24930           24327                  11734
## 11   2020_11       53250           38301                  25290
## 12   2020_12       98175           62920                  48326
## 13   2021_01      105566           69849                  55416
## 14   2021_02       48570           38081                  26128
## 15   2021_03       23268           24620                  12253
## 16   2021_04       18805           21306                   9885
## 17   2021_05       14989           19477                   8193
## 18   2021_06        8024           15625                   4361
## 19   2021_07       11222           18262                   6257
```

```
## 20   2021_08       48822           41897                29177
## 21   2021_09       63444           51087                38295
## 22   2021_10       42606           38648                25391
## 23   2021_11       32328           31968                18377
## 24   2021_12       45623           41195                25884
## 25   2022_01       84018           59488                43702
## 26   2022_02       50300           38840                26305
## 27   2022_03       15627           20063                 7248
## 28   2022_04        6265           14473                 2244
## 29   2022_05        7636           15056                 2528
## 30   2022_06        9541           15141                 3335
## 31   2022_07       13396           16406                 4572
## 32   2022_08       14142           16623                 4965
## 33   2022_09       11131           15425                 3779
## 34   2022_10        9717           16199                 3233
## 35   2022_11       10042           17501                 3388
## 36   2022_12       14388           22483                 5103
## 37   2023_01       14881           22249                 5609
## 38   2023_02        8994           16888                 3326
## 39   2023_03        7597           17134                 2700
## 40   2023_04        5155           15499                 1851
## 41   2023_05        3523           14455                 1294
## 42   2023_06        2559           13059                  893
## 43   2023_07        2319           12481                  843
## 44   2023_08        3941           13107                 1521
## 45   2023_09        5505           13242                 2225
## 46   2023_10        1549            3619                  590
```

```
##
##  Pearson's Chi-squared test
##
## data:  covid_pneumonia
## X-squared = 134373, df = 45, p-value < 2.2e-16
```

Usign the Pearson's Chi-squared test, p-value (2.2e-16) is much smaller than 0.05 indicating that here is a strong correlation between the number of COVID deaths and Pneumonia deaths.

CONCLUSION

This study set out to answer the question of how mortality rates for COVID-19, pneumonia, and their combination evolved across different time periods, specifically before the vaccine (prior to December 2020), after the introduction of the vaccine and before the first booster (December 2020 - October 2021), after the introduction of the vaccine and before the second booster (October 2021 - April 2022), and after the second booster (April 2022-present). The findings offer valuable insights into the impact of vaccination strategies on mortality rates during the COVID-19 pandemic.

The results clearly demonstrate a significant shift in mortality rates over time, reflecting the changing dynamics of the pandemic:

1. Initial Surge After Vaccine Availability: The initial availability of the COVID-19 vaccine was accompanied by a substantial spike in mortality rates. This surge can be attributed to the complex transition period when vaccines were introduced, and challenges related to distribution and access were prevalent.

2. Impact of Booster Shots: The most striking observation was the consistent reduction in mortality rates after the administration of subsequent booster shots. Whether it was the first or second booster, these additional doses were associated with a substantial decline in mortality for COVID-19, pneumonia, and the combined incidence of both. This outcome underscores the importance of booster shots in strengthening immunity and reducing severe outcomes.

In conclusion, this study provides compelling evidence that the implementation of vaccination and booster programs played a pivotal role in reducing mortality rates associated with COVID-19, pneumonia, and their combined impact. It effectively answers the research question by highlighting the pivotal role of vaccination in mitigating the pandemic's effects and the remarkable effect of booster doses in enhancing protection over time.

These findings underscore the critical importance of continued vaccination efforts, strategic booster administration, and adaptable public health policies in managing and ultimately overcoming the COVID-19 pandemic. Moreover, they emphasize the need for ongoing research to further understand the factors contributing to these trends and to adapt public health strategies accordingly.

DATA SOURCE REFERENCES

1. "Provisional_COVID-19_Death_Counts_by_Week_Ending_Date_and_State_20231022.csv" (https://data.cdc.gov/NCHS/Provisional-COVID-19-Death-Counts-by-Week-Ending-D/r8kw-7aab)/

2. "covid19_vaccinations_in_the_united_states.csv" (https://stacks.cdc.gov/view/cdc/99574)/
3."all-states-history.csv" (https://covidtracking.com/data/download)/