Statistics and Applications {ISSN 2452-7395(online)}
Volume 18, No. 1, 2020 (New Series), pp 181-196

Critical Community Size for COVID-19: A Model Based Approach for Strategic Lockdown Policy

Sarmistha Das[1], Pramit Ghosh[2], Bandana Sen[3], Saumyadipta Pyne[4,5] and Indranil Mukhopadhyay[1]

[1]*Human Genetics Unit, Indian Statistical Institute, Kolkata, West Bengal, India*
[2]*Purulia Medical College, Purulia, West Bengal, India*
[3]*All India Institute of Hygiene & Public Health, Kolkata, West Bengal, India*
[4]*Public Health Dynamics Lab, and Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA*
[5]*Health Analytics Network, Pennsylvania, USA*

Received: 24 May 2020; Revised: 06 June 2020; Accepted: 08 June 2020

Abstract

Among the U.S. cities hit by the 1918 Spanish flu, social distancing played a pivotal role in flattening the pandemic curve. Similarly, to fight against COVID-19, restrictive mass quarantine or lockdown has been implemented as the most important controlling measure. India has already enforced a lockdown of 10 weeks and is extending the period depending on the current disease scenario. However, the idea that, if the susceptible population drops below certain threshold, the infection would naturally die out in small communities after a fixed time (following the outbreak), unless the disease is reintroduced from outside, was proposed by M. S. Bartlett in 1957. This threshold was termed as Critical Community Size (CCS).

We propose an Susceptible-Exposed-Infected-Recovered (SEIR) model that explains COVID-19 disease dynamics. Using our model, we have calculated state-specific Temporary Eradication of Spread Time (TEST) and CCS that would essentially determine the ideal number of lockdown days required and the size of quarantined population. With the given state-wise rates of death, recovery and other parameters, we have identified that, if at a place the total number of susceptible population drops below CCS, infection will cease to exist after a period of expected time to extinction (TTE), unless it is re-introduced from outside. The expected TTE suggests that the disease might take a long time to fade away from the human population in absence of pharmaceutical interventions. But we find that the disease might subside substantially after TEST. This would imply lockdown phases as much as TEST could be sufficient to contain COVID-19.

Key words: Critical community size; COVID-19; lockdown; quarantine; SEIR model; pandemic curve flattening.

AMS Subject Classifications: 00A05

Corresponding Author: Indranil Mukhopadhyay
Email: indranil@isical.ac.in

1. Introduction

In the face of COVID-19 pandemic, many countries have implemented restrictive mass quarantines or lockdown as the primary controlling measure to confine the number of secondary transmissions of the disease within countries. In absence of any specific medical treatment to treat the disease, patients are generally given only supportive care. Given the rapid Phase 3 transmission of the disease, health care systems of even developed countries are starting to face challenges within a week or two. Therefore, to prevent stage 4 transmission of the disease, along with many other countries India, which is densely populated, has resorted to complete lockdown already for more than 10 weeks and it is still counting. Available data confirms that the pandemic has already affected more than five million people in around 215 countries till date and already claimed more than 0.3 million lives across the world within approximately three months. After World Health Organisation (WHO) declared the outbreak as a pandemic, many countries initiated partial to complete lockdown as was done in some provinces of China after the outbreak started. By the end of March, one-third of the global population was under some form of lockdown.

Many countries implemented variable number of lockdown days, but none has come up with any magic figure for the ideal period of lockdown. No clear-cut guideline or rationale behind the number of lockdown days has been announced by any country or WHO till date to the best our knowledge. The initial phase of lockdown of 2-4 weeks was determined mostly on trial and error basis. The prediction on the number of trial lockdown days was possibly and partially based on the fact that an affected individual could be contagious in the first 14 days of contracting the disease and also on the information of the number of known positive cases at the time of taking decision.

The idea of quarantining a small group of people after an epidemic outbreak to arrest the disease dates back to 1950s when English statistician M.S. Bartlett introduced the term 'critical community size'. Probably the idea of such mathematical development was driven by the lessons of social distancing taught by the 1918 flu pandemic or Spanish flu. The cities with strong social distancing measures, successfully delayed its peak in deaths and maintained lower death rate (Markel et al., 2007). The flattening of 1918 flu pandemic curve that took approximately 24 weeks, was disrupted and the cities witnessed sharp increase in deaths when restrictions were temporarily relaxed after 8-10 weeks.

Bartlett (1957, 1960) proposed the idea that if the susceptible population is below some threshold, the infection is as likely as not to die out after a period of time (after the epidemic outbreak) in small communities, unless the disease is reintroduced from outside. Bartlett termed this threshold as Critical Community Size (CCS). Otherwise speaking, in absence of pharmaceutical interventions if the susceptible population that is quarantined together falls below CCS, the infection would die out from the population after a period of time unless the disease is re-introduced from outside. In the present context, CCS could guide government/health policy makers with an objective strategy of lockdown period as opposed to subjective trial and error phases of lockdown.

After an epidemic outbreak in a community, the infection persists long enough to engulf the entire susceptible population. Local extinction of the disease could be possible if the

susceptible population gets depleted. In large communities, the tendency of eventual damp down of the recurrent epidemics is balanced by random variability. But in small communities the infection would die out when the number of susceptible falls below a certain threshold, which is the CCS. Only a limited number of works (Nåsell, 2005; Anderson and Britton, 2000) including our work (under review) are available on CCS, may be because it involves complicated calculations even for simplest mathematical model viz. SI (S: Susceptible, I: Infected) model. However, since the actual extent of an epidemic can be assessed only retrospectively, it is essential to calculate the CCS for COVID-19 based on a realistic model that depends on the parameters which could be determined for a specific locality.

We propose an SEIR (S: Susceptible, E: Exposed, I: Infected, R: Recovered) model to explain the disease dynamics of COVID-19. We have derived with evidence the rationale behind the importance and extent of the lockdown period and also the number of people who could safely stay together in this lockdown phase. In absence of much prior knowledge on the disease, we have to rely on the mathematical predictions to combat the virus. In this article, we provide a cautionary note from the mathematical deductions, that this pandemic might take a very long time to fade away, in absence of any pharmaceutical interventions. Our work resonates the latest updates from WHO executive director stating "this virus may never go away". WHO also mentions that it may remain in the community as another endemic virus like human immunodeficiency virus (HIV). To have less disease transmission, WHO also stresses on enforcing withdrawal of the lockdown only when the day-to-day number of COVID-19 cases reaches the lowest possible level; otherwise, the transmission may accelerate (https://www.aninews.in/news/world/europe/who-executive-director-says-coronavirus-may-never-go-away20200514012424/).

But there is always a ray of hope. Apart from the fact that, we may learn much from Spanish flu, SARS, and MERS outbreaks, our deduction suggests that there should be no reason to panic as the lockdown, if properly followed, could contain the disease. Although we have to bear the burden of slow economic recovery or even a recession, the COVID-19 epidemic could be controlled and hopefully it would not cause a more severe public health emergency in the near future.

2. Methods

We propose an SEIR model to explain the dynamics of COVID-19 infection. The entire population is divided into four compartments. These compartments are mutually exclusive in the sense that no person can belong to more than one compartment at any time point. The four compartments are: susceptible individuals (S), individuals with and without symptoms of the disease but not yet tested positive for COVID-19 (E), infected individuals who are clinically tested positive (I), and individuals who are known to have recovered from the disease (R). Note that an individual belonging to class E may transmit the disease during

the incubation period. Under this situation, we consider the model as:

$$\frac{dS}{dt} = \Lambda - \beta\left(I(t) + \phi E(t)\right)\frac{S(t)}{N} - \mu S(t) \tag{1}$$

$$\frac{dE}{dt} = \beta\left(I(t) + \phi E(t)\right)\frac{S(t)}{N} - (\gamma + \mu)E(t) \tag{2}$$

$$\frac{dI}{dt} = \gamma E(t) - (\delta + \mu + d)I(t) \tag{3}$$

$$\frac{dR}{dt} = \delta I(t) - \mu R(t) \tag{4}$$

Here β (or $\beta\phi$) represents the contact rate for COVID-19 transmission from infected (or exposed) to susceptible individuals, an individual in E moves to I at the rate γ, δ is the recovery rate, d is death rate due to the disease and μ is the natural death rate in the population. Moreover, $\Lambda = \mu N(t)$ where $N(t)$ is the population size at time t.

Next we calculate the basic reproduction number (R_0) defined as the expected number of secondary cases produced by a single infection in a completely susceptible population. We calculate R_0 for the above model using next generation matrix $G = FV^{-1}$, where, $F = \left[\frac{\partial F_i(x_0)}{\partial x_j}\right]$ and $V^{-1} = \left[\frac{\partial V_i(x_0)}{\partial x_j}\right]$. Here, F_is are the new infections in the system, while V_i denotes the transfer of infections from one compartment to another and x_0 is the disease-free equilibrium state (section 2.2). In our model, $F = \begin{bmatrix} \beta\phi\frac{S}{N} & \beta\frac{S}{N} \\ 0 & 0 \end{bmatrix}$ and $V^{-1} = \begin{bmatrix} \gamma+\mu & 0 \\ -\gamma & \delta+\mu+d \end{bmatrix}$. R_0 is defined as the maximum eigen value of the matrix G. Based on the above model R_0 will be:

$$R_0 = \frac{\beta(\phi(\delta + \mu + d) + \gamma)}{(\gamma + \mu)(\delta + \mu + d)} \tag{5}$$

2.1. Stochastic model and quasi-stationarity

First we note the nature of transition and the respective transition rates from one compartment to another (Table 1).

We construct the fully stochastic version of the model in (1)-(4) using the transition rates in Table 1. Denoting $s = S/N$, $e = E/N$, $i = I/N$, $r = R/N$, the Kolmogorov forward equations for this process can be written as follows:

$$\begin{aligned} p'_{s,e,i,r}(t) =& \lambda_1 p_{s-1,e,i,r} + \lambda_2 p_{s+1,e,i,r} + \lambda_3 p_{s+1,e-1,i,r} + \lambda_4 p_{s+1,e,i-1,r} \\ &+ \lambda_5 p_{s,e+1,i,r} + \lambda_6 p_{s,e+1,i-1,r} + \lambda_7 p_{s,e,i+1,r-1} + \lambda_8 p_{s,e,i+1,r} \\ &+ \lambda_9 p_{s,e,i,r+1} - \kappa(s,e,i,r) p_{s,e,i,r} \end{aligned} \tag{6}$$

where $\kappa(s,e,i,r) = \sum_{j=1}^{9} \lambda_j(s,e,i,r)$.

We use Kolmogorov forward equations in order to find the expected time to extinction (TTE) and evaluate CCS based on our model. Now, conditioning on non-extinction, we

Table 1: Chart of transition rates

Event	**Transition**	**Transition rate**
Immigration of Susceptibles	$(s,e,i,r) \to (s+1,e,i,r)$	$\lambda_1 = \lambda_1(s,e,i,r) = \mu N$
Death of Susceptibles	$(s,e,i,r) \to (s-1,e,i,r)$	$\lambda_2 = \lambda_2(s,e,i,r) = \mu s$
Susceptible (S) to Exposed (E)	$(s,e,i,r) \to (s-1,e+1,i,r)$	$\lambda_3 = \lambda_3(s,e,i,r) = \beta\phi se/N$
Susceptible (S) to Infected (I)	$(s,e,i,r) \to (s-1,e,i+1,r)$	$\lambda_4 = \lambda_4(s,e,i,r) = \beta si/N$
Death of Exposed	$(s,e,i,r) \to (s,e-1,i,r)$	$\lambda_5 = \lambda_5(s,e,i,r) = \mu e$
Exposed (E) to Infected (I)	$(s,e,i,r) \to (s,e-1,i+1,r)$	$\lambda_6 = \lambda_6(s,e,i,r) = \gamma e$
Infected (I) to Recovered (R)	$(s,e,i,r) \to (s,e,i-1,r+1)$	$\lambda_7 = \lambda_7(s,e,i,r) = \delta i$
Death of Infected	$(s,e,i,r) \to (s,e,i-1,r)$	$\lambda_8 = \lambda_8(s,e,i,r) = (\mu+d)i$
Death of Recovered	$(s,e,i,r) \to (s,e,i,r-1)$	$\lambda_9 = \lambda_9(s,e,i,r) = \mu r$

have,

$$q_{s,e,i,r}(t) = P\left[S(t)=s, E(t)=e, I(t)=i, R(t)=r | E(t)\neq 0, I(t)\neq 0\right] = \frac{p_{s,e,i,r}(t)}{1-p_{\bullet 00\bullet}(t)}$$

where $p_{\bullet 00\bullet}(t) = \sum_{s=0}^{\infty}\sum_{r=0}^{\infty} P\left[S(t)=s, E(t)=e, I(t)=i, R(t)=r\right] = \sum_{s=0}^{\infty}\sum_{r=0}^{\infty} p_{s,0,0,r}(t)$.
Now, differentiating $q_{s,i,c,a}(t)$ with respect to t, we have,

$$q'_{s,e,i,r}(t) = \frac{p'_{s,e,i,r}(t)}{1-p_{\bullet 00\bullet}(t)} + \frac{p_{s,e,i,r}(t)}{\left(1-p_{\bullet 00\bullet}(t)\right)^2}.p'_{\bullet 00\bullet}(t). \tag{7}$$

Now, from (6), we have, after simplification,

$$p_{\bullet 00\bullet}(t) = \mu p_{\bullet 1\,0\,\bullet}(t) + (\delta+\mu+d)p_{\bullet 0\,1\,\bullet}(t) = p_{\bullet}^{(d,\mu,\delta)}(t) \;\; (say). \tag{8}$$

From (7-8) we have,

$$q'_{s,e,i,r}(t) = \frac{p'_{s,e,i,r}(t)}{1-p_{\bullet 00\bullet}(t)} + \frac{p_{s,e,i,r}(t)}{\left(1-p_{\bullet 00\bullet}(t)\right)}.q_{\bullet}^{(d,\mu,\delta)}(t) \;\text{ where } q_{\bullet}^{(d,\mu)}(t) = \frac{p_{\bullet}^{(d,\mu,\delta)}(t)}{1-p_{\bullet 0\,0\,0}(t)} \tag{9}$$

Now, $q'_{s,e,i,r}(t) = 0$

$$\implies p'_{s,e,i,r}(t) = -\frac{p_{s,e,i,r}(t)}{\left(1-p_{\bullet 0\,0\,\bullet}(t)\right)}.q_{\bullet}^{(d,\mu,\delta)}(t)(1-p_{\bullet 0\,0\,\bullet}(t)) = -q_{\bullet}^{(d,\mu,\delta)}(t)p_{s,e,i,r}(t)$$

$$\implies p_{s,e,i,r}(t) = ce^{-q_{\bullet}^{(d,\mu,\delta)}(t).t} = q_{\bullet}^{(d,\mu,\delta)}(0)e^{-q_{\bullet}^{(d,\mu,\delta)}(t).t} \tag{10}$$

Let τ_Q be the TTE when the initial distribution equals the quasi-stationarity distribution [Nåsell, 2005]. Hence for stationary distribution,

$$E(\tau_Q) = \frac{1}{q_{\bullet}^{(d,\mu,\delta)}}. \tag{11}$$

2.2. Equilibrium points

The disease-free equilibrium is obtained as: $\Sigma_0 = (S^0, I^0, C^0, A^0) = \Big(\frac{\Lambda}{\mu}, 0, 0, 0\Big)$.

To find the other endemic equilibrium, if exists, we put $N = N(0)$, $x_1(t) = S(t)/N$, $x_2(t) = E(t)/N$, $x_3(t) = I(t)/N$, and $x_4(t) = R(t)/N$. Then equilibrium point is obtained by equating the first differentiation to zero, i.e.

$$x_1'(t) = \mu - \beta\Big(x_3(t) + \phi x_2(t)\Big)x_1(t) - \mu x_1(t) = 0 \tag{12}$$

$$x_2'(t) = \beta\Big(x_3(t) + \phi x_2(t)\Big)x_1(t) - (\gamma + \mu)x_2(t) = 0 \tag{13}$$

$$x_3'(t) = \gamma x_2(t) - (\delta + \mu + d)x_3(t) = 0 \tag{14}$$

$$x_4'(t) = \delta x_3(t) - \mu x_4(t) = 0 \tag{15}$$

For simplicity we use the notations: $x_j(t) = x_j$ for $j = 1, \ldots, 4$.

Then solving (12) - (15), we have the endemic equilibrium as:

$$\hat{x}_1 = \frac{(\gamma + \mu)(\delta + \mu + d)}{\beta(\gamma + \phi(\delta + \mu + d))} = \frac{1}{R_0} \tag{16}$$

$$\hat{x}_2 = \frac{\mu(1 - \hat{x}_1)}{\gamma + \mu} = \frac{\mu}{\gamma + \mu}(1 - \frac{1}{R_0}) \tag{17}$$

$$\hat{x}_3 = \frac{\gamma\mu(1 - \hat{x}_1)}{(\gamma + \mu)(\delta + \mu + d)} = \frac{\gamma\mu}{(\gamma + \mu)(\delta + \mu + d)}(1 - \frac{1}{R_0}) \tag{18}$$

$$\hat{x}_4 = \frac{\gamma\delta(1 - \hat{x}_1)}{(\gamma + \mu)(\delta + \mu + d)} = \frac{\gamma\delta}{(\gamma + \mu)(\delta + \mu + d)}(1 - \frac{1}{R_0}) \tag{19}$$

2.3. Diffusion approximation

Stationary distribution of an epidemic process may be approximated with a specified multivariate normal distribution using Ornstein-Uhlenbeck process when the population size N is very large and $R_0 \geq 1$. This approximation is valid only in absence of any infection. We derive an approximate distribution of the quasi-stationarity by limiting Ornstein-Uhlenbeck process (Nåsell, 2005). We consider a diffusion approximation to the stochastic version of SEIR model.

Let the changes in the scaled state variables x_1, x_2, x_3, and x_4 during the time interval be denoted by δx_1, δx_2, δx_3, and δx_4 respectively, where $\delta x_i(t) = x_i(t + \delta t) - x_i(t)$, $i = 1, 2, 3, 4$.

Under the assumptions of the original process on sequence of transitions, we evaluate the mean vector and variance-covariance matrix for δx_i $(i = 1, 2, 3, 4)$ during the time interval $(t, t + \delta t)$ as follows.

First assume that we are in the state (S, E, I, R). Then the possible transitions from this state are:
(a) S increases by 1 at the rate μ

(b) S decreases by 1 at the rate μS
(c) S decreases by 1 and E increases by 1 at the rate $\beta\phi SE/N + \beta SI/N$
(d) E decreases by 1 and I increases by 1 at the rate γE
(e) E decreases by 1 at the rate μE
(f) I decreases by 1 at the rate $(\mu + d)I$
(g) I decreases by 1 and R increases by 1 at the rate δI
(h) R decrease by 1 at the rate μR.

The random variable δx_1 equals $\frac{1}{N}$ in case (1), $-\frac{1}{N}$ in cases (2), (3), and 0 in other cases. Similarly, δx_2 equals $\frac{1}{N}$ in case (3), $-\frac{1}{N}$ in cases (4), (5), and 0 in other cases. δx_3 equals $\frac{1}{N}$ in case (4), $-\frac{1}{N}$ in cases (6), (7), and 0 in other cases. δx_4 equals $\frac{1}{N}$ in case (7), $-\frac{1}{N}$ in case (8), and 0 in other cases.

$$\text{Then, } E(\delta \boldsymbol{x}) = b(\boldsymbol{x})\delta t + o(\delta t)$$

$$\text{where} \quad b(\boldsymbol{x}) = \begin{pmatrix} \mu - \beta(x_3 + \phi x_2)x_1 - \mu x_1 \\ \beta(x_3 + \phi x_2)x_1 - (\gamma + \mu)x_2 \\ \gamma x_2 - (\delta + \mu + d)x_3 \\ \delta x_3 - \mu x_4 \end{pmatrix} \tag{20}$$

Now to derive the variance-covariance matrix we find the Jacobian matrix of $b(\boldsymbol{x})$ at point $\boldsymbol{x}$,

$$B(\boldsymbol{x}) = \frac{\partial b(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} -\beta(x_3 + \phi x_2) - \mu & -\beta\phi x_1 & -\beta x_1 & 0 \\ \beta(x_3 + \phi x_2) & \beta\phi x_1 - (\gamma + \mu) & \beta x_1 & 0 \\ 0 & \gamma & -(\delta + \mu + d) & 0 \\ 0 & 0 & \delta & -\mu \end{pmatrix}$$

Approximating $B(\boldsymbol{x})$ at equilibrium point $\hat{\boldsymbol{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)$ by $B(\hat{\boldsymbol{x}})$, we get,

$$B(\hat{\boldsymbol{x}}) = \begin{pmatrix} -\mu R_0 & -\beta\phi x_1 & -\beta x_1 & 0 \\ \mu(R_0 - 1) & \beta\phi x_1 - (\gamma + \mu) & \beta x_1 & 0 \\ 0 & \gamma & -(\delta + \mu + d) & 0 \\ 0 & 0 & \delta & -\mu \end{pmatrix}$$

Therefore, variance-covariance matrix of $\delta \boldsymbol{x} = (\delta x_1, \delta x_2, \delta x_3, \delta x_4)'$ is,
$V(\delta \boldsymbol{x}) = \frac{1}{N} S(\boldsymbol{x})\delta t + o(\delta t)$ where,

$$S(\boldsymbol{x}) = \frac{1}{N} \begin{pmatrix} \beta(x_3 + \phi x_2)x_1 + \frac{\mu}{N} + \mu x_1 & -\beta(x_3 + \phi x_2)x_1 & 0 & 0 \\ -\beta(x_3 + \phi x_2)x_1 & (\gamma + \mu)x_2 + \beta(x_3 + \phi x_2)x_1 & -\gamma x_2 & 0 \\ 0 & -\gamma x_2 & (\delta + \mu + d)x_3 + \gamma x_2 & -\delta x_3 \\ 0 & 0 & -\delta x_3 & \delta x_3 + \mu x_4 \end{pmatrix}$$

Again approximating $S(\boldsymbol{x})$ by $S(\hat{\boldsymbol{x}})$, where $\hat{\boldsymbol{x}}$ is the equilibrium point, we obtain,

$$S(\hat{\boldsymbol{x}}) = \frac{1}{N} \begin{pmatrix} \frac{\mu}{N} + \mu & -\mu(1 - \frac{1}{R_0}) & 0 & 0 \\ -\mu(1 - \frac{1}{R_0}) & 2\mu(1 - \frac{1}{R_0}) & \frac{-\mu\gamma}{\gamma+\mu}(1 - \frac{1}{R_0}) & 0 \\ 0 & \frac{-\mu\gamma}{\gamma+\mu}(1 - \frac{1}{R_0}) & 2\frac{\mu\gamma}{\gamma+\mu}(1 - \frac{1}{R_0}) & \frac{-\delta\mu\gamma}{(\gamma+\mu)(\delta+\mu+d)}(1 - \frac{1}{R_0}) \\ 0 & 0 & \frac{-\delta\mu\gamma}{(\gamma+\mu)(\delta+\mu+d)}(1 - \frac{1}{R_0}) & 2\frac{-\delta\mu\gamma}{(\gamma+\mu)(\delta+\mu+d)}(1 - \frac{1}{R_0}) \end{pmatrix}$$

For large N, the process $\sqrt{N}(\boldsymbol{x}(t) - \hat{\boldsymbol{x}})$ is approximated by a multivariate Ornstein-Uhlenbeck (O-U) process with a local drift matrix $B(\hat{\boldsymbol{x}})$ and local variance-covariance matrix $S(\hat{\boldsymbol{x}})$.

The stationary distribution of this O-U process approximates the quasi stationary distribution. It is approximately normal with mean zero and variance-covariance matrix Σ, where Σ is obtained by solving

$$B(\hat{\boldsymbol{x}})\Sigma + \Sigma B'(\hat{\boldsymbol{x}}) = -S(\hat{\boldsymbol{x}}). \tag{21}$$

Exact analytical solution for Σ is not straightforward (Anderson and Britton, 2000). Since we are interested in calculating the CCS, we can easily solve the equation (21) numerically given the parameter values and the equilibrium point.

Let σ_{ij} be the solution for the (i,j)th element of Σ, where $i, j = 1, \ldots, 4$. Diffusion approximation guides us to consider the joint distribution of $x_1(t), x_2(t), x_3(t), x_4(t)$ as four-variate normal distribution with appropriate mean and variance-covariance matrix i.e.

$$\sqrt{N}(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}) \sim N_4(\mathbf{0}, \Sigma), \text{ with } \boldsymbol{x}(t) = (x_1(t), x_2(t), x_3(t), x_4(t))', \hat{\boldsymbol{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4),$$

$$\text{and } \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{pmatrix} \tag{22}$$

An approximation for quasi-stationary distribution is obtained from truncated multivariate normal distribution. Thus in order to evaluate expected time to extinction and subsequently the CCS, we use results from conditional truncated multivariate normal distribution. Now define $\mu_2^* = \hat{x_2} + \frac{\sigma_{23}}{\sigma_{33}}(x_3 - \hat{x_3})$, $\sigma_{22}^* = \sigma_{22} - \frac{\sigma_{23}^2}{\sigma_{33}^2}$.

To calculate $p_{\bullet 10 \bullet}$ (or $p_{\bullet 01 \bullet}$ or $p_{\bullet 00 \bullet}$) first note that for large N, $\sqrt{N}(\boldsymbol{x} - \hat{\boldsymbol{x}})$ approximately follows a four-variate multivariate normal distribution with mean zero and covariate matrix Σ, as obtained form equation (22). Now, we shall show that these terms contain product of $\frac{\phi(\nu)}{\Phi(\nu)}$ terms. Since N is unknown, we cannot evaluate its values exactly. Thus we use another approximation to $\frac{\phi(\nu)}{\Phi(\nu)}$ based on a logistic function only to make the calculation relatively simple. Using the idea that $|\sigma(\beta x) - \Phi(x)|$ is minimum when $\beta = \frac{16x}{15}$ (Birnbaum, 1963; Haley, 1952)) and putting $\sigma(z) = \frac{1}{1+e^{-z}}$ and $\beta = \frac{16}{15}\frac{\pi}{\sqrt{3}}$, we approximate $\frac{\phi(.)}{\Phi(.)}$ as,

$$\begin{aligned} \frac{\phi(\nu)}{\Phi(\nu)} &= \frac{\phi(\nu)}{\int_{-\infty}^{\nu} \phi(x)dx} = \frac{\beta\phi(\nu)}{\int_{-\infty}^{\beta\nu} \phi(\frac{y}{\beta})dy} \approx \frac{\beta\phi(\nu)}{\sigma(\beta\nu)} \text{ (Williams, 2005)} \\ &= \beta\phi(\nu)(1 + e^{-\beta\nu}) \approx \beta\Big[\frac{1 + cos(\nu)}{2\pi}\Big](1 + e^{-\beta\nu}) \text{ (Raab, 1961)} \\ &\approx \beta\frac{1 + cos(\nu)}{2\pi} \text{as, } \nu \to \infty \text{ as, } N \to \infty \end{aligned}$$

Therefore, for $y > 0$, $\Phi(y+h) - \Phi(y) \approx h.\phi(y)$, we obtain

$$
\begin{aligned}
p_{\bullet 10\bullet} &= \sum_{s=0}^{\infty}\sum_{r=0}^{\infty} P(S=s, E=1, I=0, R=r) = P(E=1, I=0) \\
&\approx P(0.5 < Nx_2(t) \le 1, 0 \le Nx_3(t) \le 0.5) \\
&\approx P(0.5 < Nx_2(t) \le 1 | 0 \le Nx_3(t) \le 0.5).P(0 \le Nx_3(t) \le 0.5) \\
&= \frac{\Phi(\frac{\sqrt{N}(\frac{1}{N}-\mu_2^*)}{\sqrt{\sigma_{22}^*}}) - \Phi(\frac{\sqrt{N}(\frac{1}{2N}-\mu_2^*)}{\sqrt{\sigma_{22}^*}})}{1-\Phi(\frac{\sqrt{N}(0-\mu_2^*)}{\sqrt{\sigma_{22}^*}})} . \frac{\Phi(\frac{\sqrt{N}(\frac{1}{2N}-\hat{x_3})}{\sqrt{\sigma_{33}}}) - \Phi(\frac{\sqrt{N}(0-\hat{x_3})}{\sqrt{\sigma_{33}}})}{1-\Phi(\frac{\sqrt{N}(0-\hat{x_3})}{\sqrt{\sigma_{33}}})} \\
&\approx \frac{\frac{0.5}{N}}{\frac{\sqrt{\sigma_{22}^*}}{\sqrt{N}}} \frac{\phi(\frac{\sqrt{N}(\frac{1}{2N}-\mu_2^*)}{\sqrt{\sigma_{22}^*}})}{\Phi(\frac{\sqrt{N}\mu_2^*}{\sqrt{\sigma_{22}^*}})} . \frac{\frac{0.5}{N}}{\frac{\sqrt{\sigma_{33}}}{\sqrt{N}}} \frac{\phi(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})}{\Phi(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})} \\
&\approx \frac{1}{2\sqrt{N}} \frac{1}{\sqrt{\sigma_{22}^*}} \beta \frac{1+cos(\frac{\sqrt{N}\mu_2^*}{\sqrt{\sigma_{22}^*}})}{2\pi} . \frac{1}{2\sqrt{N}} \frac{1}{\sqrt{\sigma_{33}}} \beta \frac{1+cos(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})}{2\pi}
\end{aligned}
$$

Thus we have,

$$
\begin{aligned}
p_{\bullet 10\bullet} &= \frac{1}{\sqrt{\sigma_{22}^*}} \frac{1}{2\sqrt{N}} \beta \frac{1+cos(\frac{\sqrt{N}\mu_2^*}{\sqrt{\sigma_{22}^*}})}{2\pi} \frac{1}{\sqrt{\sigma_{33}}} \frac{1}{2\sqrt{N}} \beta \frac{1+\cos(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})}{2\pi} \text{ when } x_3 = 0 \\
p_{\bullet 01\bullet} &= \frac{1}{\sqrt{\sigma_{22}^*}} \frac{1}{2\sqrt{N}} \beta \frac{1+\cos(\frac{\sqrt{N}\mu_2^*}{\sqrt{\sigma_{22}^*}})}{2\pi} \frac{1}{\sqrt{\sigma_{33}}} \frac{1}{2\sqrt{N}} \beta \frac{1+\cos(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})}{2\pi} \text{ when } x_3 = 1 \\
p_{\bullet 00\bullet} &= \frac{1}{\sqrt{\sigma_{22}^*}} \frac{1}{2\sqrt{N}} \beta \frac{1+\cos(\frac{\sqrt{N}\mu_2^*}{\sqrt{\sigma_{22}^*}})}{2\pi} \frac{1}{\sqrt{\sigma_{33}}} \frac{1}{2\sqrt{N}} \beta \frac{1+\cos(\frac{\sqrt{N}\hat{x_3}}{\sqrt{\sigma_{33}}})}{2\pi} \text{ when } x_3 = 0
\end{aligned}
$$

Once we find $q_{\bullet}^{(d,\mu,\delta)}$, we have an expression for expected time to extinction $\hat{E}(\tau_Q)$ using (11). Clearly, $\hat{E}(\tau_Q)$ will be a function of N. However, N is unknown. To obtain this N which is nothing but CCS, we equate median time to extinction with the quasi-period ($\hat{T_0}$) (Nåsell, 2005). The quasi-period is obtained as $\hat{T_0} = \frac{2\pi}{\theta}$ where θ is the angular frequency. The angular frequency is determined by linearisation about the critical point that corresponds to the endemic infection level (Dietz, 1975). The value of N from $E(\tau_Q)log2 = \hat{T_0}$ will be the CCS value (Nåsell, 2005).

We find the quasi-period of the oscillation about the critical point using linearisation method (Dietz, 1975). Note that for our model, the linearised system about the equilibrium point $\hat{\boldsymbol{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)'$ can be written as:

$$
\frac{d\boldsymbol{x}^*}{dt} = \begin{pmatrix}
-\frac{\beta\mu(1-\frac{1}{R_0})}{\gamma+\mu}(\phi + \frac{\gamma}{\delta+\mu+d}) & -\frac{\beta\phi}{R_0} & -\frac{\beta}{R_0} & 0 \\
\frac{\beta\mu(1-\frac{1}{R_0})}{\gamma+\mu}(\phi + \frac{\gamma}{\delta+\mu+d}) & \frac{\beta\phi}{R_0} - (\gamma+\mu) & \frac{\beta}{R_0} & 0 \\
0 & \gamma & -(\delta+\mu+d) & 0 \\
0 & 0 & \delta & -\mu
\end{pmatrix} \boldsymbol{x}^* \qquad (23)
$$

where $\boldsymbol{x}^* = \boldsymbol{x} - \hat{\boldsymbol{x}}$.

Now we can find the eigen values of the matrix in (23) and find the angular frequency, provided there are imaginary roots. Putting the values of the parameters, we can find the angular frequency (θ). The quasi-period is obtained as $\hat{T}_0 = \frac{2\pi}{\theta}$, which is independent of N.

From the relation $\hat{E}(\tau_Q)\log 2 = \hat{T}_0$ (Nåsell, 2005), we can solve for N, which is the CCS. Since we are dealing with a system consisting of more than two equations, the calculations become very complicated. Hence we find an approximate value of CCS numerically.

In a nutshell, our method at first develops a fully stochastic model corresponding to the deterministic model (1)-(4); then assuming quasi-stationarity and non-extinction of infection, expected time to extinction TTE ($\equiv E(\tau_Q)$) of the disease is derived. $E(\tau_Q)$ involves some probability terms that we evaluate using diffusion approximation of the scaled state variables (S, E, I, R). τ_Q is a function of the CCS. We derive quasi-period $\hat{T}_0$ in terms of angular frequency that is obtained using linearised system at equilibrium points. Then using the relation $E(\tau_Q)\log 2 = \hat{T}_0$ (Nåsell, 2005), we could finally evaluate the CCS for the disease dynamics of COVID-19.

3. Results

For COVID-19 transmission, we have calculated the CCS and TTE of the disease based on our proposed SEIR model. We note that the value of the CCS is approximate as we have applied some mathematical approximation while applying diffusion approximation to find the quasi-stationary distribution. The value of CCS for a community or a country, depends on its parameters which we deduce from the available information on COVID-19 till date.

We apply our method to different states in India. However, this is a general method and can be applied to any country or locality provided the values of the parameters are available. Actual fatality rate due to any epidemic could only be calculated after the epidemic gets over. But in the middle of the pandemic, it is difficult to assess. So we determine the state-specific death rate (d) at four time points at an interval of seven days based on the number of deaths in the duration of May $15-21$, May $8-14$, May $1-7$, and April $24-30$ and the total number of newly infected individuals during 7 days prior to these dates respectively. From hereon, we denote the four time points as T1, T2, T3, and T4 respectively.

Different countries have implemented varying criteria of discharging COVID-19 patients from hospitals making the actual recovery rate very difficult to calculate amid the pandemic. It is yet unknown whether all the discharged patients have fully recovered from the disease or some of them would get sick again, shortly afterwards. So we have assumed the recovery rate (δ) again at four time points at an interval of seven days based on the number of recovered patients during May $15-21$, May $8-14$, May $1-7$, and April $24-30$ and the total number of newly infected individuals during dates April 24-May 7, April $17-30$, April $10-23$, and April $3-16$ respectively. These Indian state-specific numbers for newly infected cases, death, and recovery at four time points are obtained from https://api.covid19india.org/.

Another very tricky and state dependent parameter is the rate of detection of positive cases from among the exposed pool of people, *i.e.* percentage of exposed people that are

actually tested to be COVID-19 positive. In absence of enough manpower and testing kits in this dire situation, we will not be able to know the actual proportion (γ) of the exposed who could later on become a COVID-19 patient. For calculating γ, we obtained daily state-wise test positivity rate (TPR) (https://api.covid19india.org/) that is the rate at which the exposed individuals are tested and reported to be infected daily. We calculate 7-day average TPR (using geometric mean) for each state at three time points mentioned above. Next we calculated at each time point the geometric mean of 7-day average TPRs from all states and took the maximum value as γ. We obtain, $\gamma \approx 0.04$. The rationale behind taking maximum value is due to the fact that, in India the number of tests done per million is 1823 (as on May 21, 2020), which is much less than many other countries (https://www.worldometers.info/coronavirus/). It is possible that if we had enough tests, the actual TPR could be different, rather higher. The per day rate of natural death that stabilises the population under normal scenario is $\mu = \frac{1}{70}\frac{1}{365} = 0.0000391$ (assuming average longevity of an Indian is 70 years). In absence of actual contact rate (β), we have assumed $\beta = 1.1$ (Senapati et al., 2020). Another difficult parameter to obtain is the contact rate for COVID-19 transmission from exposed to susceptible individuals ($\beta\phi$). We calculated a range of $\beta\phi$ values for all Indian states at all time points and assumed the most common rate. We obtained $\beta\phi \approx 0.0011$.

Table 2: CCS and TEST for Indian states at different time points

	T4			T3			T2			T1		
State	CCS	R_0	TEST	CCS	R_0	TEST	CCS	R_0	TEST	CCS	R_0	TEST
DL	230	5.081	19-23	350	2.186	16-19	380	1.836	15-19	720	1.563	18-21
GJ	170	2.389	13-16	350	2.121	16-19	280	1.782	14-17	70	2.575	9-13
JK	490	2.140	17-20	160	2.320	13-16	240	2.201	14-17	170	1.926	12-15
KA	40	2.416	6-10	460	1.908	17-20	380	2.808	17-20	140	2.518	12-16
MP	320	3.615	17-19	190	1.785	12-15	180	1.664	11-15	70	2.475	9-12
MH	190	2.992	14-17	170	3.317	14-17	60	2.733	9-12	260	2.166	14-18
RJ	590	2.401	19-22	670	1.616	18-21	260	2.088	14-18	310	1.913	15-18
TN	410	2.076	16-19	260	3.168	16-18	550	1.682	17-20	3380	1.037	23-26
TG	270	3.131	16-19	30	2.199	1-7	740	1.402	17-21	270	1.916	14-18
UP	570	2.124	18-21	470	1.715	16-20	450	1.855	17-20	950	1.483	19-23
WB	50	3.644	9-12	400	1.573	15-19	2510	1.179	22-26	70	2.667	10-13
PB	60	4.053	11-13	180	3.436	14-17	160	4.183	15-17			
HR	260	2.764	15-18	230	4.035	16-18				360	1.625	15-18
BR							160	2.072	12-15	470	2.632	18-21
CH				400	3.147	17-20	270	3.457	16-19			
OR				160	2.124	12-16						
UT				140	3.690	14-16						
AP	90	2.340	10-13	970	1.285	18-22						
KL	80	1.784	7-11									
AS										950	1.278	18-22
HP										2330	1.178	15-19

TEST (in weeks) gives a range of lockdown period across different Indian states at all time points; DL: Delhi, GJ: Gujarat, HR: Haryana, JK: Jammu and Kashmir, KA: Karnataka, MP: Madhya Pradesh, MH: Maharastra, PB: Punjab, RJ: Rajasthan, TN: Tamil Nadu, TG: Telangana, UP: Uttar Pradesh, WB: West Bengal, AP: Andhra Pradesh, BR: Bihar, CH: Chandigarh, OR: Odisha, UT: Uttarakhand, KL: Kerala

Table 3: TEST with minimum CCS across different time points for Indian states

States	min CCS	$T4_{lower}$	$T3_{lower}$	$T2_{lower}$	$T1_{lower}$	$T4_{upper}$	$T3_{upper}$	$T2_{upper}$	$T1_{upper}$
DL	230	16	14	13	12	19	17	16	15
GJ	170	13	13	12	9	16	16	15	13
JK	160	12	13	13	12	15	16	16	15
KA	40	6	3	7	6	10	8	10	10
MP	180	14	12	11	9	17	15	15	12
MH	60	9	10	9	7	12	13	12	11
RJ	260	15	13	14	14	18	17	18	17
TN	260	14	16	13	9	17	18	17	13
TG	30	6	1	1	1	9	7	3	6
UP	450	17	16	17	15	20	19	20	19
WB	50	9	2	1	8	12	8	5	11
PB	60	11	10	11	1	13	13	13	4
HR	230	15	16	6	12	18	18	10	16
BR	160		6	12	13		9	15	16
CH	270		16	16			19	19	
OR	160		12	6	2		16	10	6
UT	140		14			16			
AP	90	10	5			13	9		
KL	160	7				11			

$T1_{lower}$ (or $T2_{lower}$ or $T3_{lower}$): minimum TEST (in weeks) required at time point T1 (or T2 or T3), $T1_{upper}$ (or $T2_{upper}$ or $T3_{upper}$): maximum TEST (in weeks) required at time point T1 (or T2 or T3); DL: Delhi, GJ: Gujarat, HR: Haryana, JK: Jammu and Kashmir, KA: Karnataka, MP: Madhya Pradesh, MH: Maharastra, PB: Punjab, RJ: Rajasthan, TN: Tamil Nadu, TG: Telangana, UP: Uttar Pradesh, WB: West Bengal, AP: Andhra Pradesh, BR: Bihar, CH: Chandigarh, OR: Odisha, UT: Uttarakhand, KL: Kerala

Table 4: CCS and TEST for India at different time points

	δ	d	CCS	R_0	TEST
T4	0.4799	0.0327	620	2.1793	23-27
T3	0.4502	0.0321	470	2.3142	21-25
T2	0.4717	0.0305	690	2.2243	23-28
T1	0.5203	0.0270	380	2.0439	19-24

We have computed state-wise CCS at time points T1, T2, T3, and T4 (Table 2). For a few states, data were missing at some or all time points and so, CCS could not be obtained. While calculating CCS we observed that the expected TTE in absence of specific treatment or vaccine, is very large. But the most interesting observation from our study is that complete lockdowns or restrictive quarantines for a definite period might eradicate the disease almost completely. We term this period as Temporary Eradication of Spread Time (TEST) for the disease, which is immensely less than the expected TTE for the disease. Although the

disease might continue to exist for a very long time on the planet in absence of pharmaceutical interventions, soothing part from our deduction is that the virus might not be able to create any havoc on its return. Our study thus provides a rationale behind the determination of the lockdown period in different Indian states going through the catastrophic effect of the pandemic. This work may aid public health workers to strategise lockdown policies.

For example, we find in Table 2, CCS of Delhi (DL) at time T2 is 380 and TEST is $15-19$. This would mean that based on the demographic figures corresponding to time T2, if the susceptible population (or community size of quarantined people) of DL is below 380, the infection will subside substantially after around $15-19$ weeks of mass quarantine/ restrictive lockdown, unless it is re-introduced from outside. TEST for DL across time points T1, T2, T3, and T4 suggests that R_0 is decreasing. But to understand whether DL or the other states are improving from the lockdown or not, we need to note Table 3. In Table 3, we find for DL if the susceptible population is below minimum CCS value among all time points, both lower and upper limit of TEST decreases over time. This suggests DL is improving in the sense that the number of lockdown period is decreasing over time. If we observe that TEST is increasing over time, it would suggest apart from the fact that lockdown should be increased in those states, the level of infection is increasing.

Using the demographic data for India (`https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series`), we find the overall CCS and TEST for India at time points T1, T2, T3, and T4 (Table 4). To compare the infection status of India, we obtained TEST at all time points with minimum CCS. We observe that if the susceptible population is below minimum CCS value (which is 380) among T1, T2, T3, and T4, TEST is almost $19-24$ weeks at all time points. So, it suggests India might have to wait at least another 5-6 months for the pandemic curve to flatten while maintaining maximum possible social distancing norms and in some situations complete lockdown, in absence of specific treatment.

4. Discussion

As things stand at present, the number of COVID-19 cases from many developed countries have surpassed those of China, from where this infection had originated. In such a dire situation, it is very difficult to propose any quintessential lockdown period specific to any country or state. The whole world is struggling to obtain unbiased data to predict on the pandemic. At the same time many questions arise in our minds as "Will the implemented number of lockdown days eradicate the virus?" or "Will it come soon again after the lockdown is over?" or like "How long should the lockdown be continued for the pandemic curve to flatten?" In this scenario of utter dilemma, with the available world-wide data, we provide state-wise estimates of the ideal lockdown phases using our proposed mathematical model for the Indian states. To the best of our knowledge, any guideline for country-wise mathematical prediction of lockdown days is not available till date. So, as the famous British statistician George E. P. Box pointed out, "All models are wrong but some are useful", we only hope that our deductions will provide some helpful suggestions to the policy-makers and public health practitioners, while we are all affected in the pandemic to varying degrees.

Our work suggests that if people are quarantined in limited groups presented by state-

specific CCS, the disease might become contained after the corresponding expected TTE, unless the disease is re-introduced from outside. We observe that although TTE in this case is a very long time (in absence of pharmaceutical interventions), the infection would subside almost completely after TEST. WHO also predicts that COVID-19 virus might continue to stay among us like the HIV. This fact matches with our observation. The TEST that we observed for Indian states or India as a whole shows a trajectory similar to Spanish flu virus. Markel et al. (2007) observed that during the 1918 Spanish flu, the overall deaths in the US cities roughly occurs over a period of 24 weeks. COVID-19 appears to have roughly a similar timeline as that of the Spanish flu. We also observe that as the contact rate (β) of infected with the susceptible increases, R_0 increases steeply. This would mean that if the lockdown is withdrawn before the infection level becomes substantially low, a second wave of infection may hit the society. During the Spanish flu, the cities that had terminated lockdown before the infection was substantially contained, witnessed another abruptly increasing death peak after a short while. In this direction, our work suggests that a lockdown should be taken very seriously to fight against COVID-19 pandemic. This paper provides evidence of the fact that even after the lockdown phase, the disease may recur but it is not expected to create a comparable pandemic situation. Presently, where contact tracing of the infected individuals can lead to tracing down of the exposed individuals, we suggest their quarantining in different feasible groups of sizes not exceeding the state-specific CCS, so that after TEST (as specified for each state) the disease may subside substantially unless any infection is re-introduced from outside. We understand that there could be some flexibility in the lockdown implementation strategies owing to the mathematical approximations made in our calculation of CCS and TEST of the disease. We note here that our model is robust to these approximations.

Another undeniable consequence of the current pandemic is the great negative impact on the economy. This is further magnified by the near paralysed state of transactions in many sectors due to lockdown. Moreover, there is a fear among the general population that if any infection recurs, it might lead to another round of spread of the disease. This fear, which is not unrealistic, may extend the lockdown further. However, after the scheduled lockdown period, if any individual gets infected and a few others get exposed to that person, we need to check whether the total number of such individuals is less than the CCS. If so, the group needs to be quarantined in smaller feasible groups to protect the rest of the population. A newly exposed group, if larger than the CCS, may be quarantined in separate, local subgroups of size no larger than CCS. Moreover, such localised lockdown or quarantine should help in preserving somewhat the daily flow of life and livelihood, and might thereby prevent, at least to some extent, the economy from being further weakened.

However, lockdown in its truest sense may not be feasible in a vast and diverse country like India. Therefore, a strategy of localised and limited lockdowns of objectively identified selected high risk population might be a cost-effective option compared to a generalised "blanket" lockdown. This would imply comprehensive screening for cases and thorough tracing of contacts. So, our take-home message during the still unfolding COVID-19 pandemic is that, till the end of TEST, we must be vigilant and careful. With any further onset of COVID-19 cases in the future, we should follow the quarantine guideline as objectively and

humanely as possible. Like any other epidemic, COVID-19 has the tendency to recur but it might not create any alarming pandemic in the future provided we keep a vigilant eye on our hygiene and have vaccinations and/or treatments. Surely, the realities on the ground − involving human life and death − are much more complex than any model can possibly ever capture. We humbly present the findings of our model as possible instruments of guidance in order to supplement relevant public policies based on ethics and ground realities.

Acknowledgments

SD and IM acknowledges the support by the NIH Fogarty International Center Award #R25 TW009717.

References

Bartlett, M.S. (1957). Measles periodicity and community size. *Journal of the Royal Statistical Society Series A (General).* **120(1)**, 48-70.

Bartlett, M.S. (1960). The critical community size for measles in the United States. *Journal of the Royal Statistical Society: Series A (General).* **123(1)**, 37-44.

Markel, H., Lipman, H.B., Navarro, J.A., Sloan, A., Michalsen, J.R., Stern, A.M., et al. (2007). Non-pharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. *JAMA.* **298(6)**, 644-654.

Nåsell, I. (2005). A new look at the critical community size for childhood infections. *Theoretical population biology.* **67(3)**, 203-216.

Andersson, H. and Britton, T. (2000). Stochastic epidemics in dynamic populations: quasi-stationarity and extinction. *Journal of mathematical biology.* **41(6)**, 559-580.

Dietz, K. (1975). Transmission and control of arbovirus diseases. *Epidemiology.* 104-121.

Birnbaum, A. and Dudman, J., (1963). Logistic order statistics. *The Annals of Mathematical Statistics*, pp.658-663.

Haley, D.C. (1952). Estimation of the dosage mortality relationship when the dose is subject to error *No. TR15, STANFORD UNIV CA APPLIED MATHEMATICS AND STATISTICS LABS.*

Williams, D., Liao, X., Xue, Y. and Carin, L. (2005, August). Incomplete-data classification using logistic regression. *InProceedings of the 22nd International Conference on Machine learning*, pp. 972-979.

Raab, D.H. and Green, E.H. (1961). A cosine approximation to the normal distribution. *Psychometrika* **26(4)**, pp.447-450.

Senapati, A., Rana, S., Das, T. and Chattopadhyay, J. (2020). Impact of intervention on the spread of COVID-19 in India: A model based study. *arxiv.org/abs/2004.04950.*

APPENDIX

Table 5 gives the values of δ and d for four time points considered in our calculation. 'NA' indicates that these parameters cannot be calculated based on the available data. As for example, if there is no death during the period under consideration, the value of d would be zero, as in case of Orissa. For such cases we did not provide CCS and TEST.

Table 5: δ and d values for different time pints

	T4		T3		T2		T1	
State	δ	d	δ	d	δ	d	δ	d
DL	0.212	0.005	0.505	0.004	0.594	0.014	0.700	0.016
GJ	0.422	0.043	0.464	0.061	0.59	0.037	0.395	0.036
JK	0.508	0.012	0.476	0.003	0.5	0.006	0.554	0.024
KA	0.44	0.02	0.552	0.032	0.376	0.019	0.427	0.014
MP	0.264	0.042	0.587	0.038	0.643	0.028	0.430	0.019
MH	0.336	0.035	0.302	0.032	0.378	0.028	0.488	0.025
RJ	0.443	0.02	0.664	0.028	0.515	0.018	0.569	0.013
TN	0.528	0.008	0.34	0.009	0.656	0.008	1.085	0.004
TG	0.348	0.006	0.503	0.003	0.766	0.033	0.553	0.029
UP	0.51	0.015	0.635	0.016	0.585	0.017	0.725	0.029
WB	0.253	0.051	0.487	0.224	0.896	0.059	0.389	0.027
PB	0.253	0.02	0.294	0.028	0.261	0.003	NA	NA
HR	0.392	0.01	0.25	0.024	NA	NA	0.682	0.006
BR	NA	NA	NA	NA	0.532	0.005	0.418	0.003
CH	NA	NA	0.333	0.019	0.302	0.019	NA	NA
OR	NA	NA	0.512	0.012	NA	NA	NA	NA
UT	NA	NA	0.25	0.05	NA	NA	NA	NA
AP	0.468	0.008	0.866	0.008	NA	NA	NA	NA
KL	0.615	0.011	NA	NA	NA	NA	NA	NA
AS	NA	NA	NA	NA	NA	NA	0.833	0.045
HP	NA	NA	NA	NA	NA	NA	3.333	0.029