COVID-19 in India: Prediction of the Evolution of Coronavirus using Epidemiological Modeling

Muthusami Rathinasamy[1], Saritha K[2], and A Bharathi[3]

[1]Dr Mahalingam College of Engineering and Technology
[2]PA College of Engineering and Technology
[3]Bannari Amman Institute of Technology

July 2, 2020

Abstract

Coronavirus Disease 2019 (COVID-19) Pandemic is a big threat for all countries' health systems. In India, the first case of coronavirus occurred in Kerala on 30 January 2020, followed by 2 more until 2 February 2020, but all three were cured, according to the World Health Organization (WHO) India situation report-2. Since this, no single case of coronavirus has been recorded across the country for a month. Nevertheless, from 2 March onwards, the number of cases rose on a regular basis. As of 21 June 2020, 410,461 confirmed cases and 13,254 total deaths, as stated in the World Health Organization (WHO) India situation report-21. This research presents significant findings about the early outbreak of COVID-19 in India. Due to the recent rapid rise in new cases of COVID-19, the pre-evolution of pandemic coronavirus is a pre-eminence in India. The susceptible-infectious-recovered (SIR) model was developed to estimate the reproductive number R0 at the early stage of the outbreak of COVID-19 and to evaluate this outbreak with available data on confirmed, deaths and recovered cases in India from 02 February 2020 to 26 June 2020.

Abstract

Coronavirus Disease 2019 (COVID-19) Pandemic is a big threat for all countries' health systems. In India, the first case of coronavirus occurred in Kerala on 30 January 2020, followed by 2 more until 2 February 2020, but all three were cured, according to the World Health Organization (WHO) India situation report-2. Since this, no single case of coronavirus has been recorded across the country for a month. Nevertheless, from 2 March onwards, the number of cases rose on a regular basis. As of 21 June 2020, 410,461 confirmed cases and 13,254 total deaths, as stated in the World Health Organization (WHO) India situation report-21. This research presents significant findings about the early outbreak of COVID-19 in India. Due to the recent rapid rise in new cases of COVID-19, the pre-evolution of pandemic coronavirus is a pre-eminence in India. The susceptible-infectious-recovered (SIR) model was developed to estimate the reproductive number R_0 at the early stage of the outbreak of COVID-19 and to evaluate this outbreak with available data on confirmed, deaths and recovered cases in India from 02 February 2020 to 26 June 2020.

Keywords: COVID-19; SIR model; Reproductive number; Prediction; India

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a major pathogen of the fast expanding coronavirus disease outbreak (COVID-19) originating in the city of Wuhan, Hubei Province, China (Qun, 2020). More than 9.8 million people worldwide have been affected and in India, with 528,859 reported cases, 4th place, with a global percentage of 5.40, 8th place in 16,095 deaths, with a global percentage of 3.00 and 4th position in 321,723 cases recovered as of 28 June 2020 (WHO, 2020, Coronavirus disease 2019 (COVID-19) situation reports - 160). Despite no alternatives and an inevitable possibility of developing a vaccine, the Government of India had to concentrate on inhibiting the 5-phase national lockdown of virus

control from 24 March to 30 June 2020 (Wikipedia, COVID-19 pandemic lockdown in India, 2020). The goal of the lockdown is to reduce the progression of coronavirus. On the one side, this lockdown could never continue forever due to its impact on people's everyday lives and health and on the economy of the country, and on the other side, relaxing the lockdown would possibly lead to an increase in the number of diseases. During this stage, the issue that seems very important to both people and the government is the spread of the disease as it reduces and the time will be right to drop the lockdown. In this paper, we modeled the coronavirus spread to address this issue. Predictions can be used to shape government actions to contain the virus and to try to prepare for what's going to happen sensitively and economically. In the subsequent part of the paper, we initially presented a model that is being used to examine the outbreak of Coronavirus in India. We also briefly address and demonstrate how to estimate an important epidemiological metric, the reproduction number. We then use our model to examine the spread of disease if there is no public health intervention.

There are various epidemiological models (Hethcote, 2009), but we have used one of the most well-known ones, the susceptible-infectious-recovered (SIR model). The underlying concept behind the SIR model of transferable chronic diseases is that there are three groups also called compartments denoted as S, I and R. S: those who are safe but susceptible to disease (i.e., at risk of contamination). In the beginning of the pandemic, S is that the population as a whole is not infected by the virus; I: Infectious people (and thus infected); R: individuals who have been contaminated but have either recovered or died, they're no longer contagious. These groups grow over time as the virus evolves in the population as: S reduces once people are infected and transfers to contagious group I; As people recover or die, they move from the infected group I to the recovered group R. In order to model the composition of the outbreak, we require three differential equations to portray the rate of change in each group, which are getting to know by: β, the rate of infection that controls the transition between S and I; γ, the rate of removal or recovery that governs the change between I and R. Expressly, this gives the following:

$\frac{dS}{dt} = -\frac{\beta I \Sigma}{}$

englishN (1)

$\frac{dI}{dt} = \frac{\beta I \Sigma}{}$

englishN-γI(2)

$\frac{dR}{dt} = \gamma I$ (3)

The equation (Eq. 1) states that the amount of susceptible people (S) decreases with the amount of actively infected people, where new contaminated cases are the results of the infection rate (β) multiplied by the amount of susceptible people (S) in contact with infectious individuals (I). The equation (Eq. 2) expresses that the quantity of infectious people (I) increments with the newly contaminated people (β I S), minus the previously infected people who recovered (i.e., γ I which is the evacuation rate γ multiplied by the infectious people I). The equation (Eq. 3) expresses that the recovered group (R) increments with the quantity of people who were infectious and who either recovered or died (γ I).

The epidemic is developing as follows: S is equal to the whole population before the outbreak of the disease starts, because no one has immunity; at the start of the epidemic, as soon as the first person is infected, S decreases by 1 and I increase by 1; this first infectious human contaminant (before recovery or death) was found to be susceptible to other individuals; the cycle persists, with newly infected individuals who, in effect, harm other vulnerable people before they recover, which is shown visually in Figure 1.

Figure 1. SIR model Source: Wikipedia

Next, in order to solve these differential equations and to find the optimum values for the unknown parameters, the β and γ used R packages on the Windows 10 operating system to fit the SIR model to the data of India (Churches, 2020). Subsequently, the residual sum of squares (RSS) was determined as indicated in

Eq. 4 which minimizes the sum of the squared differences between I(t), which is the amount of people in the contagious compartment I at time t, and the associated number of cases as expected by our model $\hat{I}$ (t).

$$\text{RSS } (\beta, \gamma) = \sum_t \left(I\left(t\right) - \hat{I}\left(t\right)\right)^2 (4)$$

In order to fit the model of incidence data for India, we considered the population of India to be 1,352,642,280 as of November 2019 according to Wikipedia and the daily cumulative incidence extracted from the {coronavirus} R package established by Rami Krispin from 2 February 2020 to 26 June 2020. Finally, the SIR model fitted with the available data by finding the values for β and γ which minimize the residual sum of squares between the observed cumulative incidence (observed in India) and the predicted cumulative incidence (predicted by our model) as β is 0.5403282 and γ is 0.4596718. β controls the transition between S and I (i.e., susceptible and infectious) and γ controls the transition between I and R (i.e., infectious and recovered).

Figure 2. Show the number of cases observed that match the number of confirmed cases expected by our model based on data of India from 02 February 2020 to 26 June 2020

Figure 2 shows that the number of observed cases fits the number of confirmed cases predicted by our model. The fact that both phenomena converge suggests that the pandemic is obviously at an exponential stage in India. More data will be needed to see if this phenomenon is sustained in the long term. Figure 3 displays the log-linear plot converting the scale into log, which is easier to understand in terms of the difference between the observed and the predicted number of confirmed cases, and also displays how the number of confirmed cases observed varies from the exponential trend. The plot shows that the number of confirmed cases remained below what would have been expected in the exponential phase at the beginning of the pandemic and until 4 March. In particular, the number of confirmed cases remained constant in 1 cases from February 16 to March 4. From 5 March to 26 June, the number of confirmed cases continued to increase at a rate close to the exponential rate.

Figure 3. Displays the difference between the number of reported cases and the estimated number as a log-linear plot based on data of India from 02 February 2020 to 26 June 2020

The basic reproduction number R_0, as stated in Eq. 5, also referred to as the basic reproduction ratio, is then determined using the SIR model, which is closely related to β and γ, which was estimated to be 1.175465.

$$R_0 = \frac{\beta}{\gamma} \ (5)$$

The reproduction number is the average number of susceptible individuals who are infected by each infectious person (Thompson, 2019; Yuan, 2020; Zhang, 2020; Fu-Chang, 2020). In other words, the reproduction number refers to the number of healthful people who are affected by the number of vulnerable people. When $R_0 > 1$ the disease starts to spread across the whole population, but not when $R_0 < 1$. Typically, the higher the value of R_0, the more difficult it is to manage the disease and the greater the risk of a pandemic. The R_0 of 1.2 being lower is due to the fact that the number of confirmed cases remained stable and equal to 1 at the start of the pandemic. It indicates that, on average, 1.2 people are infected for each infected person in India. For dynamic systems, the proportion of the population needed to be effectively immunized to prevent a sustained spread of the disease, known as the "herd immunity threshold," must be greater than $1 - \frac{1}{R_0}$ (Fine, 2011). The reproduction number of 1.2 we have estimated indicates that, given formula 1- (1/1.2), 16.7% of the population should be immunized to avoid the spread of the infection. With a population of approximately 1.4 billion in India, this translates into approximately 0.2 billion people. It is pertinent by using our model, tailored to the available data from 02 February 2020 to 26 June 2020, 146 days on reported cases in India, to see what would happen if the epidemic were to continue, without any public health intervention.

Figure 4. Displays cases of susceptible, recovered, observed and infectious fitted to data from India between 02 February 2020 and 26 June 2020

Despite these forecasts, as seen in the figure 4, with the exact same settings and no interference whatsoever

to restrict the spread of the pandemic, the peak in India is predicted to be reached by the beginning of August. Approximately 15,881,175 people will have been diagnosed by that time, which corresponds to about 3,176,235 serious injuries, about 952,870 people in need of intensive care and up to 714,652 deaths. We understand at this point why these strict containment measures and regulations are being enforced in India.

Conclusion

This resulted in a visual comparison of the fitted and observed cumulative incidence in India, showing that the COVID-19 pandemic is clearly at an exponential stage in India in terms of the number of confirmed cases. We then discussed what is the reproduction number and the results of the reproduction. In the end, our model was used to evaluate the Coronavirus epidemic if there was no public health action at all. Under this typical case, the peak of COVID-19 in India is anticipated to cross about 15,881,175 infected people and around 714,652 deaths by early August 2020. Perhaps these worrying optimistic forecasts emphasize the importance of stringent public health measures taken by governments and the need for people to take such safety measures to reduce the spread of the virus in India.

Ethical statement

We declare that ethical statement is not applicable.

Conflict of interests

There was no conflict of interest with others.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Churches (2020, Feb. 18). Tim Churches Health Data Science Blog: Analysing COVID-19 (2019-nCoV) outbreak data with R - part 1. Retrieved from https://timchurches.github.io/blog/posts/2020-02-18-analysing-covid-19-2019-ncov-outbreak-data-with-r-part-1/

Fine, P., Eames, K., Heymann, D., (2011). "Herd Immunity": A Rough Guide. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America. 52. 911-6. 10.1093/cid/cir007.

Fu-Chang, H., & Fang-Yu, W. (2020). The Estimated Time-Varying Reproduction Numbers during the Ongoing Epidemic of the Coronavirus Disease 2019 (COVID-19) in China. medRxiv. 10.1101/2020.04.11.20061838.

Hethcote, H.W. (2009). The basic epidemiology models: models, expressions for R0, parameter estimation, and applications. In Mathematical understanding of infectious disease dynamics, Lecture notes series, Institute for mathematical sciences, National University of Singapore, vol. 16, chapter 1.

Qun Li., Xuhua Guan., Peng Wu., Xiaoye Wang., Lei Zhou., et al. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N Eng J Med.* 382. 10.1056/NEJMoa2001316.

Thompson, R., Stockwin, J., Van Gaalen, R., Polonsky, J., Kamvar, Z., Demarsh, P., Dahlqwist, E., Li, S., Miguel, E., Jombart, T., et al. (2019). Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics* 29:100356

WHO. (2020). Coronavirus disease 2019 (COVID-19) situation reports - 160. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200628-covid-19-sitrep-160.pdf?sfvrsn=2fe1c658_2 (Accessed 28 June 2020)

Wikipedia. COVID-19 pandemic lockdown in India.

https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdown_in_India (Accessed 26 June 2020)

Yuan, J., Li, M., Lv, G., & Lu, Z. (2020). Monitoring Transmissibility and Mortality of COVID-19 in Europe. *Infect Dis* . 95. 10.1016/j.ijid.2020.03.050.

Zhang, S., Diao, M., Yu, W., Pei, Lei., Lin, Z., & Chen, D. (2020). Estimation of the reproductive number of Novel Coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. *Infect Dis* . 93. 10.1016/j.ijid.2020.02.033.

Figure 1. SIR model

Figure 2. Show the number of cases observed that match the number of confirmed cases expected by our model based on data of India from 02 February 2020 to 26 June 2020

Figure 3. Displays the difference between the number of reported cases and the estimated number as a log-linear plot based on data of India from 02 February 2020 to 26 June 2020

Figure 4. Displays cases of susceptible, recovered, observed and infectious fitted to data from India between 02 February 2020 and 26 June 2020
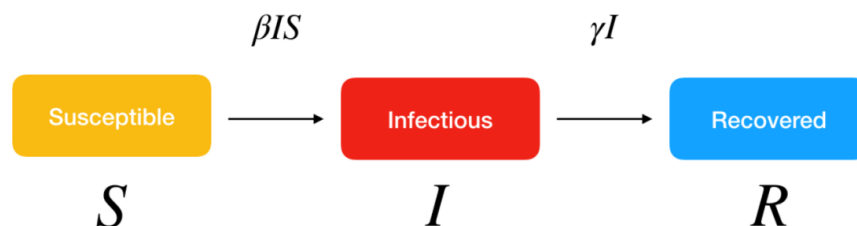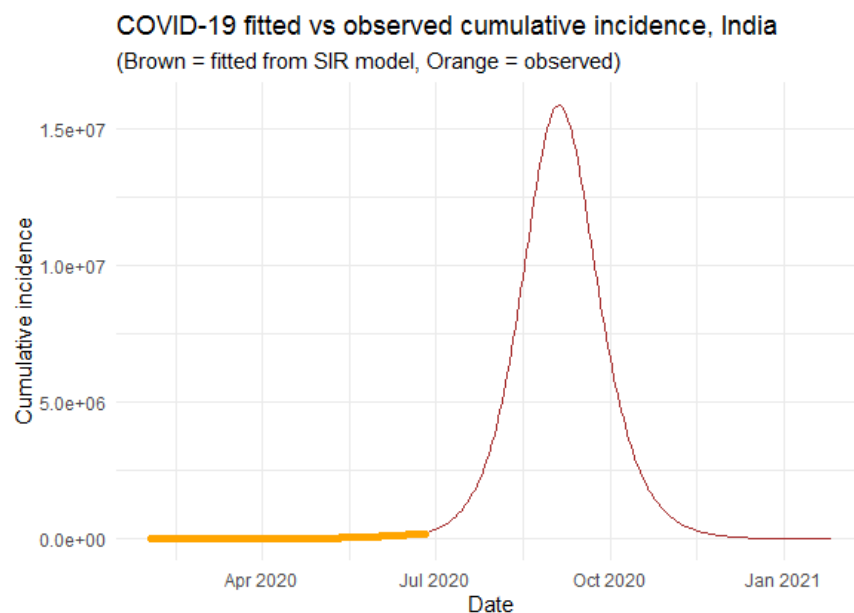

Figure 1. SIR model Source: Wikipedia

Figure 2. Show the number of cases observed that match the number of confirmed cases expected by our model based on data of India from 02 February 2020 to 26 June 2020

COVID-19 fitted vs observed cumulative incidence, India
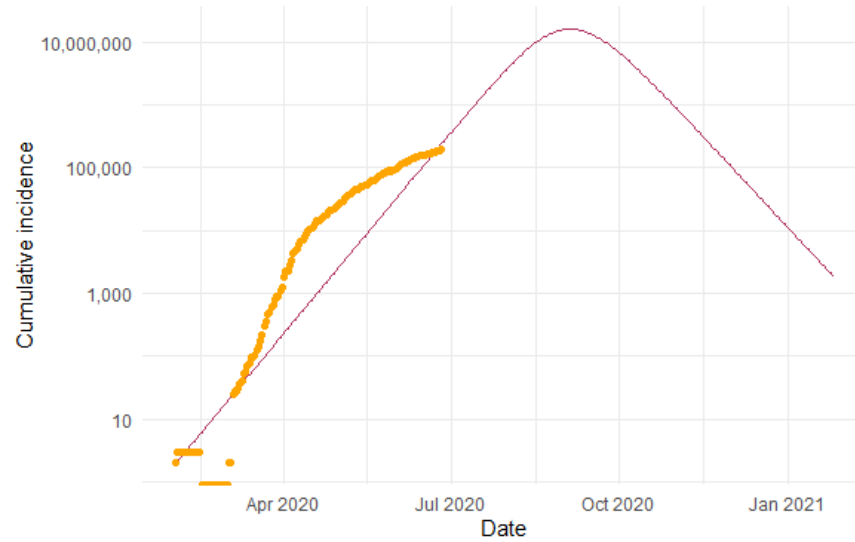### (Maroon = fitted from SIR model, Orange = observed)

Figure 3. Displays the difference between the number of reported cases and the estimated number as a log-linear plot based on data of India from 02 February 2020 to 26 June 2020

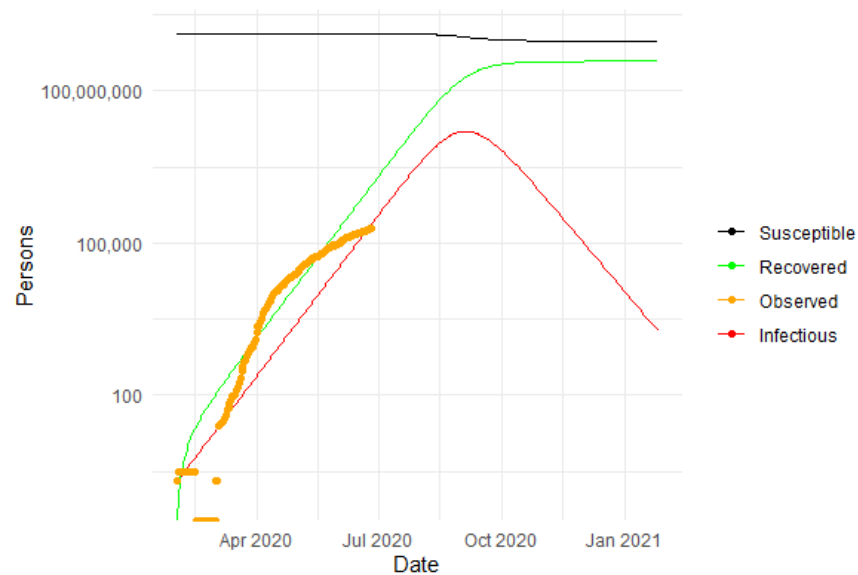## COVID-19 fitted vs observed cumulative incidence, India

Figure 4. Displays cases of susceptible, recovered, observed and infectious fitted to data from India between 02 February 2020 and 26 June 2020

6