

RESEARCH

Open Access



Real-time forecasting of COVID-19 spread according to protective behavior and vaccination: autoregressive integrated moving average models

Chieh Cheng¹, Wei-Ming Jiang², Byron Fan³, Yu-Chieh Cheng², Ya-Ting Hsu², Hsiao-Yu Wu², Hsiao-Han Chang¹ and Hsiao-Hui Tsou^{2,4*}

Abstract

Background Mathematical and statistical models are used to predict trends in epidemic spread and determine the effectiveness of control measures. Automatic regressive integrated moving average (ARIMA) models are used for time-series forecasting, but only few models of the 2019 coronavirus disease (COVID-19) pandemic have incorporated protective behaviors or vaccination, known to be effective for pandemic control.

Methods To improve the accuracy of prediction, we applied newly developed ARIMA models with predictors (mask wearing, avoiding going out, and vaccination) to forecast weekly COVID-19 case growth rates in Canada, France, Italy, and Israel between January 2021 and March 2022. The open-source data was sourced from the YouGov survey and Our World in Data. Prediction performance was evaluated using the root mean square error (RMSE) and the corrected Akaike information criterion (AICc).

Results A model with mask wearing and vaccination variables performed best for the pandemic period in which the Alpha and Delta viral variants were predominant (before November 2021). A model using only past case growth rates as autoregressive predictors performed best for the Omicron period (after December 2021). The models suggested that protective behaviors and vaccination are associated with the reduction of COVID-19 case growth rates, with booster vaccine coverage playing a particularly vital role during the Omicron period. For example, each unit increase in mask wearing and avoiding going out significantly reduced the case growth rate during the Alpha/Delta period in Canada (−0.81 and −0.54, respectively; both $p < 0.05$). In the Omicron period, each unit increase in the number of booster doses resulted in a significant reduction of the case growth rate in Canada (−0.03), Israel (−0.12), Italy (−0.02), and France (−0.03); all $p < 0.05$.

Conclusions The key findings of this study are incorporating behavior and vaccination as predictors led to accurate predictions and highlighted their significant role in controlling the pandemic. These models are easily interpretable and can be embedded in a “real-time” schedule with weekly data updates. They can support timely decision making about policies to control dynamically changing epidemics.

Keywords COVID-19, Forecasting, regARIMA, Vaccines, Nonpharmaceutical intervention

*Correspondence:

Hsiao-Hui Tsou
tsouhh@nhri.edu.tw

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

On March 11, 2020, the World Health Organization declared the global outbreak of 2019 coronavirus disease (COVID-19) to be a pandemic [1]. Over the course of the pandemic, variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have caused waves of COVID-19 cases all over the world, posing significant healthcare, economic, and social challenges. As of March 31, 2022, SARS-CoV-2 infection was reported to have been detected in 529.97 million individuals and to have caused the deaths of 6.32 million individuals [2].

Governments have attempted to control the spread of SARS-CoV-2 infection by mandating and promoting protective behavior [3]. Facemask use, stay-at-home orders, social distancing, and lockdowns were mandated in Italy [4], France [5], Canada [6], and Israel [7], the countries of interest in the present work. Such protective behaviors have been shown to control the COVID-19 pandemic, before and after the administration of vaccines for all SARS-CoV-2 variants [8, 9]. For example, mask wearing combined with social distancing effectively flattened the epidemic curve [10]. Vaccine administration has been documented to be integral to stopping epidemic spread, and the safety and efficacy of current vaccines have been proven [11–15].

Mathematical and statistical models enabling the real-time forecasting of the paths of COVID-19 epidemics and estimation of the impacts of protective behaviors, including vaccination, have also been key for outbreak control. Such forecasting is an essential part of public health policy, guiding the timely implementation of vector control operations and mitigating outbreak risks. It has been used successfully during pandemics/epidemics of influenza (H1N1-2009) [16], dengue fever [17, 18], and Ebola [19].

Several mathematical methods and machine learning models are applied to forecast future epidemic trends [20] and explore the COVID-19 transmission factors. The first approach is the regression model. Bo Y et al. [3] used the generalized linear mixed model to estimate the effectiveness of non-pharmaceutical interventions (NPIs) for containing COVID-19. Amuedo-Dorantes C et al. [21] examined the impact of NPIs on COVID-19 mortality rates using linear regression. Rustagi V et al. [22] performed linear regression to analyze the effects of vaccinations. Alshogran OY et al. [23] used univariable and multivariable regression to indicate that the COVID-19 fatality rate exhibited a positive correlation with the percentage of individuals aged over 60 years. However, previous studies appear to have not aligned with the assumption of independent errors required for regression analysis, which indicates that the regression model is not very suitable for time series data.

The second approach is machine learning. Several models have been developed for forecasting purposes. Yeşilkanat CM [24] used a random forest machine learning algorithm to forecast global daily cases. Ballı S [25] indicated that the support vector machine method performed well in analyzing the temporal patterns of cumulative COVID-19 data. Kumar Y et al. [26] predicted ten countries' pandemics by deep learning. Shetty RP et al. [27] used an artificial neural network to predict COVID-19 cases in one state of India. Chimmula VKR et al. [28] used long short-term memory forecast Canada transmission. Though machine learning is a popular technique, there are some weaknesses. For example, being a “block box model” lacks interpretation [29] and demands substantial volumes of data [20].

The third approach is time series forecasting models, such as autoregressive model, moving average model, autoregressive moving average model, and autoregressive integrated moving average (ARIMA) model. These models ascertain the stationary condition of the data. Alzaharani SI et al. [30] indicated ARIMA model outperformed other time series forecasting models. The ARIMA model can explain the complex autocorrelation found in datasets and be applied quickly in a wide variety of fields, including economics [31], manufacturing [32], and public health [33]. It has been used recently to predict the spread of COVID-19 in many countries [34–36]. Modified ARIMA models have been developed, with current examples including a hybrid ARIMA–wavelet-based forecasting model [37], a dynamic hybrid with a modified susceptible–exposed–infected–recovered–dead model [38], a hybrid ARIMA–discrete wavelet decomposition model [39], and a regression model with ARIMA errors (regARIMA) [40, 41] which have been applied to forecast pandemic trends and interpret additional factors as external regressors.

A review of the ARIMA model literature through February 4, 2023, suggests that only human mobility has been used to predict COVID-19 transmission before universal vaccination implementation in such models [42–46], and that vaccination was included as a predictor in only one study [47]. With increasing vaccination coverage and SARS-CoV-2 evolution, a comprehensive set of factors should be examined and vaccination and other protective behaviors should be included as predictors to improve dynamic forecasting.

With the objective of capturing the real-time dynamics of the pandemic and promptly addressing the associated challenges, we aimed to develop an innovative ARIMA model designed to provide accurate predictions while also effectively identify the critical determinants. In this longitudinal study, we used newly developed real-time ARIMA models to improve the accuracy of COVID-19

trend prediction, with the inclusion of vaccine coverage and protective behaviors as adjustment and predictive variables, respectively. We compared predictor effects on outbreak progression in the pre- and post-vaccine phases in four countries. Our models can help members of the public, policymakers, and healthcare professionals gain a deeper understanding of the COVID-19 pandemic, assess trends in the spread of viruses in real time, and determine the effectiveness of outbreak control strategies.

Method

Study areas

This study was conducted with data from France, Italy, Canada, and Israel. The first three countries were selected due to the similarity of their socioeconomic characteristics [i.e., real gross domestic product per capita] [48] and high baseline COVID-19 risk level (based on the obesity prevalence, proportion of the population aged > 65 years, and international arrivals) [49, 50]. These countries also experienced a significant first wave of COVID-19 in the first half of 2020, followed by waves in the Alpha & Delta (January 1–November 25, 2021 and Omicron (November 26, 2021–March 29, 2022) periods [2]. Moreover, these countries have different types of government that implemented various policy-driven responses, enabling the examination of a broad scope of COVID-19 control strategies [33]. Israel was chosen because of its leading vaccination promotion and implementation efficiency [51].

Research design

Figure 1 shows the research design. We used three primary datasets: epidemiological variables, protective behavior, and vaccine coverage. The epidemiological variables [COVID-19 case growth rate, number of intensive care unit (ICU) patients with COVID-19 per million, and number of deaths due to COVID-19 per million] served as dependent variables. Protective behaviors (mask wearing and avoiding going out) and vaccine coverage (proportions of the total national population vaccinated fully and having received booster doses) served as independent variables. We examined and compared the Alpha & Delta and Omicron periods.

In most ARIMA studies, large proportions of data are allocated to training and model building. Data series are categorized as in-sample and out-of-sample, at ratios such as 70:30 [52, 53], 80:20 [54], and 90:10 [55]. In this study, we allocated 70% of the data for training and 30% for testing. We also categorized data as in-sample and out-of-sample at ratios of 80:20 and 90:10 for sensitivity analyses to evaluate the robustness of our models in different ratios. We developed six regARIMA models with lags on protective behaviors and vaccine coverage, which profoundly affected the development of the dependent

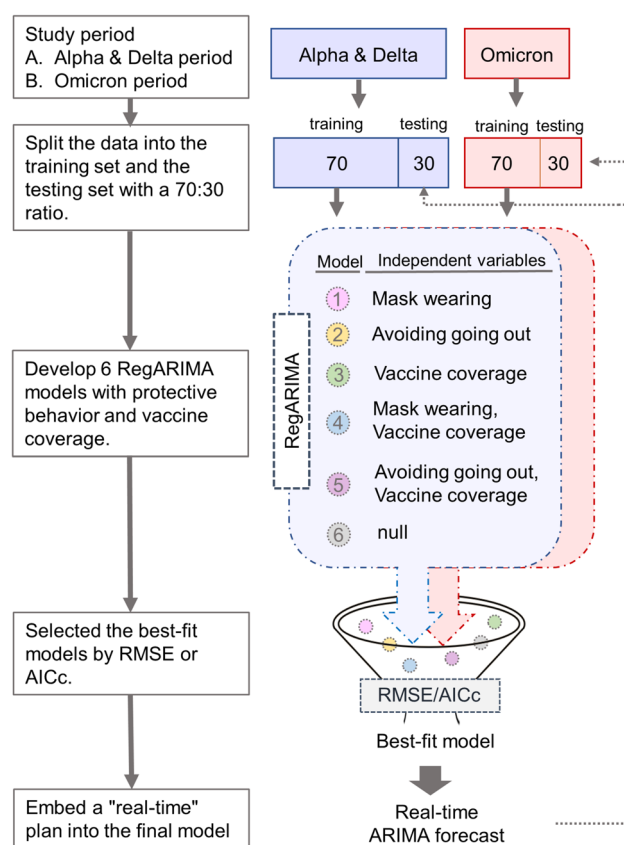


Fig. 1 Flowchart of Reg-ARIMA model development for real-time COVID-19 forecasting

epidemiological variables [56]. The best-fitting models for each period were selected according to the root mean square error (RMSE) and corrected Akaike information criterion (AICc) and applied for the real-time forecasting of COVID-19 outbreaks and conditions.

Data sources

The period of analysis was January 2021–March 2022. We retrieved epidemiological and vaccine coverage data from Our World in Data [2]. The database is public, updated in real time, and contains epidemic information collected from governments around the world and news agencies. The data on protective behaviors from the Imperial College of London's YouGov COVID-19 Behavior Tracker Data Hub [57], which is a global, anonymized, longitudinal, and weekly web-based survey. YouGov ensures the data is representative through randomly selected and stratified samples. The small deviations were corrected by post-stratification weights [58]. YouGov data has been extensively utilized in numerous studies to capture perceptions and responses of individuals toward COVID-19 [59–62]. The

utilization of weekly data is justified by multiple factors. Firstly, it enables the incorporation of the weekend effect and reduces the need for excessive adjustments to COVID-19 case numbers, making it a preferable option compared to daily data [63]. Secondly, the survey data is updated weekly, ensuring consistency in the chosen time unit. Third, given that many government policies are made on a weekly basis [64]. Taking these factors into account, emphasizing a weekly frequency becomes crucial for effective forecasting.

Dependent variables

The COVID-19 case growth rate, rather than the number of new cases per million, was selected because it better reflects epidemic trends and meets the stationary condition for time-dependent trends, thereby enabling more accurate prediction [42, 43]. Furthermore, the variance of the series is stabilized by this taking of a logarithmic approach [63]. We calculated the log weekly case growth rate (Y_t) using the method of Karaiyanov et al. [6]:

$$Y_t = \Delta \ln(C_t) = \ln C_t - \ln C_{t-1},$$

Where C_t is the number of cases in week t ; C_{t-1} is the number of cases in the last week. The case growth rates in the four countries are shown in Fig. 2.

The two other variables are the weekly numbers of ICU patients with COVID-19 and deaths due to COVID-19 per million which were calculated as the

sum of the 7-day numbers of ICU patients and deaths, respectively.

Protective behaviors

Protective behaviors were measured by asking the question “Thinking about the last 7 days, how often have you taken the following measures to protect yourself or others from coronavirus (COVID-19)? As a reminder, please exclude any measures that you have already taken for reasons other than coronavirus (COVID-19)” [65]. For mask wearing and avoiding going out, we used the response statements “Worn a face mask outside your home (e.g., when on public transport, going to a supermarket, going to the main road)” and “Avoided going out in general,” respectively. Possible responses were “Always,” “Frequently,” “Sometimes,” “Rarely,” and “Not at all,” which we back-coded so that higher scores reflected greater adherence to protective behavior policies [66], taking the values of 4, 3, 2, 1, and 0, respectively. We calculated 7-day averages to obtain weekly data.

Vaccine coverage

The proportion of the population vaccinated fully, defined following Tsou et al. [50] as the receipt of two vaccine doses, was calculated as:

$$V_t = \frac{V_{1t}}{2P_t},$$

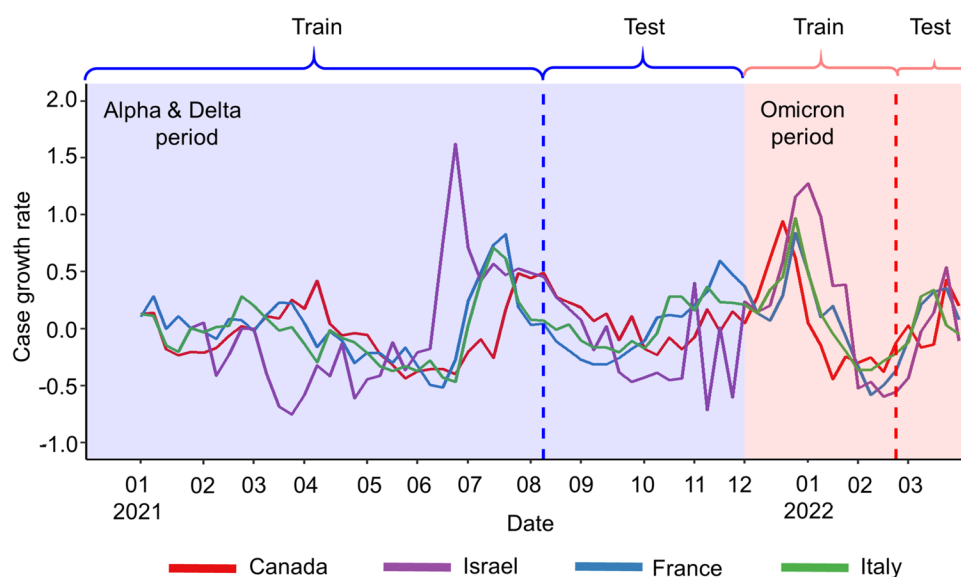


Fig. 2 Weekly case growth rates. The Alpha & Delta (blue-shaded area) and the Omicron period (pink-shaded area) started from January 2021 to November 2021 and December 2021 to March 2022, respectively. Train-test ratio was 70:30, split by dash line

Where V_t is the total number of doses that the country administered and P_t is the country's total population in week t .

The booster dose variable was used for adjustment during the Omicron period due to the flattening of the cumulative fully vaccinated population curve (S1 Fig.) [67]. It was defined as the ratio of the total number of booster COVID-19 vaccine doses administered per 100 people to the country's total population.

regARIMA model

The ARIMA (p,d,q) mathematical model was developed by Box and Jenkins. The autoregressive (AR) part uses past values to predict future values. The integration (I) uses the difference to make the time series stationarity. The moving average (MA) part sums the past error value. The p, d, and q are the orders of the autoregressive, integrated, and moving average. There are some steps for model selection. First, Box-Cox transformation is used to stabilize the variance. Second, the unit-root test is used to sure the time series stationarity. Third, autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to select the order. Then, try the models we select and use AICc and RMSE to choose the best-performed model. Last, check the residuals are white noise [68].

The regARIMA time-series analysis model used in this study combines the benefits of ARIMA and linear regression, capturing autocorrelations in the data and enabling the inclusion of exogenous variables to improve forecast performance [40]. Additionally, its output is easier to interpret than is that of ARIMA models with explanatory variables (ARIMAX), which belong to the same hybrid model family [68], and regARIMA model use has been proposed for complex research objects [69]. Thus, regARIMA is the most appropriate model for our study.

In models 1–3, mask wearing, avoiding going out, and vaccine coverage, respectively, served as independent variables. Models 4 and 5 were adjusted for vaccine coverage and included mask wearing and avoiding going out, respectively. Model 6 was a base model developed for comparison. The model formulae were:

$$Y_t = C + n_t + \begin{cases} \beta_1 M_{t-l} & \text{model (1)} \\ \beta_1 G_{t-l} & \text{(2)} \\ \beta_1 V_{(f,b)t-l} & \text{(3)} \\ \beta_1 M_{t-l} + \beta_2 V_{(f,b)t-l} & \text{(4)} \\ \beta_1 G_{t-l} + \beta_2 V_{(f,b)t-l} & \text{(5)} \end{cases},$$

$$n_t = \varphi_1 n_{t-1} + \dots + \varphi_p n_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

Where C is a constant, n_t represents ARIMA errors, β is estimates of the coefficient, and M_{t-l} , G_{t-l} , $V_{(f,b)t-l}$, and

$V_{(b)t-l}$ represent mask wearing, avoiding going out, full vaccination, and booster doses in $t-l$ week, respectively, (l is the time lag, $l=2$ for the case growth rate and number of ICU patients per million and $l=4$ for the number of deaths per million) [6, 70]. When incorporating the predictors (protective behavior and vaccination data), a comprehensive literature review is conducted as a preliminary step to understand the time lag between the predictors and the outcomes. The notation φ_i is the parameter of autoregression for $i=1, \dots, p$, where p is the number of lag observations in the model, also known as the lag order. Moreover, θ_j is the parameter of moving average for $j=1, \dots, q$, where q is the order of the moving average. The notation ε_t is a random error or residual term for the t th week. Models 1 to 5 listed in Eq. (1)–(5) are regARIMA, while model 6 is an ARIMA model that estimates values using autoregressive without any predictors.

Evaluation of model performance

We used the RMSE as a measure of the model's predictive performance in the Alpha & Delta period, as in several previous studies [17, 55, 71, 72]:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n [Y_t - \hat{Y}_t]^2},$$

where Y_t is the actual number of dependent variables at time t , \hat{Y}_t is the forecast dependent variable value, and n is the sample size. Lower values reflect more accurate forecasting.

We used the AICc, recommended for the analysis of short time series [40], to evaluate model performance in the Omicron period because the sample size was < 20 . The AICc was defined as:

$$\text{AICc} = -2\log(L) + 2k + \frac{2k(k+1)}{n-k-1},$$

Where L is the likelihood function, k is the total number of parameters, and n is the sample size. Lower values reflect better model performance.

Real-time forecasting of the case growth rate

For real-time forecasting with the best-fitting model, the data were updated every 1–2 weeks and the model was refit with the newly updated dataset. The model was run using the auto.arima function of the "forecast" package in R 4.0.2 [73].

Results

Model estimates

Figure 3A and B displays the relationship between the case growth rates and protective behaviors during the

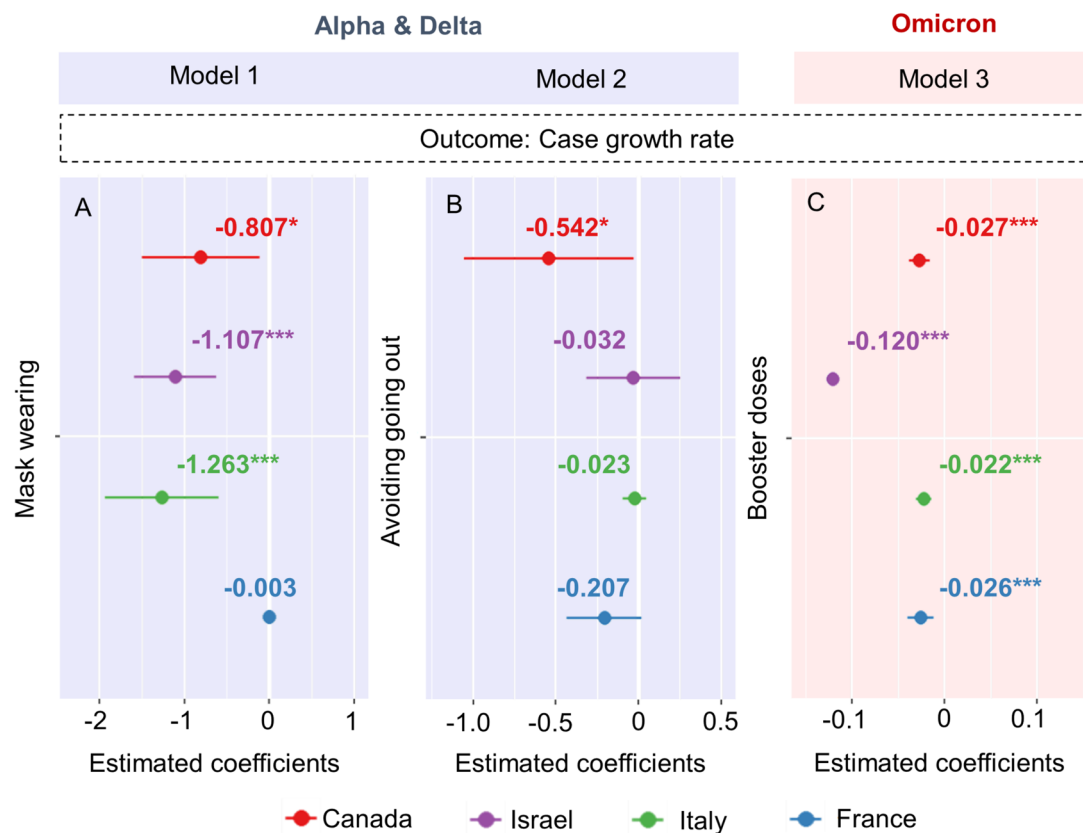


Fig. 3 Relationships between predictors and the case growth rate in the Alpha & Delta (A and B for Model 1 and Model 2, respectively) and Omicron (C for Model 3) periods. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Alpha & Delta period. The estimates produced by all models with and without adjustment for vaccine coverage are provided in S1 Table. S1 Table also shows the p , d , and q obtained from ARIMA models. All coefficient for the protective behaviors were negative, indicating that the case growth rate decreased with adherence to protective behavior policies. For example, each unit increase in mask wearing and avoiding going out significantly reduced the case growth rate in Canada (-0.807 and -0.542 , respectively; both $p < 0.05$; Fig. 3A). It is equivalent to a 2.24- and 1.72- fold reduction in the number of infected people compared to the original number during that week, respectively. In the Omicron period, each unit increase in the number of booster doses resulted in a significant reduction of the case growth rate in Canada (-0.027), Israel (-0.120), Italy (-0.022), and France (-0.026 ; all $p < 0.05$; Fig. 3C, S2 Table). It is equivalent to a 1.03-, 1.13-, 1.02-, and 1.03-fold reduction in the number of infected people compared to the original number during that week, respectively. S2 Table also shows the orders of p , d , and q obtained from ARIMA models.

Figure 4 shows the effects of full vaccination on the epidemiological variables in the Alpha & Delta period.

Full vaccination did not reduce the weekly case growth rate, but significantly reduced the numbers of ICU patients with COVID-19 and deaths due to COVID-19 per million in most countries. S1 Fig. shows the effects of full vaccination on the epidemiological variables in the Omicron period.

Model selection

Model performance is quantified in Table 1. RMSE values for the Alpha & Delta period in almost all of the countries were lowest with model 4, indicating that this model had the best fit. According to this model, which was adjusted for full vaccination, each point increase in mask wearing was associated with a reduction in the case growth rate in Italy (-1.268 , $p < 0.05$), Israel (-1.101 , $p < 0.05$), Canada (-0.701 , $p > 0.05$), and France (-0.020 , $p > 0.05$). For the Omicron period, AICc values were lowest with model 6, which was thus designated the best-fitting model.

The results were obtained with the data split at a 70:30 ratio. Similar results were obtained with an 80:20 and a 90:10 split (S3 – S6 Table).

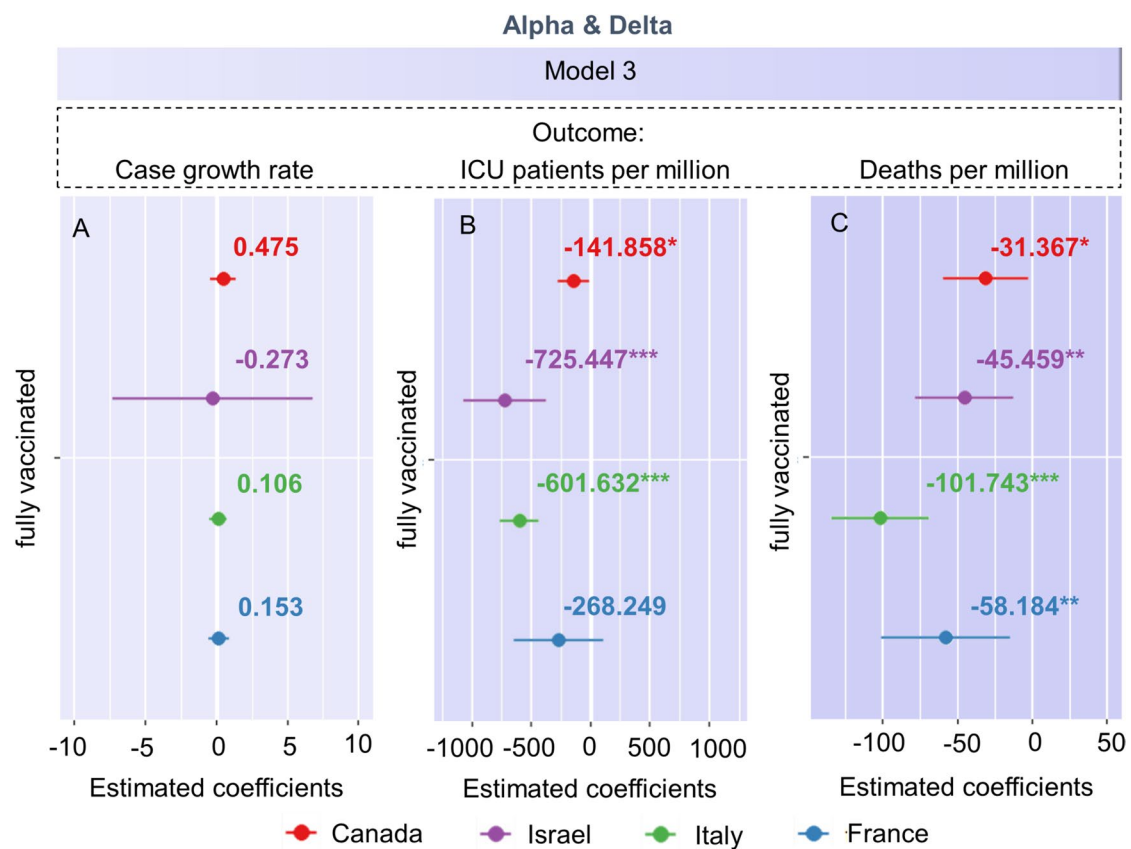


Fig. 4 Effects of full vaccination. Effects of full vaccination on the (A) case growth rate, (B) number of ICU patients with COVID-19 per million, and (C) number of deaths due to COVID-19 per million in the Alpha & Delta period. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1 Predictive performance of the regARIMA models

Outcome: weekly case growth rate							
period	country	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Alpha & Delta	Canada						
	RMSE	0.1369	0.1400	0.1447	0.1351	0.1418	0.1479
	Israel						
	RMSE	0.2688	0.3448	0.3613	0.2687	0.3437	0.3614
	Italy						
	RMSE	0.1391	0.1520	0.1527	0.1391	0.1498	0.1530
Omicron	France						
	RMSE	0.1646	0.1583	0.1643	0.1635	0.1624	0.1647
	Canada						
	AICc	12.26	8.07	11.04	10.47	12.71	5.55
	Israel						
	AICc	15.14	21.19	18.51	18.59	19.99	13.66
	Italy						
	AICc	3.09	5.99	8.06	7.71	9.6	4.68
	France						
	AICc	9.5	9.13	13.08	12.67	13.57	6.35

RMSE Root mean square error

AICc Corrected Akaike information criterion

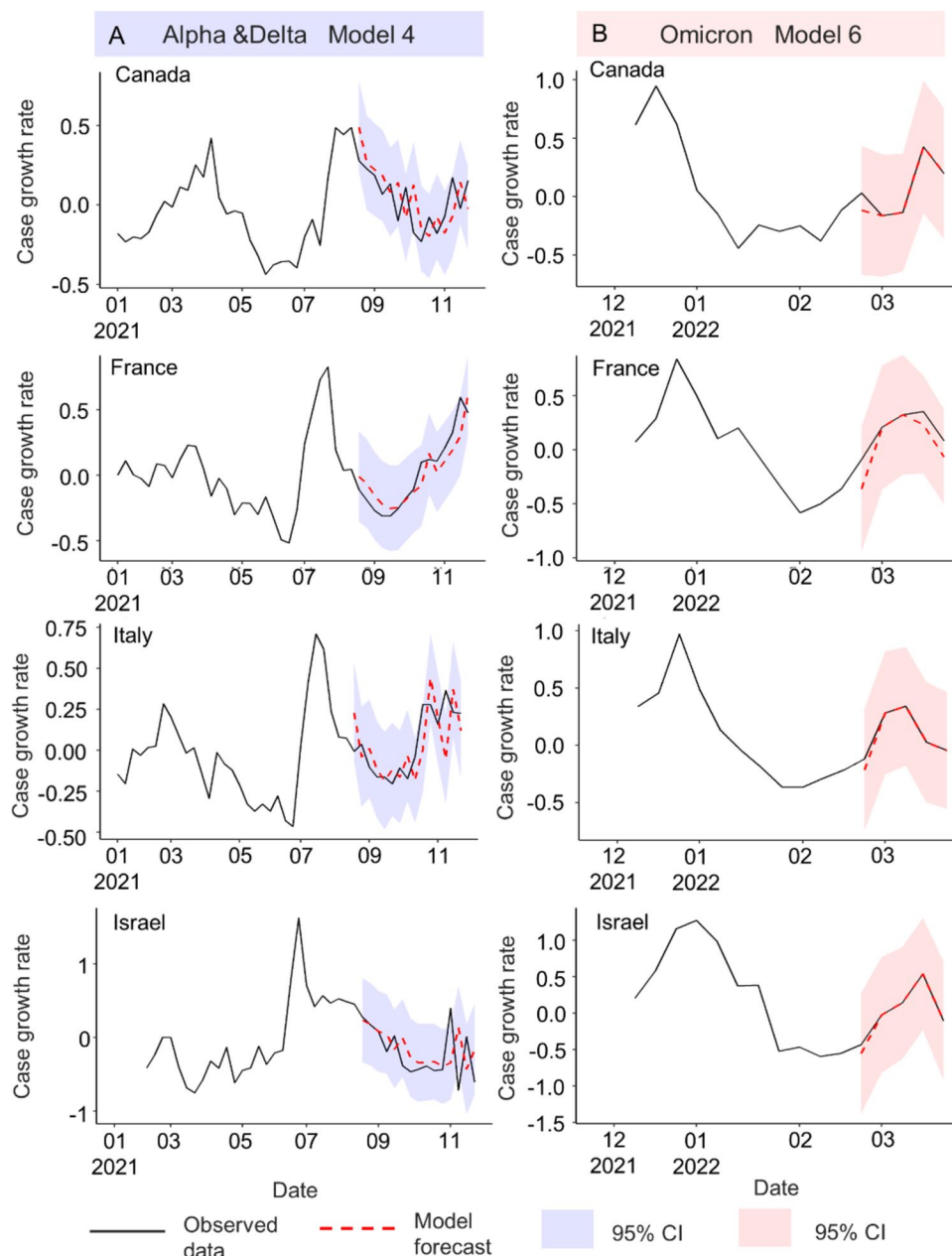


Fig. 5 Real-time forecasts of the case growth rate in the Alpha & Delta (A) and Omicron (B) periods

Real-time forecasts

The results of real-time forecasting of the case growth rates in the Alpha & Delta and Omicron periods using models 4 and 6, respectively, are shown in Fig. 5. The starting dates for prediction (first dates in the testing sets) were November 29, 2021 and February 28, 2022, respectively. The forecast obtained with model 4 was slightly delayed, whereas that obtained with model 6 was synchronized with the observed data. Overall, however, the models enabled accurate prediction.

Discussion

In this study, we developed and tested novel regARIMA models that incorporate self-reported protective behaviors and vaccine coverage for the forecasting of COVID-19 trends. The models showed good accuracy and can play important roles in the prevention of epidemic spread.

During the Alpha & Delta period of the COVID-19 pandemic, our models 1 and 2 capture the importance of mask wearing and avoiding going out during this period,

which both reduce the case growth rate. The results can be supported by other research. Chu et al. [74] indicated that face mask use can reduce the infection risk by a systematic meta-analysis. McGrail et al. [75] find that the government implementing social distancing policies can reduce the COVID-19 spread rate through the statistical modelling. Furthermore, the risks of hospitalization and mortality due to COVID-19 were more severe than those of seasonal influenza [76, 77], thus increasing people's willingness to engage in protective behavior. The high protective behavior frequency also makes us capture the effect of mask wearing and avoiding going out.

The model suggests that vaccination coverage has a limited impact on the case growth rate. We believe that vaccine hesitancy plays a significant role in the selected countries. For example, a global study indicates Italy and France's vaccine acceptance rates are lower than 60% [78]. However, full vaccination significantly reduced the number of ICU patients and deaths attributable to COVID-19, in agreement with previous findings [79–81]. Bajema et al. [79] studied the effectiveness of the vaccine, they reported that the vaccine effectiveness of prevent COVID-19 hospitalization was 80% and 95% in adults aged ≥ 65 and 18–64 respectively. Victora et al. [81] investigated the vaccine on the association of deaths. The results indicated that vaccination significantly declines mortality in elderly people.

In the Omicron period, we observed that high booster dose rates significantly reduced the case growth rate. These results may be attributable in part to governments' strong promotion of vaccine boosters due to the highly infectious nature of the Omicron variant early in this period. Full vaccination coverage in the countries of interest in this study exceeded 90% in January 2022.

In contrast to the Alpha & Delta period, it was observed that the efficacy of mask-wearing and adherence to stay-at-home orders in containing the pandemic has dwindled during the Omicron stage. This study conjectures that some reasons may cause the result. First, the Omicron variant of SARS-CoV-2 is more infectious than the Delta variant but is associated with lower hospitalization and mortality rates [82]. Second, according to the risk compensation hypothesis, the protective behaviors will decrease in vaccinated individuals. For example, Liang et al. [83] find that population mobility increases as vaccine coverage increases. In light of the factors mentioned above and the prolonged duration of the pandemic, compliance fatigue has become increasingly prevalent among individuals, leading to the growing desire to restore pre-pandemic norms.

Correspondingly, in late January 2022, the governments of France [84] and Canada [85] pledged to begin coexisting with the pandemic, Italy [86] stepped up its rollout

of vaccinations and vaccine passports, and Israel [87] ended a long-running state of emergency due to the pandemic. Living with COVID-19 requires flexibility in policy implementation to protect societies, economies, and individuals' mental health while containing the spread of the disease.

This study also presented dynamic epidemic time series forecasting tools to analyze COVID-19. Our model can accurately predict and interpret external variables. Overall, all models' performances are well. We observed that certain exogenous regressors did not demonstrate better performance in comparison to the univariate model. This result is consistent with the findings of Nassiri et al. [44] who suggest accounting other variables may enhance the predictive capability.

Our models have many advantages and can help government agencies and public health professionals to fight COVID-19. First, they are easy to understand, intuitive, and straightforward. Second, their real-time nature is more suitable for the prevention of epidemics, which are characterized by constant dynamic changes. The models support decision-making about policies as new variants or outbreaks emerge with easy scenario simulation.

The ARIMA model has been widely utilized and demonstrated remarkable success in forecasting the progression of infectious diseases [88, 89]. Building upon this, the regARIMA model, which we have proposed by integrating exogenous variables, not only exhibits improved accuracy but also holds great promise for further applications in the relevant fields. Moreover, despite the number of countries included in our study being limited due to the considerable time and effort required for data access and processing, we made diligent efforts to comprehensively account for diversity in geography, culture, and COVID policy. Consequently, we are confident that our findings possess significant potential for generalizability and offer valuable insights, particularly within high socio-economic settings.

This study has some limitations. Given that the data on protective behaviours relied on self-report measures, it is conceivable that there may be accuracy biases stemming from social desirability or imperfect memory. Moreover, YouGov claims that individuals experiencing severe symptoms, those who have been hospitalized, and other challenging-to-reach groups may be inadequately represented within the sample. Furthermore, it is important to acknowledge that the prediction provided by this model is specifically tailored to SARS-CoV-2 variants. Given the constantly evolving nature of respiratory viruses and their distinct characteristics, it is crucial to involve incorporating diverse exogenous variables or exploring alternative modelling methodologies accordingly to better encapsulate the intricate dynamics and complexities

inherent in different circumstances for giving more reliable predictions. In addition, the algorithm of fully vaccination is constrained by the types of data it has access to, which may not necessarily fully represent real-world conditions accurately.

Conclusions

Our real-time model incorporating human behavior (wearing masks and avoiding going outside) and vaccination (fully vaccinated and given booster doses) variables performed well in two periods of the COVID-19 pandemic (Alpha & Delta and Omicron). Furthermore, by leveraging the model, our findings assisted in identifying and quantifying the significant determinants, such as human behaviour and vaccination, that play a crucial role in containing pandemics. Our development model can provide a reference for public health departments to formulate policies to deal with new variants of COVID-19 or emerging infectious diseases.

Abbreviations

ARIMA	Automatic regressive integrated moving average
COVID-19	2019 coronavirus disease
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
regARIMA	Regression model with ARIMA errors
ICU	Intensive care unit
RMSE	Root mean square error
AICc	Corrected Akaike information criterion
AR	Autoregressive
I	Integration
MA	Moving average
ACF	Autocorrelation function
PACF	Partial autocorrelation function
ARIMAX	ARIMA models with explanatory variables

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-023-16419-8>.

Additional file 1: S1 Fig. Effects of full vaccination on the (A) case growth rate, (B) number of ICU patients with COVID-19 per million, and (C) number of deaths due to COVID-19 per million in the Omicron period. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **S2 Fig.** The Alpha & Delta (blue shaded area) and the Omicron periods (pink shaded area) based on the vaccine coverage for (a) full vaccination and (b) booster doses in Canada (red line), France (green line), Israel (blue line), and Italy (purple line). **S1 Appendix.** ARIMA model. **S2 Appendix.** case growth rate and first differences. **S1 Table.** Coefficient estimates of all models with/without adjusting for vaccine coverage in the Alpha & Delta period. **S2 Table.** Coefficient estimates of all models with/without adjusting by vaccine coverage in the Omicron period. (split at a 70:30 ratio). **S3 Table.** Coefficient estimates of all models with/without adjusting for vaccine coverage in the Alpha & Delta period (split at an 80:20 ratio). **S4 Table.** Coefficient estimates of all models with/without adjusting for vaccine coverage in the Alpha & Delta period (split at a 90:10 ratio). **S5 Table.** Coefficient estimates of all models with/without adjusting by vaccine coverage in the Omicron period. (split at an 80:20 ratio). **S6 Table.** Coefficient estimates of all models with/without adjusting by vaccine coverage in the Omicron period (split at a 90:10 ratio).

Acknowledgments

Chieh Cheng carried out her thesis research under the auspices of the Graduate Program of Biotechnology in Medicine, National Tsing Hua University and National Health Research Institutes. The authors thank Dr. Cheng-Hsiung Chiang of the National Health Research Institutes, Taiwan, for his insight and expertise on the issue of model overfitting. The authors thank Mr. Yi-Kai Chen, Ms. Fang-Jing Lee, and Mr. Yung-An Jang of the Institute of Population Health Sciences, the National Health Research Institutes, Taiwan, for their help with data management and language improvement. Their expertise greatly assisted in the success of this research.

Authors' contributions

All authors contribute significantly to the conduct of this study and have approved the submission of this paper for review by your Journal in its present form. The authors' contribution is listed below: C.C. and H.H.T. contributed significantly to the study's conception and design. C.C. and H.H.T. contributed to the drafting of the manuscript. C.C., Y.T.H., and H.Y.W. contributed to data preparation. C.C. and H.H.T. contributed to data validation. C.C., W.M.J., Y.C.C., and H.H.T. provided statistical analysis. C.C., W.M.J., B.F. H.H.C., and H.H.T. contributed to the interpretation of the data. C.C., W.M.J., B.F. Y.C.C., Y.T.H., H.Y.W., H.H.C., and H.H.T. provided revisions of the manuscript. H.Y., W.J.C., H.H.C., and H.H.T. provided administrative, technical, and material support. The research was supervised by H.H.T.

Funding

This study was supported by grants PH-111-PP-02, PH-112-PP-02, PH-112-GP-02, and PH-112-GP-12 from the National Health Research Institutes, a nonprofit foundation dedicated to medical research and improved healthcare in Taiwan.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. These raw data are openly available for public access at YouGov data [<https://github.com/YouGov-Data/covid-19-tracker/tree/fe0142dbf23b30755df537bacce5c6d687c52fdd>] and Our World in Data [<https://github.com/owid/covid-19-data/tree/master/public/data>]. R code is available at the GitHub repository [https://github.com/ChiehCheng/ARIMA_forecast].

Declarations

Ethics approval and consent to participate

The institutional review board of the National Health Research Institutes approved this study (EC1101202-E). Human participants or live animals are not involved in this study. The data we used in this study are from open data source such as "Our World in Data" and "YouGov data", which are publicly available and de-identified, so ethics approval and informed consent were not needed. The methods utilized in this study adhered to the relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

All authors declare no conflicts of interest, including employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications, and travel grants.

Author details

¹Department of Life Science & Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan. ²Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan, Miaoli County 350, Taiwan. ³Brown University, RI, Providence, USA. ⁴Graduate Institute of Biostatistics, College of Public Health, China Medical University, Taichung, Taiwan.

Received: 4 May 2023 Accepted: 29 July 2023

Published online: 08 August 2023

References

- Coronavirus disease (COVID-19) pandemic. Cited 2023 Jan 31. Available from: <https://www.who.int/europe/emergencies/situations/covid-19>.
- Mathieu E, Ritchie H, Rod  s-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus Pandemic (COVID-19). Our World Data. 2020. Cited 2023 Feb 9; Available from: <https://ourworldindata.org/coronavirus>.
- Bo Y, Guo C, Lin C, Zeng Y, Li HB, Zhang Y, et al. Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. *Int J Infect Dis*. 2021;1(102):247–53.
- Ferrante P. The first 2 years of COVID-19 in Italy: Incidence, lethality, and health policies. *Front Public Health*. 2022;10. Cited 2023 Feb 9. Available from: <https://www.frontiersin.org/articles/10.3389/fpubh.2022.986743/full>.
- Desson Z, Weller E, McMeekin P, Ammi M. An analysis of the policy responses to the COVID-19 pandemic in France, Belgium, and Canada. *Health Policy Technol*. 2020;9(4):430–46.
- Karaivanov Alexander, Lu SE, Shigeoka H, Chen C, Pamplona S. Face masks, public policies and slowing the spread of COVID-19: evidence from Canada. *J Health Econ*. 2021;78:102475.
- Utamura M, Koizumi M, Kirikami S. Novel deterministic epidemic model considering mass vaccination and lockdown against coronavirus disease 2019 spread in Israel: a numerical study. *Biol Methods Protoc*. 2022;7(1):bpac023.
- Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gaven  ciak T, et al. Inferring the effectiveness of government interventions against COVID-19. *Science*. 2021;371(6531):eabd9338.
- Li H, Wang L, Zhang M, Lu Y, Wang W. Effects of vaccination and non-pharmaceutical interventions and their lag times on the COVID-19 pandemic: comparison of eight countries. *PLoS Negl Trop Dis*. 2022;16(1):e0010101 (Kamel MG, editor).
- Li T, Liu Y, Li M, Qian X, Dai SY. Mask or no mask for COVID-19: a public health and market study. *PLoS One*. 2020;15(8): e0237691.
- Levin-Zamir D. Communication, health literacy and a systems approach for mitigating the COVID-19 pandemic: the case for massive vaccine roll-out in Israel. *J Health Commun*. 2020;25(10):816–8.
- Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med*. 2020;383(27):2603–15.
- Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *Lancet Lond Engl*. 2021;397(10287):1819–29.
- Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet*. 2021;397(10269):99–111.
- Harder T, Koch J, Vygen-Bonnet S, K  lper-Schiek W, Pilic A, Reda S, et al. Efficacy and effectiveness of COVID-19 vaccines against SARS-CoV-2 infection: interim results of a living systematic review, 1 January to 14 May 2021. *Eurosurveillance*. 2021;26(28):2100563.
- Tizzoni M, Bajardi P, Poletto C, Ramasco JJ, Balcan D, Gon  alves B, et al. Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm. *BMC Med*. 2012;10(1):165.
- Cheng YC, Lee FJ, Hsu YT, Slud EV, Hsiung CA, Chen CH, et al. Real-time dengue forecast for outbreak alerts in Southern Taiwan. *PLoS Negl Trop Dis*. 2020;14(7):e0008434 (Weinberger DM, editor).
- Shi Y, Liu X, Kok SY, Rajarethinam J, Liang S, Yap G, et al. Three-month real-time dengue forecast models: an early warning system for outbreak alerts and policy decision support in Singapore. *Environ Health Perspect*. 2016;124(9):1369–75.
- Chowell G, Luo R, Sun K, Roosa K, Tariq A, Viboud C. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. *Epidemics*. 2020;1(30):100379.
- Rahimi I, Chen F, Gandomi AH. A review on COVID-19 forecasting models. *Neural Comput Appl*. 2021; Cited 2023 Jun 20. <https://doi.org/10.1007/s00521-020-05626-8>.
- Amuedo-Dorantes C, Kaushal N, Muchow AN. Timing of social distancing policies and COVID-19 mortality: county-level evidence from the U.S. *J Popul Econ*. 2021;34(4):1445–72.
- Rustagi V, Bajaj M, Tanvi, Singh P, Aggarwal R, AlAjmi MF, et al. Analyzing the Effect of Vaccination Over COVID Cases and Deaths in Asian Countries Using Machine Learning Models. *Front Cell Infect Microbiol*. 2022;11. Cited 2023 Jun 20. Available from: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.806265..>
- Alshogran OY, Altawalbeh SM, Al-Azzam SI, Karasneh R. Predictors of Covid-19 case fatality rate: an ecological study. *Ann Med Surg*. 2021;1(65):102319.
- Ye  ilkana  t CM. Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. *Chaos Solitons Fractals*. 2020;1(140):110210.
- Balli S. Data analysis of Covid-19 pandemic and short-term cumulative case forecasting using machine learning time series methods. *Chaos Solitons Fractals*. 2021;1(142):110512.
- Kumar Y, Koul A, Kaur S, Hu YC. Machine learning and deep learning based time series prediction and forecasting of ten nations' COVID-19 pandemic. *SN Comput Sci*. 2022;4(1):91.
- Shetty RP, Pai PS. Forecasting of COVID 19 cases in Karnataka State using Artificial Neural Network (ANN). *J Inst Eng India Ser B*. 2021;102(6):1201–11.
- Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos Solitons Fractals*. 2020;1(135):109864.
- Jalalkamali A, Moradi M, Moradi N. Application of several artificial intelligence models and ARIMAX model for forecasting drought using the Standardized Precipitation Index. *Int J Environ Sci Technol*. 2015;12(4):1201–10.
- Alzahrani SI, Aljamaan IA, Al-Fakh EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *J Infect Public Health*. 2020;13(7):914–9.
- Challa ML, Malepati V, Kolu SNR. S&P BSE Sensex and S&P BSE IT return forecasting using ARIMA. *Financ Innov*. 2020;6(1):47.
- Sen P, Roy M, Pal P. Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy*. 2016;1(116):1031–8.
- Anwar MY, Lewnard JA, Parikh S, Pitzer VE. Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. *Malar J*. 2016;15(1):566.
- Fang L, Wang D, Pan G. Analysis and estimation of COVID-19 spreading in Russia based on ARIMA model. *Sn Compr Clin Med*. 2020;2(12):2521–7.
- Ilie OD, Cojocariu RO, Ciobica A, Timofte SI, Mavroudis I, Doroftei B. Forecasting the spreading of COVID-19 across Nine countries from Europe, Asia, and the American continents using the ARIMA Models. *Microorganisms*. 2020;8(8):1158.
- To  a G, Atalay B, Toksari MD. COVID-19 prevalence forecasting using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN): Case of Turkey. *J Infect Public Health*. 2021;14(7):811–6.
- Chakraborty T, Ghosh I. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis. *Chaos Solitons Fractals*. 2020;1(135):109850.
- Ala  raj M, Majdalawieh M, Nizamuddin N. Modeling and forecasting of COVID-19 using a hybrid dynamic model based on SEIRD with ARIMA corrections. *Infect Dis Model*. 2021;6:98–111.
- Singh S, Parmar KS, Kumar J, Makkhan SJS. Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos Solitons Fractals*. 2020;1(135):109866.
- Rob J Hyndman, George Athanasopoulos. Forecasting: Principles and Practice. 2nd edition. Melbourne, Australia: OTexts; 2018. Cited 2022 Dec 23. Available from: <https://otexts.com/fpp2>.
- Sun W, Schm  cker JD, Nakao S. Restrictive and stimulative impacts of COVID-19 policies on activity trends: a case study of Kyoto. *Transp Res Interdiscip Perspect*. 2022;13:100551.
- Garc  a-Cremades S, Morales-Garc  a J, Hern  ndez-Sanjaime R, Mart  nez-Esp   a R, Bueno-Crespo A, Hern  ndez-Orallo E, et al. Improving prediction of COVID-19 evolution by fusing epidemiological and mobility data. *Sci Rep*. 2021;11(1):15173.
- da Silva TT, Francisquini R, Nascimento MCV. Meteorological and human mobility data on predicting COVID-19 cases by a novel hybrid

- decomposition method with anomaly detection analysis: a case study in the capitals of Brazil. *Expert Syst Appl.* 2021;15(182):115190.
44. Nassiri H, Mohammadpour SI, Dahaghin M. How do the smart travel ban policy and intercity travel pattern affect COVID-19 trends? Lessons learned from Iran. *PLOS One.* 2022;17(10):e0276276.
 45. Paternina-Caicedo A, Alvis-Guzmán N, Dueñas C, Narvaez J, Smith AD, De la Hoz-Restrepo F. Impact of mobility restrictions on the dynamics of transmission of COVID-19 in Colombian cities. *Int Health.* 2022;14(3):332–5.
 46. Ahdika A, Primandari AH, Adlin FN. Considering the temporal interdependence of human mobility and COVID-19 concerning Indonesia's large-scale social distancing policies. *Qual Quant.* 2022. <https://doi.org/10.1007/s11135-022-01497-4>. Cited 2023 Feb 7.
 47. Masti V, Veerendranath V, Kumar HM. Multivariate Covid-19 Forecasting with Vaccinations as a factor: the case of India and USA. In: 2022 IEEE Region 10 Symposium (TENSymp). 2022. p. 1–6.
 48. The World Factbook - The World Factbook. Cited 2023 Feb 20. Available from: <https://www.cia.gov/the-world-factbook/>.
 49. Economist Intelligence Unit. How well have OECD countries responded to the coronavirus crisis? . Cited 2023 Feb 20. Available from: <https://www.eiu.com/n/campaigns/oecd-countries-responded-to-the-coronavirus-crisis/>.
 50. Tsou HH, Kuo SC, Lin YH, Hsiung CA, Chiou HY, Chen WJ, et al. A comprehensive evaluation of COVID-19 policies and outcomes in 50 countries and territories. *Sci Rep.* 2022;12(1):8802.
 51. McKee M, Rajan S. What can we learn from Israel's rapid roll out of COVID 19 vaccination? *Isr J Health Policy Res.* 2021;10(1):5.
 52. generator metatags. Forecasting Pneumonia Toddler Mortality Using Comparative Model ARIMA and Multilayer Perceptron | Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi). 2022 Cited 2023 Feb 9; Available from: <http://www.jurnal.iain.or.id/index.php/RESTI/article/view/4106>.
 53. Banerjee SC, Banerjee S, Rai P. Vaccine Supply Optimization and Forecasting using Random Forest and ARIMA Models. In: 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT). 2022. 1–5.
 54. Kumar M, Gupta S, Kumar K, Sachdeva M. Spreading of COVID-19 In India, Italy, Japan, Spain, UK, US: a prediction using ARIMA and LSTM model. *Digit Gov Res Pract.* 2020;1(4):24:1–24:9.
 55. Hernandez-Matamoros A, Fujita H, Hayashi T, Perez-Meana H. Forecasting of COVID19 per regions using ARIMA models and polynomial functions. *Appl Soft Comput.* 2020;1(96):106610.
 56. Zaki N, Mohamed EA. The estimations of the COVID-19 incubation period: a scoping reviews of the literature. *J Infect Public Health.* 2021;14(5):638–46.
 57. Jones, Sarah P. Imperial College London. Imperial College London Big Data Analytical Unit and YouGov Plc. Cited 2023 Feb 9. Available from: <https://www.imperial.ac.uk/global-health-innovation/what-we-do/our-response-to-covid-19/covid-19-behaviour-tracker/>.
 58. Aknin LB, Andretti B, Goldszmidt R, Helliwell JF, Petherick A, Neve JED, et al. Policy stringency and mental health during the COVID-19 pandemic: a longitudinal analysis of data from 15 countries. *Lancet Public Health.* 2022;7(5):e417–26.
 59. Petherick A, Goldszmidt R, Andrade EB, Furst R, Hale T, Pott A, et al. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nat Hum Behav.* 2021;5(9):1145–60.
 60. Hwang YS, Jo HS. Associations Between Compliance With Non-pharmaceutical Interventions and Social-distancing Policies in Korea During the COVID-19 Pandemic. *J Prev Med Public Health Yebang Uihakhoe Chi.* 2021;54(4):230–7.
 61. Goldszmidt R, Petherick A, Andrade EB, Hale T, Furst R, Phillips T, et al. Protective behaviors against COVID-19 by individual vaccination status in 12 countries during the pandemic. *JAMA Netw Open.* 2021;4(10):e2131137.
 62. Olivera Mesa D, Hogan AB, Watson OJ, Charles GD, Hauck K, Ghani AC, et al. Modelling the impact of vaccine hesitancy in prolonging the need for Non-Pharmaceutical Interventions to control the COVID-19 pandemic. *Commun Med.* 2022;2(1):1–8.
 63. Ekinci A. Modelling and forecasting of growth rate of new COVID-19 cases in top nine affected countries: Considering conditional variance and asymmetric effect. *Chaos Solitons Fractals.* 2021;1(151):111227.
 64. Nikolopoulos K, Punia S, Schäfers A, Tsinopoulos C, Vasilakis C. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *Eur J Oper Res.* 2021;290(1):99–115.
 65. Imperial College London YouGov Covid 19 Behaviour Tracker Data Hub. YouGov Data; 2023. Cited 2023 Mar 11. Available from: <https://github.com/YouGov-Data/covid-19-tracker/blob/fe0142dbf23b30755df537bacce5c6d687c52fdd/22%200401%20English%20questionnaire%20full%20v1.0.pdf>.
 66. Tan CM, Owuamalam CK, Ng PK. Stay at home, protect the NHS and save lives! Confidence in government moderates the negative effects of staying at home on mental health. *Personal Individ Differ.* 2021;1(179):110948.
 67. Visual FT, team DJ. Covid-19 vaccine tracker: the global race to vaccinate. Cited 2022 Dec 30. Available from: <https://ig.ft.com/coronavirus-vaccine-tracker/>.
 68. Rob J Hyndman - The ARIMAX model muddle. 2010 Cited 2023 Feb 12. Available from: <https://robjhyndman.com/hyndsight/arimax/>.
 69. Rodionov DG, Konnikov EA, Nasrutdinov MN. A Transformation of the approach to evaluating a region's investment attractiveness as a consequence of the COVID-19 pandemic. *Economies.* 2021;9(2):59.
 70. Ritter M, Ott DVM, Paul F, Haynes JD, Ritter K. COVID-19: a simple statistical model for predicting intensive care unit load in exponential phases of the disease. *Sci Rep.* 2021;11(1):5018.
 71. Bianchi L, Jarrett J, Hanumara R. Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *Coll Bus Fac Publ.* 1998; Available from: https://digitalcommons.uri.edu/cba_facpubs/86.
 72. Serrà J, Arcos JLI. An empirical evaluation of similarity measures for time series classification. *Knowl-Based Syst.* 2014;67:305–14.
 73. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. *J Stat Softw.* 2008;29(27):1–22.
 74. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet.* 2020;395(10242):1973–87.
 75. McGrail DJ, Dai J, McAndrews KM, Kalluri R. Enacting national social distancing policies corresponds with dramatic reduction in COVID19 infection rates. *PLOS One.* 2020;15(7):e0236619.
 76. Beatty K, Hamilton V, Kavanagh PM. Just a bad flu? Tackling the “info-demic” in Ireland through a comparative analysis of hospitalised cases of COVID-19 and influenza. *Public Health.* 2021;1(194):19–24.
 77. Jacek C, Karolina S, Orzeł A, Frączek M, Tomasz Z. Comparison of the clinical differences between COVID-19, SARS, influenza, and the common cold: a systematic literature review. *Adv Clin Exp Med.* 2021;30(1):109–14.
 78. Sallam M. COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. *Vaccines.* 2021;9(2):160.
 79. Bajema KL, Dahl RM, Prill MM, Meites E, Rodriguez-Barradas MC, Marconi VC, et al. Effectiveness of COVID-19 mRNA vaccines against COVID-19-Associated Hospitalization - five veterans affairs Medical Centers, United States, February 1–August 6, 2021. *MMWR Morb Mortal Wkly Rep.* 2021;70(37):1294–9.
 80. Tenforde MW, Self WH, Adams K, Gaglani M, Ginde AA, McNeal T, et al. Association between mRNA vaccination and COVID-19 Hospitalization and disease severity. *JAMA.* 2021;326(20):2043–54.
 81. Victora CG, Castro MC, Gurzenda S, Medeiros AC, França GVA, Barros AJD. Estimating the early impact of vaccination against COVID-19 on deaths among elderly people in Brazil: Analyses of routinely-collected data on vaccine coverage and mortality. *eClinicalMedicine.* 2021;38:101036.
 82. Chaguza C, Coppi A, Earnest R, Ferguson D, Kerantzas N, Warner F, et al. Rapid emergence of SARS-CoV-2 Omicron variant is associated with an infection advantage over Delta in vaccinated persons. *Med N Y N.* 2022;3(5):325–334.e4.
 83. Liang LL, Le HM, Wu CY, Sher CY, McGuire A. Human mobility increased with vaccine coverage and attenuated the protection of COVID-19 vaccination: a longitudinal study of 107 countries. *J Glob Health.* 2023;13:06009.
 84. Lock S, Ambrose T, Gayle D, Belam M, Ambrose SL (now); T, Belam (earlier) M. France to ease restrictions; Austrian parliament approves mandatory vaccination – as it happened. *The Guardian.* 2022 . Cited 2023 Feb 2; Available from: <https://www.theguardian.com/world/live/2022/jan/20/covid-news-live-who-recommends-lifting-international-travel-bans-and-abolishing-proof-of-vaccination-requirements>.

85. Davidson S, Producer Ctvn ca MD, Contact F]. Toronto. 2022. Ontario's top doctor says it's time to learn to live with COVID-19. Cited 2023 Feb 2. Available from: <https://toronto.ctvnews.ca/ontario-s-top-doctor-says-it-s-time-to-learn-to-live-with-covid-19-1.5757373>.
86. How do Italy's Covid-19 rules change in February?. The Local Italy. 2022 cited 2023 Feb 2. Available from: <https://www.thelocal.it/20220127/how-do-italys-covid-19-rules-change-from-february-1st/>.
87. Knesset passes "COVID law," ending state of emergency in place for most of pandemic | The Times of Israel. Cited 2023 Feb 2. Available from: <https://www.timesofisrael.com/mks-pass-covid-law-ending-state-of-emergency-in-place-since-start-of-pandemic/>.
88. Earnest A, Chen MI, Ng D, Sin LY. Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore. *BMC Health Serv Res*. 2005;5(1):36.
89. He Z, Tao H. Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: a nine-year retrospective study. *Int J Infect Dis*. 2018;1(74):61–70.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

