Is this beginning or the end of COVID-19 battle in India? A data-driven mathematical model-based analysis of outbreak

Anupam Singh[1*], Jhilik Dey[2], Shivam Bhardwaj[3]

[1]Centre for Data Analytics (SVSM), Shri Vaishnav Vidyapeeth Vishwavidyalaya (SVVV University), Indore, India. Email: anupam.nitb@gmail.com

[2]Cell Biology and Physiology Division, Indian Institute of Chemical Biology, Kolkata, India.

[3]Department of Management, Sage University, Indore, India

*Corresponding author

Abstract

The India has experienced early and harshest lockdown in response to the outbreak from 25th March 2020. However, an accurate estimation of progression of spread of the infection and level of preparedness to combat this disease are urgently needed. Using a data-based mathematical model, our study has made predictions on the number of cases that are expected to rise in India till 14th June 2020. The epidemiological data of daily cases have been utilized from 25th March (i.e., the first day of lockdown) to 23rd April 2020. In the study, we have stimulated two possible scenarios (optimistic and pessimistic) for the prediction. As per the optimistic approach of modeling, COVID-19 may end in the first week of June 2020 with a total of 77,900 infected cases including 2,442 fatalities. However, the results under the pessimistic scenario are bit scary. The results show that a total of 283,300 infected cases with 10,180 fatalities till 14th June. To win the battle, 10 weeks of complete lockdown is much needed at least in the infected states and the union territories of India. Alternatively, the isolation of clusters (hotspot regions) is required if India wants resume some essential activities.

Keywords: Mathematical model; Data Analysis; coronavirus; COVID-19; Prediction; Epidemiology

Introduction

The sudden outbreak of COVID-19 which is originated in Wuhan (China) has engulfed almost the whole world. On 11 March 2020, WHO declared Novel coronavirus disease (COVID-19)

outbreak as a pandemic and urged all countries to take necessary actions for treating, identifying active cases, and preventing transmission to save people's lives. The worldwide spread has created panic amongst the people and hence it has become a challenge for governments to control and mitigate COVID-19 as soon as possible. As of 23rd April 2020, there have been 2,549,632 confirmed cases with 175,825 deaths (World Health Organization).

Coronavirus is one of the families of various viruses that are responsible for respiratory diseases like cough and cold, severe acute respiratory syndrome (SARS), and Middle East Respiratory Syndrome (MERS). A zoonotic coronavirus (CoVs), provisionally known as 2019 novel betacoronavirus (2019- nCoV) or the severe acute respiratory syndrome coronavirus 2(SARS-CoV-2) has originated and identified in Wuhan city of China's Hubei Province and it is rapidly spreading to the rest of the world [8]. As in late December 2019, at Wuhan, some patients were admitted to the hospital and were diagnosed with severe pneumonia of unknown sets of reasons associated with clinical symptoms of high fever, dry cough, and dyspnea. Many of those clusters of people were linked to the Huanan seafood and wet animal market there [9]. Following activating surveillance and further etiologic and epidemiologic investigations of their respiratory samples, China informed this outbreak to the World Health Organization (WHO) on 31st December 2019 [13]. From the throat swab samples of the affected patients, the causative agent was identified by the Chinese Centre for Disease Control and Prevention (CCDC) on 7th January 2020 and the virus was named as Severe Acute Respiratory Syndrome Coronavirus 2(SARS-CoV-2) by the International Committee on Taxonomy of Viruses (ICTV). Thenceforth, the disease was named as Coronavirus Disease 2019 (COVID-19) by WHO on 11th February 2020(World Health Organization, WHO Director-General's Remarks at the Media briefing on 2019-nCoV on 11 February 2020, (2020) https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020).

The virus (SARS-CoV-2)

CoVs are large enveloped positively single-stranded RNA viruses that cause respiratory illness in animals as well as humans ranging from a mild cold to life-threatening pneumonia. In the Coronaviridae family, seven contagious virus affects human, four are endemic respiratory virus responsible for causing 15-30% of common cold infection(mild respiratory illness), two of them i.e. SARS-CoV-1 and MERS are considered to be epidemic while SARS-CoV-2 is now has been

considered as pandemic[7]. Studies from genetic sequencing of SARS-CoV-2 have shown 50% similarity with MERS-CoV and more than 80% similarity with SARS-CoV[16,9]. Due to a high level of genomic similarity it can be suggested that both SARS-CoV and SARS-CoV-2 have originated from the bat which acts as their natural primary host [3].

The disease (COVID-19)

The third CoV outbreak in the last two decades is the COVID-19 where a novel CoV (nCoV) has appeared in the human population by crossing species after Severe Acute Respiratory Syndrome CoV(SARS-CoV-1) outbreak in south China, 2002 and the Middle East Respiratory Syndrome CoV(MERS-CoV) outbreak in Saudi Arabia, 2012 [4]. The transmission of infection occurs through airborne droplets produced during coughing and sneezing by the affected person as well as asymptomatic persons both before and after the arrival of symptoms [6]. The incubation period of the virus ranges from 2-14 days which is a quiet longer duration available for the transmission of infection. Clinical symptoms associated with COVID-19 are generally sore throat, cough, fever, fatigue, and dyspnea. Sometimes, in some patients (often persons with comorbidities and elderly people), following these aforesaid symptoms after nearly 7 to 10 days, the disease may progress to Acute Respiratory Distress Syndrome(ARDS), pneumonia, respiratory collapse and multiple organ failure leading to death[13]. Hence, in some people there is a mild disease while in others it may progress lethal while some always remain asymptomatic. The case fatality rate of SARS-CoV-1 is nearly 9.6% while that of MERS-CoV is around 34% [4]. But for the COVID-19 fatality rate of the cases is approximately ranged between 2-3% which is less than that of the previous two outbreaks i.e. SARS-CoV and MERS-CoV but it has the highest transmission competence with less pathogenesis [13].

COVID -19 Outbreaks in India

India started the fight when the government evacuated 767 people through airlifting. Foreign travelers were the main reason behind the spread of viruses. The first case of COVID-19 in India was detected on 30th January 2020. A student who returned from Wuhan city was found to be coronavirus positive. Another patient who returned from Italy was found to be infected with this virus. Later on few Italian tourists and their guide were detected positive in Rajasthan. Since then there occurred a gradual rise in the number of cases in all regions of India. Initially New Delhi

and Maharashtra were the states with the highest rate of infection. As on 23rd April 2020 the total number of confirmed cases reported is 23,039 with 721 deceased and 5012 recovered cases (COVID-19 tracker India, available at http://www.covid19india.org, Accessed on 19th April 2020). India is a highly dense country of 1.35 billion populations with limitations in the proper health care system and infrastructure. Hence, it is at a higher potential risk of infection than the other developed countries. So, India to a greater extent has to aim at non-compulsive community intervention with people's commitment and assurance at enough lucidity [5]. These preventive and control strategies might help in the mitigation of infection by hindering person to person transmission at the community level as well as national level. To fight against COVID, the Indian government has taken unprecedented measures like lockdown, cancellation of Visas, flights, and trains, which creates hope of successful control of the situation [15].

Methods and Models

2.1 Exponential Model

In epidemiology, researchers use different mathematical models to estimate the outbreak. The selection of a model depends on the number of factors of that region or country such as current growth rate of disease, demography (e.g., age and population density), weather or humidity changes, immunity, healthcare facilities, and government interventions [1][11]. Studies suggest that most pandemics follow exponential growth in the initial stage and then flatten out [10]. Different researchers have used different mathematical models to predict the spread of COVID-19. The current study assesses the suitability of each model.

The exponential model is good for estimating the outbreak in the short run of the first phase as the pandemic spread exponentially. Thus, if the number of infected cases (I) over time (t) with infection rate (β) is known, the spread can be calculated by,

$$\frac{dI}{dt} = \beta I \qquad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Integrating above equation 1, we get-

$$I(t) = I_0\, e^{\beta t} \qquad \ldots\ldots\ldots\ldots\ldots (2)$$

Where I_0 is a constant and can be obtained by fitting with the data. And, I is the cumulative number of cases including individuals recovered and diseased. A J-shape is formed in the exponential growth phase of a pandemic. The exponential growth model fails to predict decay (downward movement) and the plateau of the curve.

As of 19th April 2020, India is in the 2nd phase of disease transmission, where total cases of infections are increasing steadily. Thus, exponential growth will not be able to predict accurately.

2.2 Logistic Model

Unlike the exponential growth model, this model predicts the growth of the disease spread under the constraints. We mean constraints in terms of redefining the number of susceptible after the interventions imposed by the government (especially, restriction People's mobility, strict lockdown, and dedicated medical facilities). Thus, change in the number of infected (I) over time (t) is calculated by-

$$\frac{dI}{dt} = \beta I(1 - \frac{I}{S}) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where I is total number of accumulated infected cases, $\beta>0$ infection rate, and S is susceptible people.

If $I(0)=I_0>0$ is initial number of infected cases.

$$I(t) = \frac{S}{1+Ae^{(-\beta t)}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Where $A= (S/I_0)-1$ and I_0 is initial number of invectives. And S is number of susceptible people in the population.

2.3 SEIR Model

The widely used framework for predicting the population growth or disease spread is 'Susceptible-Infectious-Removed' (SIR). However, in pandemic (COVID-19 cases) like situation; the 'Susceptible-Exposed-Infectious-Removed' (SEIR) model has better accuracy as the model accommodates many factors which affect the growth of spread. Given the present situation in India after a number of interventions by the government, we formulated the model equation as follows-

$$\frac{dS}{dt} = -\beta I \quad \text{.....} \quad (5)$$

$$\frac{dE}{dt} = \beta I - \delta E \quad \text{.....} \quad (6)$$

$$\frac{dI}{dt} = \delta E - \gamma I - \mu I \quad \text{.....} \quad (7)$$

$$\frac{dR}{dt} = \lambda I \quad \text{.....} \quad (8)$$

$$\frac{dD}{dt} = kI \quad \text{.....} \quad (9)$$

Where S is susceptible individuals; E is infected individuals but asymptomatic; I is infected and symptomatic individuals. The λ and k are the rate of recovery and death. dR and dD are cases of recoveries and deaths over time (t), respectively. We numerically simulated the SEIR model to project forward. The following parameters (Table 1) have been used in this study.

Table 1: Parameters of the study

Parameter	**Notation**	**Value**	**Remarks**	**Reference**
Initial population size	No	350 million	Constant	Census data
Initial susceptible population	So	0.5No	Constant	Assumed
Transmission rate	β	{0.80, 1.10}/day	A stepwise function	Fitted
Mean latent period	δ	3	Constant	Wu et al. (2020)
Mean infectious period	γ	5	Constant	Wu et al. (2020)

Note:

1. *Total urban population of the top 20 infected Indian states considered as initial population.*

2. *To estimate total susceptible, 0.50% of total urban population of the top 20 highly infected Indian states was taken.*

Data

We obtained daily updates of the number of reported confirmed cases for the COVID-19 pandemic across states in India from the Ministry of Health and Family Welfare website (www.mohfw.gov.in/index.html) and COVID-19 India (www.covid19india.org/). The data was

collected from the first day of lockdown (e.g., March 25, 2020) to April 23, 2020 for our analysis. We projected the progression of virus spread for the next 42 days from the end of the 2nd phase of lockdown (e.g., 3rd May). We regressed the current growth up to 3rd May and then projected the growth from 4th May to 14th June considering the governments' interventions. Given the incubation period of 14 days for COVID-19, the duration was divided into three cycles of 14 days each (14*3=42 days) in our analysis.

Scenario Analysis

The coming weeks are going to be tough for India. The government is mulling over the actions to be taken after the lockdown 2.0, which is ending on 3rd May. The government has the challenge of managing two contradicting things simultaneously; one is to fight the pandemic by the virtue of isolation and social distancing, and the other one is to the keep wheel of the economy moving. A well thought decision will be very crucial for the people and the country as a whole. Under this circumstance, we simulated the data in two scenarios (optimistic and pessimistic). The fact under each is scenario is analyzed for the predictions.

1. **A minimum of 70 days complete** lockdown with no essential activities can prove most effective to combat the COVID-19 outbreak in India. According to Richard Horton (editor in chief, The Lancet), the Indian government's decision of lifting the lockdown will prove worse because the second wave of the disease will prove to be more detrimental. The investment and efforts of the early lockdown will go in vain if the lockdown is withdrawn. With 10 weeks of successful lockdown in Wuhan, China proved to be effective for containing the spread.
2. **Lockdown for 70 days with resuming the essential activities** can prove effective in controlling and preventing infection spread to some extent. But after 3rd May 2020, the permission to restart some essential activities may lead to spread to disease further. There are chances that infection may spread among the workers first and then to the people who are in their contacts and the proximity. The government, in this case, must be watchful.
3. **Withdrawal of lockdown-** if India rushes to exit the lockdown, there will be a devastating situation. No contact tracing will be possible, and the situation will be out

of control. One can't imagine the situation when the transmission goes into the community with limitations of per capita health care facilities in the country.

Results and Discussion

The rate of COVID-19 transmission in India is low as compared to countries like the United States, Spain, Italy, France, etc. So far, India is in a better position in the battle as compared to many other countries where spread was almost equal once but now total infected cases in those countries have gone up much fold (as on 23rd April). Given the very complex demography (e.g., population density, diversity, belief system), it raised challenges to the government in imposing the sudden restrictions on people to curb the virus spread. However, the Indian government managed to take bold and timely actions such as lockdown and mobility restriction to fight against COVID-19. Thus, considering the current situation in India, projection using a linear growth model is found suitable for this study. Now that India is resuming certain economic activities, it poses further speculations if the cases may rise again due to people's mobility and gathering. Under this circumstance, we considered two scenarios in our analysis; optimistic, and pessimistic. The optimistic scenario indicates that the spread of the virus follows the current trend with a complete lockdown in the infected states. The pessimistic scenario indicates that the growth of spread would be with a higher transmission rate due to people's non-compliance of social distancing norms due to removal or relaxation of lockdown.

Figure 1 shows the actual daily cases. From the curve, it is evident that the daily new cases are increasing steadily. Figure 2 depicts the progression of the outbreak in both optimistic and pessimistic scenarios. The optimistic models are statically significant at $p<.01$ with R^2 values of 91.8% for linear, and 85% for growth model, respectively. Under this scenario daily new cases increase and reach a peak on in the second week of May. Thereafter, downward movement of the curve starts, and likely to touch the bottom on 14th June with approximately 40 new cases daily. Thus, according to optimistic models, the COVID-19 pandemic in India will end in the 3rd week of June 2020 with a total of 77,900 cases.

Figure 1: Actual daily and total cases from 25/March/2020 to 23/April/2020

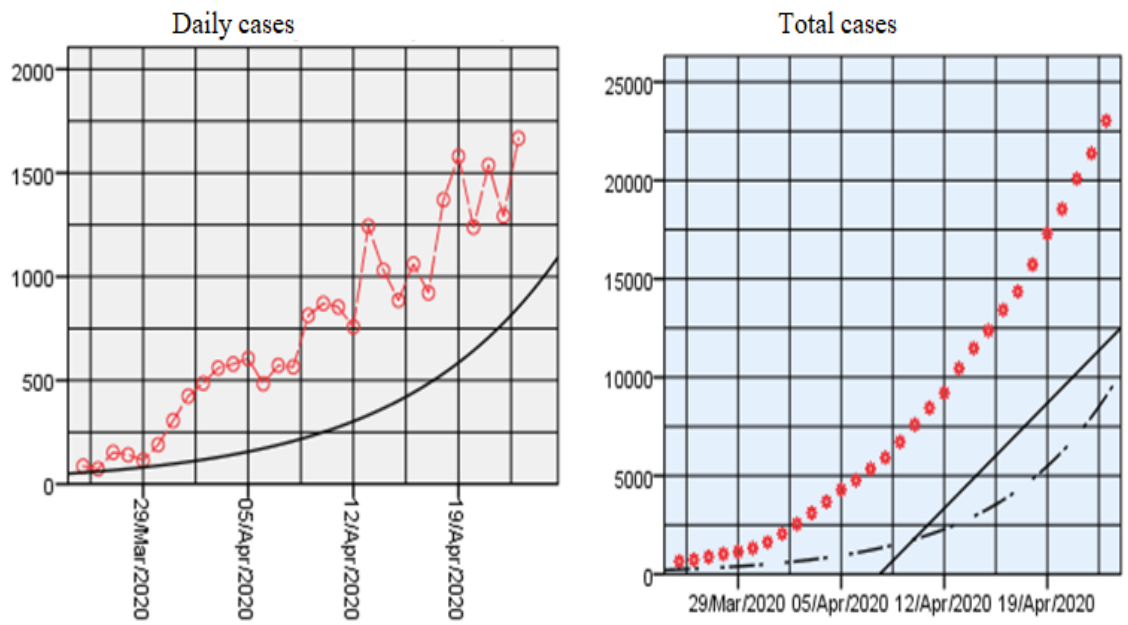
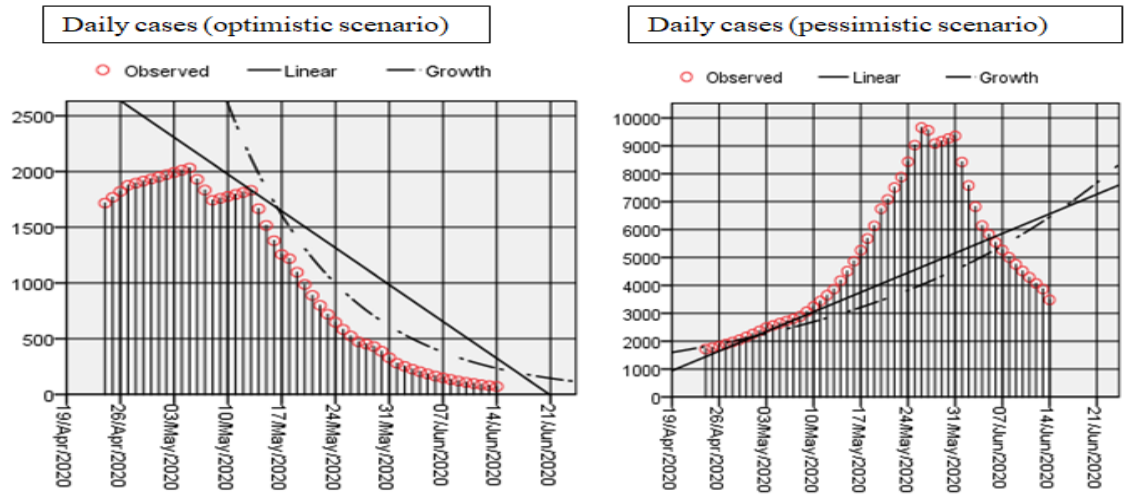


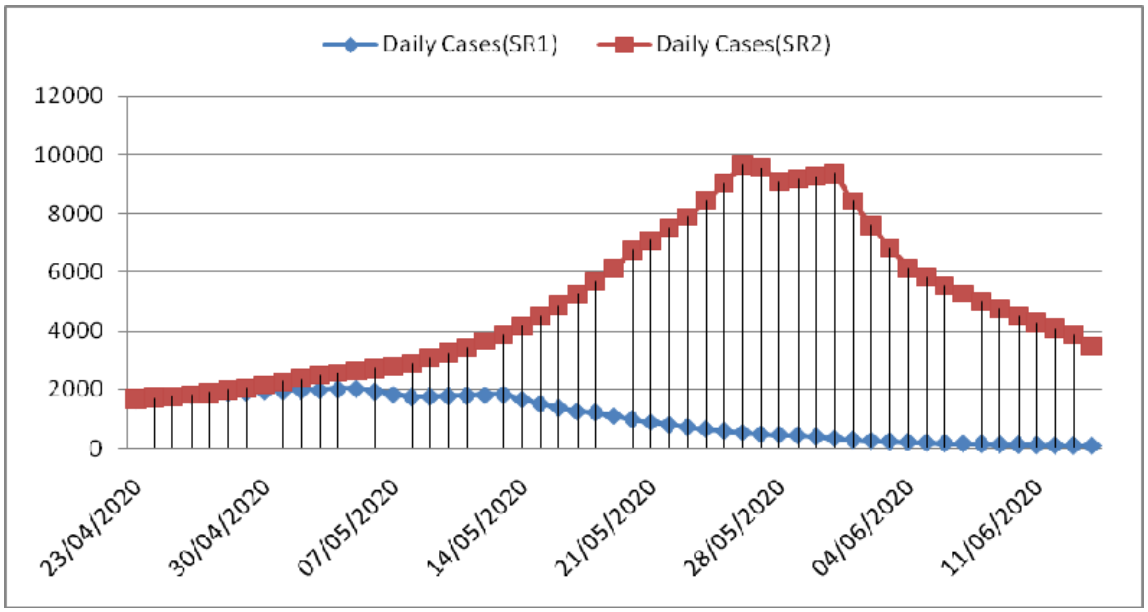
Figure 2: Prediction on daily new cases till



The pessimistic scenario considers the worst possible cases in a certain outbreak situation. According to the results, the pessimistic models are statically significant at $p<.01$ with R^2 values

49% for linear and 56% for the growth model. As per these models (linear and growth), the daily new cases will reach a peak between 24th to 31st May (Figure 2, pessimistic scenario) with total cases of 207,780 (Figure 4). As of 14th June, 283,348 individuals will be infected with COVID-19. Under this scenario the end of the pandemic is expected to be delayed by 4-5 weeks and hence will likely to vanish in the second week of July 2020.

Figure 3: Daily new cases (Optimistic vs. Pessimistic scenario)



SR1- Scenario 1 (optimistic scenario), SR2- Scenario 2 (pessimistic scenario)

Figure 4: Projection of total COVID-19 Cases in India (Optimistic vs. Pessimistic scenario)

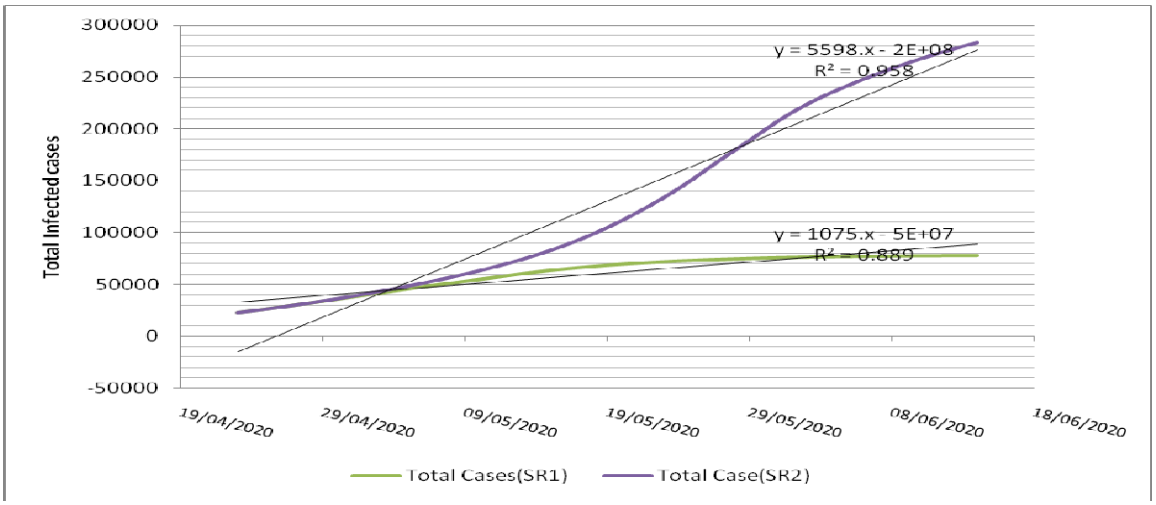
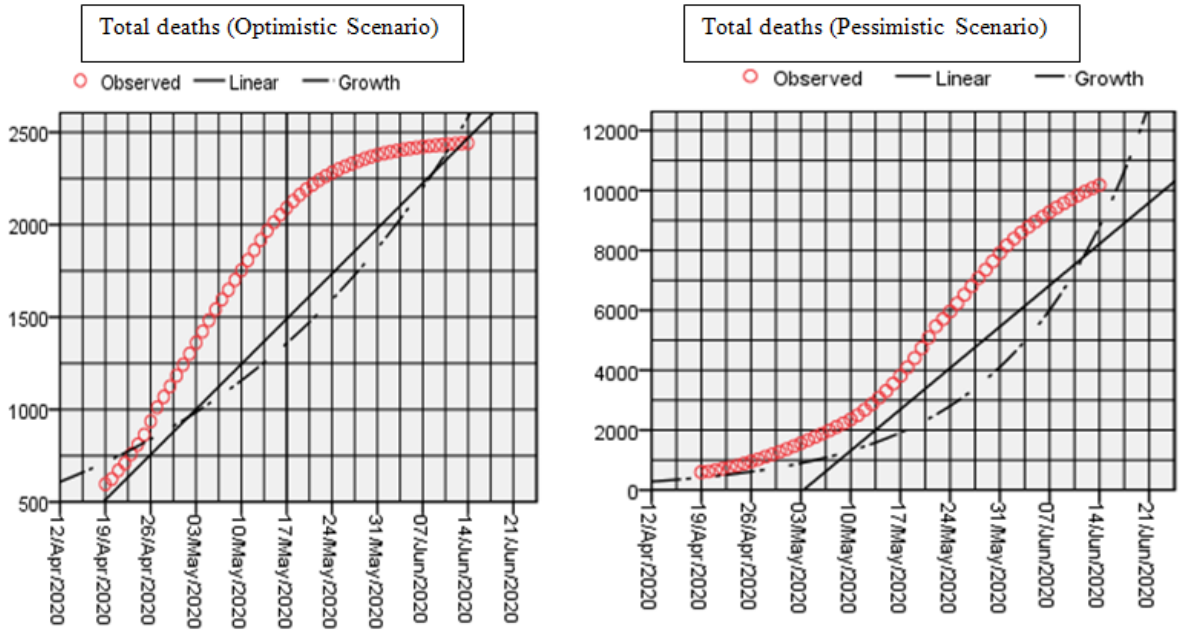


Figure 5 presents a projection of the total death toll in India till 14th June 2020. As per the results obtained, in the optimistic scenario, total fatality reaches to 2,442. While in the case of a pessimistic scenario, the total fatality would go up to 10,180.

Figure 5: Projection of total deaths in India



Conclusion

We have estimated the spread of COVID-19 utilizing mathematical models. The data were taken from the very early days of lockdown to specifically examine the impact of lockdown in curbing the disease outbreak. In the current study, we have used different forecasting models under two scenarios, the first scenario of optimistic approach depicting the progression of infection in the presence of government interventions like social distancing, complete lockdown, public awareness, travel restrictions, and another scenario of a pessimistic approach of the infection progression without (or less) such interventions. By validating all analyses done, if India can manage interventions implemented, the daily infected cases have gradually restrained from mid-May, 2020. Hence, under an optimistic approach, the total number of cases is expected to come down from the 3rd week of May 2020. Results under optimistic conditions confirm the total 77,900 infected cases till 14th June. However, the results of the pessimistic scenario estimation

are horrifying. The results show that the spread will be quite high in the mid of May. The total case, till 14th June will reach 283,300. Thus, the complete lockdown (at least in the infected Indian states) and travel restrictions are much needed in the fight against a highly contagious virus like COVID-19 to avoid the pessimistic scenario in the country. It should further be noted that in both scenarios we assumed no local transmission. For a country like India with high population density and poor health care systems, it would be quite impossible to mitigate outbreak without any of such serious laws and steps enacted by both state and central government. The six weeks of lockdown may not be sufficient. Additionally, with more rigid lockdown mass screening, isolation, enforced quarantine and proper medical treatment should be executed meticulously for pandemic mitigation.

These predicting models thus provide an insight into the epidemiological status in the country by forecasting the potential propensity of a disease outbreak. Thus on the analytical basis, it would help in the execution as well as manipulation of all current interventions and approaches to combat COVID- 19 in India.

Our present model signifies the prediction capability of COVID-19 growth in India by considering the effects of interventions (both government and self) as preventive and control measures undertaken since the time of spread. But these models are devoid of the impact of certain biological factors like differential immunity, previous and present health status of person i.e. comorbidities associated and passive immunization status. The physical factors like temperature, humidity, their influence have not been considered in the modeling.

Implications and suggestions

According to Boston Consulting Group (BCG) analysis (source, John Hopkins University), countries that implemented complete lockdown are in a better position in the combat with the COVID. India has also taken the early steps such as lockdowns, travel restrictions, and identifications of infected zones, etc.

If a lockdown is fully or partially withdrawn in a premature state of an outbreak, there will be higher chances of the exponential growth of spread. Complete lockdown signifies a gradual decrease in transmission rate with time due to more contact inhibition. To date, the only effective way is to do proper mass scale random testing along with testing suspects. The aggressive

contact tracing efforts are much needed to minimize the fluidity of the spread. A strict lockdown will be helpful for India in breaking the chain and contact tracing.

According to the world's renowned virologist, Ian Lipkin, it is difficult to predict the end of the pandemic in India. However, he advised India to follow the two strategies to win the battle; first is the aggressive testing, and second is the isolation of clusters (hotspots region). If these steps are taken seriously, India may go with an option of relaxation of essential activities, and with some restrictions; people can resume their daily activities. So, contact tracing is most essential in this scenario. Discovering the contact history of the infected or suspected person is an urgent measure that has to be undertaken. The identification of the most affected areas (hotspots) and isolating them completely will be effective in further prevention. There are moderate chances of infection spread in this scenario with a capacity of control by the government to some extent.

Conflict of interest

No conflict of interest to declare.

Funding source

None.

Ethical approval

Approval was not required.

References

[1] Araujo MB, Naimi B. Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate. medRxiv 2020; doi.org/10.1101/2020.03.12.20034728.

[2] Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. J Travel Med. 2020; 27(2), 1–3. doi: 10.1093/jtm/taaa008.

[3] Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol 2019 181–92. Available from: http://dx.doi.org/10.1038/s41579-018-0118-9.

[4] Daga MK, Kumar, N., Aarthi, J., Mawari, G., Garg S, Rohatgi, I. From SARS-CoV to Coronavirus Disease,(2019).(COVID-19)-A Brief Review 2019; 6(4),1–9, https://doi.org/10.24321/2349.7181.201917

[5] Ebrahim SH, Ahmed QA, Gozzer E. Covid-19 and community mitigation strategies in a pandemic,2020; 1066(March):1–2, https://doi.org/10.1136/bmj.m1066

[6] Fraction, R.E. Function VS. Correspondence Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. 2020; 2019–20, https://www.nejm.org/doi/pdf/10.1056/NEJMc2001468

[7] Gorbalenya AE, Baker SC, Baric RS, Groot RJ De, Gulyaeva AA, Haagmans BL, et al. The species and its viruses – a statement of the Coronavirus Study Group. Biorxiv (Cold Spring Harb Lab. 2020; 1–15. Available from: https://www.biorxiv.org/content/10.1101/2020.02.07.937862v1.full

[8] Heymann DL. A novel coronavirus outbreak of global health concern. 2020; 395:15–8. Available from: doi: https://doi.org/10.1016/S0140-6736(20)30185-9

[9] Lu, H., Stratton C.W. Outbreak of pneumonia of unknown etiology in Wuhan, China□: The mystery and the miracle. 2020; 401–2,doi: 10.1002/jmv.25678

[10] Ma J, Dushoff J, Bolker BM, Earn DJD. Estimating Initial Epidemic Growth Rates. Bull Math Biol. 2014; 76(1):245–60. doi: 10.1007/s11538-013-9918-2

[11] Ma Y, Zhao Y, Liu J, He X, Wang B, Fu S, et al. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. Sci Total Environ. 2020; 724, 138226. doi.org/10.1016/j.scitotenv.2020.138201

[13] Singhal, T. A Review of Corona virus Disease-2019 (COVID-19), The Indian Journal of Pediatrics, 2020; 87(April), 281–6. 10.1007/s12098-020-03263-6

[14] Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. 2019; 1–13. https://doi.org/10.3390/jcm9020462

[15] Tanne J. H., Hayasaki E., Zastrow M., Pulla P., Smith P., Rada A.,Garcia et al. Covid-19: how doctors and healthcare systems are tackling coronavirus worldwide 2020; BMJ, https://doi.org/10.1136/bmj.m1090

[16] Tong Y, Ren R, Med M, Leung KSM, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia new england journal 2020;1199–207, DOI: 10.1056/NEJMoa2001316