Mathematical Biosciences and Engineering

MBE, 18 (6): 9775–9786.
DOI: 10.3934/mbe.2021479
Received: 18 August 2021
Accepted: 25 October 2021
Published: 05 November 2021

http://www.aimspress.com/journal/MBE

Research article

Forecast of the COVID-19 trend in India: A simple modelling approach

Haitao Song [1], Guihong Fan [2], Shi Zhao [3], Huaichen Li[4], Qihua Huang [5] and Daihai He [6,*]

[1] Complex Systems Research Center, Shanxi University, Taiyuan 030006, China
[2] Department of Mathematics, Columbus State University, Columbus 31907, USA
[3] JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China
[4] Department of Pulmonary and Critical Care Medicine, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China
[5] School of Mathematical and Statistical Sciences, Southwest University, Chongqing 400715, China
[6] Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China

* **Correspondence:** Email: daihai.he@gmail.com.

Abstract: By February 2021, the overall impact of the COVID-19 pandemic in India had been relatively mild in terms of total reported cases and deaths. Surprisingly, the second wave in early April becomes devastating and attracts worldwide attention. Multiple factors (e.g., Delta variants with increased transmissibility) could have driven the rapid growth of the epidemic in India and led to a large number of deaths within a short period. We aim to reconstruct the transmission rate, estimate the infection fatality rate and forecast the epidemic size. We download the reported COVID-19 mortality data in India and formulate a simple mathematical model with a flexible transmission rate. We use iterated filtering to fit our model to deaths data. We forecast the infection attack rate in a month ahead. Our model simulation matched the reported deaths well and is reasonably close to the results of the serological study. We forecast that the infection attack rate (IAR) could have reached 43% by July 24, 2021, under the current trend. Our estimated infection fatality rate is about 0.07%. Under the current trend, the IAR will likely reach a level of 43% by July 24, 2021. Our estimated infection fatality rate appears unusually low, which could be due to a low case to infection ratio reported in previous study. Our approach is readily applicable in other countries and with other types of data (e.g., excess deaths).

Keywords: COVID-19; India; mathematical modelling; iterated filtering; forecast

1. Introduction

The COVID-19 pandemic has lasted for more than 1.5 years now. The severity of the impact of COVID-19 varied wildly across nations [1]. While the pandemic started to slow down in most developed countries due to the combined effects of high infection attack rate and high vaccination coverage, the impact in India had been mild in terms of the total number of deaths and cases per capita, until April 2021. An ongoing second wave of the COVID-19 devastates India 'unexpectedly' since early April 2021.

The unfold of the COVID-19 in India is as follows. On January 30, 2020, the first case of COVID-19 was reported in Thrissur, India [2]. On March 12, 2020, the first COVID-19 fatality in India was reported [3]. On March 31, 2020, a Tablighi Jamaat religious congregation event in Delhi resulted in a quarantine of 22,000 people. In March 2020, the cluster-containment campaign was carried out. However, by March 15, 2020, only 10% of the testing capacity had been used per day in India. By early July 2020 [4], the seroprevalence reached 57% or 16% among inhabitants in three slums or other areas of Mumbai, which is the commercial capital of India. These high infection attack rates (IAR, the proportion of the population got infected) caught wide media attention. A supermodel for COVID-19 progression developed by a government-appointed committee [5] estimated that the COVID-19 (first wave) peaked in October 2020 and it would be under control by February 2021. Unexpectedly, a new variant, named Lineage B.1.617 (later as Delta variant for sub-lineage B.1.617.2), was detected in India in October 2020 [6,7] and is blamed as one of the factors for the second wave. In October 2020, B.1.617.3 (the first subline-age of this variant) was first detected in India. Subsequently, B.1.617.1 and B.1.617.2 were detected in December 2020 [8]. From February 2021, new cases of B.1.617 increase rapidly in India [6]. From the global initiative on sharing all influenza databases [9], B.1.617 contributed to 63.6% infections of COVID-19 in India in April 2021.

The early stage of COVID-19 epidemic mitigation in India was promising. As reported in [10], it had the highest number of daily tests in the world by the third quarter of the year 2020, although the cases and deaths per capita are not particularly high. The vaccination program was also initiated on 16 January 2021. By February 2021, the confirmed cases of COVID-19 had dropped to 9,000 people per day. However, the pandemic changed dramatically thereafter. A major second wave of the COVID-19 emerged in India in early April 2021. By the end of April, an average of 300,000 daily new cases and 2000 daily deaths were reported. Multiple factors (e.g., Delta variant) could have caused the rapid expansion of the epidemic in India and caused a large number of deaths [11,12]. As of May 3, 2021, nearly 20 million confirmed cases and 218,959 deaths were reported in total. During the period of the second wave in India, test positivity has increased dramatically from 2% on March 1 to 22% on May 1.

Through serological surveys, the sero-prevalence among people ten years or older was estimated at 6.6% and a cumulative 74.3 million people had been infected with COVID-19 by August 2020 [13]. However, only 3,621,245 cases and 64,469 deaths were reported by August 2020 [10]. A survey [14] conducted by the Indian Council of Medical Research (ICMR) found that 21.4% of 28,589 surveyed people above 18 years old had been infected by Feb. 4, 2021. These serological studies found a much higher IAR than reported cases, namely a very low ascertainment rate or infection to case ratio. Therefore, one needs to take into account these serological studies into modelling to yield reasonable IAR estimation and useful forecast for pandemic mitigation.

Mathematical modelling can be successfully used to explain and predict the spread of infectious

diseases [15,16]. In this work, we adopt a simple compartmental epidemic model. The model assumes a time-dependent transmission rate due to the changes in human behavior and the implementation of control measures. Thus, models for forecast need to incorporate time-varying transmission rates. Other epidemic parameters, including the generation interval, the reporting rate, and the infection fatality rate may vary as well. But for the sake of simplicity of the model and identification issues (the change might not be identified via such model fitting), we focus on the time-varying transmission rate and assume other parameters to be constant.

2. Materials and methods

We download daily reported COVID-19 cases and deaths from [10]. To have a better understanding of the historical seasonality of respiratory diseases, we also download weekly reported influenza cases from [17]. We adopt a susceptible-exposed-infectious-hospitalized-death-recovered (SEIHDR) model with a time-varying transmission rate:

$$\dot{S} = -\frac{\beta SI}{N},$$

$$\dot{E} = \frac{\beta SI}{N} - \sigma E,$$

$$\dot{I} = \sigma E - \gamma I,$$

$$\dot{H} = \theta\gamma I - \kappa H,$$

$$\dot{D} = \pi\kappa H,$$

$$\dot{R} = (1-\theta)\gamma I + (1-\pi)\kappa H.$$

Here, the compartments *S*, *E*, *I*, *H*, *D*, and *R* denote susceptible, exposed, infectious, hospitalized (including severe case but not hospitalized), total death, and recovered individuals, respectively. Here we used 'hospitalized' to denote a medium class (e.g., severe cases) that has a probability to die. This medium class serves as a delay class to account for the delay from loss of infection to death. Parameter $\beta(t)$ is the transmission rate, σ is the infectiousness emergence rate, γ is the infectiousness disappearance rate,κ is the removed rate. We set σ, γ and κ at 0.5 per day, 1/3 per day and 1/14 per day, respectively, such that the mean latent period, infectious period, and mean delay from loss of infection to death are 2 days, 3 days, 14 days (i.e., reciprocal of the rate, these are theoretical values when the time step size is infinitely small. In practice when the time step in integration is 1 day, the three durations are slightly longer than 2 days, 3 days, and 14 days). These are in line with observations [18–21]. The choice of σ and γ are due to the restriction that the sum of the mean latent period and the mean infection period equals the mean generation time, which is about 5 days [22]. θ is the ratio of *H* cases out of all infected cases and π is the proportion of deaths out of *H* cases. All parameters are constant except $\beta(t)$ being time-varying. Since we did not fit the model to the daily hospitalized cases, the exact definition of *H* class is not relevant (which is a medium class to take into account the delay between loss of infectiousness to death). We consider that $\theta \ll 1$ and $\pi \ll 1$. The product $\pi\theta$ equals the infection fatality rate. These two cannot be disentangled with purely death data and without additional reliable data (e.g., hospitalization). Thus, we fix π at 0.01 and estimate θ. The result is insensitive to the value choice of π. Our main model includes a latent period. However, for the sensitivity analysis, we also test a model with a longer

infectious period at 5 days but no latent period, thus the generation time is still 5 days.

Based on previous works [23–25], we set $\beta(t) = \exp(\text{cubic spline with } n_m \text{ nodes})$, i.e., an exponential cubic spline with n_m nodes evenly distributed over the study period. Using standard model selection technique, we found n_m=9 attains the smallest the second-order Akaike Information Criterion (AICc). Thus n_m=9 is used in this work. The time step size is one day and D_t is the daily number of deaths. The reported deaths were defined as C_t with

$$C_t \sim \text{NegativeBinomial}(\text{mean} = D_t, \text{variance} = D_t(1 + \tau D_t)).$$

Here, the parameter τ denotes the overdispersion and accounts for the measurement noise due to surveillance and heterogeneity among individuals.

The basic reproductive number $\mathcal{R}_0(t) = \beta(t)/\gamma$ is a function of time. We fit the model to reported deaths with state-of-the-art iterated filtering [26,27], and assume that death report is relatively insensitive to testing policy, testing effort, and testing availability. Although an even better way would be to use the excess deaths [28], our modelling approach is readily applicable to excess deaths.

We download daily death data in India from the World Health Organization website for the period from March 12, 2020, to July 24, 2021. We divide the time interval of data into two parts: the part before May 15, 2021 is called the training part, and the pat after May 15, 2021 is called the testing part. We fit our model to the whole time series (including both training and testing parts), but we set the testing part to be N.A. (not applicable). This operation will force the log-likelihood contribution from the testing part to be zero. Namely, the testing part data play no role in the fitting. For sensitivity analysis, we test the cut-off date at May 25, 2021, instead of May 15, 2021.

Our cubic spline spans over whole time interval of data (both training part and forecast part). Thus, the transmission rate over the training part is estimated from the data, the transmission rate over the forecast part is a natural extension of that over the training part. We can simulate our model and yield both the training part and the testing part (ie, the forecast part), since we have the initial state (end state of the training part) and transmission rate (natural extension) for the forecast part.

Occasionally, there are abrupt data points that may be due to retrospective checking of late reported deaths in previous days. Denote x_i as the reported deaths on days i. If $x_i > 1.5(x_{i-1} + 50)$, we call x_i abnormal data. We replace x_i with $0.25(x_{i-3} + x_{i-2} + x_{i-1} + x_i)$, and allocate the extra deaths to all days proportional to their current value. This smoothing only affects five data points, therefore will not affect our fitting and forecast.

3. Results

Figure 1 plots daily reported COVID-19 cases and deaths in India, compared with reported daily (converted from weekly by simply divided by 7) influenza laboratory confirmation (mean level) in the previous five years and the current year. Influenza persisted yearly in the previous five years with peaks in March and April, which could be associated with religious holidays and/or large gatherings (e.g., elections). For instance, Holi (a popular ancient Hindu festival, also known as the Festival of Love) is on March 29, and Haridwar Kunbh Mela is celebrated in April 2021. At the same time, the elections in India in 2021 were held in April 2021. The cohort of these large gatherings and celebrations is a driving factor of the spread of COVID-19 and causes a dramatic increase of new cases and deaths in India. We show the raw case-fatality-ratio using the brown dashed curve. The raw case-fatality-ratio is defined as the ratio of cumulative COVID-19 deaths divided by the cumulative cases, where we ignore the delay between deaths reports and cases reports. The impact of vaccination

on the raw case-fatality-ratio at the population level is not yet evident.

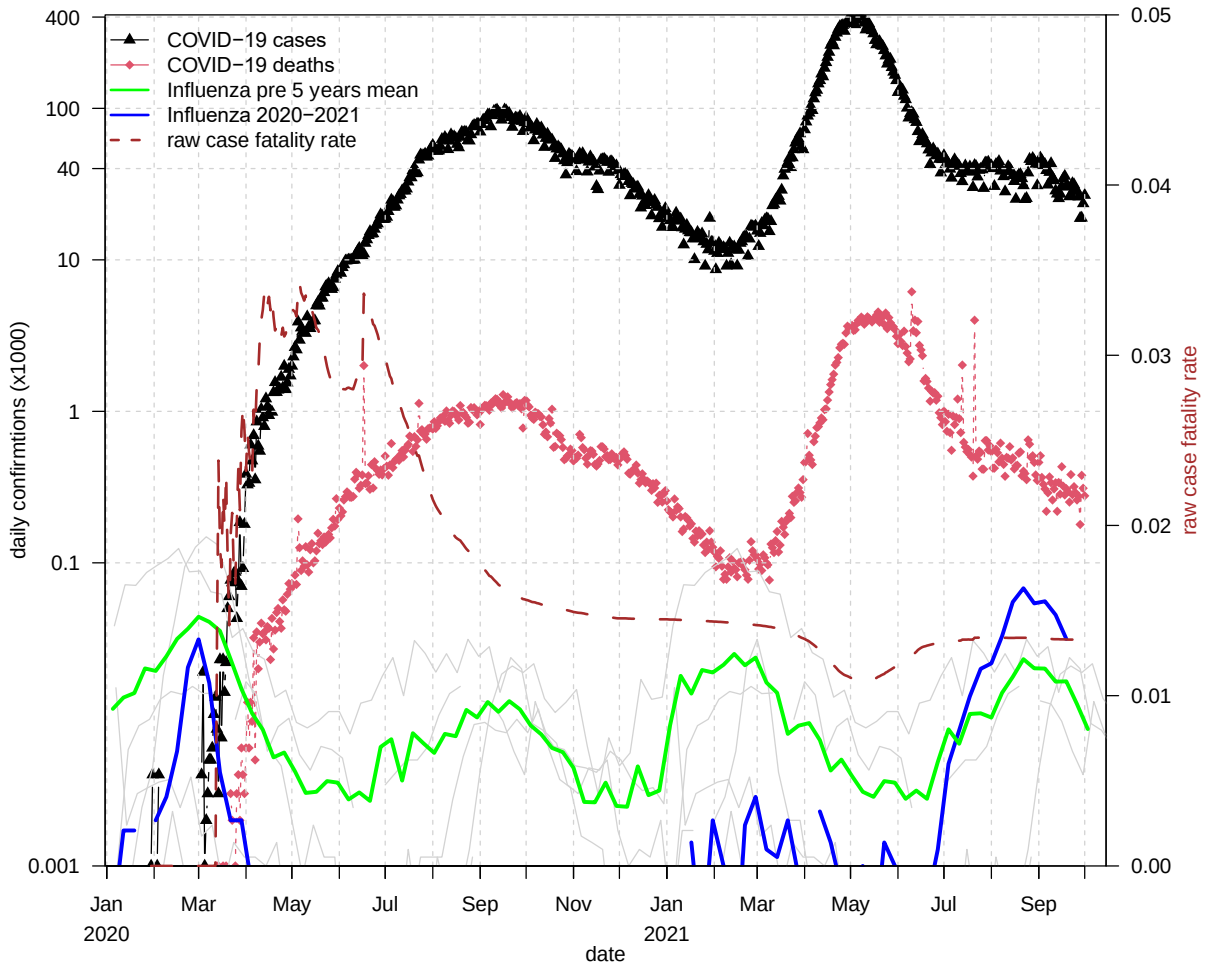

Figure 1. Reported daily Covid-19 cases and deaths in India, compared with reported mean daily (converted from weekly by simply divided by 7) influenza laboratory confirmations in the previous five years (green thin curve) and the current year (blue bold curve). Influenza persists through a year in the previous five years with a peaking in March-April, which could be associated with religious holidays and/or large gatherings (e.g., elections). The brown dashed curve indicates the raw case-fatality-ratio, i.e., cumulative deaths divided by cumulative cases (where we omitted the time delay between deaths and case reporting).

In Figure 2, we compare the population standardized report of COVID-19 deaths in four countries. The daily death per capita in India is still at a relatively low level compared to the other three countries (Figure 2). However, from the beginning of April 2021, the rapid elevation of the second wave of COVID-19 in India is devastating with 300,000 daily new cases for 14 consecutive days and 3,000 daily deaths for 5 consecutive days. Such a dramatic increase in daily cases attracted worldwide attention and global concern.

Figure 2 also shows that the vaccination coverage rate is low in India. The vaccination program was launched on January 16, 2021 in India, and 4 million doses per day were administered by April 2021. However, the overall coverage of vaccines at the population level is still low. Thus, we omit

the effect of vaccination in our model. The effect of vaccination is to decrease the effective reproductive number by reducing the susceptible pool. Some states in India were unable to begin vaccination due to the shortage of vaccination supplies. The shortage of vaccination supplies could make the epidemic of COVID-19 in India even worse.

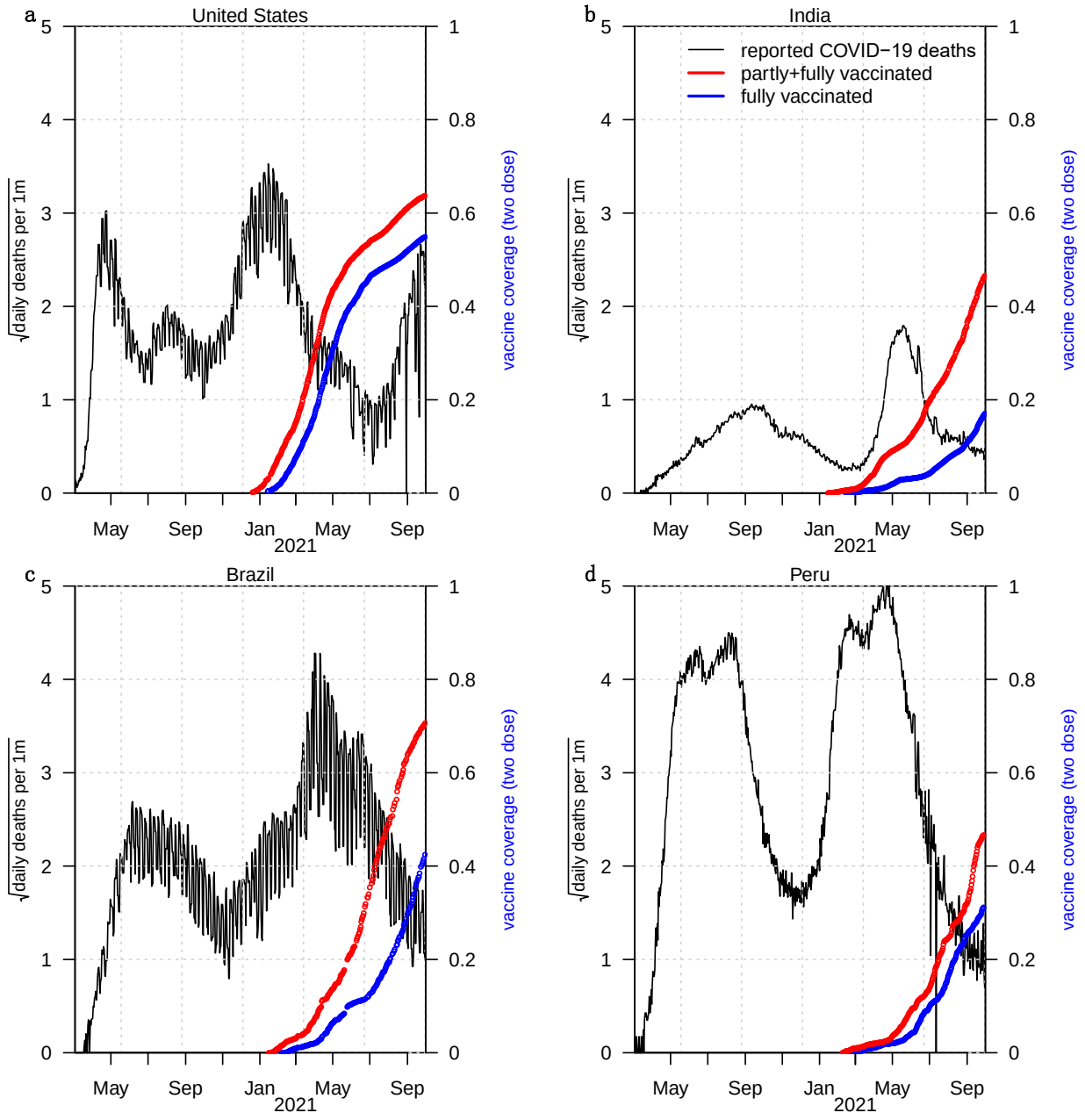

Figure 2. Comparison of the population standardized report of Covid-19 deaths (black curve) in four countries and vaccination coverage, fully vaccination (i.e., two-doses for most vaccines or one-dose for Johnson and Johnson's Janssen Vaccine, or one-dose after confirmed infection) per capita as blue circle, and vaccinated (both one dose and two dose) per capita as red circles. The vaccinated population (both partly and fully) in India reached 50% in September, 2021. But in our study period (see later), a training period before May 15, 2021, the vaccinated population on y reached 10.13%, thus we omit the vaccination impact in our model for simplicity.

Figure 3 shows the fitting and forecast results in four scenarios: with a latent period (a,b) or without a latent period (c,d), cut-off at May 15, 2021 (a,c) vs cut-off at May 25, 2021. Our model simulation well matched the reported deaths in each scenario, although the later the cut-off, the better the forecast. We call the scenario in panel (b) our focal scenario. The reconstructed transmission rate (in the unit of $\mathcal{R}_0(t) = \beta(t)/\gamma$, blue dashed curve) showed a decreasing trend in March-April 2020 and a wave pattern in March-April 2021 with a peak on April 1, 2021. The daily reported deaths will peak around May 15, 2021. The infection attack rate reached43% by Jul 24, 2021, under the current trend. We showed the infection attack rate and the effective reproductive number over time in Figure 4. The effective reproductive number is $\mathcal{R}_e(t) = \mathcal{R}_0(t)S(t)$. Our estimated IAR is about 10% by Feb 2021, which is lower than the estimates of the serological study of ICMR 21.4% [14]. However, our estimated IAR is about 5.4% by August 2020, which is close to the serological study of 6.6% [13].

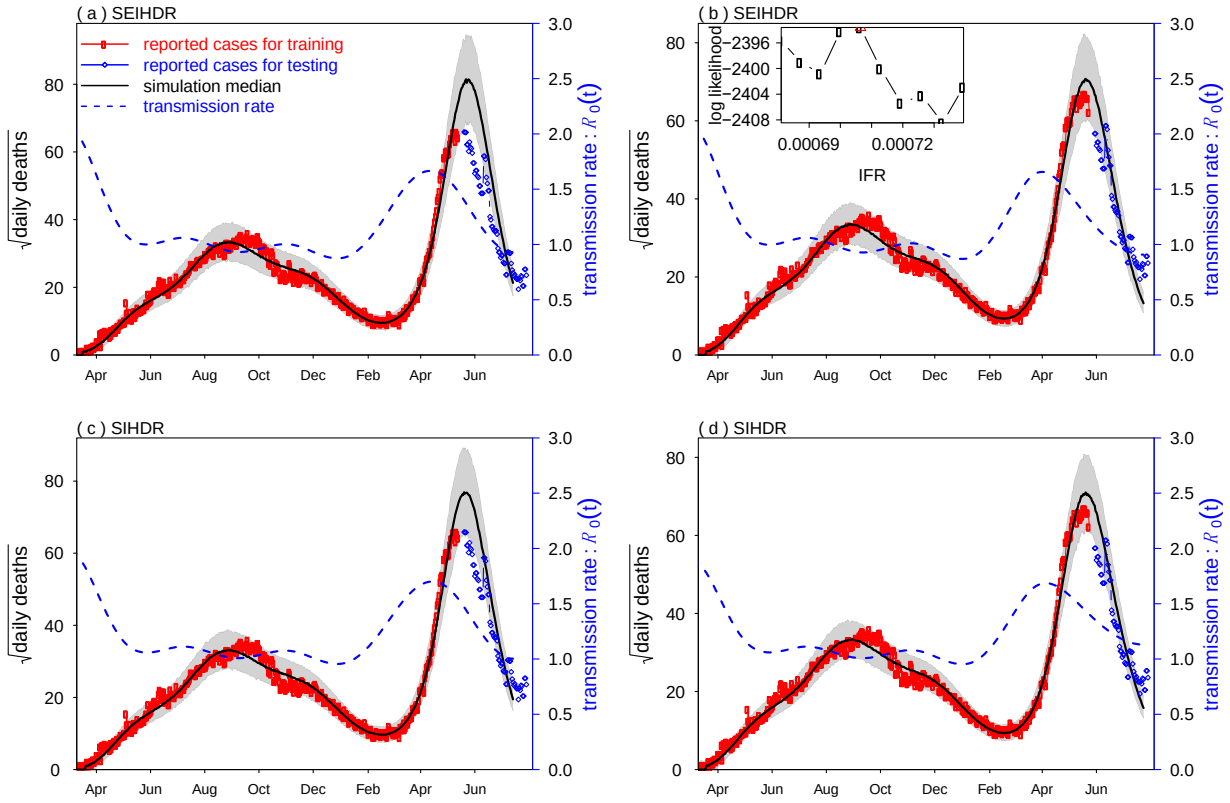

Figure 3. Best fitting results with the maximum log-likelihood in four scenarios. Fitting an SEIHDR model in panels (a,b) or a reduced SIHDR without the latent period in panels (c,d) to the reported death in India with a flexible transmission rate. In panels (a,c), the cut-off date for training and testing parts is May 15, 2021. In panels (b,d), the cut-off date is May 25, 2021. The red circles (/blue diamonds) denote the training part (/ the testing part) of daily reported COVID-19 deaths. The black curve (/shaded region) denotes the median (/ the 95% range) of 1000 model simulations. The blue dashed curve denotes the reconstructed transmission rate in the unit of $\mathcal{R}_0(t)$, which is different from the effective reproductive number $\mathcal{R}_e(t)$, and the relationship between the two is $\mathcal{R}_e(t) = \mathcal{R}_e(t)S(t)$. The inset panel of (b) showed the profile of log-likelihood as a function of the infection fatality rate (including ratio due to under-reporting of deaths).

We note that the transmission rate (in terms of $\mathcal{R}_0(t)$) never reach 0.5 or below during the whole time before the second wave (most time around 1). In our fitting in Figure 3, we assume the lower bound of $\mathcal{R}_0(t)$) after May 2021, in the short term of the future, it is bounded by the lower bound before May 2021. The fact that before the second wave, the lower bound of $\mathcal{R}_0(\mathrm{t})$ was around 1, reflected the stringent level (or compliance level) of control in India. Thus, it is probably unrealistic to expect $\mathcal{R}_0(\mathrm{t})$ may go to much lower than that in the short term of the future.

The testing data (blue diamond in Figure 3) largely fall in the 95% CI of the simulation in all four scenarios, although the result in the focal scenario in Figure 3b is slightly better than the other three scenarios.

We performed the sensitive analysis in Figures 3 and 4 by considering multiple scenarios. Our method is insensitive to the inclusion of a latent class. And slightly extending the training period improves the forecast performance. We provide more sensitive analysis (e.g., Partial rank correlation coefficients, PRCC, test in supplementary.).

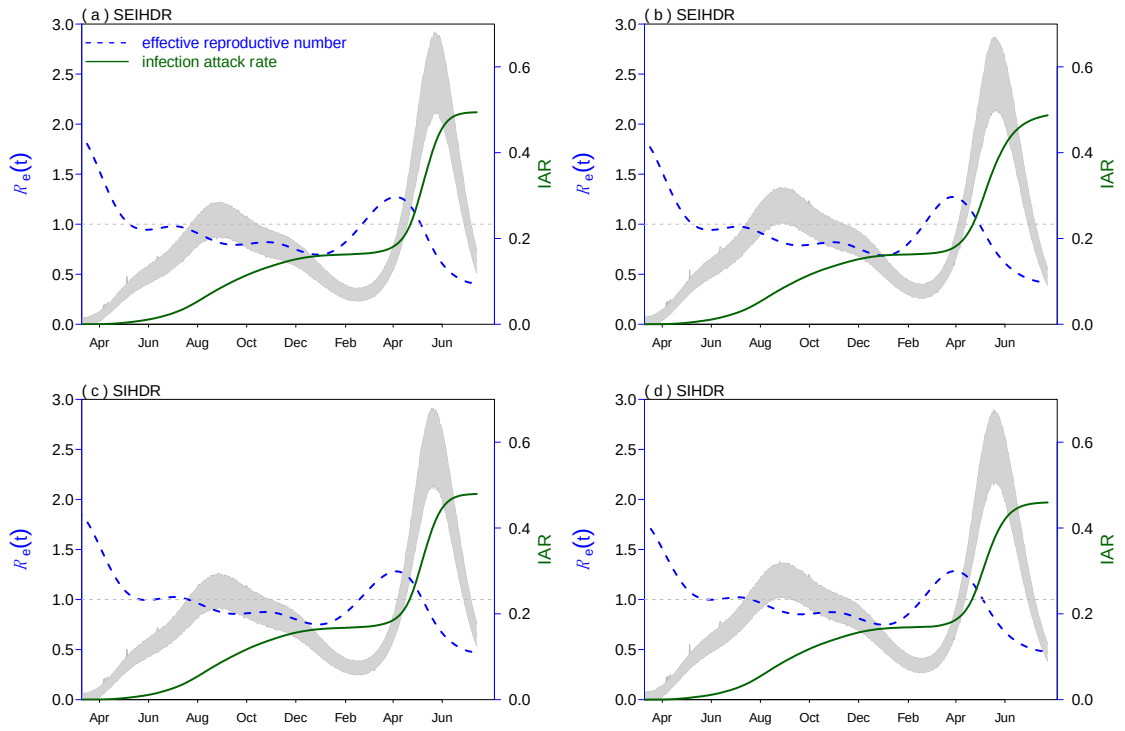

Figure 4. The reconstructed effective reproductive number (blue dashed curve) and infection attack rate (dark green solid curve) as functions of time. The shaded regions showed a 95% range of 1000 model simulations. These are the best results as in the four scenarios in Figure 3.

4. Discussion and conclusion

We proposed a simple model approach for modelling and forecasting the COVID-19 epidemic in India. The method can be readily applied to other countries. When the forecast part is relatively short, this method should work well. Hence, the reconstructed transmission rate naturally contains a part over the forecast period, which shows a decreasing trend in India. This decreasing trend mimics the

decreasing trend in April 2020. This decreasing trend means that the current spreading of COVID-19 in India has already started to slow down (e.g., on the second-order derivative).

The objective of this study is to model and forecast the transmission of COVID-19 in India. The estimated infection attack rate (IAR) could reach 43% by July 24, 2021. Huang et al. [29] predicted that 33,470,999 (upper bound) confirmed cases would be reported by May 31, 2021, which means a 2.4% infection attack rate (case/population). This is obviously low. If a 1/10 (or 1/20) reporting rate is considered, then the IAR becomes 24% (or 48%). This is largely reasonable. Namely, a reporting rate needs to be assumed. In our case, we explicitly modeled infection fatality rate (IFR) and assumed that the death data is relatively reliable. The estimated infection fatality rate is about 0.07%, which is significantly lower than a biologically reasonable level of 0.6%. For instance, Russell et al. [30] estimated an IFR of 0.6% in China. This implies that the COVID-19 death could be under-reported by a factor of 1/10 [31]. Thus, these estimates of IFR relied on the confidence of the COVID-19 death data. If death data are under-reported by a factor, then the IFR could be underestimated by the same factor.

The long-term prediction is difficult, due to the ever-changing reality like human behavioral change and governmental action. Here we proposed a simple modeling approach. We assumed that the transmission rate in the forecast part is a natural extension of the training part. Both parts constitute a whole (exp-) cubic spline function. Ranjan et al. [32] estimated the peak for the second wave to occur in mid-May 2021 with a daily count exceeding 0.35 million. However, on April 30, 2021, a total of 400,000 daily new cases and 3,500 daily deaths were reported, and there are 300,000 daily new cases for 14 consecutive days and 3,000 daily deaths for 5 consecutive days. The prediction in [32] is apparently lower than reality. According to the seroprevalence results [6], Murhekar et al. [33] estimated a cumulative 74.3 million infections while 3,621,245 cases were reported in India by August 2020, i.e., a factor 1 to 20. Thus, under-reporting of COVID-19 death could have caused a low IFR estimate here.

The strength of this work includes that we fit a simple model to the reported death which is believed insensitive to COVID-19 testing efforts compared to reported cases. Mild cases will less likely to be tested. Thus, fluctuation in the reported cases likely contains variation in testing efforts. We assume the transmission rate in the forecast period is a natural extension of that over the training part. Both the first and second derivatives are continuous in the transition point between training and forecast. This approach is novel. We compared our estimated IAR to the reported serological studies. To our knowledge, very few studies have done such a comparison. We argue that it is important that the model output is largely in line with the serological study at different time points.

The limitation of this work includes: we assume all parameters are constant except for the transmission rate; the model is for the whole country while we ignore the heterogeneity across regions. We only relied on the reported data and adopted a non-mechanistic cubic spline function type of transmission rate. Alternatively, one could consider explicitly incorporating all kinds of control measures (e.g., google mobility matrix, etc.) We ignored the effects of vaccination since the coverage was low (by May 15, 2021, 10.13% was vaccinated at least one dose) in India by far. Our model can be extended to cover a longer period and a vaccination component can be added (see [34]).

We argue that multiple factors drove the rapid growth of the second wave and caused a large number of deaths within a very short period in India. We found that the reproductive number has never been significantly lower than 1 in India which reflected that the stringent level (or compliance level) of control in India is less ideal. The fitting procedure is explained in method and more detail (with code) can be found in previous works [25,35]. In Figures 2 and 3, we showed the square root of

deaths (or deaths per million), thus one can see more clearly the detail in the troughs (i.e., small numbers) of the waves as well as the peaks (i.e., large numbers).

Many serological studies found that a very large proportion of the population has been infected by early 2021, while the reported death per capita is low due to under-reporting, or other unknown mechanisms. One can easily see that the IFR is very low in India using the deaths toll divided by the estimates IAR through serological study. Banaji [36] found a lower bound IFR in Mumbai at 0.1% with consideration of excess deaths (additional to reported COVID-deaths), which is not very far from our estimates here. Here we lack excess death data. Actually, one does not need a model to reach this conclusion. Given the total deaths 428309, and the infection attack rate of 43% (which is in line or even lower than many serological studies), $428309/(1.4\times10^9\times0.43) = 0.071\%$. This means our estimate is consistent. Either other unknown mechanisms or severe under-reporting is responsible for this low IFR.

Acknowledgments

We would like to thank the anonymous referees very much for their valuable comments and suggestions. The authors are supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (HKU C7123-20G), the Fund Program for the Scientific Activities of Selected Returned Overseas Professionals in Shanxi Province (20200001), the National Natural Science Foundation of China (12171291), the Key Research and Development Project in Shanxi Province (202003D31011/GZ), the Program for the Outstanding Innovative Teams (OIT) of Higher Learning Institutions of Shanxi, the Shanxi Scholarship Council of China (HGKY2019004), and the Scientific and Technological Innovation Programs (STIP) of Higher Education Institutions in Shanxi (2019L0082).

Conflict of interest

The authors declare that they have no competing interests.

References

1. J. Rui, Y. Niu, Q. Wang, W. Zhang, Z. Chen, Z. Zhao, et al., Containing the transmission of COVID-19: A modelling study in 160 countries, *Front. Med.*, **8** (2021), 701836. doi: 10.3389/fmed.2021.701836.
2. M. A. Andrews, B. Areekal, K. R. Rajesh, J. Krishnan, R. Suryakala, B. Krishnan, et al., First confirmed case of COVID-19 infection in India: A case report, *Indian J. Med. Res.*, **151** (2020), 490–492. doi: 10.4103/ijmr.IJMR_2131_20.
3. S. Kaushik, S. Kaushik, Y. Sharma, R. Kumar, J. P. Yadav, The Indian perspective of COVID-19 outbreak, *Virus. Dis.*, **31** (2020), 146–153. doi: 10.1007/s13337-020-00587-x.
4. S. Biswas, *India coronavirus:* More than half of Mumbai slum-dwellers had Covid-19, 2020. Available from: https://www.bbc.com/news/world-asia-india-53576653.
5. *India TV News Desk*, India can control COVID-19 by Feb 2021, festive laxity can lead to 26 lakh cases in a month: Govt panel, 2020. Available from: https://www.indiatvnews.com/news/india/coronavirus-peak-in-india-over-can-be-controlled-by-february-2021-says-govt-appointed-panel-vidyasagar-657984.

6. *Cov-lineages.org,* PANGO lineages Lineage B.1.617. Available from: https://cov-lineages.org/lineages/lineage_B.1.617.html.
7. A. Rambaut, E. C. Holmes, A. O'Toole, V. Hill, J. T. McCrone, C. Ruis, et al., A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology, *Nat. Microbiol.*, **5** (2020), 1403–1407. doi: 10.1038/s41564-020-0770-5.
8. *World Health Organization (WHO),* Weekly epidemiological update on COVID-19-27 April 2021. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---27-april-2021.
9. J. P. Sahoo, A. P. Mishra, K. C. Samal, Triple Mutant Bengal Strain (B. 1.618) of Coronavirus and the Worst COVID Outbreak in India, *Biotica Res. Today*, **3** (2021), 261–265.
10. *World Health Organization (WHO),* WHO Coronavirus (COVID-19) Dashboard. Available from: https://covid19.who.int/.
11. S. Biswas, *India coronavirus:* Delta plus India: Scientists say too early to tell risk of Covid-19 variant. BBC News, 2021. Available from: https://www.bbc.com/news/world-asia-india-57564560.
12. H. Ellyatt, *Health and science:* After being ravaged by the delta Covid variant, how is India doing now? CNBC, 2021. Available from: https://www.cnbc.com/2021/07/23/coronavirus-how-india-is-doing-now-after-delta-variant-spread.html.
13. M. Murhekar, T. Bhatnagar, S. Selvaraju, V. Saravanakumar, J. W. V. Thangaraj, N. Shah, et al., SARS-CoV-2 antibody prevalence in India: findings from the second nationwide household serosurvey, August-September 2020, *Available at SSRN,* 2020. doi: https://dx.doi.org/10.2139/ssrn.3715460.
14. J. Sriram, *Podcast, Coronavirus update:* ICMR's sero-surcey, the peak of COVID-19 in India, and new cases reported in Beijing. Available from: https://www.thehindu.com/podcast/coronavirus-update-icmrs-sero-survey-the-peak-of-covid-19-in-india-and-new-cases-reported-in-beijing/article31837397.ece.
15. D. J. D. Earn, P. Rohani, B. M. Bolker, B. T. Grenfell, A simple model for complex dynamical transitions in epidemics, *Science*, **287** (2000), 667–670. doi: 10.1126/science.287.5453.667.
16. M. J. Keeling, P. Rohani, *Modeling infectious diseases in humans and animals*, Princeton university press, 2011. doi: 10.1515/9781400841035.
17. *World Health Organization (WHO),* FluNet. Available from: https://www.who.int/tools/flunet.
18. Y. M. Bar-On, A. Flamholz, R. Phillips, R. Milo, Science Forum: SARS-CoV-2 (COVID-19) by the numbers, *eLife*, **9** (2020), e57309. doi: 10.7554/eLife.57309.
19. J. Griffin, M. Casey, A. Collins, K. Hunt, D. McEvoy, A. Byrne, et al., Rapid review of available evidence on the serial interval and generation time of COVID-19, *BMJ Open*, **10** (2020), e040263. doi: 10.1136/bmjopen-2020-040263.
20. A. Svensson, A note on generation times in epidemic models, *Math. Biosci.*, **208** (2007), 300–311. doi: 10.1016/j.mbs.2006.10.010.
21. W. M. de Souza, L. F. Buss, D. da Silva Candido, J. P. Carrera, S. Li, A. E. Zarebskial, et al., Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil, *Nat. Hum. Behav.*, **4** (2020), 856–865. doi: 10.1038/s41562-020-0928-4.
22. X. Tang, S. S. Musa, S. Zhao, S. J. Mei, D. H. He, Using proper mean generation intervals in modeling of COVID-19, *Front. Public Health*, **9** (2021), 691262. doi: 10.3389/fpubh.2021.691262.

23. D. He, S. Zhao, Q. Lin, S. S. Mus, L. Stone, New estimates of the Zika virus epidemic attack rate in Northeastern Brazil from 2015 to 2016: A modelling analysis based on Guillain-Barré Syndrome (GBS) surveillance data, *PLOS Neglect. Trop. D.,* **14** (2020), e0007502. doi: 10.1371/journal.pntd.0007502.
24. L. Stone, D. He, S. Lehnstaedt, Y. Artzy-Randrup, Extraordinary curtailment of massive typhus epidemic in the Warsaw Ghetto, *Sci. Adv.,* **6** (2020), eabc0927. doi: 10.1126/sciadv.abc0927.
25. S. Zhao, L. Stone, D. Gao, D. H. He, Modelling the large-scale yellow fever outbreak in Luanda, Angola, and the impact of vaccination, *PLOS Neglect. Trop. D.,* **12** (2018), e0006158. doi: 10.1371/journal.pntd.0006158.
26. D. He, E. L. Ionides, A. A. King, Plug-and-play inference for disease dynamics: Measles in large and small populations as a case study, *J. R. Soc. Interface,* **7** (2010), 271–283. doi: 10.1098/rsif.2009.0151.
27. E. L. Ionides, C. Bretó, A. A. King, Inference for nonlinear dynamical systems, *Proc. Natl. Acad. Sci.*, **103** (2006), 18438–18443. doi: 10.1073/pnas.0603181103.
28. *Centers for Disease Control and Prevention (CDC)*, Excess Deaths Associated with COVID-19. Available from: https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm.
29. S. Huang, J. G. Xu, F. Wei, Z. Peng, Z. Jin, Y. M. Xu, et al., The epidemic in India remains on an upward trend, which may cause total of 200 million cases in the world 2021. Available from: https://riph.nankai.edu.cn/2021/0502/c23093a355929/pagem.htm.
30. T. W. Russell, J. Hellewell, C. I. Jarvis, K. van Zandvoort, S. Abbott, R. Ratnayake, et al., Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020, *Eurosurveillance*, **25** (2020), 2000256. doi: 10.2807/1560-7917.ES.2020.25.12.2000256.
31. M. Krishman, COVID in *India:* Is India counting all its Covid-19 deaths?, 2021. Available from: https://www.rfi.fr/en/international/20210722-is-india-counting-all-its-covid-19-deaths.
32. R. Ranjan, A. Sharma, M. K. Verma, Characterization of the Second Wave of COVID-19 in India, medRxiv, 2021. doi: 10.1101/2021.04.17.21255665.
33. M. V. Murhekar, T. Bhatnagar, J. W. V. Thangaraj, V. Saravanakumar, M. S. Kumar, S. Selvaraju, et al., SARS-CoV-2 seroprevalence among the general population and healthcare workers in India, December 2020–January 2021, *Int. J. Infect. Dis.,* **108** (2021), 145–155. doi: 10.1016/j.ijid.2021.05.040.
34. H. Song, G. Fan, Y. Liu, x. Y. Wang, D. H. He, The second wave of COVID-19 in South and Southeast Asia and vaccination effects, Research Square, 2021. doi: 10.21203/rs.3.rs-877949/v1.
35. Q. Lin, A. P. Y. Chiu, S. Zhao, D. H. He, Modeling the spread of Middle East respiratory syndrome coronavirus in Saudi Arabia, *Stat. Methods Med. Res.*, **27** (2018), 1968–1978. doi: 10.1177/0962280217746442.
36. M. Banaji, Estimating COVID-19 infection fatality rate in Mumbai during 2020, medRxiv, 2021. doi: 10.1101/2021.04.08.21255101.