Nonlinear Dyn (2022) 109:77–90
https://doi.org/10.1007/s11071-022-07473-9

ORIGINAL PAPER

Data analysis and prediction of the COVID-19 outbreak in the first and second waves for top 5 affected countries in the world

Ashabul Hoque · Abdul Malek · K. M. Rukhsad Asif Zaman

Received: 6 November 2021 / Accepted: 20 April 2022 / Published online: 7 May 2022

Abstract In this paper, we introduce a SEIATR compartmental model to analyze and predict the COVID-19 outbreak in the Top 5 affected countries in the world, namely the USA, India, Brazil, France, and Russia. The officially confirmed cases and death due to COVID-19 from the day of the official confirmation to June 30, 2021 are considered for each country. Primarily, we use the data to make a comparison between the cumulative cases and deaths due to COVID-19 among these five different countries. This analysis allows us to infer the key parameters associated with the dynamics of the disease for these five different countries. For example, the analysis reveals that the infection rate is much higher in the USA, Brazil, and France compared to that of India and Russia, while the recovery rate is found almost the same for these countries. Further, the death rate is measured higher in Brazil as opposed to India, where it is found much lower among the remaining countries. We then use the SEIART compartmental model to characterize the first and second waves of these countries, as well as to investigate and identify the influential model parameters and nature of the virus transmissibility in respective countries. Besides estimating the time-dependent reproduction number (Rt) for these countries, we also use the model to predict the peak size and the time occurring peak in respective countries. The analysis demonstrates that COVID-19 was observed to be much more infectious in the second wave than the first wave in all countries except France. The results also demonstrate that the epidemic took off very quickly in the USA, India, and Brazil compared to two other countries considered in this study. Furthermore, the prediction of the epidemic peak size and time produced by our model provides a very good agreement with the officially confirmed cases data for all countries expect Brazil.

Keywords COVID-19 · SEIATR model · Reproduction number · Second wave · Peak period · Epidemic evolution

A. Hoque (✉) · A. Malek
Department of Mathematics, University of Rajshahi, Rajshahi 6205, Bangladesh
e-mail: ashabulh@yahoo.com

A. Malek
e-mail: malekbio@gmail.com

K. M. R. A. Zaman
Department of Public Health, North South University, Dhaka, Bangladesh
e-mail: arnob88dr@gmail.com

1 Introduction

1.1 COVID-19 and its characteristics

The coronavirus disease (COVID-19) from China has spread globally since January 2020 and has become a pandemic. Currently, there are around 222 countries

that report laboratory-confirmed cases across the World (WHO). Cumulative data of confirmed cases (182,688,695), recovered (167,288,083), and deaths (3,956,008) are taken from the beginning to June 30, 2021 (Johns Hopkins University). An infectious disease is the episode of an illness that is not generally expected in a particular group of people, geographical region, or time. Due to the relatively new nature of this disease, proper control measures and therapeutic interventions are still under development, which in turn is creating tremendous tension and panic around the World. Not only is the COVID-19 pandemic threatening our social and personal life and the broader aspects, such as our economy, health, and development in both the national and global sense. Due to uncertainties of the disease, investigators have used several models to forecast the characteristics of transmission parameters, primary reproduction number (R_0), time-depended reproduction number (R_t), the time of peak (t^p), etc.

1.2 Related works

The researches on mathematical modeling are playing a key role in understanding the epidemics, and it may help to predict the intensity of pandemics in the early stages. This also demonstrates a significant role in making the right decision during outbreak control [1]. In this regard, several researchers have developed mathematical models for the COVID-19 epidemic [2–13]. Very recently, Kuddus and Rahman [2] have used the improved SLIR model with nonlinear incidence, and they have observed that the transmission rate of each parameter had a significant impact on COVID-19. Lobato et al. [3] proposed a dynamic data segmentation approach to provide reasonable estimates for all parameters. A three-party differential game model including epidemic prevention and risk coefficient was proposed by [14], and results were presented based on theoretical and numerical analysis. Toda [4] estimated the COVID-19 transmission rates for several countries based on the SIR model for regular data of confirmed cases. Tang et al. [15] proposed a compartmental model with a clinical progress compartment and epidemiological compartment, and they showed that the isolation compartment could successfully reduce the transmission hazard. Biswas et al. [16] formulated a deterministic compartmental model to estimate the model parameters and compared these against the reported data. Wu et al. [5] used a four-compartment (SEIR) model to explain the transmission rates and forecast the countrywide and worldwide feast of the COVID-19 epidemic based on published data from December 31, 2019, to January 28, 2020. They also determined the basic reproduction number for COVID-19, and it was nearly 2.68 for China. COVID-19 outbreaks on human-to-human transmission based on the computational modeling of probable epidemic trajectories were estimated by Imai et al. [17]. They focused that the control actions needed to block over 60% of transmission to control the outbreak effectively. Ahmed et al. [18] developed a SEIR time-fractional model to investigate the nature of coronavirus in Pakistan and discussed the stability analysis. Pedersen et al. [19] proposed the SIQR model to discuss the dynamics of COVID-19 in Italy. Hoertel et al. [20] represented a stochastic agent-based microsimulation model of COVID-19 to examine the impact of mask-wearing, physical distancing, and shielding individuals and showed that those were slowing the spread of the epidemic and reducing the mortality rate.

In epidemiology, the average basic reproduction number (R_0) is defined as the average number of secondary cases that would be generated by a primary infectious disease in a susceptible population [21]. Determining R_0 is often stimulating for the involvement of numerous factors and a deficiency of unbiased data. In most cases, secondary infections cannot be estimated precisely, particularly for COVID-19, where asymptomatic patients are barely recognized. There are many techniques to calculate R_0 [22], some of them agree with each other, and some are developed based on the secondary infections. In general, if R_0 is greater than 1, the disease has a spreading potential with a tendency to increase the number of new cases. On the other hand, the value of R_0 is equal and below 1 indicates that it is a threshold disease, and the infection will spread slowly, and the disease dies out, respectively. To effectively eliminate an epidemic disease from a population, R_0 needs to be less than unity. It is thus of interest to estimate the value of R_0 for emerging diseases.

Estimating the reproduction number in the early stages of the spread of the COVID-19 have helped to understand essential aspects of the pandemic. In this context, Marimuthu et al. [12] estimated the introductory reproduction number rate (R_0) at 1.379 using the

exponential growth method for India. They also estimated R_0 with CI at different states (1.450 for Maharashtra, 1.444 for Gujarat, 1.297 for Delhi, and 1.405 for Tamil Nadu) in India and observed the disease is pandemic type. Read et al. [23] determined the value of R_0 as 3.1 fitting data with the SEIR model and Poisson-distributed theory with daily time increments. Beenstock and Dai [24] computed the values of the effective reproduction number regularly in several countries applying the perpetual register technique. Hong et al. [25] considered a dynamical epidemiology model developing overtime in which the effective reproduction number was estimated subject to the stochastic shocks. Chaves et al. [26] estimated the R_0 for COVID-19 based on the susceptible–infectious–removed model, and they observed the value of R_0 in the range of 2.58–2.43. Hong and Li [27] estimated the time-dependent reproduction number based on the Poisson model with transmission and removal rates and reported probable arbitrary errors. Khosravi et al. [28] determined the value of R_0 by a maximum-likelihood method which was 2.7 for the COVID-19 pandemic in the first 14 days and reduced to 1.13 by 42 days in Iran.

Some relevant works [7, 15, 29] applied the various techniques to predict the peak time and size. According to [7], the epidemic peak and size were estimated for COVID-19 by using confirmed data, and he expected that it could reach the early-middle summer in Japan. According to Tang et al. [15], the predicted peak could be found within two weeks from January 23, 2020 (China), but this result did not satisfy the accurate picture. Zahiri et al. [29] determined the predictions of the COVID-19 pandemic, such as the actual number of victims, infection rate, and peak time in Iran. Ranjan et al. [30] studied the COVID-19 spreads in the ongoing second wave in India and its several cities up to April 19, 2021. They learned the dynamic advancements of the epidemic parameters from the beginning of the outbreak.

Recently, many researchers have studied mathematical modeling of COVID-19 for second waves (Iftimie et al., [31]; Vasconcelos et al., [32]; Salyer et al., [33]; Lobato et al., [3]). The so-called second wave of COVID-19 is characterized by an expressive increase in the number of confirm cases after a significant drop in the number of new infections during the first wave. In this context, Iftimie et al. [31] discussed the comparative study of first and second waves of COVID-19 in Spain and showed that the fatality rate was lower in the case of second wave. Vasconcelos et al. [32] introduced a generalized logistic model with time-dependent parameters to analyze the dynamics of the COVID-19 fatality curves for several countries and the intensity of the standard second wave was less than the first wave. Salyer et al. [33] studied the first and second waves of the COVID-19 pandemic in Africa and showed that the second wave appeared to more aggressive compared to first wave. These studies have made significant contributions to investigate the COVID-19 in different aspects.

1.3 Study regions and objective

Though COVID-19 affected many countries at different magnitudes, the TOP 5 countries based on WHO till June 30, 2021, are considered in this study. The present study focuses on the six-compartmental model, with treatment individuals, to estimate the impact of treatment and isolation on the COVID-19 pandemic. Our aim is to determine the parameters of the SEIATR model that best describe the behavior of the first and second waves in USA, India, Brazil, France, and Russia. This study also will focus the nonlinear dynamics of the system which will give more information on the dynamics of the epidemic system.

2 Models and data

2.1 Physical descriptions

SIR model and its modified version, such as Susceptible-Exposed-Infectious-Recovered (SEIR) model, were widely used to analyze and investigate the COVID-19 outbreak [34–37]. Most of these models were developed or investigated based on the assumption that the transmission and removal rates were constant, which might not hold authenticity [27]. In the present model (SEIATR), we subdivide the total number of population into six deterministic compartments: the susceptible (denoted by S), the exposed (denoted by E), the infected (denoted by I), the asymptomatic (represented by A), the treatment (denoted by T) and the recovered (denoted by R). The diagram of SEIATR in Fig. 1 (Corona Tracker) shows how an entity moves through each compartment in the

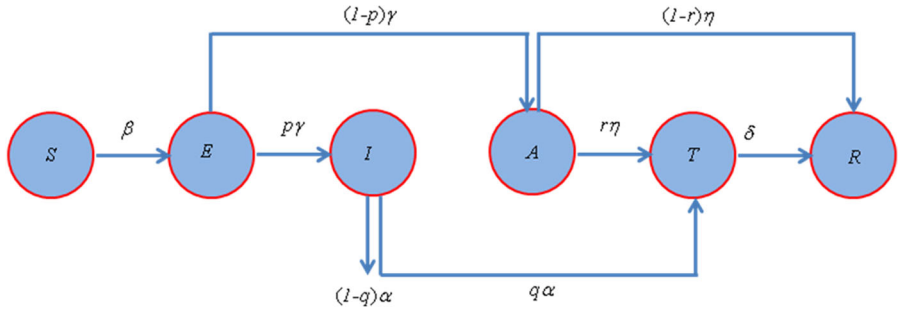


Fig. 1 The figure represents a diagram of the model described by the Eqs. (1–6). Interventions including intensive contact tracing followed by asymptomatic (A) and treatment (T) are indicated

model. In this study, the natural recruitment and mortality rates are not considered, because these are negligible compared to the total population. But the death rate due to COVID-19 is considered to be constant in this study.

There are three main parameters to estimate the new infective individuals across the hazard scale affected by the COVID-19. These are: (1) Transmission rate (β), (2) Average incubation period ($1/\gamma$), and (3) Fraction p of exposing compartment (E) proceeds to the infective compartment (I). Commonly, the incubation period ($1/\gamma$) is taken to be constant for the sake of simplicity in each country (Table 1).

2.2 Model equations

We introduce the following SEIATR compartmental model (modification of SEIR well-known model) for forecasting of the COVID-19 outbreak of Top 5 affected countries in the World. The SEIATR model equations are described in below:

$$\frac{dS(t)}{dt} = -\beta(t)S(t)(I(t) + \varepsilon_1(t)A(t) + \varepsilon_2(t)T(t)) \quad (1)$$

$$\frac{dE(t)}{dt} = \beta(t)S(t)(I(t) + \varepsilon_1(t)A(t) + \varepsilon_2(t)T(t)) - p\gamma(t)E(t) - (1-p)\gamma(t)E(t) \quad (2)$$

$$\frac{dI(t)}{dt} = p\gamma(t)E(t) - (1-q)\alpha(t)I(t) - q\alpha(t)I(t) \quad (3)$$

Table 1 Parameter values for model Eqs. (1–6)

Country/wave	Values of the parameters									
	β/million	$\varepsilon 1$	$\varepsilon 2$	p	γ	α	f/million	ε	r	δ
First wave										
USA	202.80 (95% CI, 180.68–224.91)	0.40	0.20	0.12	1/6	0.9	3.75	0.5	0.05	0.627
India	20.67(95%CI: 18.62 22.72)	0.40	0.20	0.22	1/6	0.9	0.33	0.5	0.05	0.972
Brazil	104.43(95% CI: 93.63, 115.24)	0.40	0.20	0.17	1/6	0.9	3.32	0.5	0.05	0.883
France	22.54(95%CI: 13.03, 32.05)	0.40	0.20	0.35	1/6	0.9	2.92	0.5	0.05	0.945
Russia	31.75(95%CI: 28.41 35.09)	0.40	0.20	0.20	1/6	0.9	0.71	0.5	0.05	0.871
Second wave										
USA	322.29(95%CI:295.55, 349.03)	0.40	0.20	0.15	1/6	0.9	4.81	0.5	0.05	0.492
India	110.20(95%CI: 93.66 126.74)	0.40	0.20	0.23	1/6	0.9	1.35	0.5	0.05	0.645
Brazil	247.76(95%CI:233.63, 261.90)	0.40	0.20	0.35	1/6	0.9	6.81	0.5	0.05	0.645
France	228.96(95%CI:193.31, 264.61	0.40	0.20	0.18	1/6	0.9	3.21	0.5	0.05	0.965
Russia	98.66(95%CI: 92.93 104.39)	0.40	0.20	0.12	1/6	0.9	2.54	0.5	0.05	0.699
References	*	**	**	*	***	**	*	**	**	*

*Worldometer, **Estimated and ***WHO

$$\frac{dA(t)}{dt} = (1-p)\gamma(t)E(t) - (1-r)\eta(t)A(t) - r\eta(t)A(t) \quad (4)$$

$$\frac{dT(t)}{dt} = q\alpha(t)I(t) + r\eta(t)A(t) - \delta(t)T(t) \quad (5)$$

$$\frac{dR(t)}{dt} = \delta(t)T(t) + (1-r)\eta(t)A(t) \quad (6)$$

$$\frac{dN(t)}{dt} = -(1-q)\alpha(t)I(t) \quad (7)$$

The rate of change of susceptive is described in Eq. (1) which is influenced by infected, expose and treatment compartments, where $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are the relative horizontal transmissions of the diseases due to asymptomatic and treatment individuals, respectively. This amount of individuals goes to expose compartment at a rate $\beta(t)$ at time t. The rate of change of exposed individuals is represented by Eq. (2) which is reduced at a rate $\gamma(t)$ at t. p-fraction of the expose compartment goes to the infected compartment and the remaining fraction *(1-p)* proceed to the asymptomatic compartment. The rate of change infected individuals is given by Eq. (3) which is increased by the individuals of expose compartment at a rate of $\gamma(t)$ and reduce at a rate $\alpha(t)$ to the asymptomatic and treatment compartments, respectively. The parameters q and r are fractions of the infective and asymptomatic populations that proceed to the treatment compartment at a rate $\alpha\ (t)$ and $\eta\ (t)$, respectively, and $1/\delta\ (t)$ is the average treatment period. The rate of change of asymptomatic and treatment compartments are presented in Eq. (4) and Eq. (5), respectively. The rate of change of removal compartment is obtained in Eq. (6) which is increased at a rate of $\eta(t)$ and $\delta(t)$ from the asymptomatic and treatment compartments.

For solving Eqs. (1) to (7), the following initial conditions are required:

$$S(0) = S_0,\ E(0) \geq 0,\ I(0) = I_0,\ A(0) \geq 0,\ T(0) \geq 0,\ R(0) \geq 0 \quad (8)$$

and

$$S(t) + E(t) + I(t) + A(t) + T(t) + R(t) = N = S_0 + I_0. \quad (9)$$

For disease-free equilibrium, an equilibrium solution of model Eqs. (1–6) is obtained with $E = I = 0$. Then by solving the equilibrium system, it has been found $A = T = R = 0$ and let $S = S_0$. Thus, disease-free equilibrium is $(S_0, 0, 0, 0, 0, 0)$.

Diseases compartments are found from the system using Eqs. (2–5) as

$$\frac{dE}{dt} = \beta S(I + \varepsilon_1 A + \varepsilon_2 T) - p\gamma E - (1-p)\gamma E \quad (10)$$

$$\frac{dI}{dt} = p\gamma E - (1-q)\alpha I - q\alpha I \quad (11)$$

$$\frac{dA}{dt} = (1-p)\gamma E - (1-r)\eta A - r\eta A \quad (12)$$

$$\frac{dT}{dt} = q\alpha I + r\eta A - \delta T \quad (13)$$

In general, the above model equations can be expressed in the following tensor form

$$\frac{\partial x_i}{\partial t} = f_i(x_i, y_i) - v_i(x_i, y_i), \quad (14)$$

where f_i denotes the secondary infection rate that increases the i th disease compartment and v_i shows the other progression or transition rate (such as death, recovery) that decreases the i th disease compartment. Mathematically, the new infection matrix f and transition matrix v can be written, respectively,

$$f = \begin{bmatrix} \beta S(I + \varepsilon_1 A + \varepsilon_2 T) \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } v = \begin{bmatrix} \gamma E \\ -p\gamma E + \alpha I \\ -(1-p)\gamma E + \eta A \\ -q\alpha I + r\eta A + \delta T \end{bmatrix}, \quad (15)$$

The time-dependent reproduction number (R_t) can be estimated by the method of next-generation matrix $k = F/V_1$ suggested by van den Driessche et al. [38], and it displays the positive eigenvalue (R_t) of the matrix k at the disease-free equilibrium. Applying the above technique, it is found as

$$F = \frac{\partial f_i}{\partial x_j}(0, y_0) \text{ and} V_1 = \frac{\partial v_i}{\partial x_j}(0, y_0) \quad (16)$$

where

$$F = \begin{bmatrix} 0 & \beta S_0 & \beta\varepsilon_1 S_0 & \beta\varepsilon_2 S_0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and}$$

$$V_1 = \begin{bmatrix} \gamma & 0 & 0 & 0 \\ -p\gamma & \alpha & 0 & 0 \\ -(1-p)\gamma & 0 & \eta & 0 \\ 0 & -q\alpha & -r\eta & \delta \end{bmatrix}.$$

Following van den Driessche et al. [38], the time-varying reproduction number (R_t) is then found based on the spectral radius or positive eigenvalue of FV^{-1}.

After simplification, the time-dependent reproduction number can be derived as

$$R_t(t) = \beta(t)S_0\left[\frac{p}{\alpha(t)} + \frac{\varepsilon_1(t)(1-p)}{\eta(t)} + \frac{\varepsilon_2(t)(r+pq-pr)}{\delta(t)}\right] \tag{17}$$

The time-varying reproduction number (R_t) is an important term that provides a real-time picture of an outbreak [12]. R_t is proportional to the contact rate for the SIR model and will vary according to the local situation. In contrast, for the SEIATR model, R_t can be obtained as the spectral radius of a next-generation matrix of the transmission of the diseases, and the time-varying reproduction number is determined for a disease-free equilibrium [39].

2.3 Computational method

COVID-19 breaks out all over the World, and various public health measures, treatments with some drugs, immunity, guard, etc., are only the control measures for this virus yet. To predict and determine the intensity of the COVID-19 pandemic, model Eqs. (1–6) are applied as the fundamental equations and are solved numerically with the help of the Runge–Kutta method. The incidence data between starting date for each country and June 30, 2021, have been considered to get the best-fit parameters in the proposed model. The initial values of susceptible, exposed, infectious, and recovered cases are taken for each country based on the human population and confirmed cases data from WHO, Worldometer, and Johns Hopkins University's database. Based on the WHO, the incubation period of COVID-19 is about 6 days, hence $\sigma = 1/6$ was set up in the calculation. The values of the several parameters of the proposed model are taken for COVID-19 from several published papers, reports, and reviews of preceding, which are stated in Table 1. The outbreak data mentioned in Table 1 would usually have dual uncertainty. These are: (a) asymptomatic infected people could spread the infection and (b) insufficient diagnostic tests. This investigation has accumulated such delay to a SEIATR proposed model to predict the epidemic peak and size.

2.4 Data

The COVID-19 has been thoroughly observed by several recognized societies such as the World Health Organization (WHO), Worldometer, and Johns Hopkins University's database. Currently, there are around 222 countries that reported laboratory-confirmed cases for COVID-19 across the World (WHO). Although it was better to include the most affected countries in our analysis, it was tedious to work with such big data, model limitation, data initialization, parameter selection, data accuracy, and keeping track to achieve our objectives. That is why the work has kept analyzing with Top 5 affected countries, namely USA, India, Brazil, France, and Russia, where the first case in each country was reported on January 21, 2020; January 30, 2020; February 26, 2020; January 24, 2020; and January 31, 2020, respectively (Worldometer). For analyzing the COVID-19 pandemic, the data were taken from the starting date to June 30, 2021, for each country. The governmental websites of many countries also have been following these numbers starting from several time points. The websites mentioned above have become valuable capitals to help advance the thoughtfulness of the feast of the virus.

3 Results and discussion

3.1 Cumulative reported cases

In Fig. 2, the cumulative confirmed cases, deaths, and recoveries for COVID-19 are compared among the Top 5 affected countries. These countries are most affected by the maximum number of cumulative points as of June 30, 2021. The USA has leading cumulative cases and death among the countries shown, with about 34.52 million and 0.62 million, respectively. In Fig. 2, it can be seen that the outbreak

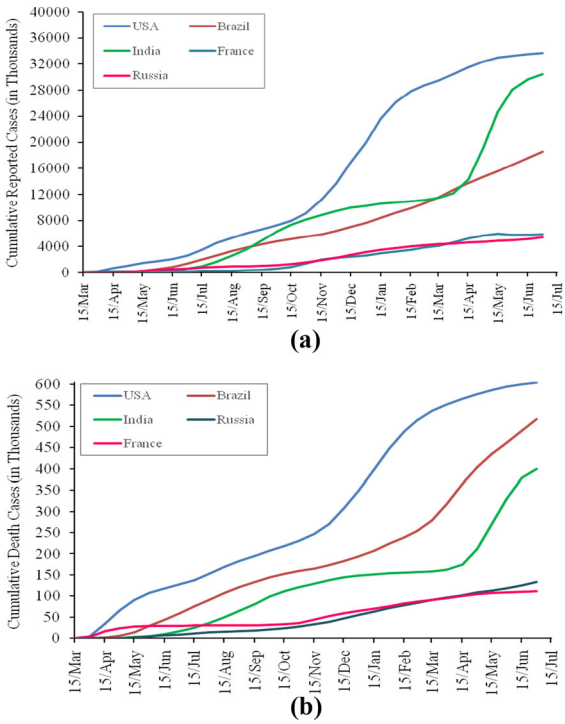


Fig. 2 Comparisons among the cumulative **a** confirmed cases and **b** deaths due to COVID-19 among the Top 5 affected countries in the world. The *x*-axis denotes the number of months since the first laboratory-confirmed case in the representative countries

started to move in the USA, India, and Brazil, with a relatively fast spread of the disease. The spreading patterns are almost the same for France and Russia. Furthermore, many deaths are observed in the USA, Brazil, and India, where fewer deaths can be seen both in France and Russia. In addition, the data of recovery for the USA is not available from the Worldometer (from 14/12/2020 to ongoing), so we could not compare here.

The infected, death, and recovery rates are compared in Fig. 3. It can be seen that the infected rates are much higher in the case of the USA, Brazil, and France, and the recovery rates are almost the same for all the countries. Furthermore, the death rate is measured higher in Brazil, whereas it is found lower in India.

3.2 Model-implied reproduction number

Figure 4 demonstrates the variations in daily confirmed cases and the normalized time-varying

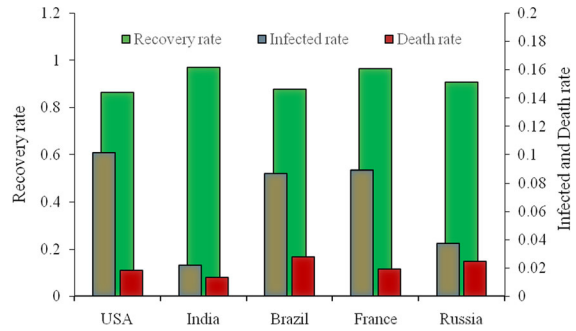
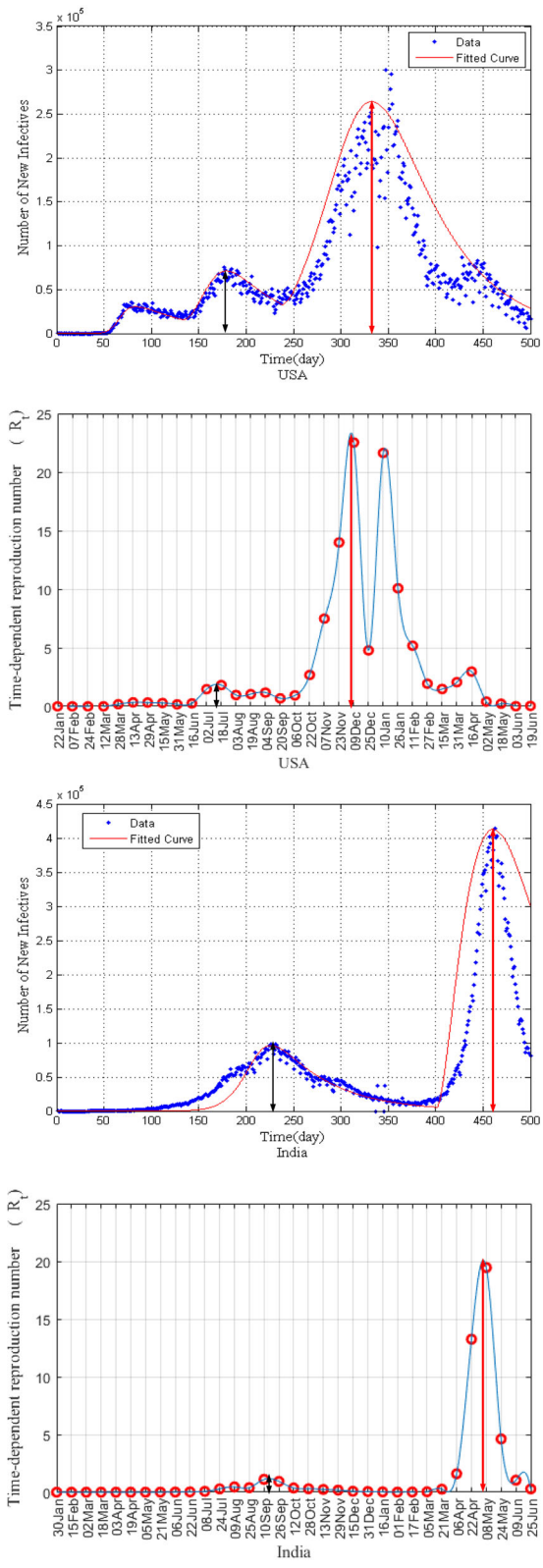


Fig. 3 Graphs for COVID-19 infected, recovery and death rate. Infected rate = Infectious cases/Total population; Death rate = Total death/Infectious cases; Recovery rate = Total recovery/Infectious cases

reproduction number (R_t) for the Top 5 affected countries. The curves show the growing trends of the epidemic for each country, and virtually all the countries showed a significant change in (R_t) from starting dates to June 30, 2021. R_t was estimated in a 15 days interval from the starting date, taking the reported median incubation period of 6 days. In Fig. 4, the (R_t) remains almost constant from starting date to 3 to 6 months for the Top 5 countries suggesting a stabilization of the transmissibility rate. R_t starts increasing and reaches a maximum of around 1.85, 1.12, 9.33, 1.44, and 0.22 in the first wave for the USA, India, Brazil, France, and Russia, respectively. After this relatively long quiet interval, R_t started rising maximum 22.24, 19.46, 27.11, 10.32, and 1.55 in the USA, India, Brazil, France, and Russia, respectively, which can be considered an indicator of the arrival date of the second wave. In Fig. 4, the high reproduction number with Brazil at the beginning and India on May 05, 2021, indicates that the infectious disease transmission was not well managed. In addition, our results show good agreement with those determined by Khailaie et al. [40] using the SECIR model. It is concluded that the proposed model is capable of obtaining good estimations for the model parameters and, subsequently, for R_t, as illustrated in Fig. 5.

On the other hand, the average basic reproduction numbers (R_0) of (R_t) are calculated based on the arithmetic mean formula for the Top 5 affected countries. The values of average R_0 are measured 3.24 (95% CI: 2.289, 4.2), 6.75 (95% CI: 5.5, 8.0), 1.37 (95% CI: 0.70, 204), 1.57 (95% CI: 1.14, 2.01), and 0.37 (95% CI: 0.30, 0.45) for the USA, Brazil, India, France and Russia, respectively. The smallest



◀**Fig. 4** Shows the variations in daily confirmed cases and time-dependent reproduction number R_t for Top 5 affected countries

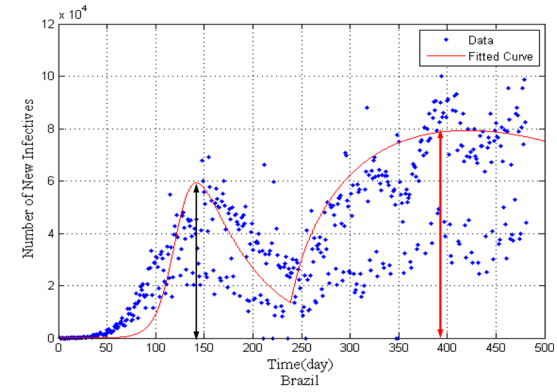
R_0 is observed in Russia (0.37), while the highest is in Brazil (6.75). From these estimations of R_0, it can be concluded that COVID-19 has an epidemic nature ($R_0 >> 1$) for all countries. The estimated R_0 is an important factor in infectious disease epidemiology because $1 - 1/R_0$ could reduce the force of infection to eliminate the disease outbreak. For example, at $R_0 = 4.40$ this fraction is 77%, but at $R_0 = 5.40$ this fraction is 81%. With the increase in R_0, the fraction $1 - 1/R_0$ increases, which may help decide to force the pandemic's control.

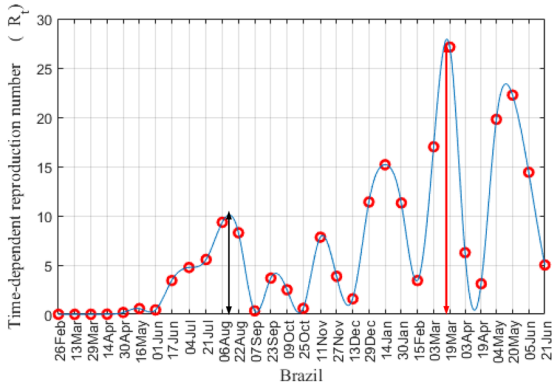
3.3 Peak prediction analysis

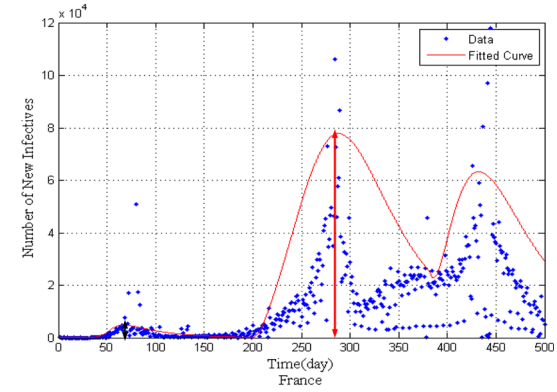
We define the epidemic peak by t^P in the first wave and assume it will occur within 36 months from the beginning of the COVID-19 pandemic. As an example, we discuss the procedure to estimate the epidemic peak and size for the USA. For identification rate $p = 0.12$, we obtain Fig. 5 on the long time behavior for $\beta = 224.92$, 202.80, and 180.68 per million. Under the current situation, it is seen in Fig. 6 that the predictable epidemic peak is $t^p = 6.3$ months (95% CI, 6.0–6.6) for the USA. That is, starting from January 21 ($t = 0$), the estimated infected individuals (peak size) is $(I(t)) = 74{,}500$, while the real peak size was 62,000 (Table 2) for USA and the estimated epidemic peak period is $t^P = 6.3$ month (July 30) with the uncertainty range is from $t^p = 6.0$ month (July 21) to $t^p = 6.6$ month (August 9) (Fig. 6). We obtain an estimation of $\beta = 202.80$ (95% CI, 180.68–224.92) for $p = 0.12$, where the parameter p is calculated so that the curve approaches fit with the real data in Fig. 6. Similarly, as shown in Table 2, the epidemic peaks were estimated for India, Brazil, France, and Russia.

3.4 Characterization of first and second waves and model results

The epidemic peaks and sizes have been investigated for the USA, Brazil, India, France, and Russia mentioned in Table 2 and Fig. 4. It can be seen that the estimated peak sizes of infectious individuals are closed to the actual dimensions of data (Fig. 4, Table 2)







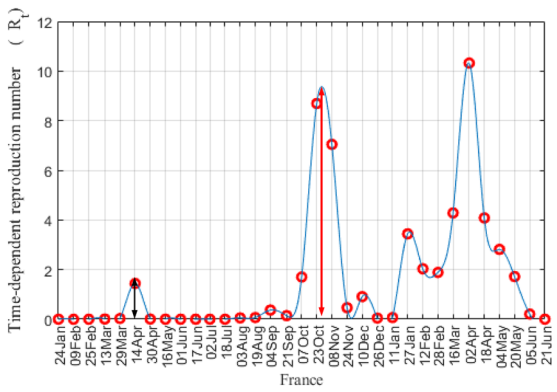
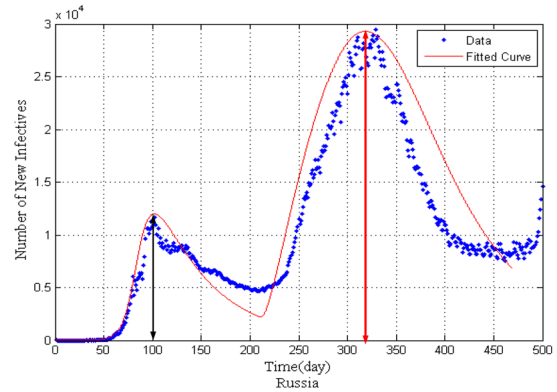


Fig. 4 continued



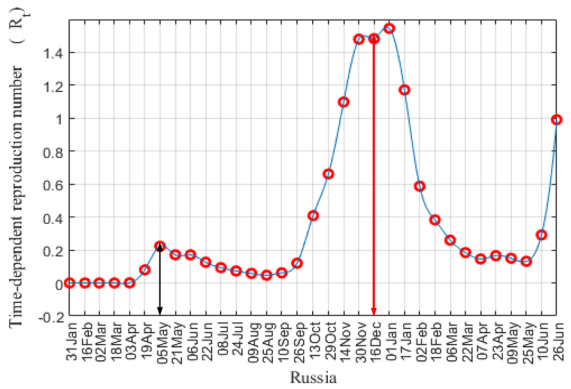


Fig. 4 continued

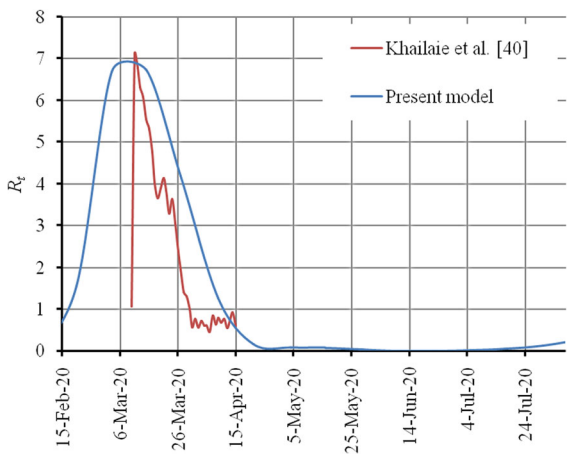


Fig. 5 Comparisons of the time-depedent reproduction number of the COVID-19 in USA, in the period between February 15 and November 25, 2020. The values of R_t are calculated using Eq. (17)

for all the countries in the case of first and second wave data. The second wave of COVID-19 has hit the Top 5 affected countries very hard, with the daily confirm

issues are reaching almost triple the first peak values for the USA, India, and Russia. Estimating peak sizes of infectious individuals fit on recent data demonstrate that the infection case is much higher than the first wave for all the Top 5 affected countries. The actual data of Brazil fluctuates near the second wave, and the calculated result of the confirmed case shows a slight difference from the existing data in the case of France (Fig. 4). It is noted that the third wave also appears in France. The accuracy of the data accumulated from Worldometer and Johns Hopkins University's database is a matter of trust for representing the entire study.

3.5 Evolutions of the system with parameters

The system's evolutions with parameters, such as susceptible, exposed, infected, asymptomatic, treatment, and recovered, are shown in Fig. 7, which shows how all the system classes affect the dynamics of an epidemic. Mainly, all the individual compartment increases first except susceptible *S*. After that; these converge to zero except the recovered compartment *R*. The patterns of the curves of the evolution parameters are all most similar types of the USA, India, Brazil, France, and Russia, so for the convenient, we have discussed these only for the USA and Russia. Since there is no evidence that the recovered compartment is resistant to the virus, this makes the system hard to converge. Furthermore, it is also seen that the asymptomatic and expose parameters are very effective compared to other parameters, and both link to zero slowly. Similar evolutions were done by He et al. [41].

3.6 Nonlinearilty of the transmission dynamics

Nonlinearity during the entire course of a pandemic disease arises due to a number of factors. For example, seasonality, age structure of a population, variation in immunity strength across different countries, mass isolation imposed by the policy makers, self-regulated awareness of the community when the spread of the virus hikes and the rate of virus mutation, to name a few [41–43]. The fluctuations in transmission observed in different countries during the course of the pandemic are presumably due to one or more of these factors. Similar nonlinear effects are observed in other biological phenomena. For example, in the case of a secondary or tertiary infection caused by a different strain of the dengue virus. In most cases, the primary infection causes a mild self-limiting illness in the majority of individuals [44–46]. However, aberrant immune response in the initial stages of infection in the case of secondary or tertiary infection caused by different strains of the virus often leads to severe dengue by inducing a vascular leak and excessive inflammation due to high levels of inflammatory cytokines. Similar phenomena are also observed in within-host dynamics of the SARS-COV transmission between cells caused by different strains of the virus [47].

4 Conclusions

In this paper, a SEIATR model was developed for the COVID-19 pandemic, and the model parameters were estimated by using the Runge–Kutta method. Using the estimated parameter values, the predictions for both the first and second wave for five different countries were produced. In addition, the model was used to investigate how the parameters affect the dynamics of the disease. The obtained results of epidemiological insights are briefly summarized below:

(1) The cumulative confirmed cases and deaths due to COVID-19 considering 15 days interval is presented and then a comparison is made among five different countries. The officially confirmed cases and death due to COVID-19 from the day of the official confirmation to June 30, 2021 are considered for each country. It was observed that in the case of USA, India, and Brazil, the epidemic spread took off very quickly. Furthermore, higher cumulative deaths were observed in the USA, Brazil, and India. The intensity of COVID-19 infections and deaths patterns were found almost the same for France and Russia.

(2) The time-dependent reproduction number (R_t) was estimated for all five countries, and it suggested that the COVID-19 was much more infectious in the second wave than the first wave. Further, the estimated value of R_0 in Russia was found to be as lowest as 0.37 (95% CI: 0.30, 0.45) while in Brazil it was found as high as 6.75 (95% CI: 5.5, 8.0).

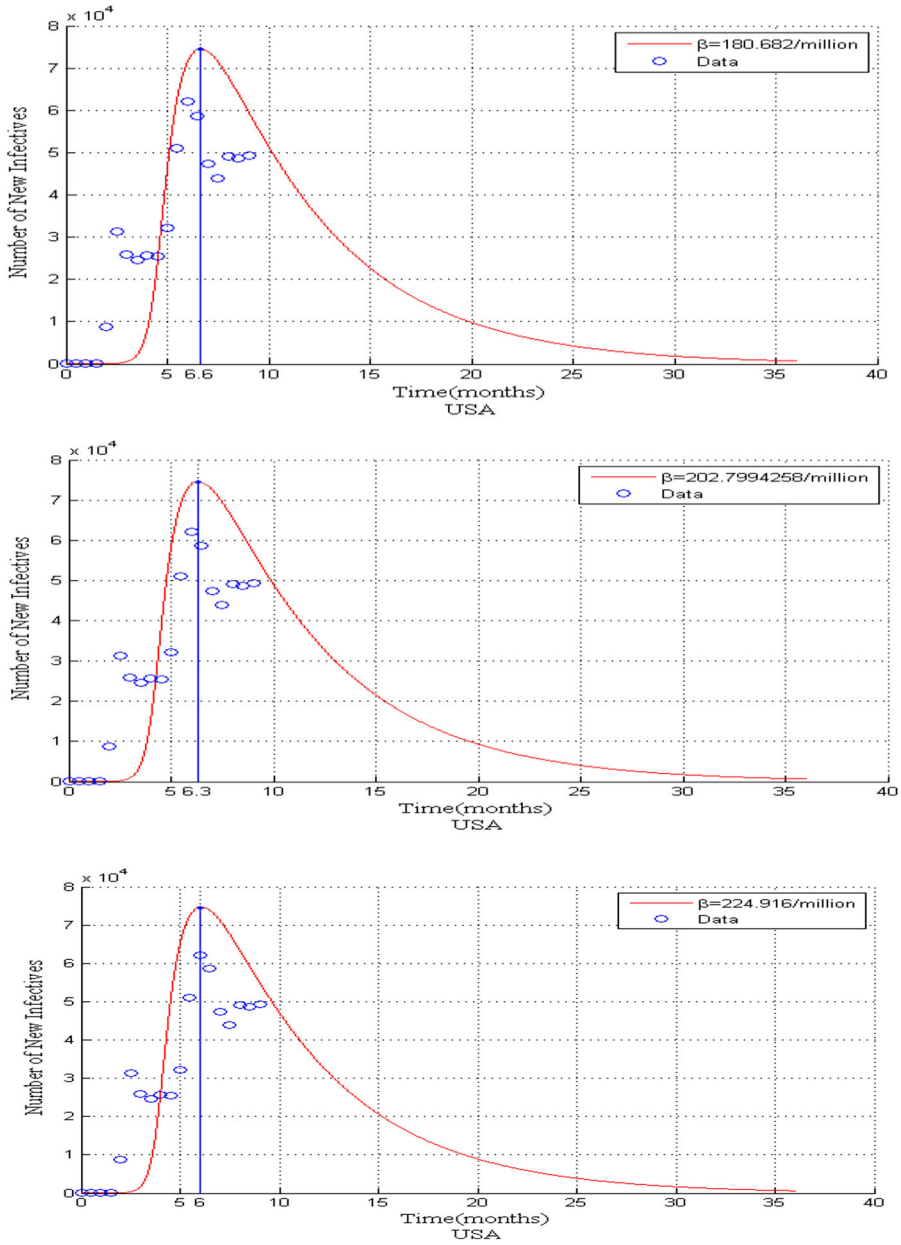


Fig. 6 Time-variation of the infective individuals' for the USA are identified at peak time t^p in a month for identification rate $p = 0.12$. The blue circles show active reported cases. The red lines are the model results, and the vertical blue lines show the epidemic peak. The values of β are 180.68, 202.80, and 224.92, respectively

(3) The peak sizes and times of the confirmed cases satisfied the model results for five different affected countries. It did not mean that our analyses had provided exact but erroneous estimations for all five countries instead of more minor errors obtained for France.

(4) The patterns of the curves of the epidemic evolution parameters were determined almost similar types for the USA and Russia. It was observed that all the compartments except the recovered compartment were converged to zero.

(5) Finally, the high reproduction number suggested that the outbreak might be more severe than that has been officially reported. We obtained peak time and size when a pandemic had already happened for the first and second waves. However, the pandemic peak time and size for a third or so on the waves are yet to be learned.

Table 2 Estimated COVID-19 pandemic peak times and peak sizes

Country	First identified date	First wave				Second wave			
		t^p month	Peak period	Estimated peak size	Real peak size	t^p month	Peak period	Estimated peak size	Real peak size
USA	21-01- 2020	6.3	30 July, 20	74,500	62,000	11.03	17 Dec, 20	254,000	239,900
India	30-01-2020	7.5	11 Sept., 20	98,000	97,570	15.4	05 May, 21	412,520	412,431
Brazil	26-02-2020	5.1	30 July, 20	70,000	51,000	14.26	28 Apr, 21	79,726	80,000
France	24-01-2020	2.3	01 April, 20	4820	4844	9.63	07 Nov, 21	86,000	86,655
Russia	02-02-2020	4.3	12 June, 20	12,000	10,500	10.53	11 Dec, 21	29,000	28,206

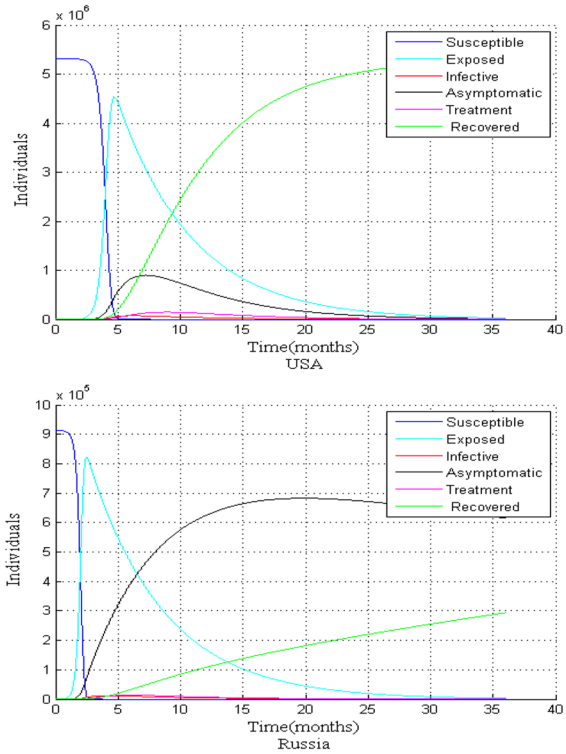


Fig. 7 Evolutions of the system with parameters of the Model Eqs. (1–6). The parameters p and δ are used 0.12 and 0.627 for the USA and 0.20 and 0.871 for Russia, respectively

Overall, our findings suggest that the dramatical reduction in total outbreak size and peak prevalence is possible by decreasing the transmission rate. In future work in this direction, the proposed model can be expanded by considering some other compartments as well as can be used in investigating optimal control strategies. Besides, this research work may become useful for health professionals and policy makers to take necessary measures against the future impact of coronavirus even when the vaccine is available. The insights of the fluctuations across different countries and different waves of the epidemic may help to understand different scenarios of the epandemic, as well as to define policies to successfully mitigate the disease spread.

A sophisticated study considering the temperature condition of the country, demographic information, the strength of public health facilities and the governmental interventions taken to suppress the transmission may provide more useful information in tackling and designing more effective control strategies against any future epidemic threat. Our work should encourage any such future works in this direction. Further, estimated epidemic parameters during the first and second waves, as well as, the insights of the fluctuations across different countries and different waves of the epidemic are expected to be useful in post Covid-19 analysis for respective countries' policy makers in designing more effective control strategies in future.

Acknowledgments The authors would like to thank the anonymous reviewers and the Honorable editor for their careful reading, useful comments and constructive suggestions for the improvement of the manuscript of the present research work.

Authors' contributions Ashabul Hoque contributed to formal analysis, writing—review and editing, supervision. Abdul Malek contributed to model development, numerical solution and graph preparation. KM Rukhsad Asif Zaman contributed to data analysis and validation.

Funding No funding was received for this study.

Availability of data and material All the data used in this manuscript can be found on WHO) Worldometer, and Johns Hopkins University's database.

Data accessibility Data can be retrieved from https://github.com/CSSEGISandData/COVID-19.

Declarations

Conflict of interest The author declares no conflict of interest.

References

1. Yousef, A., Taghdir, M., Sepandi, M.: The estimate of the basic reproduction number for novel coronavirus disease (COVID-19): A systematic review and meta-analysis. J. Prev. Med. Public Health (2020). https://doi.org/10.3961/jpmph.20.076
2. Kuddus, M.A., Rahman, A.: Analysis of COVID-19 using a modified SLIR model with nonlinear incidence. Res. Phys. **27**, 104478 (2021)
3. Lobato, F.S., Libotte, G.B., Platt, G.M.: Mathematical modelling of the second wave of COVID-19infections using deterministic and stochastic SIDR models. Nonlinear Dyn. **106**, 1359–1373 (2021). https://doi.org/10.1007/s11071-021-06680-0
4. Toda, A.A.: Susceptible-Infected-Recovered (SIR) dynamics of COVID-19 and economic impact. CEPR Covid Econ. **1**, 43–63 (2020)
5. Wu, J.T., Leung, K., Leung, G.M.: Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet **395**(10225), 689–697 (2020)
6. Yang, C., Wang, J.: A mathematical model for the novel coronavirus epidemic in Wuhan, China. Math. Bios. Eng. **17**(3), 2708–2724 (2020)
7. Kuniya, T.: Prediction of the epidemic peak of coronavirus disease in Japan. J. Clin. Med. **9**(3), 789 (2020). https://doi.org/10.3390/jcm9030789
8. Shim, E., Tariq, A., Choi, W., Lee, Y., Chowel, G.: Transmission potential and severity of COVID-19 in South Korea. Int. J. Infect. Dis. **93**, 339–344 (2020)
9. Remuzzi, A., Rmuzzi, G.: COVID-19 and Italy: what next? Lancet **395**, 1225–1228 (2020)
10. Stephen, A.L., Grantzet, K.H.: The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med. https://doi.org/10.7326/M20-0504
11. Liu, Y., Gayle, A.A., Wilder-Smith, A., Rocklöv, J.: The reproductive number of COVID-19 is higher compared to SARS coronavirus. J. Trav. Med. (2020). https://doi.org/10.1093/jtm/taaa021
12. Marimuthu, S., Joya, M., Malavikaa, B., Nadaraja, A., Asirvathamb, E.S., Jeyaseelana, L.: Modelling of reproduction number for COVID-19 in India and high incidence States. Clin. Epid. Global Health (2020). https://doi.org/10.1016/j.cegh.2020.06.012
13. Yu, X.: Modeling return of the epidemic: Impact of population structure, asymptomatic infection, case importation and personal contacts. Trav. Med. Infect Dis. (2020). https://doi.org/10.1016/j.tmaid.2020.101858
14. Yin, S., Zhang, N.: Prevention schemes for future pandemic cases: mathematical model and experience of interurban multiagent COVID-19 epidemic prevention. Nonlinear Dyn. **104**, 2865–2900 (2021). https://doi.org/10.1007/s11071-021-06385-4.(0123456789)
15. Tang, B., Wang, X., Li, Q., Bragazzi, N.L., Tang, S., Xiao, Y.: Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions. J. Clin. Med. **9**, 462 (2020)
16. Biswas, S.K., Ghosh, J.K., Sarkar, S., Ghosh, U.: COVID-19 pandemic in India: a mathematical model study. Nonlinear Dyn. **102**, 537–553 (2020)
17. Imai, N., Cori, A., Dorigatti, I., Baguelin, M., Donnelly, C.A., Riley, S.: Report 3: transmissibility of 2019-nCoV. (2020). https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news–wuhan-coronavirus/
18. Ahmad, Z., Arif, M., Ali, F., Khan, I., Nisar, K.S.: A report on COVID-19 epidemic in Pakistan using SEIR fractional model. Sci. Rep. **10**, 22268 (2020). https://doi.org/10.1038/s41598-020-79405-9
19. Pedersen, M.G., Meneghini, M.: Quantifying undetected COVID-19 cases and effects of containment measures in Italy: predicting phase 2 dynamics. https://doi.org/10.13140/RG.2.2.11753.85600
20. Hoertel, N., Blachier, M., Blanco, C., Olfson, M., Massetti, M., Rico, M.S., Limosin, F., Leleu, H.: A stochastic agent-based model of the SARS-CoV-2 epidemic in France. Nat. Med. **26**, 1417–1421 (2020)
21. Dietz, K.: The estimation of the basic reproduction number for infectious diseases. Stat. Methods Med. Res. **2**(1), 23–41 (1993)
22. Breban, R., Vardavas, R., Blower, S.: Theory versus data: how to calculate R0? PLoS ONE **2**(3), e282 (2007)
23. Read, J.M., Bridgen, J.R.E., Cummings, D.A.T., Ho, A., Jewell, C.P.: Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv (2020). https://doi.org/10.1101/2020.01.23.20018549
24. Beenstock, M., Dai, X.: The natural and unnatural histories of Covid-19 contagion. CEPR Covid Econ-10. **10**, 87–115 (2020)
25. Hong, H., Wang, N., Yang, J.: Implications of stochastic transmission rates for managing pandemic risks. NBER Working Paper 27218 (2020)
26. Chaves, L.F., Hurtado, L.A., Rojas, M.R., Friberg, M.D., Rodriguez, R.M., Avila-Aguero, M.L.: Covid-19 basic reproduction number and assessment of initial suppression policies in Costa Rica. Math. Model Nat. Phenom. (2020). https://doi.org/10.1051/mmnp/2020019
27. Hong, H.G., Li, Y.: Estimation of time-varying reproduction numbers underlying epidemiological processes: a new statistical tool for the COVID-19 pandemic. PLoS ONE **15**(7), e0236464 (2020). https://doi.org/10.1371/journal.pone.0236464
28. Khosravi, A., Chaman, R., Rohani-Rasaf, M., Zare, F., Mehravaran, S., Emamian, M.H.: The basic reproduction number and prediction of the epidemic size of the novel

coronavirus (COVID-19) in Shahroud, Iran. Epid Infect **148**, e115 (2020). https://doi.org/10.1017/S0950268820001247
29. Zahiri, A.P., Nasab, S.R., Roohi, E.: Prediction of peak and termination of novel coronavirus Covid-19 epidemic in Iran. medRxiv. (2020). https://doi.org/10.1101/2020.03.29.20046532
30. Ranjan, R., Sharma, A., Verma, M.K.: Characterization of the second wave of COVID-19 in India. 2021. https://doi.org/10.1101/2021.04.17.21255665
31. Iftimie, S., Lo'pez-Azcona, A.F., Vallverdu', I., Herna'ndez-Flix, S., de Febrer, G., Parra, S.: First and second waves of coronavirus disease-19: A comparative study in hospitalized patients in Reus, Spain. PLoS ONE **16**(3), e0248029 (2021). https://doi.org/10.1371/journal.pone.0248029
32. Vasconcelos, G.L., Brum, A.A., Almeida, F.A.G., Macedo, A.M.S., Duarte-Filho, G.C., Ospina, R.: Standard and anomalous second waves in the COVID-19 pandemic. medRxiv. (2021). https://doi.org/10.1101/2021.01.31.21250867
33. Salyer, S.J., Maeda, J., Sembuch, S., Kebede, Y., Tshangela, A., Ouma, A.O., Nkengasong, J.: The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. Lancet **397**, 1265–1275 (2021). https://doi.org/10.1016/S0140-6736(21)00632-2
34. Nesteruk, I.: Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China. medRxiv 2020
35. Chen, Y., Cheng, J., Jiang, Y., Liu, K.: A time delay dynamical model for outbreak of 2019-nCoV and the parameter identification. (2020) arXiv:00418
36. Peng, L., Yang W., Zhang, D., Zhuge, C., Hong, L.: Epidemic analysis of COVID-19 in China by dynamical modeling. (2020). arXiv:06563
37. Xu, C., Dong, Y., Yu, X., Wang, H., Tsamlag, L., Zhang, S., Chang, R., Wang, Z., Yu, Y., Long, R., Wang, Y., Xu, G., Shen, T., Wang, S., Zhang, X., Wang, H., Cai, Y.: Estimation of reproduction numbers of COVID-19 in typical countries and epidemic trends under different prevention and control scenarios. Front. Med. (2020). https://doi.org/10.1007/s11684-020-0787-4
38. van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Bios. **180**, 29–48 (2002)
39. Arino, J., Brauer, F., van den Driessche, P., Watmough, J., Wu, J.: A final size relation for epidemic models. Math. Bios. Eng. **4**(2), 159 (2007)
40. Khailaie, S., Mitra, T., Bandyopadhyay, A., Schips, M., Mascheroni, P., Vanella, P., Lange, B., Binder, S., Meyer-Hermann, M.: Development of the reproduction number from coronavirus SARS-CoV-2 case data in Germany and implications for political measures. BMC Med. **19**, 32 (2021). https://doi.org/10.1186/s12916-020-01884-4
41. He, H., Peng, Y., Sun, K.: SEIR modeling of the COVID-19 and its dynamics. Nonlinear Dyn. (2020). https://doi.org/10.1007/s11071-020-05743-y
42. Motozono, C., Toyoda, M., Zahradnik, J., Saito, A., Nasser, H., Tan, T.S., Ngare, I., Kimura, I., Uriu, K., Kosugi, Y., Yue, Y.: SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. Cell Host Microbe **29**(7), 1124–1136 (2021)
43. Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., Robertson, D.L.: SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. **19**(7), 409–424 (2021)
44. Malavige, G.N., Jeewandara, C., Ogg, G.S.: Dysfunctional innate immune responses and severe dengue. Front. Cell. Infect. Microbiol. **10**, 600 (2020)
45. Ben-Shachar, R., Koelle, K.: Minimal within-host dengue models highlight the specific roles of the immune response in primary and secondary dengue infections. J. R. Soc. Interface **12**(103), 20140886 (2015)
46. Nikin-Beers, R., Ciupe, S.M.: The role of antibody in enhancing dengue virus infection. Math. Biosci. **263**, 83–92 (2015)
47. Majumdar, P., Niyogi, S.: SARS-CoV-2 mutations: the biological trackway towards viral fitness. Epidemiol Infect. **149**, (2021)