SHORT COMMUNICATION

Forecasting the Spread of the Sixth Wave of COVID-19 Epidemic in Southern Iran: An Application of ARIMA Models

Vahid Rahmanian[1], PhD;
Mohammad Jokar[2], DVM;
Elham Mansoorian[3], MSc

[1]Department of Public Health, Torbat Jam Faculty of Medical Sciences, Torbat Jam, Iran
[2]Faculty of Veterinary Medicine, Karaj Branch, Islamic Azad University, Karaj, Iran
[3]Research Center for Social Determinants of Health, Jahrom University of Medical Sciences, Jahrom, Iran

Correspondence:
Vahid Rahmanian, PhD;
Department of Public Health, Torbat Jam Faculty of Medical Sciences, Torbat Jam, Iran
Tel: +98 9175985204
Email: vahid.rahmani1392@gmail.com
Received: 15 April 2023
Revised: 18 May 2023
Accepted: 30 June 2023

Abstract

Background: Through the fifth wave of the Covid-19 outbreak in Jahrom, the fatality and incidence of the virus increased. The quick spread of infection is one of the causes of this dreadful situation. Therefore, recognizing the future epidemic trend can be a useful instrument to decrease mortality and morbidity. This study aimed to determine the time trends and select the best model to predict the sixth wave of the COVID-19 outbreak using ARIMA models.

Methods: We used daily data of 9533 hospital cases (Suspected and PCR-confirmed COVID-19 cases) between 4th March 2020 and 31st December 2021. Nine different ARIMA models were fitted to our data. Autocorrelation functions (ACF) and partial autocorrelation (PACF) plots were used to determine model parameters. Likelihood-ratio test for comparison of the reduced and full model was used. In addition, Akaike Information Criteria (AIC) was also used to choose the final model. Data were analyzed by STATA 14 software with a significant level of 0.05.

Results: The ARIMA (3, 0, 3) model was selected among the potential models, with lower AIC (999) and MAPE (3.18%) values. This model showed that the daily number of hospitalized patients may increase from 5.85 (2.16-15.79) to 8.55 (1.47-49.48) in two months. By March 01, 2022, the predictable daily hospitalized cases could reach 468.36 (03.79-2209.88).

Conclusion: Time series models is a useful tool for predicting the hospitals' admission trend during an epidemic. Thus, they can be used as early warning models in the readiness of hospital systems during epidemics.

Please cite this article as: Rahmanian V, Jokar M, Mansoorian E. Forecasting the Spread of the Sixth Wave of COVID-19 Epidemic in Southern Iran: An Application of ARIMA Models. J Health Sci Surveillance Sys. 2023;11(Supplement 3):664-669.

Keywords: COVID-19, Predicting, Sixth wave, Spread, Time series modeling

Introduction

With the advent of the SARS-CoV-2, it first spread to Wuhan, China, and then spread to all countries shortly.[1] The World Health Organization (WHO) called it Covid-19 and a pandemic and declared a state of emergency.[2] Covid-19 is an emerging and unknown disease, and despite much research, many aspects remain unknown. WHO has stated that the most important way to control the disease is to break the chain of transmission. In this regardthe most important health recommendations were to observe personal hygiene and reduce people's exposure to each other.[3]

This pandemic of the COVID- 2019 is spreading worldwide, with 289,713,814 confirmed cases and 5,457,330 deaths globally when writing this report on 2022/01/02, 6:00 am. The Islamic Republic of Iran reported 6,266,939 confirmed cases and 131,639 deaths as of 02nd January 2022.[4] The virus has spread to all provinces of Iran, and no region has been left without confirmed virus cases. Prevention strategies (vaccination, case finding and massive testing, elective quarantine) advised and other coping methods applied, such as tracking individual's contact with confirmed patients using extensive testing and segregation of asymptomatic patients, can help control the pandemic and decrease the basic reproduction number (R_0) to less than one.[5] These strategies have had varying degrees of success across different countries.[5, 6]

Despite the approval of multiple vaccines and injection of the third and fourth booster doses of Covid-19 vaccines in many countries, the virus showed multiple mutations such as Delta (B. 1.617. 2), Lambda (C.37) and Omicron (B. 1.1. 529). It has raised many concerns for major health systems worldwide about the potential impact of existing vaccines on mutations and when the pandemic will end. Therefore, how long will this pandemic continue in IR and the world?

This study aimed to determine the time trends and select the best model to predict the sixth wave of the COVID-19 epidemic, using ARIMA models in Jahrom County.

Methods

This study used daily data of 9533 hospital patients (Suspected and PCR-confirmed COVID-19 cases) between 4th March 2020 and 31st December 2021. During the study period, the daily incidence cases were used to create an ARIMA time series analysis. In general, when dealing with data that does not exhibit seasonal patterns, the model is fitted as ARIMA (p, d, q). The parameters of the ARIMA model are as follows: p represents the autoregressive number (AR), which determines how many time periods are used to predict the current period. This parameter is determined from the partial autocorrelation functions (PACF) diagram, d: the number of differentials taken to static mean and the parameter q represents the moving averages number (MA), which takes into account the deviation of the series data from the series average over a certain number of time periods to predict the current time period. This parameter is specified from the autocorrelation functions (ACF) diagram.[7, 8]

In the next stage, the Likelihood-ratio test was used to compare the reduced and full model. AIC (lower Akaike Information Criteria) and BIC (Bayesian information criterion) were good methods among models extended by diverse lags.[9]

Then, the selected ARIMA model was used to predict the frequency of daily admission of hospitalized cases. The predicting precision was assessed by MAPE (mean absolute percentage error), via the following formula, the lower the value, the higher the precision of the fitted model:

$$MAPE = \frac{1}{N}\sum_{t=1}^{n}\frac{\text{Real cases} - \text{Forecasted cases}}{\text{Real cases}}$$

Where: N=The number of time forecast.

Data analysis was performed using Stata software (version 14) in a time series analysis package Box-Cox regression test was used to evaluate the stationary in the data variance and Dickey-Fuller test was used to confirm the stationary in the data mean. To confirm the goodness-of-fit of the selected model, the authors used the histogram of the residuals (difference between actual and predicted cases) to check the normality, and the portmanteau test was used to show the independence of the residuals. The alpha level was 0.05.

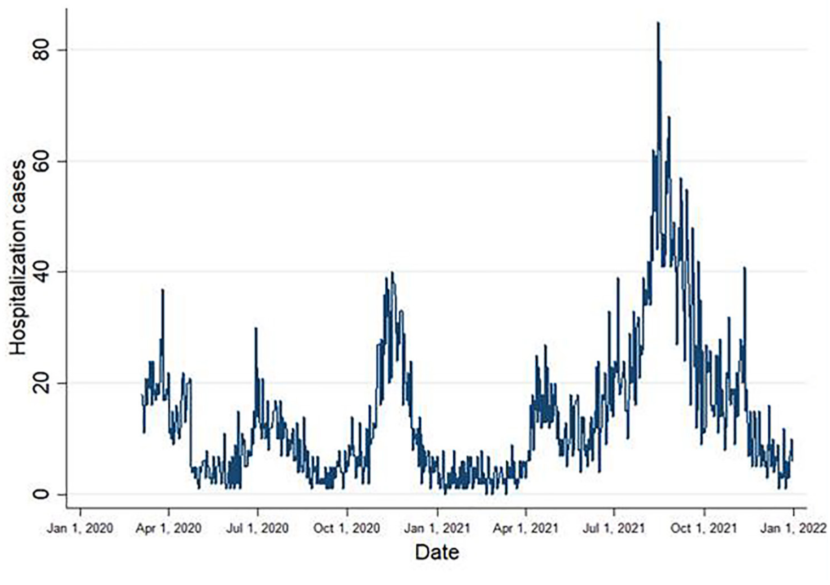


Figure 1: Trend and distribution 9533 hospital cases between 4th March 2020 and 31th December 2021in Pymaneie hospital.

Results

Figure 1 shows the trend of daily admission cases in a hospital specified for COVID-19 (Peymanie) in Jahrom County from 4th March 2020 and 31st December 2021.

In the next step, the third-order smoothing method was used to remove the series noises and the regular and justified patterns, and reduce the fluctuations in the data. This method considers the average of the earlier three data collected in three different times instead of individual data (moving average).

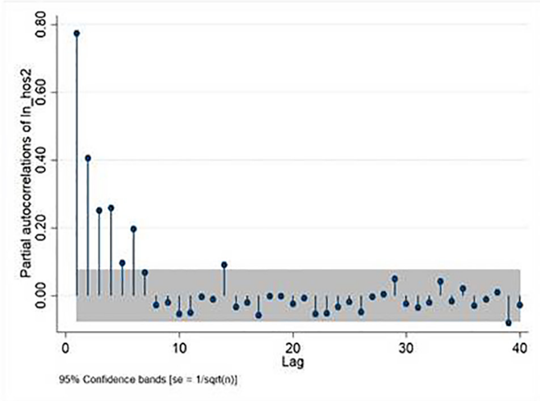
The highest number of hospitalization occurred on August 23rd 2021 (60 cases) and has been decreasing since November 15th 2021 (13 case).

ARIMA model was made using hospital case data within the study period. It showed that the series were non-stationary regarding variances (theta=0.037), and log conversion was applied. The mean was stationary in a series (Dickey-Fuller, t statistic=-9.18 P<0.001).

To describe the main parameters of the ARIMA model (p, d, q), ACF and PACF diagrams were drawn for hospital case, and a combination of MA (27) and AR (6) to the possible models was suggested (Figure 2). In these charts, the gray zone displays the 95% confidence level, and the lines that are continuous out of the limited area will be significant. According to the ACF plot, the lack of sinusoidal wave coefficients, and the lack of significant correlation in 12 and 24 lags, the disease does not have a seasonal trend.

The possible models for ARIMA of hospital cases were ARIMA (0, 0, 1); ARIMA (1, 0, 1); ARIMA (1, 0, 2), ARIMA (2, 0, 1), ARIMA (2, 0, 2); ARIMA (3, 0, 2), ARIMA (3, 0, 3), ARIMA (4, 0, 3) and ARIMA (3, 0, 4).

The authors created different models by changing the parameter to achieve the final model with the most appropriate value of parameters. Furthermore the model with fewer parameters was reduced model, and the model with more parameters was named the full model. In the next step, the models were compared in pairs with the likelihood ratio test, and finally, the best model and the parsimonious model were selected as the final model, considering the lowest AIC and BIC measures.



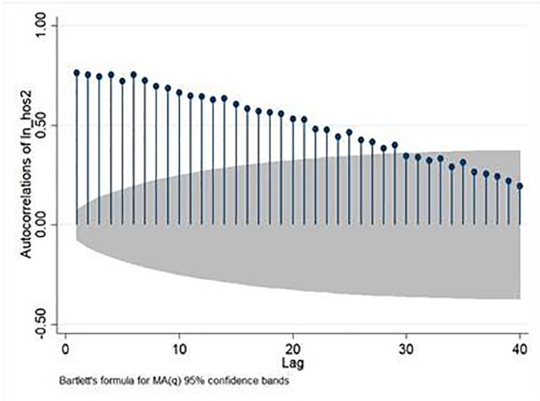


Figure 2: Autocorrelation functions (ACF) and Partial autocorrelation functions (PACF) plots of the trend in Hospital cases.

After this step, it was found that the ARIMA (3, 0, 3) model was the most suitable of all investigated models, with lower AIC (999) and BIC (1035) measures (Table 1).

In reality, for every possible model, a time series analysis was done, and the AIC quantity of the appropriate model was calculated and compared together. The P value of the Likelihood-ratio test for reduced and full models was significant (P<0.001).

Table 1: ARIMA models for cases of hospital cases.

Component	ARIMA (0, 0, 1)	ARIMA (1, 0, 1)	ARIMA (1, 0, 2)	ARIMA (2, 0, 1)	ARIMA (2, 0, 2)	ARIMA (3, 0, 2)	ARIMA (3, 0, 3)	ARIMA (3, 0, 4)	ARIMA (4, 0, 3)
Constant	2.25	2.26	2.26	2.26	2.26	2.26	2.26	0.67	2.26
L1.ar	-	0.98	0.98	0.92	0.12	1.67	2.17	0.62	0.87
L2.ar	-	-	-	-0.02	0.85	0.78	-1.84	-0.31	0.45
L3.ar	-	-	-	-	-	0.10	0.66	-	-0.52
L4.ar	-	-	-	-	-	-	-	-	0.17
L1.ma	0.51	-0.72	-0.74	-0.71	0.11	-1.44	-1.93	-0.44	-0.63
L2. ma	-	-	-0.01	-	-0.60	0.58	1.51	-0.57	-0.45
L3. ma	-	-	-	-	-	-	-0.43	0.22	0.41
L4. ma	-	-	-	-	-	-	-	0.13	-
Sigma	0.76	0.51	0.51	0.51	0.51	0.50	0.50	0.50	0.50
AIC	1543	1005	1007	1007	1008	1003	999	1002	1002
BIC	1556	1023	1030	1030	1035	1034	1035	1042	1042

AR (autoregressive), MA (moving average), AIC (Akaike information criteria), BIC (Bayesian information criterion), L (Lag), - (Not applicable)

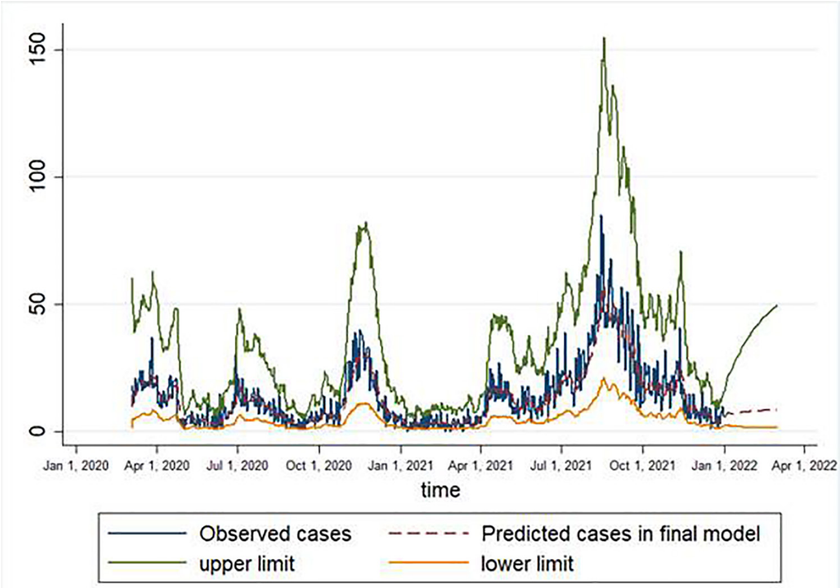


Figure 3: Observed numbers hospital cases for the period between 4th March 2020 and 31th December 2021 and 1-step ahead predicted values between 01th January 2022 to 01th March 2022 based values on ARIMA (3, 0, 3) model.

Table 2: The forecast values (95% *CI*) according to fitted models of hospital patients for the period from 01th January 2022 to 01th March 2022.

Days	Forecast Cases (95%CI)	Days	Forecast Cases (95%CI)
01Jan2022	5.85 (2.16-15.79)	01 Feb 2022	7.92 (1.59-39.32)
02Jan2022	6.28 (2.26-17.46)	02 Feb 2022	7.94 (1.58-39.81)
03Jan2022	6.90 (2.44-19.52)	03 Feb 2022	7.97 (1.57-40.28)
04Jan2022	7.38 (2.56-21.29)	04 Feb 2022	8.00 (1.57-40.75)
05Jan2022	7.52 (2.53-22.34)	05 Feb 2022	8.02 (1.56-41.22)
06Jan2022	7.38 (2.39-22.79)	06 Feb 2022	8.05 (1.55-41.67)
07Jan2022	7.13 (2.21-23.01)	07 Feb 2022	8.08 (1.55-42.10)
08Jan2022	6.96 (2.07-23.35)	08 Feb 2022	8.10 (1.54-42.53)
09Jan2022	6.94 (2.00-24.01)	09 Feb 2022	8.13 (1.53-42.94)
010Jan2022	7.04 (1.94-24.97)	10 Feb 2022	8.15 (1.53-43.35)
11Jan2022	7.20 (1.98-26.08)	11 Feb 2022	8.17 (1.52-43.75)
12Jan2022	7.34 (1.98-27.14)	12 Feb 2022	8.20 (1.52-44.13)
13Jan2022	7.41 (1.96-28.00)	13 Feb 2022	8.22 (1.52-44.51)
14Jan2022	7.43 (1.92-28.66)	14 Feb 2022	8.25 (1.51-44.88)
15Jan2022	7.41 (1.88-29.21)	15 Feb 2022	8.27 (1.51-45.25)
16Jan2022	7.39 (1.83-29-76)	16 Feb 2022	8.29 (1.50-45.60)
17Jan2022	7.41 (1.80-30.37)	17 Feb 2022	8.31 (1.50-45.94)
18Jan2022	7.45 (1.78-31.07)	18 Feb 2022	8.33 (1.50-46-28)
19Jan2022	7.57 (1.77-31.82)	19 Feb 2022	8.36 (1.49-46.61)
20Jan2022	7.56 (1.75-32.55)	20 Feb 2022	8.38 (1.49-46.93)
21Jan2022	7.60 (1.74-33.23)	21 Feb 2022	8.40 (1.49-47.24)
22Jan2022	7.64 (1.72-33.65)	22 Feb 2022	8.42 (1.49-47.54)
23Jan2022	7.66 (1.70-34-42)	23 Feb 2022	8.44 (1.48-47-84)
24Jan2022	7.68 (1.68-34.98)	24 Feb 2022	8.46 (1.48-48.13)
25Jan2022	7.70 (1.67-35-54)	25 Feb 2022	8.48 (1.48-48.42)
26Jan2022	7.73 (1.65-36.11)	26 Feb 2022	8.50 (1.48-48.69)
27Jan2022	7.72 (1.64-36.68)	27 Feb 2022	8.51 (1.48-48-.96)
28Jan2022	7.80 (1.63-37.25)	28 Feb 2022	8.53 (1.48-49.23)
29Jan2022	7.83 (1.62-37.80)	01 Mar 2022	8.55 (1.47-49.48)
30Jan2022	7.86 (1.61-38.33)	-	-
31Jan2022	7.89 (1.60-38.83)	-	-
Total	468.36 (103.79-2209.88)		

CI: Confidence interval

Figure 3 displays the model's fitted number of hospital patients from 4th March 2020 and 31st December. To assess the validity of the selected model, we displayed a fitted model with real data of hospitalization between 4th March 2020 and 31st December 2021. Then we forecasted the number of admission patients with 95% confidence interval between 1st January 2022 to 1st March 2022 based on ARIMA (3, 0, 3) model. The fitted model is fully fitted to the real data (Figure 3).

The total cumulative hospital patients admission in the next 60 days' in the hospital specified for COVID-19 in Jahrom County was predicted to be around 468.36 (CI95%:103.79-2209.88) (Table 2).

Generally, the forecasted hospital patients followed a similar pattern to real hospital patients, demonstrating that ARIMA (3, 0, 3) model could predict hospitalization cases very well.

The final model's residuals distribution was normal, as shown in Figure 4. The MAPE measure was 3.18%, and the portmanteau test for white noise was (Q=53.68, P=0.072).

Discussion

The key purpose of this study was to use a time series analysis called the ARIMA model to perceive and forecast the trend of COVID-19 in southern Iran based on daily hospitalized patients available by Jahrom University of Medical Sciences. It is important to know the duration and the trend of hospitalization when an epidemic occurs.

This study showed that the ARIMA (3, 0, 3) model can be used to predict hospital admission trends during the epidemic and complement the current surveillance system to public health policy-makers for early warning and readiness of hospital systems during epidemics. This model showed that the number of daily hospitalized cases in Jahrom may increase from 5.85 (2.16-15.79) to 8.55 (1.47-49.48) in two months. By March 01, 2022, the predictable daily hospitalized cases could reach 468.36 (03.79-2209.88). Given the upper limit of the predicted cases, it is quite clear that the health authorities of this city should adopt disease prevention and control programs to prevent and control hospitalized cases.

There is an increasing application of the ARIMA model in public health, particularly COVID-19. One study used the ARIMA (8, 1, 7) model to predict daily COVID-19 cases in Bangladesh,[10] and another study used ARIMA (1, 2, 2) to predict death cases of COVID-19 worldwide.[11]

In Iran, one study applied ARIMA (1, 0, 0) and ARIMA (1, 0, 1) for predicting the incident and death cases of COVID-19.[8] In addition, other studies in Iran have used ARIMA models to predict the trend of the Covid-19 epidemic wave.[12-14]

International public health surveillance systems must enhance the quality and capacity of health systems to respond quickly to the COVID-19 pandemic. These activities depend on the knowledge of the COVID-19 occurrence in the future and the factors associated with the its occurrence. Therefore, using models is necessary to predict the incidence of COVID-19, particularly for the regions where the risk is high and such assessment is essential to make further preventive plans.

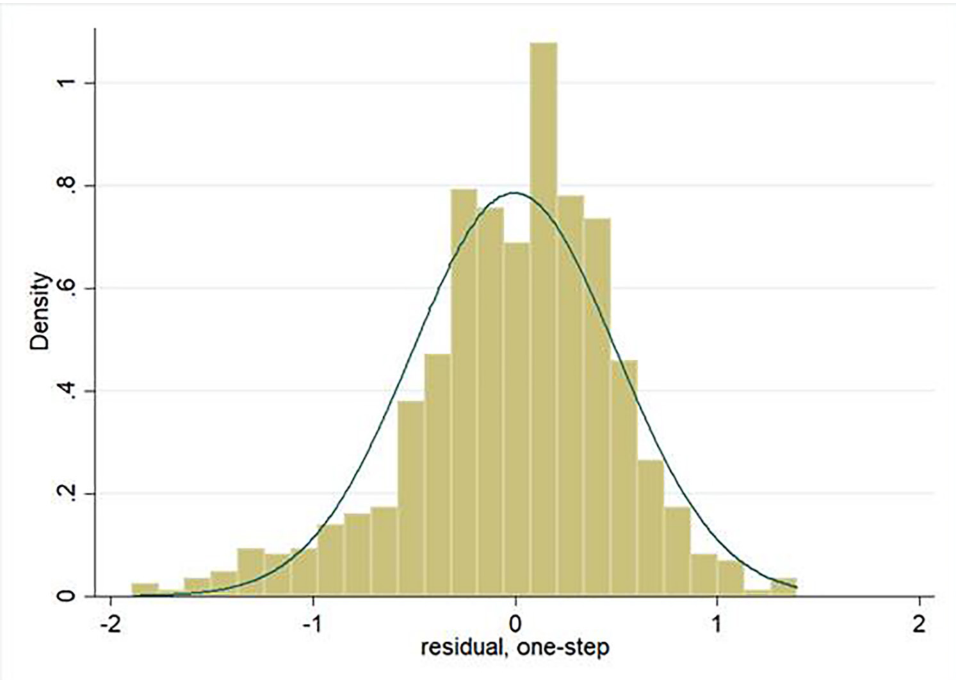


Figure 4: Histogram of the residuals in ARIMA (3, 0, 3) model.

ARIMA models are suitable for short-term prediction, so we recommend using other time series models without this limitation, like the markov switching model for long-term time series prediction in COVID-19 in future studies.[15]

Conclusion

Time series models are useful toosl for predicting the admission trend of hospitals during the epidemic. Thus, they can be used as early warning models in the readiness of hospital systems during epidemics.

Acknowledgment

The authors thank the Vice Chancellor for Research of Jahrom University of Medical Sciences for approving and financially supporting this project (Project identification code IR.JUMS.REC.1398.120). We also thank the the hospital staff for their cooperation in providing the data.

Authors' Contribution

VR, conceived and developed the idea for the article; EM and MJ prepared numerous drafts; VR contributed to the statistical analysis; MJ, VR, and EM. revised the manuscript. All the authors read and approved the final manuscript.

Conflicts of interest*:* None declared.

References

1 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. lancet 2020; 395 (10223): 497-506. doi: 10.1016/S0140-6736 (20)30183-5, PMID: 31986264 PMCID: PMC7159299.

2 Rahmanian V, Rabiee MH, Sharifi H. Case fatality rate of coronavirus disease 2019 (COVID-19) in Iran-a term of caution. Asian Pac J Trop Med 2020: 13 (7): 328-330. doi: 10.4103/1995-7645.281617.

3 WHO. Coronavirus disease 2019 (COVID-19) Situation Report –53. 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200313-sitrep-53-covid-19.pdf.

4 WHO. Coronavirus (COVID-19) Dashboard. 2 January 2022. https://covid19.who.int/.

5 Pan L, Wang J, Wang X, Ji JS, Ye D, Shen J, et al. Prevention and control of coronavirus disease 2019 (COVID-19) in public places. Environ Pollut 2022; 292: 118273. doi: 10.1016/j.envpol.2021.118273, MID: 34634404 PMCID: PMC8498926.

6 Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? lancet 2020; 395 (10228): 931-4. doi: 10.1016/S0140-6736 (20)30567-5, PMID: 32164834 PMCID: PMC7158572.

7 Rahmanian V, Bokaie S, Rahmanian K, Hosseini S, Taj Firouzeh A. Analysis of temporal trends of human brucellosis during 2013-2018 in Yazd Province to predict future trends in incidence: A time-series study using ARIMA model. Asian Pac J Trop Med 2020; 13 (4): 1-6. doi: 10.4103/1995-7645.281528.

8 Esmaeilzadeh N, Shakeri M, Esmaeilzadeh M, Rahmanian V. ARIMA models forecasting the SARS-COV-2 in the Islamic Republic of Iran. Asian Pac J Trop Med 2020; 13 (11): 521. doi: 10.4103/1995-7645.291407.

9 Dohoo IR, Martin W, Stryhn HE. Veterinary epidemiologic research. 2end ed: AVC Incorporated; 2009.

10 Kibria HB, Jyoti O, Matin A. Forecasting the spread of the third wave of COVID-19 pandemic using time series analysis in Bangladesh. Inform Med Unlocked. 2022; 28: 100815. doi: 10.1016/j.imu.2021.100815, PMID: 34961844 PMCID: PMC8694818.

11 Chaurasia V, Pal S. COVID-19 pandemic: ARIMA and regression model-based worldwide death cases predictions. SN Comput Sci 2020; 1 (5): 1-12. doi: 10.1007/s42979-020-00298-6, PMID: 33063056 PMCID: PMC7456206.

12 Moftakhar L, Seif M. The exponentially increasing rate of patients infected with COVID-19 in Iran. Arch Iran Med. 2020; 23 (4): 235-8. doi: 10.34172/aim.2020.03, PMID: 32271595.

13 Konarasinghe K. Forecasting wave-like patterns of COVID-19 daily infected cases in Iran. J New Front Healthcare Biol Scie 2021; 2 (1): 39-56.doi: 10.5281/zenodo.5071363.

14 Tran T, Pham L, Ngo Q. Forecasting epidemic spread of SARS-CoV-2 using ARIMA model (Case study: Iran). Global J Enviro Sci Manag 2020; 6 (Special Issue (Covid-19)): 1-10. doi: 10.22034/GJESM.2019.06.SI.01.

15 Rahmanian V, Bokaie S, Haghdoost A, Barouni M. Predicting cutaneous leishmaniasis using SARIMA and Markov switching models in Isfahan, Iran: A time-series study. Asian Pac J Trop Med 2021; 14 (2): 83-93. doi: 10.4103/1995-7645.306739.