Projecting the future trajectory of COVID-19 infections in India using the susceptible-infected-recovered (SIR) model

Kaushalendra Kumar, Wahengbam Bigyananda Meitei, and Abhishek Singh

An Analytical Paper for Policymakers

(स्थापना / Established in 1956)
बेहतर भविष्य के लिए क्षमता निर्माण
Capacity Building for a Better Future

International Institute for Population Sciences, Mumbai

(www.iipsindia.ac.in)

May 10, 2020

IIPS Analytical Series on COVID 19:

Paper 7: Projecting the future trajectory of COVID-19 infections in India using the susceptible-infected-recovered (SIR) model

(This work is not peer-reviewed. For any equerry, comments, suggestions and clarifications, please email the authors)

For research, and updates on Covid-19, visit:

https://www.iipsindia.ac.in/content/covid-19-information

Abstract

Background

COVID-19 infection started in India by the end of January 2020 and spread across all the state by April 2020. As on May 5, 2020, the total number of active cases was 32138, cured and discharged were 12726, deaths were 1568, and 1 person migrated. Despite the daily rise in the burden of COVID-19 infections in the country, there is limited scientific work on projecting the future trajectory of COVID-19 infections in India.

Objective

This policy paper aims to examine the future trajectory of COVID-19 in India. We first present the scenario at the national level. Then we present the scenario at national level after excluding Maharashtra, Gujarat and Delhi. Finally, we present a scenario of spread in case the lockdown is extended further by two weeks till May 17, 2020, three weeks till May 24, 2020, and four weeks till May 31, 2020.

Methods

We use data from the crowdsource patient database available on https://www.covid19india.org to model the spread of COVID-19 in India. We primarily rely on the daily incidence of COVID-19 in India as available in the database. We use susceptible-infected-recovered (SIR) model to examine the spread of COVID-19 before the lockdown period (i.e. between March 14 to March 24), during the first lockdown period (i.e. between March 25 to April 14), and during the second lockdown period (i.e. between April 15 to May 03).

Results

At the national level, the basic reproduction number (R_0) before the lockdown was 1.862. The R_0 declined to 1.455 and 1.200 during the first and second lockdown periods respectively. The combined R_0 for 1st and 2nd lockdown periods was 1.258. For a given recovery rate (0.363), as R_0 reduces from 1.862 to 1.200, highest estimated prevalence reduces from 12.9% to 1.5%. At the same time, the corresponding duration of pandemic increased from two months to more than 6 months. If the R_0 observed in second lockdown period continues, only a maximum of 1.5% population will be infected by mid of August 2020. In this scenario the prevalence is likely to reach zero by November 2020. Further our projections, based on the Government of India's recommendation to extend the lockdown to May 17, 2020, indicate that only a maximum of 0.23% of India's population is likely to be infective by October 2020.

Conclusions

Our findings indicate that the strategy of lockdown has contributed in flattening the curve of infection and is also likely to reduce the burden of infection tremendously in India. Our findings translated into numbers indicate that an estimated maximum of 3.0 million Indians are likely to be infective by October 2020 if the third lock down as announced by the Government of India is implemented with full rigor.

Key words

COVID-19; susceptible-infected-recovered (SIR) model; basic reproduction number; case recovery rate; case fatality rate; India

Projecting the future trajectory of COVID-19 infections in India using the susceptible-infected-recovered (SIR) model

Kaushalendra Kumar[1], Wahengbam Bigyananda Meitei[2] and Abhishek Singh[3]

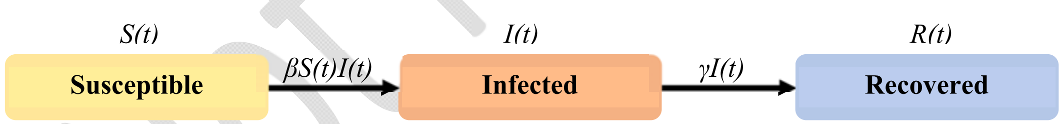
1. Assistant Professor, Department of Public Health & Mortality Studies, International Institute for Population Sciences, Mumbai, India. Email: kaushalendra.1983@gmail.com
2. PhD Scholar, International Institute for Population Sciences, Mumbai, India. Email: bigyan030294@gmail.com
3. Professor, Department of Public Health & Mortality Studies, International Institute for Population Sciences, Mumbai, India. Email: abhi_iips@yahoo.co.in

Background

COVID-19 infection started in the Wuhan city of China towards the end of 2019. Soon after it spread across all the continents. The majorly affected countries from COVID-19, in terms of both mortality and morbidity, are the European countries such as Italy, Spain, United Kingdom, France, Portugal, etc. Later on United States of America (USA) joined this group of countries and currently is world leader in number of infected and deaths from the COVID-19. In India, the spread of COVID - 19 infections started sometimes in end of January 2020 and was constant till February (Press Information Bureau 2020a). However, it started growing rapidly in March. With the experience of overwhelming burden of COVID-19 in European countries, the Government of India was quick to implement a complete nationwide lockdown on March 25 for 21 days (Press Information Bureau 2020b). Later on, on the request of a majority of states, the nationwide lockdown was extended until May 03 2020 (Press Information Bureau 2020c). As on May 5, 2020 at 8:00 am, the total number of active cases was 32138, cured and discharged were 12726, deaths were 1568, and 1 person migrated (MOHFW 2020). The three states of India that are predominantly affected by COVID-19 are Maharashtra (Total confirmed cases – 14541), Gujarat (Total confirmed cases – 5804), and Delhi (total confirmed cases – 4898) (MOHFW 2020). Madhya Pradesh, Rajasthan, Tamil Nadu, and Uttar Pradesh are the other leading states. It is apparent from the statistics that the spread of COVID-19 in India during the lockdown period has not followed the growth trajectory that some of the badly affected European countries and USA have experienced in the last couple of months (Luo 2020). Given that the lockdown appears to have containment effect on the growth trajectory of the COVID-19 in India, this working paper aims to examine the growth trajectory of COVID-19 in India using the Susceptible-Infected-Recovered (SIR) model. SIR model is one of the best suited models for projecting the spread of infectious diseases like COVID-19 where a person once recovered is not likely to become susceptible to the infection again (Kermack and McKendrick 1991). In particular, we examine the effect of complete lockdown of the country on spread of the infection. We first present the scenario at the national level. Then we present the scenario at national level after excluding Maharashtra, Gujarat and Delhi. Finally, we present a scenario of spread in case the lockdown is extended further by two weeks till May 17, 2020, three weeks till May 24, 2020, and four weeks till May 31, 2020.

Data and Methods

We use data from the crowdsource patient database available on https://www.covid19india.org to model the spread of COVID-19 in India accessed on May 4, 2020. COVID-19 INDIA is online data hub/source which compiles and updates the COVID-19 information from state press bulletins, official handles, PIB, Press Trust of India, ANI reports. We primarily rely on the daily incidence of COVID-19 in India as available in the database. We use SIR model to examine the spread of COVID-19 before the lockdown period (i.e. between March 02 to March 24), during the first lockdown period (i.e. between March 25 to April 14), and during the second lockdown period (i.e. between April 15 to May 03). The three-compartment model (Kermack and McKendrick 1991), i.e. the SIR model can be diagrammatically represented as,



where, $S(t)$ = Proportion of individual susceptible to COVID-19 at time t, $I(t)$ = Proportion of individual who have been infected by COVID-19 and are capable of infecting others at time t, and $R(t)$ = Proportion of individual who have been infected by COVID-19 and recovered at time t, such that $S(t) + I(t) + R(t) = 1$. And, β is the parameter controlling how much the disease can be transmitted. While γ is the parameter representing the rate of recovery in a particular period. The model allows us to describe the number or proportions of individuals in each compartment by solving the following ordinary differential equations,

$$\frac{dS(t)}{t} = -\beta S(t)I(t) \quad (1)$$

$$\frac{dI(t)}{t} = \beta S(t)I(t) - \gamma I(t) \qquad (2)$$

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

Several assumptions have been discussed with respect to the SIR model (Brauer and Castillo-Chavez 2012; Daley and Gani 1999). But a particular assumption that we made in the study is that once an individual is recovered, he or she will be immune to the disease. Further, we also ignored the births and deaths from the model due to the rapid pace of rate of infection. Based on the SIR model, the reproduction number is defined as,

$$R_0 = \frac{\beta}{\gamma} \qquad (4)$$

Here, R_0 is the average number of new COVID-19 cases produced by a single COVID-19 infected case over his/her infectious period.

Estimation of Time-varying Reproduction Number (R_t)

Here in this paper, we estimated the reproduction number for different phase of lockdown using the methods developed by Cori et al. (2013). It is a generic and robust method for estimating the time-varying reproduction number. Here, we assume that, if a person is infected then he or she has an infectivity profile *w(s)*, which is a probability distribution dependent on time since infection, *s*, but independent of calendar time, *t*. According to Fraser (2007) the reproduction number at time *t* i.e. R_t can be estimated by the ratio of the number of new infections generated at time *t* i.e. *I(t)*, to the total number of infections developed by the infected individuals (infectiousness of infected individuals) at time *t*, given by $\sum_{s-1}^{t} I(t-s)w(s)$ (Cori et al. 2013; Fraser 2007). Here the summation of infection incidence up to time *t-1* is weighted by infective function *w(s)* which may be approximated by the distribution of serial interval (SI) (Cori et al. 2013; Fraser 2007). Serial interval measures the time between clinical onset of symptom in infector (primary case) and clinical onset of symptom in infectee (secondary case) (Cori et al. 2013; Last 2001). Therefore, R_t is the average number of secondary cases that each infected individual would infect if the conditions remained as they were at time *t*. We assume the *w(s)* approximation i.e. SI follow a Gamma distribution with mean 4 and standard deviation 3. Studies have estimated mean (3.96 to 7.5) and standard deviation (2.9 to 4.75) of SI for COVID-19 (see Appendix Table 1) (Du et al. 2020; Hellewell et al. 2020; Kucharski et al. 2020; Li et al. 2020; Lourenco et al. 2020; Nishiura et al. 2020; Tang et al. 2020; WHO 2020; You et al. 2020; Zhao et al. 2020; Zhuang et al. 2020). Considering better immunity of Indian against COVID-19 due to almost universal immunization of BCG we have taken mean SI as 4 and standard deviation of SI as 3.

Then using the value of γ (estimated from the data) in equation (4), β is estimated and the SIR model is fitted based on the estimated R_t by solving the above differential equation. The other parameters viz. proportion of susceptible, proportion of infected and proportion of recovered are also estimated from the data (see Appendix Table. 2).

We also calculate case-recovery rate (CRR) total, CRR closed, case-fatality rate (CFR) total, and CFR closed. These are defined as

CRR total = (patients recovered / total cases) × 100
CRR closed = (patients recovered / (deaths + recovered)) × 100

CFR total = (deaths / total cases) × 100
CFR closed = (deaths / (death + recovered)) × 100

The entire analysis was carried out using R-Software. Parameters used for SIR model fitting for projection of proportion infected is given in Appendix Tables 2.

Future projection after the extension of lockdown period for two, three, and four weeks is based on the following assumption:

1) Proportion recovered, *R(t)* and R_0 are projected linearly based on cases till May 03, 2020. The R^2 for the fit till May 17, May 24, and May 31 were 0.90, 0.92, and 0.92 respectively.
2) Average recovery rate (γ) in the last one week follows logarithmic curve. Hence the recovery rate is projected using log linear model. The R^2 for projected recovery during three time points were 0.90.
3) Proportion infected, *I(t)* is taken from SIR model projection based on cases till May 03, 2020. This model gives *I(t)* for May 17, 2020. SIR fitted model on May 17, 2020 gives *I(t)* for May 24, 2020, and SIR fitted model on 24 may 2020 gives *I(t)* for May 31, 2020 (see Appendix Table 3).

Projections for males and females were carried out by taking daily reported cases by sex. For the cases for whom sex was not reported, we applied the same proportion as for those cases for whom reports of sex were available in the dataset. Parameters used for male-female projections are shown in Appendix Table 4.

Results

Figures 1 and 2 shows daily and cumulative number of infected, recovered, and deceased due to COVID-19 in India. A sharp increase in the infected persons are seen in April 2020 (**Figure 1**). However, a rapid increase in recovered from COVID-19 is seen after the second week of April. The daily gap between the infected and recovered also registered an increase during the month of April. **Figure 2** shows that the total number of infected and number recovered increased have seen an exponential increase since the third week of March; and the growth rate appears to be higher in the last two weeks of April compared with the first two weeks.

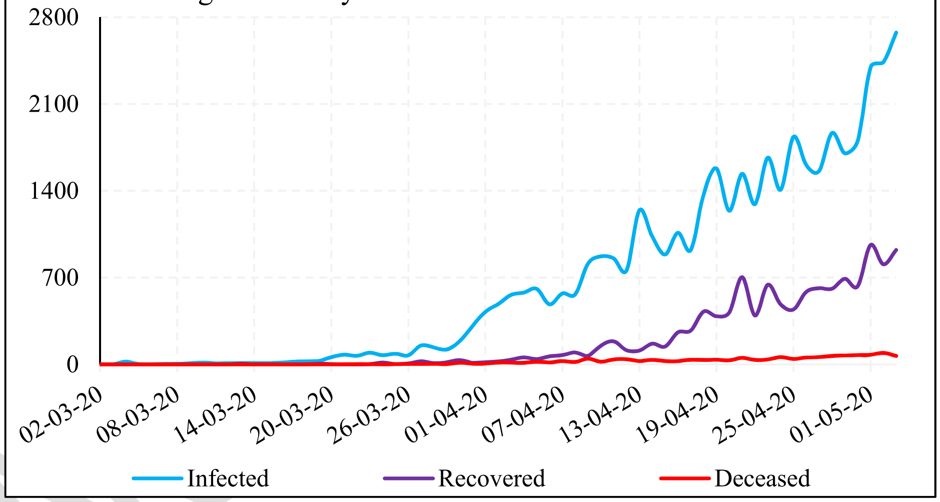


Figure 1. Daily occurrence of COVID-19 in India

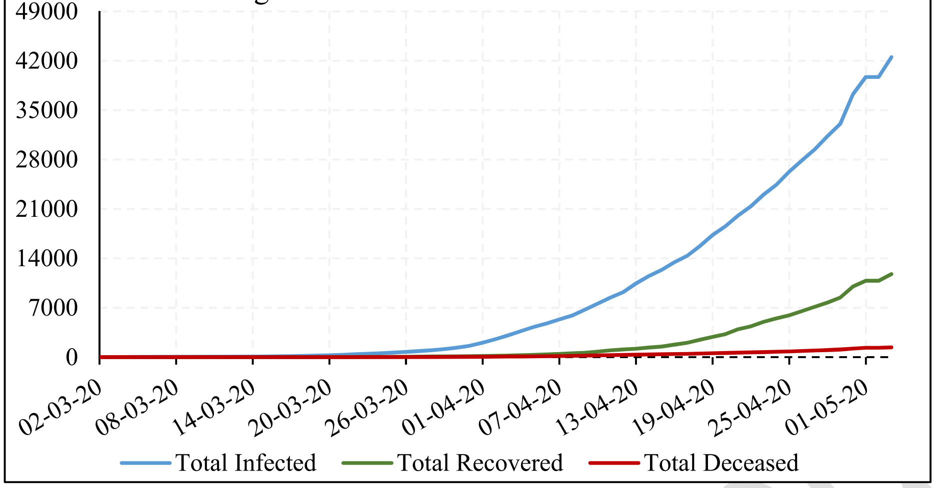


Figure 2. Total COVID-19 cases in India

Figure 3 shows the share of daily infected, recovered, and deceased of COVID-19 to the total infected, recovered, and deceased COVID-19 cases. Until the second week of April, the daily recovered as a share of total recovered was lower compared with daily infected as a share of total infected and daily deceased as a share of total deceased. But the daily recovered as a share of total recovered was highest in last 10 days of April. In the recent past, the daily deceased as a share of total deceased was higher than daily infected as a share of total infected. This indicates that the rate of infection as compared with rate of recovery and rate of death has slowed down in the last two weeks of April. Moreover, for the infection to decline, the daily infected as a share of total infected and the daily deceased as a share of total deceased should decline over time. In comparison, the daily recovered as a share to total recovered should increase over time. Notice that we do not find such trend in share of daily infected to total infected and share of daily deceased to total deceased.

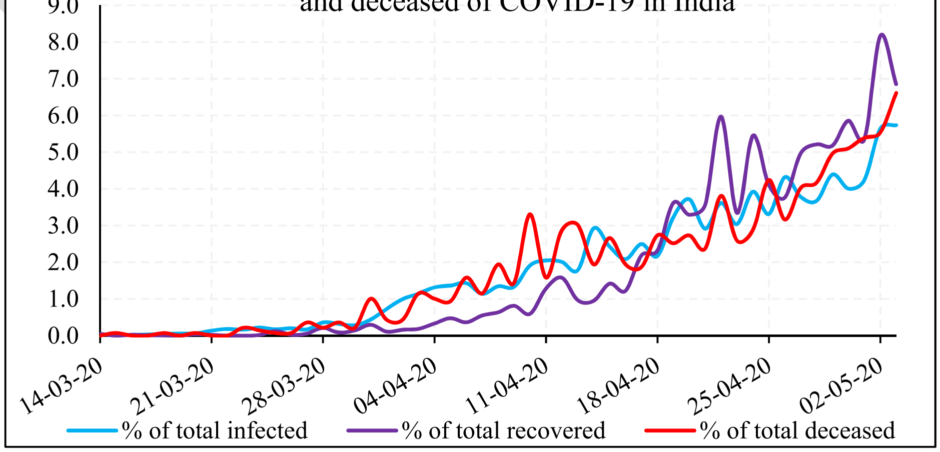


Figure 3. Percentage share of the total infected, recovered, and deceased of COVID-19 in India

Scenario in Maharashtra, Gujarat, and Delhi as a group and rest of India

Figure 4 shows the percentage share of Maharashtra (MH), Gujarat (GJ), and Delhi (DL) to India of total infected, recovered, and deceased cases. Share of total infected and total deceased continuously increased in the month of April 2020, but share of total recovered was constant around (36-39%) in the last two weeks. During March 14 to May 03, the share of total infected, total recovered, total deceased has increased from 20% to 54%, 10% to 38%, and 50% to 65% respectively.

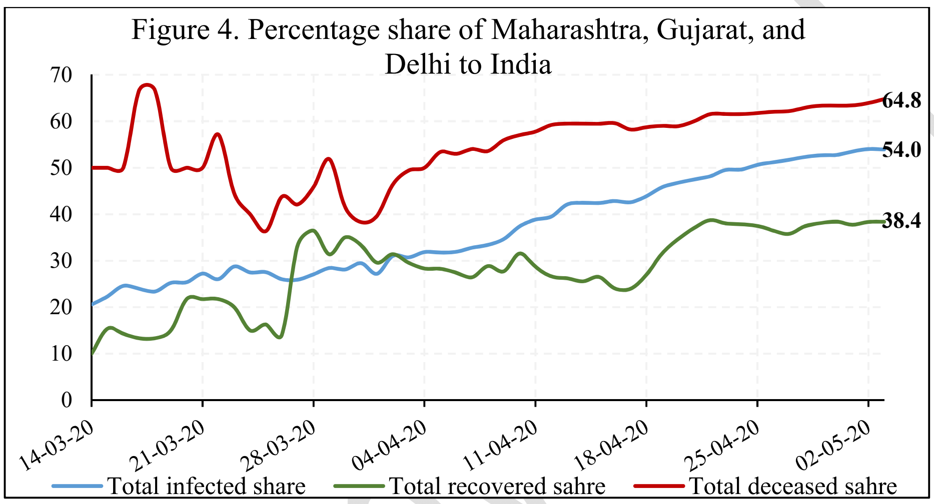


Figure 5 shows the CRR total and CRR closed for MH, GJ, and DL as a group and rest of India as the second group. Both CRR total and CRR closed are lower in MH, GJ, and DL group than in rest of India. Interestingly, CRR closed increased faster in rest of India than MH, GJ and DL group. Hence the CRR closed gap between rest of India and MH, GJ, & DL group increased consistently in the month of April. Whereas, in case of CRR total, the gap between MH, GJ, and DL group and rest of India increased in the first half of April and then became stable. It indicates that compared with rest of India either infection rate is increasing or recovery rate is constant or slower in MH, GJ, and DL.

Figure 5. Case recovery rate (CRR) total and case recovery rate closed of COVID-19, India

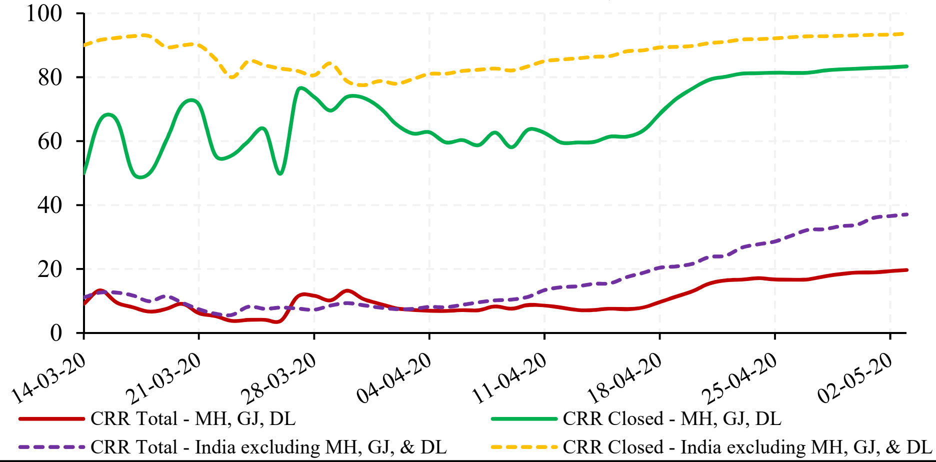
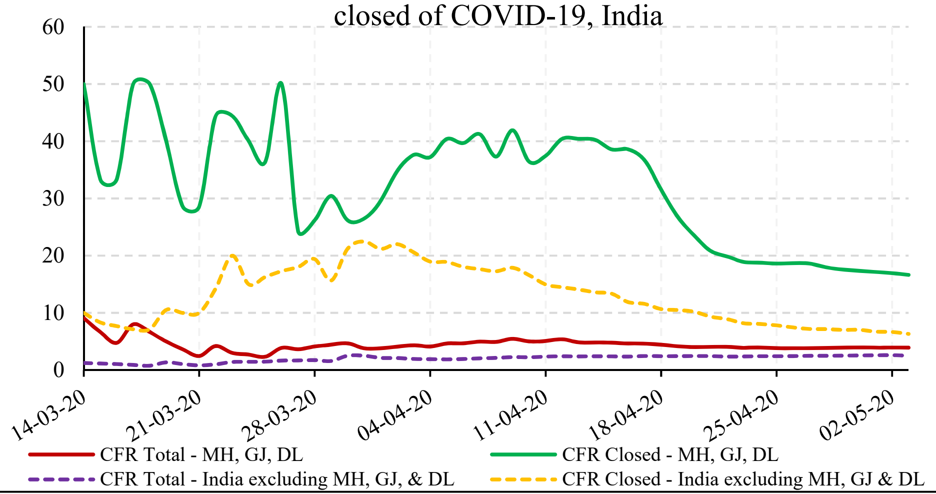


Figure 6 gives the CFR total and CFR closed for MH, GJ, and DL group and rest of India. Total and closed CFR is higher in MH, GJ, an DL compared with rest of India since the second half of March. In April, while the gap between CFR total and CFR closed consistently declined in rest of India, both show erratic pattern in MH, GJ, and DL group; both increased during first half of April, declined in the second week, and became constant in the last week. These patterns indicate that the incidence of COVID-19 is higher in MH, GJ, and DL group compared with rest of India.

Figures 4, 5, & 6 show that incidence is increasing faster than recovery in MH, GJ, DL compared with rest of India.

Figure 6. Case fatality rate (CFR) total and case fatality rate closed of COVID-19, India



Projection of growth trajectory of COVID-19 in India

Table 1 gives the estimates of basic reproduction number (R_0) and its 95% confidence intervals before lockdown period, and till end of the first lockdown period and second lockdown period, and whole lockdown period; and average recovery rate in the last one week. In India, R_0 declined from 1.862 to 1.455 till end of the first lockdown. It further declined to 1.200 till end of the second lockdown. Whereas, in India excluding MH, GJ, and DL, R_0 declined from 1.888 before lockdown 1.350 till end of the first lockdown period to 1.162 in the second lockdown period. If we consider only three states MH, GJ, and GJ together R_0 declined from 1.811 before lockdown period to 1.621 till end of the first lockdown period, 1.229 in the second lockdown period. The estimates of R_0 for whole India is always higher than India excluding MH, GJ, and DL. Whereas, estimates of R_0 for MH, GJ, and DL is always higher compared with the rest of India, barring before lockdown period. Notice that R_0 was consistently higher in MH, GJ, and DL group compared to rest of India.

Figure 7 shows the projected prevalence of proportion infected from May 03, 2020 to November 03, 2020. There appears to be a positive relationship between R_0 and proportion infected, but inverse relationship between R_0 and duration of pandemic. Here, for a given recovery rate (0.363), as R_0 reduces from 1.862 to 1.200, total infective of COVID-19 reduces from 12.9% to 1.5%. At the same time, the corresponding duration of pandemic increased from two months to more than 6 months. The infection was spreading at a much faster rate before the lockdown (R_0 = 1.862) and if the same situation would have prevailed then an estimated maximum 12.9 % of India's population would have got infected by the first week of June 2020. In that scenario, the total number of infected would have closed to zero by second week of July 2020. In comparison, if the R_0 of the first lockdown period and average recovery rate on May 03 prevails then the proportion infected would reach a maximum by third week of June 2020; and an estimated 5.5% of population would be infected. In this scenario the prevalence may reach zero by mid of August, 2020. Interestingly, if the R_0 observed in second lockdown period continues, only a maximum of 1.5% population will be infected by mid of August 2020. In this scenario the prevalence is likely to reach zero by November 2020. If we consider total lockdown period, an estimated maximum 2.3% of the total Indian population would be infected by the end of July 2020.

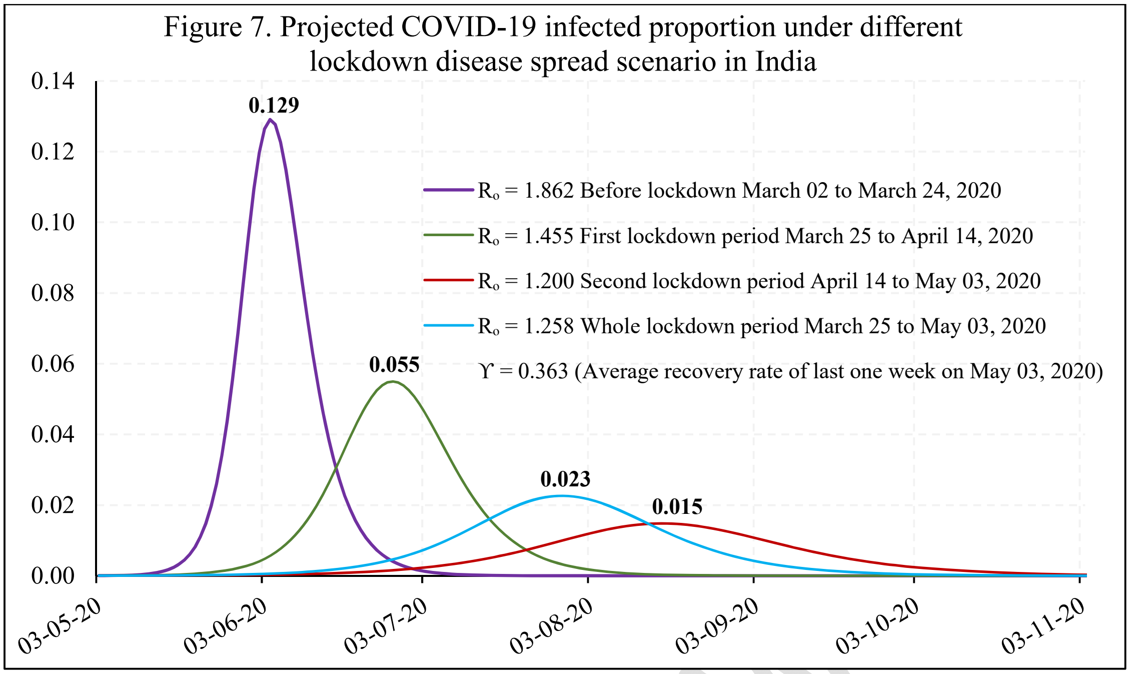


Figure 7. Projected COVID-19 infected proportion under different lockdown disease spread scenario in India

Figure 8 shows the COVID-19 infected proportions in India excluding MH, GJ, and DL from May 03 onwards. Projection based on cases occurred during March 02 to March 24 shows that an estimated maximum 13.2% of population would be infected by the third week of May, 2020; and the prevalence would have reached zero by the mid of June 2020. However, the R_0 decreased considerably during the first lockdown. If the situation during the first lockdown would have prevailed, then an estimated maximum 3.7% of the population would have got infected by the third week of June 2020. In this scenario, the prevalence would have reached zero by the end of July 2020. In the second lockdown, R_0 further declined to 1.162 in India excluding MH, GJ, and DL. If this condition continues 1% population would be COVID-19 infective by first week of August 2020, and subside to zero by October 2020. Considering the cases that occurred in the first and second lockdown together, an estimated maximum 1.7% of the population would get infected by mid of July, 2020. The prevalence in this scenario is likely to reach zero by the end of September 2020.

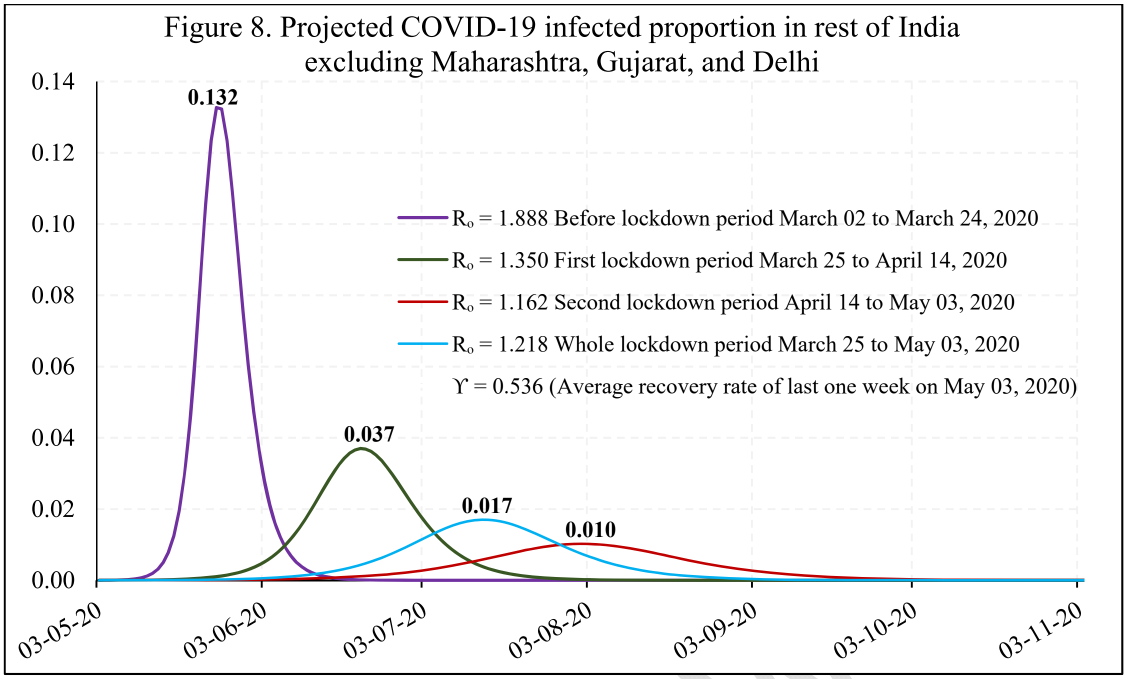
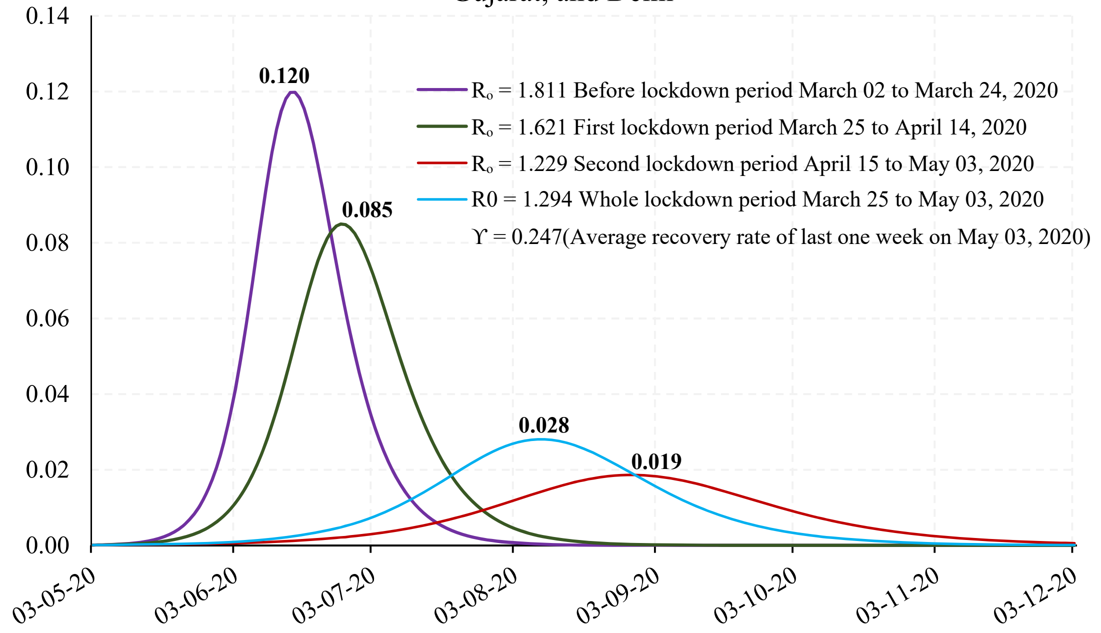
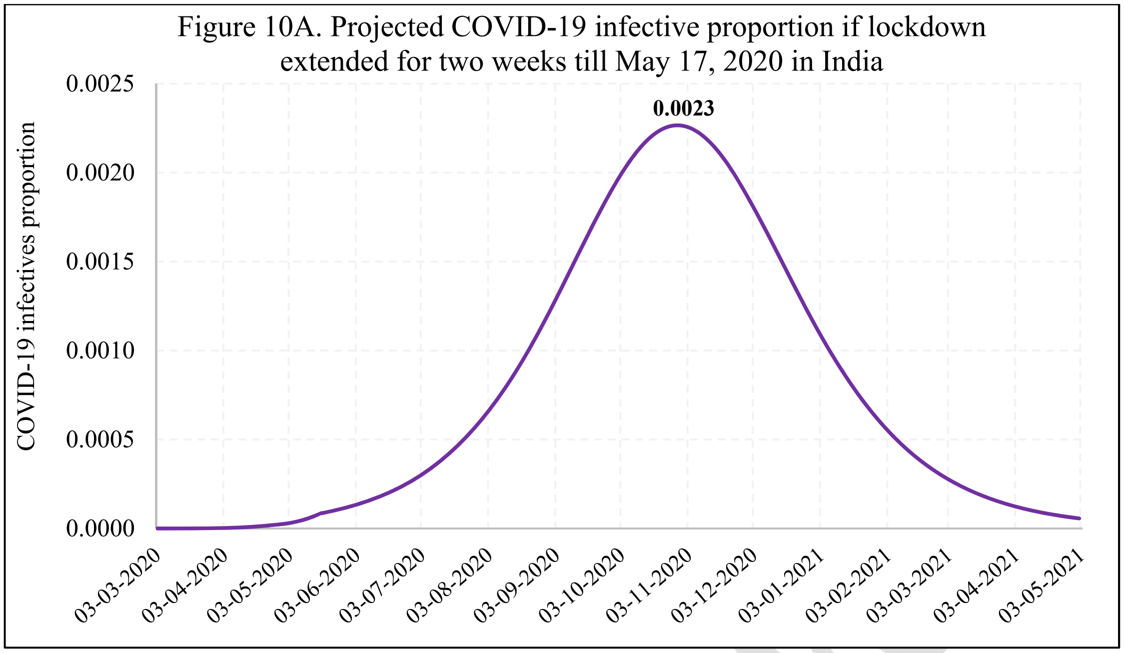


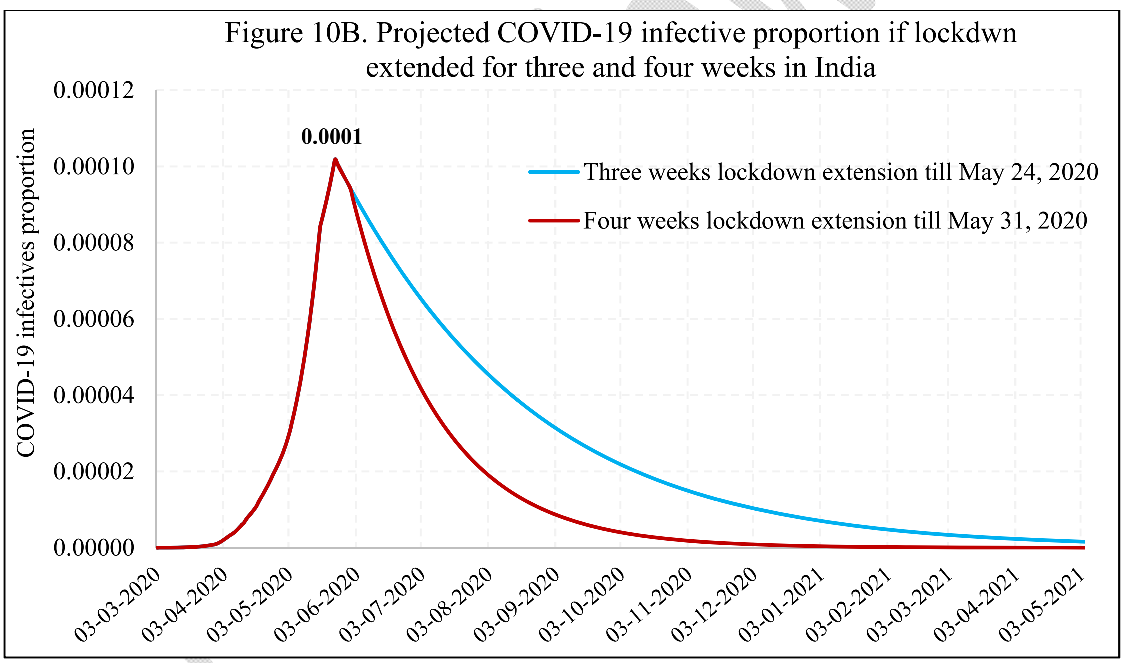
Figure 9 shows the COVID-19 infected proportions in MH, GJ, and DL from May 03 onwards. Projection based on cases occurred before the first lockdown shows that an estimated maximum 12% of the population of these stated would have got infected by mid of June 2020. The projections based on cases occurred during first lockdown indicates that an estimated maximum 8.5% of the population of these states would have got infected by the third week of June 2020. Whereas, if the R_0 as observed in second lockdown prevails, then an estimated maximum 1.9% of the population is likely to be infected by end of August 2020 and become close zero by December, 2020

Figure 9. Projected COVID-19 infected proportion in Maharashtra, Gujarat, and Delhi



Figures 10A and 10B show the projected number of COVID-19 infected cases if the lockdown is extended by 2 weeks, 3 weeks, and 4 weeks. If the lockdown is extended by two weeks until May 17 (as already announced by the Government of India) (Press Information Bureau 2020d). , an estimated maximum 0.23% of India's population is likely to get infected by the end of October 2020. In this scenario, the prevalence is likely to be zero by February 2021. If the lockdown is extended until May 24 or May 31 the R_0 is likely to become less than one (R_0 = 0.975 if lockdown extended till 24 May 2020 and R_0 = 0.945 if lockdown extended till 31 may 2020). In this scenario, COVID-19 pandemic may begin to recede and infective cases shall start to decline. As shown in Figure 10B, less than 0.01% of the India's population is likely to be infected by June or July 2020.

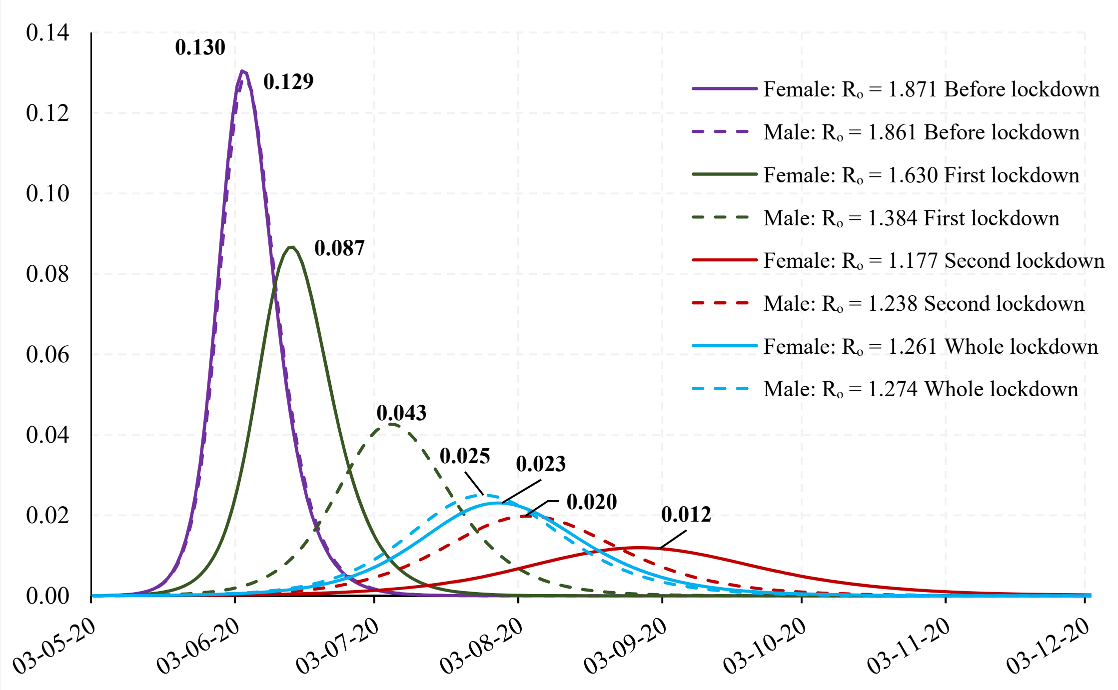




Projection of growth trajectory of COVID-19 in India by gender

The projections, based on the R_0 observed during the second lockdown, by gender suggests that more males than females are likely to be infected of COVID-19 (**Figure 11**). In this scenario, a maximum of 0.20% of male population is likely to be infected by the first week of August. Whereas, a maximum of 0.12% of female population is likely to be infected by the first week of September.

Figure 11. Projected COVID-19 infective Male and Female in India



Discussion

Using the SIR model, we projected the future trajectory of COVID-19 in India. We find a decline in estimated R_0 in the first and second lockdown. At the national level, the R_0 before the lockdown was 1.862 which declined to 1.455 during the first lockdown and further decline to 1.200 in the second lockdown. The R_0 during the total lockdown was 1.258. Interestingly, if the R_0 observed in second lockdown period continues, only a maximum of 1.5% of India's population is likely to be infected by mid of August 2020. In this scenario the prevalence is likely to reach zero by November 2020. When we estimated the R_0 for India excluding MH, GJ, and DL (the three badly affected states of India), we find a still lower R_0. In this case, the estimated R_0 during the first and second lockdowns were 1.350 and 1.162. The overall R_0 was 1.218 and the maximum population likely to be infective is 1.7%.

Our estimates of percentage infective in India is much lower compared to those obtained in some of the existing Indian studies. Singh and Adhikari (2020), using a R_0 of 2.10, reported that in absence of any social distancing or lock down the estimated infective in India are likely to be 900 million (Singh and Adhikari 2020). Estimates from Schueller et al. (2020) indicates that, with hard lockdown and continued social distancing, the peak total infections in India will be 97 million and the number of infective by September is likely to be over 1100 million

(Schueller et al. 2020). Our estimates, based on the R_0 estimated for the total lockdown and the population of India on 1st March 2020 borrowed from the Office of the Registrar General of India (ORGI 2006), lies at 30.5 million (RGI 2006). In the absence of the two lockdowns, the estimated peak population infective would have been over 171 million. These estimates clearly suggest that the two lockdowns have played a key role in reducing the number of peak infective by about 140 million. Further our projections, based on the Government of India's recommendation to extend the lockdown to May 17, 2020, indicate that only a maximum of 0.23% of India's population is likely to be infective by October 2020. Converted into numbers, only 3.0 million Indian population is likely to be infective by the end of October 2020. Coincidently, a projection, based on a statistical model 'COVID-19 Med Inventory', by a group of researchers from Jawaharlal Nehru Centre for Advanced Scientific Research (JNCAR), Indian Institute of Science-Bangalore, IIT Bombay, and Armed Forces Medical College-Pune reported that the number of positive cases is likely to touch 3.0 million (Sharma 2020).

The estimates of R_0 given in our paper is within the plausible range observed across the different countries of the world. Our estimated R_0 during the first lockdown was 1.455 and during the second lockdown was 1.200. Dwivedi et al. (2020), using a mathematical approximation, estimated an R_0 of 1.56 for the period 4-19 April 2020 (Dwivedi et al. 2020). According to Singh and Adhikari (2020), the projected R_0 ranges between 2.10 in March 2020 to 0.50 in August 2020 (Singh and Adhikari 2020). Short-term projections by Imperial College London suggests a R_t of 1.45 for India for the week starting May 03 2020 (Sangeeta Bhatia et al. 2020). Kissler et al. (2020) have shown that the maximum winter time R_0 is 2.0 and for summer time it is 1.4 (Kissler et al. 2020). Data from Wuhan China suggests that the median daily reproduction number reduced from 2.35 to 1.05 after the implementation of travel restriction, home quarantining, centralized quarantining system and treatment and improving medical resources (Inglesby 2020; Kucharski et al. 2020). The estimated R_0 for United States of America was 1.51 on April 01 2020 (Gunzler and Sehgal 2020). The time-varying reproduction number in the United States declined from 4.02 to 1.51 between March 17 and April 1, 2020 (Gunzler and Sehgal 2020). Italy recorded the reproduction number of 2.6 and 3.3 with the assumption that transmission starts from 5th Feb 2020 and 10th Feb 2020 respectively (Zhuang et al. 2020). The reproduction number in Republic of Korea was estimated at 2.6 and 3.2 (as on 5th March 2020) with the assumption transmission starts from 31st Jan. 2020 and 5th Feb 2020 respectively (Zhuang et al. 2020).

Limitations of the study may also be noted. The estimation of time-varying reproduction number is completely based on the time series of incidence of COVID-19 provided in the dataset. So, the results are likely to be affected by the accuracy and coverage of reporting. Due to limited data on cases by gender, projection of growth trajectory by gender is particularly likely to be affected. Second, as India is in early stage of the epidemic, estimation of exact value of R_0 is slightly complicated. Keeping this in mind, we have provided a 95% confidence interval to compensate for the over or under-estimation of R_0. Third, the time-varying R_0 are estimated with the assumption that the serial interval distribution remains constant throughout the epidemic. Fourth, the estimation of proportion of future infections are based on the assumption that once an individual is recovered, he or she will be immune to the disease, which may not be true for all cases. Finally, the projection after extension of lockdown for two, three, and four weeks is based on trend of SIR model parameters till May 03, 2020. Slight deviations from these trends may be observed in future.

Despite the afore-mentioned limitations, our study is perhaps the first attempt to project the future trajectory of COVID-19 in India using SIR models. Our estimates indeed suggest positive benefits of first and second lockdown in India. Moreover, our models also suggest that the third lockdown, that is currently underway, is likely to offer considerable benefits in terms of reduced infections, hospitalizations, and deaths due to COVID-19. If India is successful in strictly implementing the third lockdown, peak infective will be only 3.0 million. In case the R_0 estimated during second lockdown continues, the peak infective will be 20 million. However, a matter of concern is that COVID-19 case recovery rate in India is comparatively very low. This may be due to the early stage of the infection in the country. Government of India must try to push the CRR Total towards 100, in the absence of which the chance of re-occurrence of infection is very high.

Acknowledgement

Authors are grateful to Prof. K.S. James, Director and Senior Professor International Institute for Population Sciences for motivating the authors to conduct this study. Authors are also grateful to him for his comments/suggestions during the various stages of development of this policy paper.

Disclaimer*: This article is e-print and has not been peer-reviewed. Hence, this research should not be relied upon without context to guide clinical practice or health-related behavior and should not be considered and reported as established information without consulting multiple experts in the field.* <u>*In addition, any decision to relax or extend the nationwide lockdown must not be made based on the findings of this research.*</u>

References

Brauer, Fred and Castillo-Chavez, Carlos (2012), *Mathematical Models in Population Biology and Epidemiology* (Second edn., Texts in Applied Mathematics; New York, NY: Springer).

Cori, Anne, et al. (2013), 'A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics', *American Journal of Epidemiology,* 178 (9), 1505-12.

Daley, D. J. and Gani, J. (1999), *Epidemic Modelling: An Introduction* (Cambridge Studies in Mathematical Biology; Cambridge: Cambridge University Press).

Du, Zhanwei, et al. (2020), 'The serial interval of COVID-19 from publicly reported confirmed cases', *medRxiv*, 2020.02.19.20025452.

Mumbai, A Situational Analysis Paper for Policy Makers.

Fraser, Christophe (2007), 'Estimating individual and household reproduction numbers in an emerging epidemic', *PloS one,* 2 (8), e758-e58.

Gunzler, Douglas and Sehgal, Ashwini R. (2020), 'Time-Varying COVID-19 Reproduction Number in the United States', *medRxiv*, 2020.04.10.20060863.

Hellewell, Joel, et al. (2020), 'Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts', *The Lancet Global Health,* 8 (4), e488-e96.

Inglesby, Thomas V. (2020), 'Public Health Measures and the Reproduction Number of SARS-CoV-2', *JAMA*.

Kermack, W. O. and McKendrick, A. G. (1991), 'Contributions to the mathematical theory of epidemics—I', *Bulletin of Mathematical Biology,* 53 (1), 33-55.

Kissler, Stephen M., et al. (2020), 'Social distancing strategies for curbing the COVID-19 epidemic', *medRxiv*, 2020.03.22.20041079.

Kucharski, Adam J., et al. (2020), 'Early dynamics of transmission and control of COVID-19: a mathematical modelling study', *The Lancet Infectious Diseases,* 20 (5), 553-58.

Last, John M. (2001), *A dictionary of epidemiology.*, ed. International Epidemiological Association (Oxford, England: Oxford University Press).

Li, Qun, et al. (2020), 'Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia', *New England Journal of Medicine,* 382 (13), 1199-207.

Lourenco, Jose, et al. (2020), 'Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic', *medRxiv*, 2020.03.24.20042291.

Luo, Jianxi (2020), 'Predictive Monitoring of COVID-19', *SUTD Data-Driven Innovation Lab* (updated May 05, 20202) <https://ddi.sutd.edu.sg/>, accessed May 05.

MOHFW (2020), 'COVID-19 Statewise Status', *COVID-19 INDIA* 05 May.

Nishiura, Hiroshi, Linton, Natalie M., and Akhmetzhanov, Andrei R. (2020), 'Serial interval of novel coronavirus (COVID-19) infections', *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases,* 93, 284-86.

Press Information Bureau (2020), 'Update on Novel Coronavirus: one positive case reported in Kerala', <https://pib.gov.in/allRel.aspx>, accessed 30 JAN.

--- (2020), 'Government of India issues Orders prescribing lockdown for containment of COVID-19 Epidemic in the country', <https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1607997>, accessed 24 March.

--- (2020), 'PIB'S DAILY BULLETIN ON COVID-19', <https://pib.gov.in/PressReleasePage.aspx?PRID=1614457>, accessed 14 April.

--- (2020), 'Extension of Lockdown for a further period of Two Weeks with effect from May 4, 2020', <https://pib.gov.in/PressReleasePage.aspx?PRID=1620095>, accessed May 10.

RGI (2006), 'Population Projections for India and the States: 2001-2026', *Report of the Technical Group on Population Projections* (New Delhi: Office of the Registrar General of India, New Delhi).

Sangeeta Bhatia, et al. (2020), 'Short-term forecasts of COVID-19 deaths in multiple countries.', <https://mrc-ide.github.io/covid19-short-term-forecasts/index.html#authors>, accessed May 10.

Schueller, Emily, et al. (2020), 'COVID-19 in India: Potential Impact of the Lockdown and Other Longer-Term Policies', (Washington D.C: The Center For Disease Dynamics, Economics & Policy).

Sharma, Richa (2020), 'Projection: 38,220 deaths, 5.35 lakh COVID-19 cases in India by mid-May', *THE NEW INDIAN EXPRESS*.

Singh, Rajesh and Adhikari, Ruby (2020), 'Age-structured impact of social distancing on the COVID-19 epidemic in India', *arXiv: Populations and Evolution*.

Tang, Biao, et al. (2020), 'Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study', *medRxiv*, 2020.03.09.20033464.

WHO (2020), 'Coronavirus disease 2019 (COVID-19)', *Situation Report – 46* (Jeneva: World Health Organization).

You, Chong, et al. (2020), 'Estimation of the Time-Varying Reproduction Number of COVID-19 Outbreak in China', *medRxiv*, 2020.02.08.20021253.

Zhao, Shi, et al. (2020), 'Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020', *medRxiv*, 2020.02.21.20026559.

Zhuang, Zian, et al. (2020), 'Preliminary estimating the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea and Italy by 5 March 2020', *International Journal of Infectious Diseases*.

Table 1. Estimates of basic reproduction number (R_0) and 95% CI based on COVID-19 cases during March 02 to May 03 in India

	India	Rest of India (excluded – MH, GJ, DL)	Maharashtra, Gujarat, Delhi
Before lockdown March 02 to March 24, 2020	1.862 (1.712, 2.019)	1.888 (1.707, 2.077)	1.811 (1.548, 2.094)
First lockdown period March 25 to April 14, 2020	1.455 (1.428, 1.482)	1.350 (1.317, 1.385)	1.621 (1.574, 1.667)
Second lockdown period April 15 to May 03, 2020	1.200 (1.187, 1.214)	1.162 (1.142, 1.183)	1.229 (1.211, 1.247)
Whole lockdown period March 25 to May 03, 2020	1.258 (1.246, 1.270)	1.218 (1.200, 1.235)	1.294 (1.277, 1.311)
Average recovery rate of last one week of May 03, 2020	0.3626 (36.26%)	0.5359 (53.59%)	0.2471 (24.71%)

MH – Maharashtra, GJ – Gujarat, DL – Delhi

 # For rest of India before lockdown period is March 14 to March 24, 2020 hence whole India and rest of India may not be compared

Appendix Table 1. Different estimated values of mean and standard deviation (SD) if serial interval (SI) of COVID-19

Mean SI	SD of SI	Study
5.50		Coronavirus disease 2019 (COVID-19) Situation Report – 46
4.00	3.10	Zhao S, Gao D, Zhuang Z, et al. 2020. Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. medRxiv: 2020.02.21.20026559.
5.00	3.00	Tang B, Xia F, Bragazzi NL, et al. 2020. Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study. medRxiv: 2020.03.09.20033464.
4.50	3.10	Zhuang Z, Zhao S, Lin Q, et al. 2020. Preliminary estimating the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea and Italy by 5 March 2020. medRxiv: 2020.03.02.20030312.
4.41	3.17	You C, Deng Y, Hu W, et al. 2020. Estimation of the Time-Varying Reproduction Number of COVID-19 Outbreak in China. medRxiv: 2020.02.08.20021253.
3.96	4.75	Du Z, Xu X, Wu Y, et al. 2020. The serial interval of COVID-19 from publicly reported confirmed cases. medRxiv: 2020.02.19.20025452.
7.50	3.40	Li Q, Guan X, Wu P, et al. 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine, 382: 1199-1207.
4.70	2.90	Nishiura H, Linton NM, Akhmetzhanov AR. 2020. Serial interval of novel coronavirus (COVID-19) infections. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases, 93: 284-286.

Appendix Table 2. Parameters used for SIR model fitting as on May 03, 2020

	India	Rest of India[#]	Maharashtra, Gujarat, and Delhi
Susceptible (S)	0.999960301	0.999976927	0.999870574
Infected (I)	0.000030310	0.000016529	0.000104684
Recovered (R)[@]	0.000009389	0.000006544	0.000024742
Average recovery rate of last one week of April 30, 2020 (γ)	0.3626	0.5359	0.2471

[#]excluding Maharashtra, Gujarat, Delhi
[@]As dead are at the risk of infection hence recovered also include dead cases.

Appendix Table 3. Projected parameters used for SIR model fitting for India

	Third lockdown extension period in Whole India		
	Two weeks till May 17, 2020	Three weeks till May 24, 2020	Four weeks till May 31, 2020
Susceptible (S)	0.999901222	0.999880152	0.999884692
Infected (I)	0.000083778	0.000101848	0.0000943077
Recovered (R)[@ $]	0.00001500	0.000018	0.000021
Average recovery rate (γ)	0.42	0.435	0.45
Basic reproduction number (R_0)[$]	1.07	0.975	0.945

[@]As dead are at the risk of infection hence recovered also include dead cases.

[$]Proportion recovered (R) and basic reproduction number (R_0) are projected linearly based on cases till May 02, 2020.

Average recovery rate (γ) in last one week follows logarithmic curve hence projected log linearly.
Proportion infected (I) taken from SIR model projection based on cases till May 02, 2020 which gives I for May 17, 2020, SIR fitted model on May 17, 2020 gives I for May 24, 2020, and SIR fitted model on 24 may 2020 gives I for May 31, 2020.

Appendix Table 4. Estimates of basic reproduction number (R_0) and 95% CI based on COVID-19 cases during March 02 to May 03 in India

	Male	Female
Before lockdown March 02 to March 24, 2020	1.861 (1.673, 2.058)	1.871 (1.626, 2.132)
First lockdown period March 25 to April 14, 2020	1.384 (1.352, 1.416)	1.630 (1.576, 1.683)
Second lockdown period April 15 to May 03, 2020	1.238 (1.221, 1.255)	1.177 (1.156, 1.199)
Whole lockdown period March 25 to May 03, 2020	1.274 (1.259, 1.289)	1.261 (1.241, 1.281)
Susceptible (S)	0.999960301	
Infected (I)	0.000030310	
Recovered (R)[@]	0.000009389	
Average recovery rate of last one week of May 03, 2020	0.357 (35.7%)	

[@]As dead are at the risk of infection hence recovered also include dead cases.