ARIMA Model Analysis of COVID-19 Mortality Rate Changes: Evaluation of Pfizer Vaccine Effectiveness

Weiqun Xu

School Of Health Policy & Management, Nanjing Medical University, Nanjing, 211166, China

xuweiqunxwq@stu.njmu.edu.cn

Abstract. COVID-19, caused by a novel coronavirus called SARS-CoV-2, is an infectious disease. The virus was initially identified in late 2019 in Wuhan, Hubei Province, China, and rapidly disseminated worldwide, resulting in a widespread global pandemic and substantial loss of life. To examine whether the Pfizer vaccine had a significant effect on the reduction of mortality from CDC by analysing the difference in mortality rates before and after Pfizer vaccination of confirmed cases of COVID-19 An autoregressive summated moving average (ARIMA) was constructed based on data released by the CDC. The implicit assumption of this model is that the predictions of the model are treated as a "control group" containing only the time trend that is not affected by the vaccine, and that the experimental group is the true value affected by the vaccine. Two models, ARIMA (2,1,1) and ARIMA (10,2,3), showed an increasing difference between predicted and actual values. The mortality rate of COVID-19 had a substantial decrease following widespread administration of the Pfizer COVID-19 vaccine. This study aims to examine if the widespread distribution of the Pfizer vaccine has resulted in a decrease in the fatality rate of COVID-19 infections.

Keywords: Covid-19, Death rate, ARIMA.

1 Introduction

Since the COVID-19 epidemic various countries have suffered a major blow, with more than 1.19 million cumulative confirmed deaths in the United States according to real-time data from the CDC's Centers for Disease Control and Prevention (https://covid.cdc.gov). And according to real-time tracking data from the World Health Organisation there are nearly 800 million confirmed cases and nearly seven million deaths globally(https://coronavirus.jhu.edu/) At the beginning of the epidemic the clinical presentation of COVID-19 can be divided into approximately two categories the first is asymptomatic carriers and the second and fulminant disease characterised by severe respiratory symptoms especially acute respiratory failure, where severe symptoms are present in about 5% of COVID-19 patients and 20% of hospitalised patients requiring treatment in intensive care units. In addition, about more than 75% of COVID-19 inpatients require supplemental oxygen to sustain life due to

Q. Wu et al. (eds.), *Proceedings of the 2024 3rd International Conference on Public Service, Economic Management and Sustainable Development (PESD 2024)*, Advances in Economics, Business and Management Research 309,
https://doi.org/10.2991/978-94-6463-598-0_22

cardiopulmonary insufficiency respiratory failure. [1] In response to the severe symptoms and complications of early epidemics of COVID-19 and the high mortality rate, vaccine development has been initiated in various countries, especially in China and the United States, where the first COVID-19 vaccine was developed, and mass vaccination was initiated at almost the same time. Because for this pandemic virus, vaccination to prevent the disease and medication to treat the symptoms to improve the body's immunity is the most effective way to control the pandemic.

Many biopharmaceutical companies have begun research and development such as: Pfizer - BioNTech (Comirnaty) - mRNA vaccine developed by Pfizer in the USA in collaboration with BioNTech Germany, Moderna - mRNA-1273 vaccine developed by Moderna in the USA, AstraZeneca - mRNA-1273 vaccine developed by AstraZeneca UK in collaboration with the University of Oxford, and the mRNA-1273 vaccine developed by Pfizer in the USA in collaboration with the University of Oxford. (AstraZeneca) - Adenovirus vector vaccine developed by AstraZeneca in the UK in co-operation with the University of Oxford (Vaxzevria), Johnson & Johnson - Adenovirus vector vaccine developed by Johnson & Johnson in the USA (Janssen), Sinovac - Inactivated vaccine developed by Sinovac (CoronaVac). Sinopharm - an inactivated vaccine developed by Sinopharm (BBIBP-CorV), Bharat Biotech - an inactivated vaccine developed by Bharat Biotech (Covaxin), have all started clinical trials and published their results, and in the case of Pfizer's trial, the BNT162b2 vaccine candidate has been developed by Pfizer. BNT162b2 vaccine candidate developed by Pfizer achieved good results and verified its safety and efficacy [2].

However, many questions about the efficacy of the vaccine have arisen at a time when the neoguana outbreak has subsided, and more and more people are beginning to doubt the safety and efficacy of the neoguana vaccine. [3] Currently, researchers have developed a variety of models for COVID-19 vaccine efficacy prediction, including infectious disease dynamics models (e.g., the SEIR model), segmented regression models, Bayesian time-series models, exponential smoothing models, machine-learning methods, and so on. However, these epidemiological models are sensitive to parameters and rely on multiple assumptions, whereas machine learning methods, Bayesian time series, etc. require large amounts of data, have high complexity and uncertainty, and have ethical concerns [4]. To overcome these limitations, this study attempts to use the ARI-MA model. This model has the advantages of simple structure, wide adaptability, and accurate capture of features of time series data [5], and can effectively predict whether the COVID-19 vaccine has a direct effect on the mortality of COVID-19s [6]. Therefore, in this paper, the ARI-MA model was chosen to analyse whether mass vaccination with the Pfizer vaccine had a significant impact on the mortality rate of COVID-19 infections in the United States.

2 Research Design

2.1 Data Source

Data were published from 2020-1-11, when the counting of new coronavirus mortality began, to 2024-6-1, when the counting ceased. In this case, 2020-12-14 Pfizer vaccine

mass vaccination in the United States was taken as the time point, 2020-1-11 to 2020-12-14 as the training data, and 2020-12-14 to 2024-6-1 as the comparison of validation data.

2.2 ARIMA Model

The ARI-MA model is used to analyse and forecast random time series. The key steps in constructing this model include: first checking the smoothness of the time series, and if it is not smooth then performing a differencing operation to determine the parameter d; then estimating the autoregressive part(parameter p) and the moving average part (parameter q) using the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots; and then estimating the autoregressive and moving average part of the model (parameter q) by checking that the sequence of the model residuals is not white noise (using the ACF, PACF plots and Box-Pierce test) to ensure the validity of the model. In ARIMA it can be divided into three parts, the first being the AR, the differencing process (I), and the MA [7]. The "p" in this context represents the term "autoregressive" which denotes the aggregation of earlier observations in the model, specifically capturing a linear combination across the initial p periods. This formulation is defined as:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t \tag{1}$$

q represents the portion of the moving average, i.e., the relationship between the current value and the past white noise, which is given by the following:

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_p \varepsilon_{t-p} \tag{2}$$

Confirmation of I conversion of a non-stationary series to a stationary series of difference order dd

2.3 ADF Unit Root Test

Prior to modelling, it is necessary to do a unit root test on the data, assuming the initial hypothesis that the data is nonstationary. After putting the data into stata for the ADF test, this part can see the Table 1 shows that the p-value for the first-order differencing is 0.007, and the p-value for the second-order differencing is 0.000, both of which are below 0.1. so, the original hypothesis can be rejected. It shows that after the differential treatment this data tends to be stable and meets the modelling requirements of ARIMA

Table 1. test for weak stationarity

	t	p
Ln value	-3.574	0.0321
1st order difference	-4.069	0.0070
2nd order difference	-5.449	0.0000

3 Results and Analysis Based on Empirical Evidence

3.1 Determination of Order and Testing for Residuals

The data is basically stabilized in the time series after the first order differencing therefore it was determined that the ARIMA model with d=1

This section of the thesis primarily starts with the use of AFC and PACF plots i.e. Figure 1, to rank the mortality rate of COVID-19 infections before the Pfizer COVID-19 vaccine was introduced The models ARIMA (2,1,1) and ARIMA (10,2,3) were derived based on the AIC minimum information criterion as well as the relevant model fit indices.

PACF and ACF are used in ARIMA model to determine the order of AR and MA respectively. 17PACF can determine the order of P and ACF can determine the order of Q. The Partial Autocorrelation Function (PACF) and Autocorrelation Function (ACF) can determine the order of Q.

The basic principle of order fixing: the value of PACF or ACF of a certain order, which exceeds the boundary (i.e., the critical value), the order is significant and desirable.

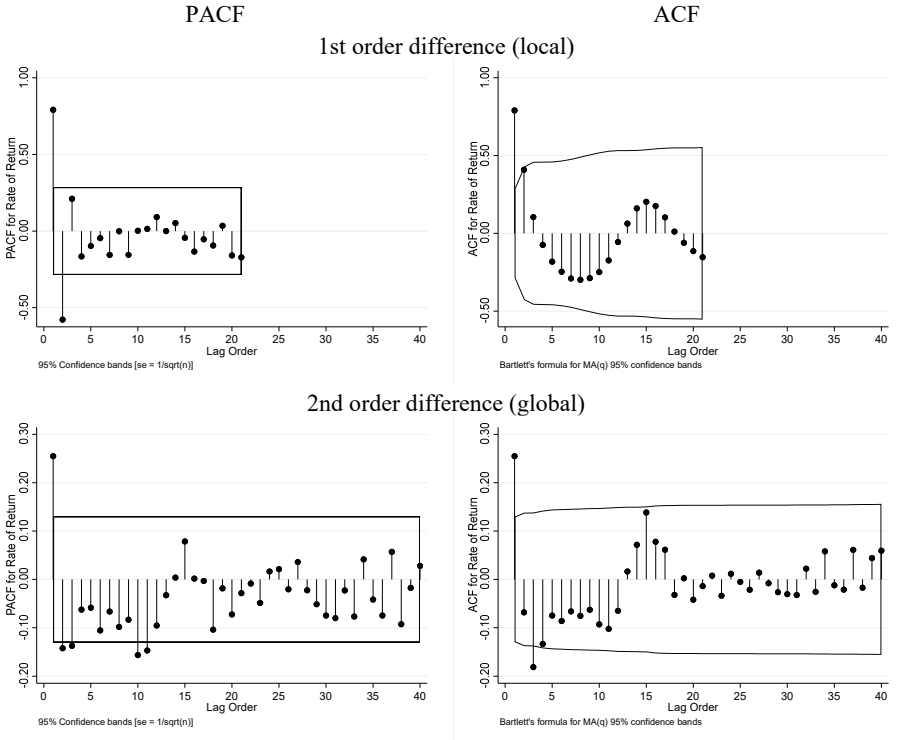

Fig. 1. ARMA (p, q) identification (Photo credit: Original)

When analysing any time series whose ACF in figure shows tail extension and PACF in figure shows truncation, in applicable model is AR model. The intended value of the parameter p is specifically determined by the lag order of the PACF truncated tail. If the partial autocorrelation function (PACF) displays a tail extension and the autocorrelation function (ACF) indicates a truncated state, then the appropriate model to use is the moving average (MA) model. In this case, the optimal value for the parameter q is equal to the lag order of the ACF truncation. Therefore, we build the ARIMA(p,d,q) model based on these two sets of data sequentially and perform the residual test shown in Table 2 below after the model construction is completed.

Table 2. Residual test

Model	Portmanteau (Q) statistic	Prob > chi2
Local - ARIMA (2,1,1)	33.8962	0.7405
Global – ARIMA (10,2,3)	14.5156	0.9999

The ARIMA model successfully passes the residual test, demonstrating that the residuals conform to the characteristics of a white noise series, which is characterised by its unpredictability. So after completing the construction of ARIMA model and residuals test, Continue utilising Stata for data forecasting and generating graphs to predict the timeframe for the Pfizer COVID-19 vaccination in the America before the mass vaccination (local) that is, 2020-12-14 data as the training data and the number of prediction after that that is the model Local-ARIMA (2,1,1), the other one is the The other is the CDC's data on the overall COVID-19 mortality rate in the US during the COVID-19 epidemic (global) ,model Global-ARIMA(10,2,3).

3.2 Model Results

Figures 2, 3, 4, and 5 provide the actual and fitted values of weekly mortality for the two different models, with graphs reflecting the differences.

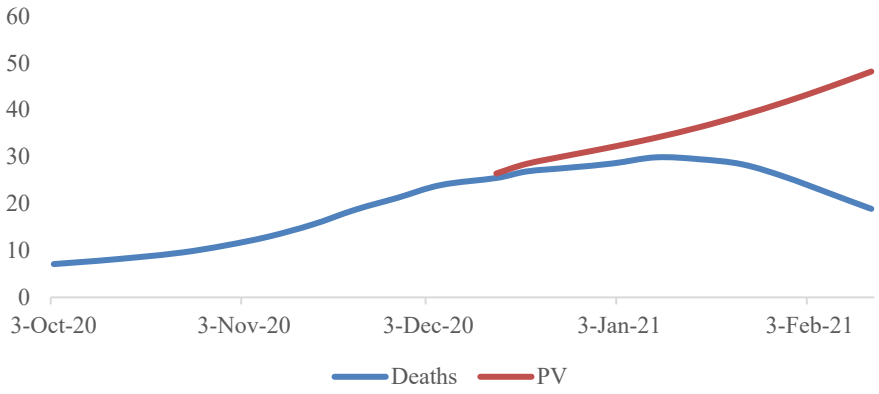

Fig. 2. Mortality before and after Covid-19, local model (Photo credit: Original)

The folded line in orange in the figure indicates the direction of the fitted mortality rate, and a table of the data based on the ARIMA (2,1,1) model shows that without mass vaccination in the United States, the mortality rate for COVID-19 infections would have been increasing and exponentially increasing [8].

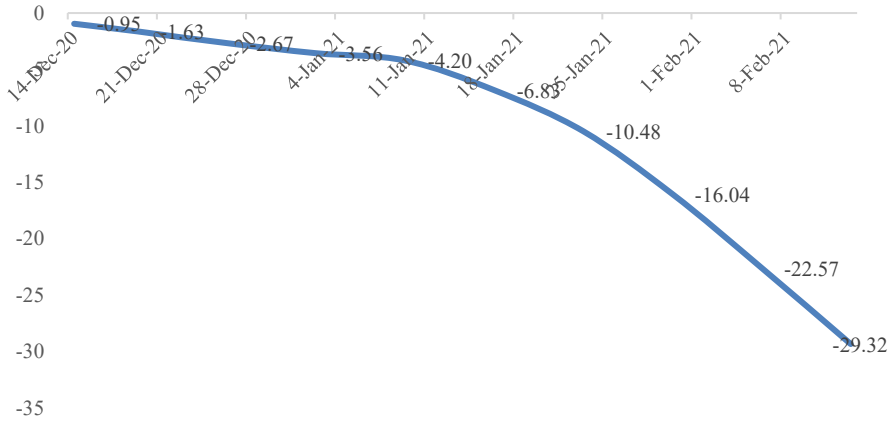

Fig. 3. The effect of Pfizer COVID-19 vaccine on mortality, local model.
Photo credit: Original

However, with the implementation of mass vaccination, the reality showed that the rate of increase in mortality began to slow down and declined significantly after peaking [9].

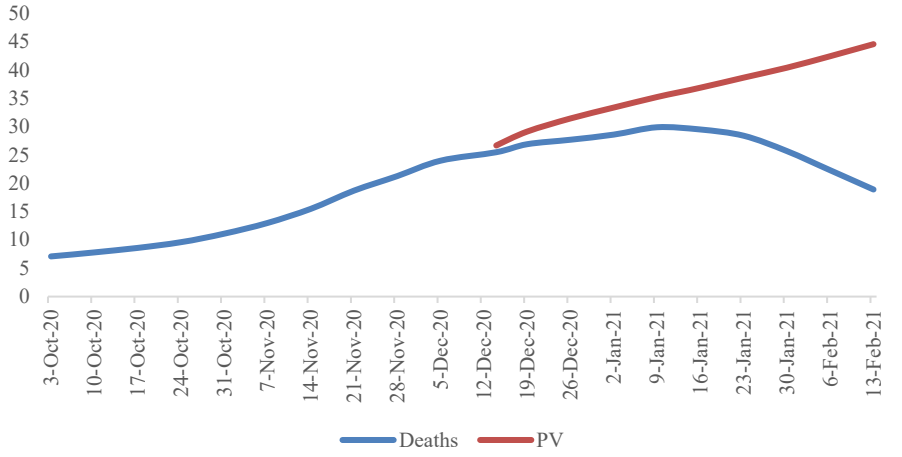

Fig. 4. Mortality before and after Covid-19, global model.
Photo credit: Original

By analysing the plotted difference graphs, it is clear to observe that the difference between the mortality rate predicted from the training data and the actual mortality rate from COVID-19 infections gradually widens, and that this widening is occurring at an increasing rate, especially in the data following mass vaccination with the vaccine. This disparity is shown even more clearly in two graphs based on the total US COVID-19 mortality rate.

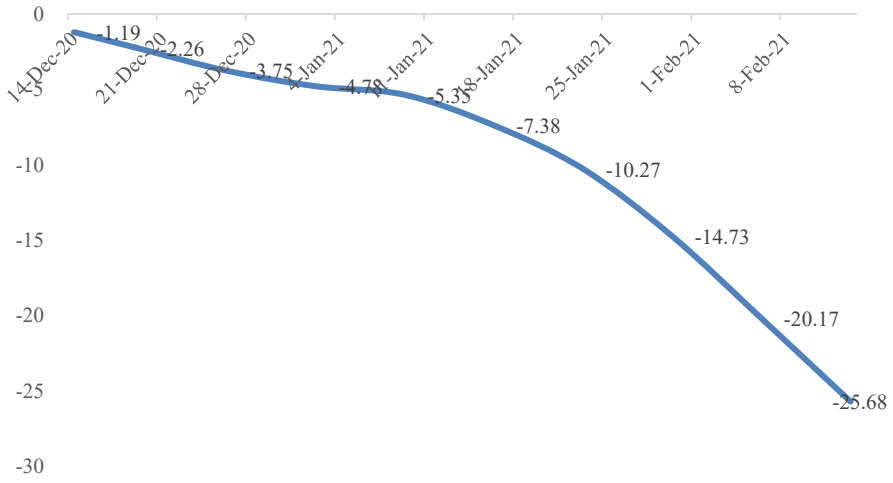

Fig. 5. The effect of Pfizer COVID-19 vaccine on mortality, global model.
Photo credit: Original

The observed gap between the actual data and the model-fitted data as the Pfizer COVID-19 vaccine becomes widely available shows the critical role of the vaccine in reducing mortality from COVID-19 infections. [10] Although this effect was not apparent in the early stages of vaccination, as vaccination coverage increased and more and more people were protected by the vaccine, the mortality rate began to decline significantly. At the same time, there was a consistent increase in the disparity between the actual data and the predicted values of the model, particularly in the later stages, with discrepancies reaching as high as 25%.Therefore, the present study clearly demonstrates that the widespread vaccination with Pfizer's COVID-19 vaccine had a significant and sustained impact on reducing the mortality rate of COVID-19 infections in the America [11].

4 Discussion

There are now two main views on the role of the COVID-19 vaccine, the first being that the majority of views suggest that the COVID-19 vaccine significantly reduces the risk of COVID-19 infection and in particular reduces the mortality rate in critically ill patients and post-infection. Many studies have also evaluated a long-lasting and

effective effect on viral immunity after vaccination, and these views are in line with those of this paper. However, it has to be noted that in some reports there are negative aspects of vaccination such as side effects after vaccination such as severe allergies and myocarditis, as well as the lack of effectiveness of vaccines against viral variant, especially against some resistant variants.

Based on the research ideas in this paper we can see that with the increase in vaccine coverage the mortality rate of COVID-19 infection in the United States has been decreasing, so the following insights can be drawn

For policy making: Firstly, this study demonstrates the importance of mass vaccination, and that governments need to maximise the effectiveness of vaccination through high vaccination coverage, in order to obtain the protective effect of the vaccine. Second, although many of the developed countries' Neoguana data monitoring organisations have ended their monitoring of actual Neoguana outbreaks, they should maintain their monitoring of vaccine coverage in less developed areas. Third, public awareness campaigns should be strengthened to reduce negative public sentiment about vaccines in the streaming media era, to increase public confidence in vaccines, and to strengthen their understanding of the need for vaccination.

From the perspectives of ordinary people and patients in society, this paper can give some insights. Firstly, vaccination is essential and a high vaccination coverage rate is necessary to achieve group protection. Secondly, patients should be more confident as the medical and statistical community is nowadays continuously focusing on the safety of vaccines and updating vaccination strategies and recommendations to maximise the protection of patients' health.

5 Conclusion

The main focus of this paper is to thoroughly analyze the vaccination on mortality fromCOVID-19 infections during the U.S COVID-19 pandemic and to provide policy makers with policy references and recommendations versus vaccination recommendations for the general public. In order to achieve this goal this study chose to use the mortality data published by the CDC to observe the difference between their actual and fitted values before and after Pfizer's COVID-19 vaccination.

In the course of this study, an ARIMA model was used to predict the direction of the mortality rate of COVID-19 epidemic infections in the United States in the absence of COVID-19 vaccination, predicting the predicted value of its mortality rate. This method is broadly more accurate for capturing time series. Predictions can be made from both the pre-mass vaccination data and the overall data, and the study showed that the mortality rate of the COVID-19 outbreak was significantly reduced after mass vaccination with the Pfizer vaccine.

There are still some shortcomings in this paper, such as there are many external factors may interfere with the prediction of ARIMA model, such as policies, the public's willingness to vaccinate, the government's vaccination policy and so on. In addition, the ARIMA model is usually based on a fixed time series and seasonality to predict the vaccination, while vaccination is a dynamic process.

As vaccination becomes more widespread, governments and organisations can track and evaluate the effects of vaccination over time to see if the vaccine provides long-term protection against mutated viruses, and more complex models that take into account a wider range of factors can be used in developed regions to study the protection of vaccines against new cases of coronavirus.

References

1. Lu, Cheng-kuan. (2021-09-01). New strategy to trace the origin of new coronaviruses to natural evolutionary processes. Science and Technology Daily, 003.
2. Yadav, T., Kumar, S., Mishra, G., & Saxena, S. K. (2023). Tracking the COVID-19 vaccines: The global landscape. Human vaccines & immunotherapeutics, 19(1), 2191577.
3. Troiano, G., & Nardi, A. (2021). Vaccine hesitancy in the era of COVID-19. Public health, 194, 245-251.
4. Shi, Liang, Zhang jianfeng, Li Wei & Yang Kun. (2022). Artificial Intelligence for Prevention and Control of Tropical Infectious Diseases. Chinese Journal of Schistosomiasis Control (05), 445-452.
5. Lou, H. R., Wang, X., Gao, Y., & Zeng, Q. (2022). Comparison of ARIMA model, DNN model and LSTM model in predicting disease burden of occupational pneumoconiosis in Tianjin, China. BMC public health, 22(1), 2167.
6. Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief, 29, 105340.
7. Yadav, T., Kumar, S., Mishra, G., & Saxena, S. K. (2023). Tracking the COVID-19 vaccines: The global landscape. Human vaccines & immunotherapeutics, 19(1), 2191577.
8. Yue, H., & Hu, T. (2021). Geographical Detector-Based Spatial Modeling of the COVID-19 Mortality Rate in the Continental United States. International journal of environmental research and public health, 18(13), 6832.
9. James, N., Menzies, M., & Radchenko, P. (2021). COVID-19 second wave mortality in Europe and the United States. Chaos (Woodbury, N.Y.), 31(3), 031105. https://doi.org/10.1063/5.0041569
10. Skowronski, D. M., & De Serres, G. (2021). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. The New England journal of medicine, 384(16), 1576-1577.
11. Tregoning, J. S., Flight, K. E., Higham, S. L., Wang, Z., & Pierce, B. F. (2021). Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. Nature reviews. Immunology, 21(10), 626-636.