Special Proceedings: ISBN#: 978-81-950383-1-2
24th Annual Conference, 23–27 February 2022; pp 69- 82

# Method Based on State-Space Epidemiological Model for Cost-Effectiveness Analysis of Non-Medical Interventions- A Study on COVID-19 in California and Florida

**Vishal Deo**[1] **and Gurprit Grover**[2]
[1]*Department of Statistics, Ramjas College, University of Delhi, Delhi, India*
[2]*Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi, India*

Received: 02 June 2022; Revised: 12 July 2022; Accepted: 15 July 2022

## Abstract

Non-medical containment measures, like quarantine, lockdown, travel restrictions, physical distancing *etc.*, are paramount towards containing the spread of a novel epidemic, especially at its initial stage when little is known about its transmission dynamics and the pathogen responsible for the infections. For these containment measures to be effective, timely identification of infectives through clinical testing is essential. To stress upon the importance of extensive random testing for breaking the chains of transmissions, we have designed a detailed framework for carrying out cost-effectiveness analysis (CEA) of extensive random testing in comparison to targeted testing (the testing policy followed by most countries). This framework can be easily extended to CEA of any other non-medical or even medical interventions for containing epidemics.

We have used the state-space susceptible-infected (quarantined/ free)-recovered-deceased model, which enables predictions of transmission dynamics in the presence of undetected cases, to forecast epidemiological parameters under the two scenarios being compared. The health outcomes have been measured in terms of the estimates of total number of deaths, and infections prevented because of the intervention. Since long-term generic health state measurement is not involved in this study, utility scores are not required for evaluating health benefits.

As a demonstration, the proposed methodology is applied to the COVID-19 data of California and Florida to carry out CEA of 'extensive random testing' over 'targeted testing' for containing the spread of the epidemic. During the period of the study, these two states were among the worst affected states in the USA, and also had very high percentages of positivity of COVID-19 tests, which raised speculations of inadequate testing capacity.

*Key words*: State-space epidemic model; Underreporting; MCMC; Cost-effectiveness analysis; Random testing; Non-medical interventions; SI(Q/F)RD model.

## 1. Introduction

Whenever we encounter an epidemic, the best medical intervention we can think about for containing its spread is a quick resort to mass vaccination of the susceptible population. However, when we face a pandemic like COVID-19, caused by the SARS-CoV-2 virus, the

Corresponding Author: Vishal Deo
Email address: vishaaldeo@gmail.com; vishal_deo@ramjas.du.ac.in

novelty of the virus puts up an arduous challenge before the scientists to develop an effective vaccine in a short span of time. Further, the necessary safety protocols underlining the testing and approval of vaccines, followed by the herculean task of manufacturing it in abundance, makes it practically impossible to get a potent vaccine within a year of the outbreak of the pandemic. Consequently, it is of paramount importance to strategically implement non-medical interventions, like physical distancing, quarantine, lockdown measures *etc.*, to minimize the spread of the infection. The rationale behind these non-medical containment measures is to break the chain of infections by bringing down the basic reproduction number/ rate, $R_0$, below one. $R_0$ is an important factor for risk assessment of any epidemic and is defined as the expected number of secondary cases that arise from a typical infectious index-case in a completely susceptible host population. When $R_0$ is less than one, an infected case is expected to produce less than one new infected. This marks the decline in the number of infecteds over time and, eventually, the epidemic dies out.

Success of any non-medical containment measure relies heavily on the ability to have sufficient testing capacity to identify and isolate the infected people. Even the strongest of the lockdown measures will fail to serve its purpose of breaking transmission chains unless it is accompanied with sufficient amount of random testing. Further, as also argued by the W.H.O, high positivity rate of testing potentially indicates insufficient testing capacity in the region (Deo and Grover (2021)). This leads to a significant amount of underreporting of cases. Significant underreporting of COVID-19 cases in various countries, including the U.S.A., has been reported by various scientific studies (Deo and Grover (2021), Wu *et al*. (2020), Lau *et al*. (2020)). Or, in other words, in the absence of sufficient testing capacity, lockdown measures can only succeed in delaying the spread of the epidemic. W.H.O has issued repeated appeals and advisories to all countries to employ extensive random testing (World Health Organisation (2020 a)). However, only a few countries showed any conviction to conduct adequate number of COVID-19 tests and confined their strategy to testing of symptomatic and high-risk people only. Citing these reasons, we have considered analysing the effectiveness of extensive random testing over targeted testing as a non-medical intervention in containing the spread of COVID-19- both in terms of effectiveness in reducing transmission rates and the associated costs. By the phrase ‘targeted testing’ we imply testing of only symptomatic and high-risk people. To perform the cost-effectiveness analysis (CEA), we have considered the case of two of the worst affected states of USA, California and Florida, which had very high percentages of positive tests. Since the level of testing, and protocols/ procedure of reporting of number of deaths vary between different state jurisdictions, the level of underreporting of deaths and cases can also be expected to vary between states. This is the reason that we have performed state-wise analyses, rather than analysing the combined data of the USA. For forecasting the transmission dynamics of the pandemic under different assumptions regarding prevalence of underreporting, we have used the state-space susceptible-infected (quarantined/ free) -recovered- deceased (SI(Q/F)RD) model given by Deo and Grover (2021). It should be noted that, although underreporting of cases can occur because of various other reasons, we have assumed that lack of sufficient testing is the primary reason for underreporting.

## 2. Methodology

To realize the objective of conducting CEA of extensive random testing against targeted testing, we propose the following sequence of steps, which are then implemented on the COVID-19 time-series data of California and Florida.

## 2.1. Predictions using the state-space SI(Q/F) RD model

The Dirichlet-Beta state-space SI(Q/F) RD model, proposed by Deo and Grover (2021), is defined as follows.

### 2.1.1. Defining transitions between different compartments of the model

The states and transitions of the compartmental set-up of the SI(Q/F)RD model can be visualised in Figure 1. Further, these transitions are quantified through the following set of differential equations.

$$\frac{d\theta_t^S}{dt} = -\left[\beta_1\theta_t^Q + \beta_2\theta_t^F\right]\theta_t^S \quad (1)$$

$$\frac{d\theta_t^I}{dt} = \left[\beta_1\theta_t^Q + \beta_2\theta_t^F\right]\theta_t^S - \gamma_1\theta_t^Q - \gamma_2\theta_t^F - d_1\theta_t^Q - d_2\theta_t^F \quad (2)$$

$$\frac{d\theta_t^R}{dt} = \gamma_1\theta_t^Q + \gamma_2\theta_t^F = \gamma\theta_t^I \ (if\ \gamma_1 = \gamma_2 = \gamma) \quad (3)$$

$$\frac{d\theta_t^D}{dt} = d_1\theta_t^Q + d_2\theta_t^F \quad (4)$$

$$where, \theta_t^Q = p_t\theta_t^I\ and\ \theta_t^F = (1-p_t)\theta_t^I, and\ \theta_t^S + \theta_t^I + \theta_t^R + \theta_t^D = 1 \quad (5)$$

Here, $\theta_t^S, \theta_t^I, \theta_t^Q, \theta_t^F, \theta_t^R$ and $\theta_t^D$ are the true but unobserved (latent) prevalence of susceptibles, infecteds, infected and quarantined, infected and free (undetected), recovered, and deceased respectively. In other words, they are the probabilities of a person being in the respective compartments at time *t*.

### 2.1.2. Dirichlet-Beta state-space formulation of the SI(Q/F) RD model

Let $\boldsymbol{\theta}_t = (\theta_t^S, \theta_t^I, \theta_t^R, \theta_t^D)^T$ be the latent population prevalence, and $\boldsymbol{f}(\boldsymbol{\theta}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{d})$ be the solution of the set of differential equations for time *t*, where the function takes the values of the vectors $\boldsymbol{\theta}_{t-1}$, $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$ , $\boldsymbol{d} = (d_1, d_2)^T$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)^T$ as the arguments. Then the Bayesian hierarchical Dirichlet-Beta state-space SI(Q/F)RD is defined as follows [Deo and Grover (2021)].

$$Y_t^I|\boldsymbol{\theta_t}, \tau \sim Beta\left(\lambda^I\theta_t^I, \lambda^I(1-\theta_t^I)\right) \quad (6)$$

$$Y_t^R|\boldsymbol{\theta_t}, \tau \sim Beta\left(\lambda^R\theta_t^R, \lambda^R(1-\theta_t^R)\right) \quad (7)$$

$$Y_t^D|\boldsymbol{\theta_t}, \tau \sim Beta\left(\lambda^D\theta_t^D, \lambda^D(1-\theta_t^D)\right) \quad (8)$$

$$\text{and, } \boldsymbol{\theta_t}|\boldsymbol{\theta_{t-1}}, \tau \sim Dirichlet(\kappa f(\boldsymbol{\theta_{t-1}}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{d})) \quad (9)$$

where, $\tau = \{\boldsymbol{\theta_0}, \kappa, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{d}, \lambda^I, \lambda^R, \lambda^D\}$, $\boldsymbol{\theta_0}$ is the baseline value of the vector $\boldsymbol{\theta_t}$, and $\lambda^I, \lambda^R, \lambda^D, \kappa > 0$ control the variances of the distributions defined in equations (6), (7), (8) and (9) respectively. Prior distributions of the model parameters are defined as follows.

$$\theta_0^I \sim Beta(1, (Y_1^I)^{-1}), \theta_0^R \sim Beta(1, (Y_1^R)^{-1}), \theta_0^D \sim Beta(1, (Y_1^D)^{-1}), \theta_0^S = 1 - \theta_0^I - \theta_0^R - \theta_0^D \quad (10)$$

$$R_i \sim LogN\left(\mu_{r_i}, \sigma_{r_i}^2\right), \sigma_{r_i}^2 = ln\left(\frac{V(R_i)+(E(R_i))^2}{(E(R_i))^2}\right) \text{ and } \mu_{r_i} = ln\left(E(R_i)\right) - \frac{\sigma_{r_i}^2}{2}, i = 1,2 \quad (11)$$

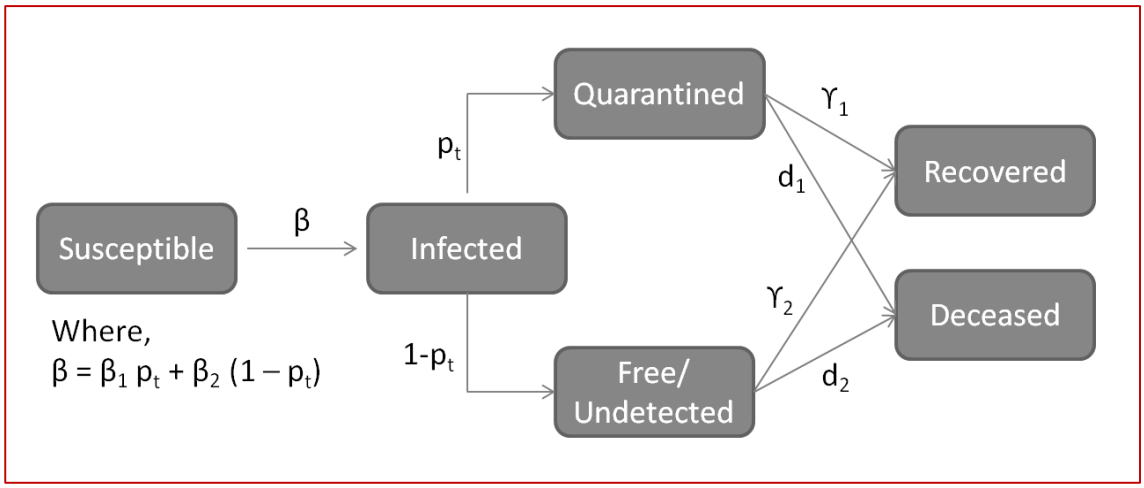
$$\gamma_i \sim LogN\left(\mu_{g_i}, \sigma_{g_i}^2\right), \sigma_{g_i}^2 = ln\left(\frac{V(\gamma_i)+(E(\gamma_i))^2}{(E(\gamma_i))^2}\right) \text{ and } \mu_{g_i} = ln\left(E(\gamma_i)\right) - \frac{\sigma_{g_i}^2}{2}, i = 1,2 \quad (12)$$

$$p_t \sim Beta\left(a_p, b_p\right), \forall t = 1,2 \ldots\ldots T \quad (13)$$

$R_1$ and $R_2$ are basic (average) reproduction rates associated with quarantined (Q) and undetected (F) infecteds, respectively. That is, $R_i = \frac{\beta_i}{(\gamma_i + d_i)}$ , $i = 1,2$.

$$\kappa \sim Gamma(a_k, b_k), \lambda^I \sim Gamma(a_I, b_I), \lambda^R \sim Gamma(a_R, b_R), \lambda^D \sim Gamma(a_D, b_D) \quad (14)$$

Elaborate procedure for the estimation of the parameters and hyper-parameters, and for forecasting using this model is outlined in Deo and Grover (2021). These procedures are used to predict the number of infections and deaths under the base intervention- targeted testing, using current estimates of underreporting based on the observed data.




*Source: Deo and Grover (2021)*

**Figure 1: SI(Q/F)RD model structure- $p_t$ is the proportion of infecteds detected and quarantined, $1\text{-}p_t$ is the proportion of infecteds who are undetected and roaming freely among the susceptible, $\beta_1$ is the transmission rate associated with quarantined infected and $\beta_2$ is the transmission rate associated with undetected infected, $\gamma_1$ and $d_1$ are rate of recovery and rate of death for quarantined cases and $\gamma_2$ and $d_2$ are rate of recovery and rate of death for undetected cases.**

## 2.2. Prediction under the assumption of extensive random testing and CEA

Extensive random testing can be expected to result in a significant rise in expenditure on the testing kits and medical personnel. However, it can play a major role in breaking the chains of transmission and hence, result in a significant reduction in the overall number of infecteds and deaths due to the COVID-19 epidemic. The outcome of CEA will tell us how much additional overall cost is required to save one additional person from getting infected, or to save one additional person from dying due to the infection. That is, CEA will be conducted in terms of the outcomes, 'infection' and 'death'. It should also be noted that, if the total duration of the epidemic is reduced drastically because of the recommended intervention 'extensive random testing', the overall expected cost may even come out to be lesser than the expected cost of using targeted testing strategy.

To derive the outcomes pertaining to the recommended intervention, *i.e.*, 'extensive random testing', following procedure is followed.

a. In terms of the SI(Q/F)RD model, the major difference between the outcomes of the two scenarios will rely on the difference in the proportion of infecteds being detected and quarantined, *i.e.*, $p_t$.
b. It will be impractical to assume that 100% infecteds can be detected using extensive random testing. This is because even popular tests like the reverse transcription polymerase chain reaction (RT-PCR), which is also recommended by the W.H.O. [World Health Organisation (2020 b)], do not have 100% sensitivity and specificity. Sensitivity and specificity may vary according to the laboratory settings and expertise levels of the medical practitioners. Different studies have reported varying levels of sensitivity and specificity of the RT-PCR test, mostly ranging from around 80% to 95% [(West *et al.* (2020), Padhye (2020), Tahamtan and Ardebili (2020)]. On a conservative note, we have assumed that the average proportion of detection of infecteds will be 80%, *i.e.*, $p_t$ will have a mean of 0.8. Instead of assigning a fixed value to $p_t$, we have assumed $p_t$ to follow Beta distribution to introduce realistic variability in the calculations. The mean of the distribution is taken as 0.8 and its variance is obtained from the results of the state-space model estimated by the method described in section 2.6.
c. To simulate a practically realistic situation, we have assumed that the extensive random testing can be applied only after first 30 days of the outbreak of the epidemic. This is because, extensive random testing requires procurement of testing kits and other logistic arrangements on a large scale, which need some time to be organized. To accommodate this assumption into calculations, the mean value of the distribution of $p_t$ for the first 30 days can be based on the average of posterior estimates of $p_t$ for the first 30 days obtained from the state-space SI(Q/F)RD model. That is, we are assuming that there will not be much difference in the outcomes and costs associated with the two interventions in the initial days of the epidemic.
d. Simulation exercise:
    i. At each $t$, $t = 1, 2, \ldots, T$, $L$ number of values are generated on the parameters $R_1, R_2, \gamma_1$ and $\gamma_2$ from their distributions defined in the equations (11) and (12). The parameters of these distributions are calculated on the basis of their posterior estimates obtained from the state-space SI(Q/F)RD model. The corresponding values on $p_t$ are simulated from its distribution defined in the previous step c. Fixed values of the death rates, $d_1$ and $d_2$, are assumed to be same, as also for the state-space SI(Q/F)RD model.
    ii. For each combination $(t, l)$, $t = 1,2,\ldots, T$ and $l = 1,2,\ldots,L$ , the respective simulated values of the parameters are used in the fourth degree Runge-Kutta approximation of the solution of the set of differential equations of the SI(Q/F)RD model to obtain $f(\boldsymbol{\theta}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{d})$, *i.e.*, the values of the latent process at time $t$ as a function of their values at time $t-1$. At the start of the iteration, the initial values of these latent process variables are assigned as the vector $\boldsymbol{\theta}_0$. The mean of the $L$ values of the latent process at a time $t$ is taken as its estimate, *i.e.*, $\widehat{\boldsymbol{\theta}}_t = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{\theta}_t^{(l)}$. Sample quantiles (0.025, 0.975) are used to obtain 95% credible intervals at each $t$.
    iii. At each $t$, $t = 1,2,\ldots,T$, $L$ values of $\lambda^I, \lambda^R$ and $\lambda^D$ are simulated from their respective Gamma distributions whose parameters are calculated from the posterior estimates of their means and variances obtained from the state-space SI(Q/F)RD model. At each combination $(t, l)$, $t = 1,2,\ldots, T$ and $l = 1,2,\ldots,L$ , using the estimate of the latent prevalence process, $\widehat{\boldsymbol{\theta}}_t$, from the previous step and the generated values of $\lambda^I, \lambda^R$ and $\lambda^D$, $(Y_t^{I(l)}, Y_t^{R(l)}, Y_t^{D(l)})$ are simulated from their respective Beta distributions. Finally, mean of these $L$ values at a time $t$ is taken as the estimate of

the observed process at *t*. These proportions can be multiplied with the total number of susceptibles (total population of the state) and rounded to obtain the estimated counts of each compartment at time *t*, $t = 1,2,\dots T$.

e. Total number of infected cases and total number of deaths, till the end of the epidemic, are calculated from the predictions for each case (interventions). These values give us the difference in outcomes (infection/ death) under two interventions. Let, ($C_1$, $D_1$) be the estimates of total number of infecteds and total number of deaths during the entire course of the epidemic for the base intervention, targeted testing, and ($C_2$, $D_2$) be the respective estimates for the recommended intervention, extensive random testing.

To obtain the estimate of total costs associated with the two interventions we will first need to estimate the total number of tests that will be conducted under the two testing strategies (interventions). For the base intervention of targeted testing, the current percentage of positivity of tests in the state can be used to obtain an estimate of the total number of tests to be conducted by the end of the epidemic. If $r_1$ is the current proportion of positive tests in the state, the estimate of total number of tests which will be conducted under the base intervention will be given as, $N_1 = \frac{Q_1}{r_1}$, where $Q_1$ is the number of infecteds who are detected and quarantined. For the second intervention of extensive random testing, the proportion of positive tests is taken as the probability that a person in the state got infected during the entire duration of the epidemic and is simply given as, $r_2 = \frac{C_2}{Total\ Population}$. Subsequently, the total number of tests under the second intervention is estimated as, $N_2 = \frac{Q_2}{r_2}$, where $Q_2$ is the number of infecteds who are detected and quarantined under the intervention extensive random testing. As an alternative, $N_2$ has also been taken as the total population, assuming that all individuals are tested (once) by the time the epidemic gets over in the state.

Let $Z$ be the per unit average cost of COVID-19 test, then the incremental cost-effectiveness ratio (ICER) is calculated as the ratio of change in cost to the change in outcome as follows,

$$\text{ICER}_{\text{inf}} = \frac{(N_2 - N_1)Z}{(C_1 - C_2)} \quad \text{and} \quad \text{ICER}_{\text{death}} = \frac{(N_2 - N_1)Z}{(D_1 - D_2)} \tag{15}$$

## 3. Implementation and Results

### 3.1. Data

In this paper, we have used the same data for conducting the CEA which was used for demonstrating the estimation and prediction methodology of state-space SI(Q/F)RD model in Deo and Grover (2021). Description of the data is provided in Table 1.

### 3.2. Estimates and predictions for the base intervention- targeted testing

Posterior estimates of the parameters of the Dirichlet-Beta state-space SI(Q/F)RD model and the predicted values of number of infecteds and deaths based on these estimates are taken from the results of Deo and Grover (2021). These results are presented in the Appendix A in the Table A.1, Table A.2, Graph A.1, and Graph A.2.

**Table 1: Data description**

| Sl. No. | Data | Source |
|---|---|---|
| **1** | Daily time-series data on total confirmed cases and total deaths for the states of California and Florida (Till 11 July 2020) | Github repository of the Centre for Systems Science and Engineering (CSSE), Johns Hopkins University, Maryland, USA [https://github.com/CSSEGISandData/COVID-19] |
| **2** | Weekly state-wise estimates of excess deaths associated with COVID-19 till 11 July 2020. | Website of CDC [https://www.cdc.gov/nchs/ nvss/vsrr/covid19/excess_deaths.html] |
| **3** | Data on rates of positivity of COVID-19 testing for the two states, California and Florida [As on 29 July 2020] | Website of Johns Hopkins University [https://coronavirus.jhu.edu/ testing/testing-positivity]. |

*Source: Deo and Grover (2021)*

### 3.3. Predictions under the assumption of extensive random testing (recommended intervention)

Once the posterior estimates of the transmission parameters are obtained from the state-space SI(Q/F)RD model, predictions of observed process under the assumption of extensive random testing are carried out using the steps outlined in section 2.2. Based on the posterior mean and standard deviation of the parameters of the state-space model, following specifications are used for conducting the required simulations to predict the transmission dynamics of the epidemic.

California:

$$R_1 \sim LogN(-0.822, 0.495), E(R_1) = 0.497, V(R_1) = 0.069;\ \beta_1 = R_1(\gamma + d_1)$$

$$R_2 \sim LogN(0.376, 0.106), E(R_2) = 1.464, V(R_2) = 0.024;\ \beta_2 = R_2(\gamma + d_2)$$

$$\gamma \sim LogN(-2.68, 0.087), E(\gamma) = 0.069, V(\gamma) = 0.00004$$

$$p_t \sim Beta(4.49, 59.61),\ E(p_t) = 0.07, V(p_t) = 0.001\ \forall t \leq 30$$

$$p_t \sim Beta(15.2, 3.8), E(p_t) = 0.8, V(p_t) = 0.008\ \forall t > 30$$

$$\lambda^I \sim Gamma(1012524.75, 1.88e-06), \lambda^R \sim Gamma(1633152.503, 1.46e-05),$$
$$\lambda^D \sim Gamma(1355.195, 0.00262)$$

Florida:

$$R_1 \sim LogN(-1.19, 0.573), E(R_1) = 0.359, V(R_1) = 0.05;\ \beta_1 = R_1(\gamma + d_1)$$

$$R_2 \sim LogN(0.476, 0.06), E(R_2) = 1.612, V(R_2) = 0.009;\ \beta_2 = R_2(\gamma + d_2)$$

$$\gamma \sim LogN(-2.77, 0.063), E(\gamma) = 0.063, V(\gamma) = 0.00002$$

$$p_t \sim Beta(1.39, 8.55),\ E(p_t) = 0.14, V(p_t) = 0.011\ \forall t \leq 30$$

$$p_t \sim Beta(41.87, 10.47), E(p_t) = 0.8, V(p_t) = 0.003 \; \forall t > 30$$

$$\lambda^I \sim Gamma(999169.436, 1.76e-06), \lambda^R \sim Gamma(1807366.511, 1.35e-05),$$
$$\lambda^D \sim Gamma(1022.341, 0.011)$$

The entire simulation exercise for this section is implemented in R programming through self-written codes. Plots of predicted values of daily number of active infected cases and cumulative deaths, along with their 95% confidence intervals, are shown in Figures 2 and 3, respectively. For a comparative assessment of the predictions of transmission trajectory of the epidemic under the two interventions, daily counts of active infecteds and cumulative number of deaths for both cases are plotted together in Figures 4 and 5.

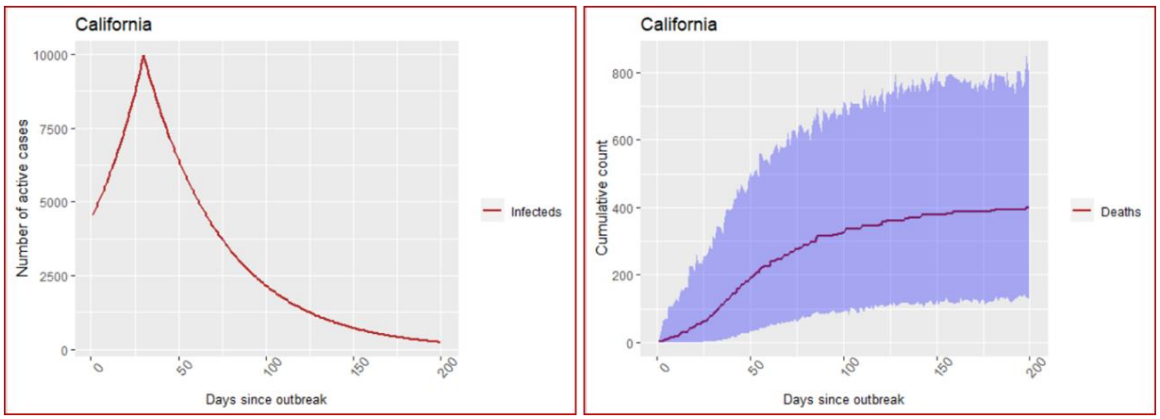



**Figure 2: California - Predictions under the assumption of extensive random testing. The blue shaded region depicts the region of 95% confidence intervals based on simulated values. The confidence region for number of infecteds is too narrow to be visible in the graph.**

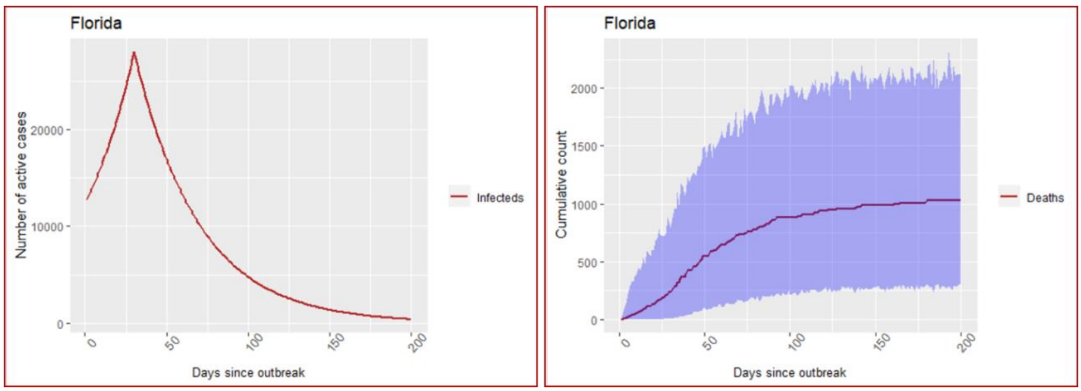



**Figure 3: Florida - Predictions under the assumption of extensive random testing. The blue shaded region depicts the region of 95% confidence intervals based on simulated values. The confidence region for number of infecteds is too narrow to be visible in the graph.**

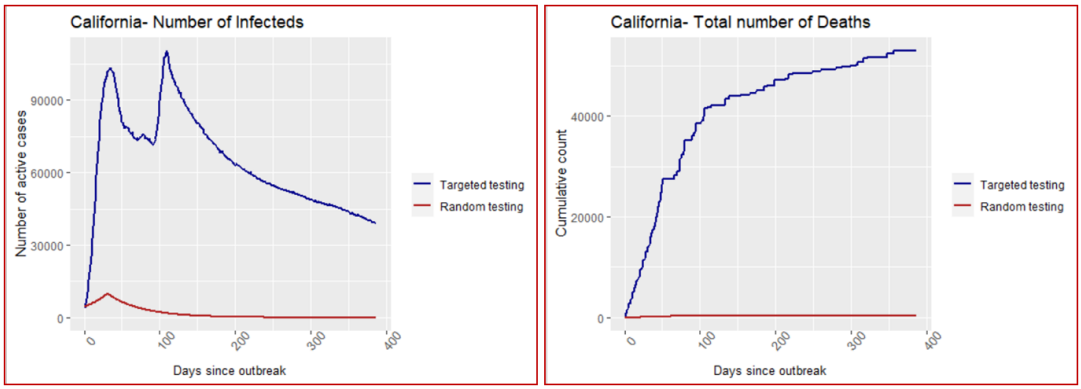



**Figure 4: California - Comparative graphs of predictions of cases under both interventions**

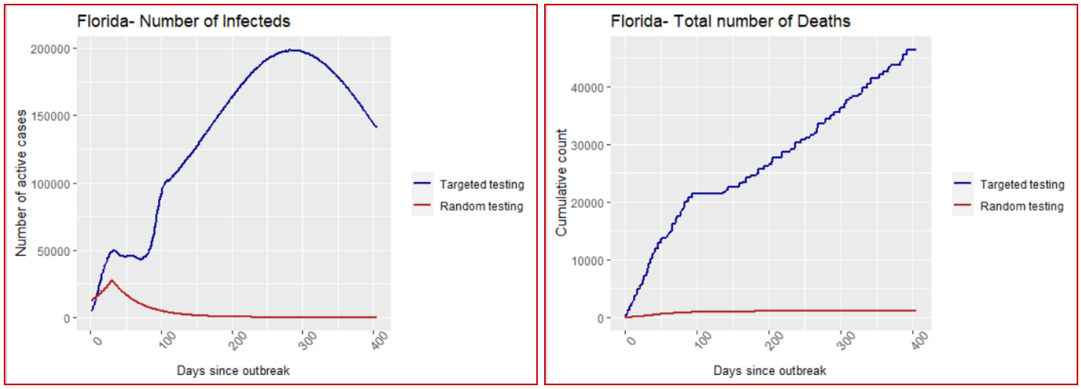



**Figure 5: Florida - Comparative graphs of predictions of cases under both interventions**

### 3.4. CEA of extensive random testing over targeted testing

To estimate the cost incremental, we first need the estimates of total number of tests to be conducted under both interventions. The rates of positivity of COVID-19 testing, as reported till 29 July 2020, were 7.47% in California and 18.96% in Florida. These percentages were taken as $r_1$ for estimating number of tests under the base intervention of targeted testing. The rates of positivity of tests under the assumption of extensive random testing, $r_2$ are obtained as $r_2 = \frac{C_2}{Total\ Population}$ for the two states and are provided in Table 2.

**Table 2: Estimates of rates of positivity of tests under extensive random testing**

| Intervention= Extensive random testing | California | Florida |
|---|---|---|
| $C_2$, total no. of infecteds by the end of the epidemic | 45,819 | 108,290 |
| Total no. of susceptibles at the start of the epidemic (Taken as total population of the state) | 39,512,223 | 21,477,737 |
| $r_2$ (considered as the probability that a person in the state got infected during the entire duration of the epidemic) | = 0.0012 (0.12%) | = 0.005 (0.5%) |

Values of cost incremental, changes in outcome measures (both in terms of number of infections and number of deaths), and ICERs are calculated by implementing these values of

$r_1$ and $r_2$ in the steps discussed in part e of section 2.2. Cost of COVID-19 test (RT-PCR) varies considerably across USA. However, leaving out some extreme cases, the average cost per unit of the RT-PCR test is around $100 in USA [Kliff (2020)]. We have used this average cost for evaluating cost incremental owing to increment in the number of tests. Results on the difference in number of tests, cost increment (or decrement), and changes in the outcomes of number of infections and number of deaths, on using the proposed intervention 'extensive random testing' over the base intervention 'targeted testing', are furnished in Table 3. Further, incremental cost effectiveness ratios associated with extensive random testing as compared to targeted testing are presented in Table 4.

**Table 3: Changes in outcomes and costs on using extensive random testing instead of targeted testing as the intervention to contain the spread of SARS-CoV-2 infections**

| State | Total Confirmed Cases | Total Detected Cases | Number of Tests | Number of Deaths |
|---|---|---|---|---|
| Predictions under the base case (Base intervention = Targeted testing) | | | | |
| California | 2384143 | 1328158 | 17779893 | 58292 |
| Florida | 4793903 | 1873747 | 9882632 | 58937 |
| Predictions under the ideal case (Recommended intervention = Extensive random testing) (Case A- Number of tests estimated using $r_2$ ) | | | | |
| California | 45819 | 41237 | 35560914 | 405 |
| Florida | 108290 | 97461 | 19329963 | 1039 |
| Predictions under the ideal case (Recommended intervention = Extensive random testing) (Case B-Assuming that everyone was tested by the end of the epidemic) | | | | |
| California | 45819 | 41237 | 39512223 | 405 |
| Florida | 108290 | 97461 | 21477737 | 1039 |
| Changes in Cost and Outcomes | | | | |
| State | Reduction in occurrence of infection | Reduction in occurrence of death | Tests incremental | Testing cost incremental ($) |
| Case A-Cost-effectiveness of extensive random testing with respect to targeted testing | | | | |
| California | 2338324 | 57887 | 17781021 | 1778102100 |
| Florida | 4685613 | 57898 | 9447331 | 944733100 |
| Case B-Cost-effectiveness of extensive random testing with respect to targeted testing | | | | |
| California | 2338324 | 57887 | 21732330 | 2173233000 |
| Florida | 4685613 | 57898 | 11595105 | 1159510500 |

**Table 4: Incremental cost effectiveness ratios associated with extensive random testing as compared to targeted testing**

| Incremental Cost Effectiveness Ratio- ICER | | |
|---|---|---|
| State | Number of extra tests per unit reduction in infection | |
| | Case A* | Case B** |
| California | 8 | 9 |
| Florida | 2 | 2 |
| | Number of extra tests per unit reduction in death | |
| | Case A | Case B |
| California | 307 | 375 |
| Florida | 163 | 200 |

| ICER (infections) | Additional testing cost per unit reduction in infection ($) | |
|---|---|---|
| | Case A | Case B |
| California | 760 | 929 |
| Florida | 202 | 247 |
| ICER (deaths) | Additional testing cost per unit reduction in death ($) | |
| | Case A | Case B |
| California | 30717 | 37543 |
| Florida | 16317 | 20027 |

*Case A- Number of tests estimated using $r_2$.
**Case B- Assuming that everyone was tested by the end of the epidemic.

## 4. Discussion

As per the posterior estimates obtained from the state-space SI(Q/F)RD model, around 43% infected cases in California and 61% infected cases in Florida, on an average, go unreported [Deo and Grover (2021)]. Further, the significantly higher posterior estimates of average reproduction number associated with undetected infecteds [California: 1.464 (sd: 0.155) and Florida: 1.612 (sd: 0.097)] as compared those for quarantined infecteds [California: 0.497 (sd: 0.262) and Florida: 0.359 (sd: 0.224)] stresses upon the necessity for conducting the CEA proposed in this study.

For both states, CEA of extensive random testing over targeted testing has yielded very strong results in favour of the former; refer Table 3 and Table 4. Citing uncertainties because of some unknown factors and leaving some space for errors in testing, even if we assume that 80% of the infecteds can be detected and quarantined using extensive random testing, a total of around 2.3 million people in California and 4.7 million people in Florida could be saved from the infection by the end of the epidemic if extensive random testing was used instead of targeted testing. Further, it is estimated that around 58 thousand deaths due to COVID-19 could be averted in each state if the states resorted to extensive random testing (after first month of outbreak) instead of targeted testing. These are huge expected gains for humanity, especially when every single life matter for us. The ICER values (in terms of number of tests) suggest that, on an average, only around 9 and 2 additional number of tests would be required in total to save one extra person from getting infected in California and Florida, respectively, by the time the epidemic ends. That is, around 760- 929 USD (California) and 202- 247 USD (Florida) additional expenditure on COVID-19 tests would be required to save every additional person from getting infected. Number of additional tests required to save one additional death from COVID-19 is estimated to be around 307- 375 for California and 163- 200 for Florida. That is, on using extensive random testing over targeted testing, one extra loss of life due to COVID-19 can be averted on an additional expenditure of around 30717- 37543 USD in California and around 16317- 20027 USD in Florida.

## 5. Conclusion

We have provided a comprehensive framework for conducting CEA of non-medical interventions for containing epidemics like COVID-19. To the best of our knowledge, there is no standard procedure available in the literature for conducting such analysis.

Results of the CEA conclude that extensive random testing, which has been strongly recommended by WHO, is significantly cost-effective over targeted testing. Since the $R_0$ values associated with quarantined infecteds in both states are estimated to be below 1, extensive

random testing, resulting in quarantining of at least 80% infecteds, is expected to result in the epidemic to end quite quickly as compared to the case of targeted testing. So, targeted testing may imply a smaller number of tests over a much longer period of time, while extensive testing means a very high number of tests for a much shorter period of time. This simple logic is corroborated by the ICER values obtained from the CEA of extensive random testing over targeted testing. For California, if the state is willing to conduct around 9 extra tests (or spend around 900 USD extra amount on testing) for saving one additional person from getting infected, or if the state is willing to conduct around 375 extra tests (or spend around 37500 USD extra amount on testing) for saving one additional person from dying due to COVID-19, extensive random testing can be considered as cost-effective over targeted testing. While for Florida, willingness to spend an extra amount of around 200 USD (2 extra tests) for saving one additional person from getting infected, or willingness to spend an extra amount of around 20,000 USD (200 extra tests) for saving one additional person from dying due to COVID-19, renders extensive random testing as cost-effective over targeted testing.

**Acknowledgements**

The authors are grateful to the editor and anonymous reviewers for their invaluable suggestions towards improving the paper.

**References**

Deo, V. and Grover, G. (2021). A new extension of state-space SIR model to account for Underreporting – An application to the COVID-19 transmission in California and Florida. *Results in Physics*, **24**, 104182. https://doi.org/10.1016/j.rinp.2021.104182

Kliff, S. (2020, June 16). Most Coronavirus Tests Cost About $100. Why Did One Cost $2,315? *The New York Times*. Retrieved from https://www.nytimes.com/2020/06/16/upshot/coronavirus-test-cost-varies-widely.html

Lau, H., Khosrawipour, T., Kocbach, P., Ichii, H., Bania, J., and Khosrawipour, V. (2020). Evaluating the Massive Underreporting and Undertesting of COVID-19 Cases in Multiple Global Epicenters. *Pulmonology*, **27(2)**, 110-115. https://doi.org/10.1016/j.pulmoe.2020.05.015.

Padhye, N. S. (2020). Reconstructed diagnostic sensitivity and specificity of the RT-PCR test for COVID-19. *MedRxiv (Preprint)*. doi:https://doi.org/10.1101/2020.04.24.20078949

Tahamtan, A., and Ardebili, A. (2020). Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Review of Molecular Diagnostics*, **20(5)**, 453-454. doi:https://doi.org/10.1080/14737159.2020.1757437

West, C. P., Montori, V. M., and Sampathkumar, P. (2020). COVID-19 Testing: The Threat of False-Negative Results. *Mayo Clinic Proceedings*, **95(6)**, 1127-1129. doi:https://doi.org/10.1016/j.mayocp.2020.04.004

World Health Organisation. (2020 a). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020*. Retrieved from https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020

World Health Organisation. (2020 b). *Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19).* Retrieved from https://www.who.int/publications-detail/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19)

Wu, S. L., Mertens, A. N., Crider, Y. S., Nguyen, A., Pokpongkiat, N. N., Djajadi, S., *et al.* (2020). Substantial Underestimation of SARS-CoV-2 Infection in the United States. *Nature Communications*, **11**, 4507. https://doi.org/10.1038/s41467-020-18272-4

## Appendix A

**Table A.1: Posterior estimates of time-invariant parameters of the state-space SI(Q/F) RD model, along with their standard deviations and 95% credible intervals- California**

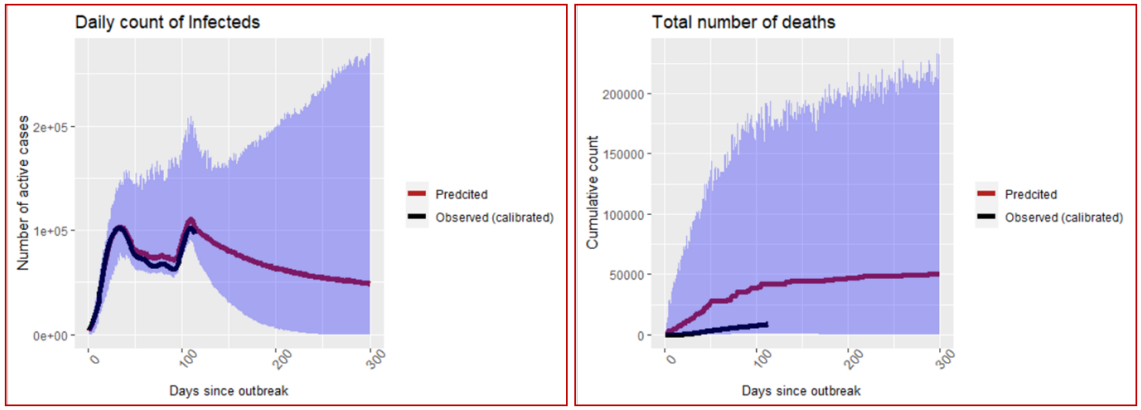
| Parameter | Posterior mean | Posterior standard deviation | 95% credible interval |
|---|---|---|---|
| $R_1$ | 0.497 | 0.262 | [0.068, 1.004] |
| $R_2$ | 1.464 | 0.155 | [1.214, 1.813] |
| $\gamma$ | 0.069 | 0.006 | [0.056, 0.081] |
| $\kappa$ | 336063.593 | 47259.956 | [243264.879, 431918.329] |
| $\lambda^D$ | 1355.195 | 718.277 | [397.588, 2632.883] |
| $\lambda^I$ | 1012524.750 | 734717.729 | [1349.955, 2006982.462] |
| $\lambda^R$ | 1633152.503 | 334437.988 | [1073803.103, 2360304.964] |
| $\hat{\beta}_1 = \hat{R}_1(\hat{\gamma} + d_1)$ | 0.035 | | |
| $\hat{\beta}_2 = \hat{R}_2(\hat{\gamma} + d_2)$ | 0.102 | | |

*Source: Deo and Grover (2021)*

**Table A.2: Posterior estimates of time-invariant parameters of the state-space SI(Q/F) RD model, along with their standard deviations and 95% credible intervals- Florida**

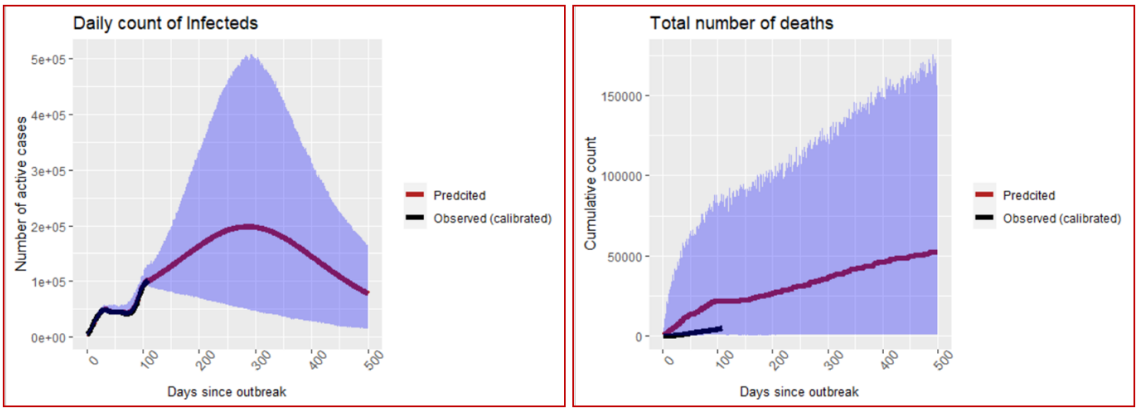
| Parameter | Posterior mean | Posterior standard deviation | 95% credible interval |
|---|---|---|---|
| $R_1$ | 0.359 | 0.224 | [0.052, 0.880] |
| $R_2$ | 1.612 | 0.097 | [1.416, 1.799] |
| $\gamma$ | 0.063 | 0.004 | [0.054, 0.071] |
| $\kappa$ | 500800.490 | 94547.445 | [327261.995, 679843.447] |
| $\lambda^D$ | 1022.341 | 303.916 | [539.044, 1629.331] |
| $\lambda^I$ | 999169.436 | 753473.727 | [4778.595, 2403835.884] |
| $\lambda^R$ | 1807366.511 | 365988.299 | [1164580.155, 2616920.665] |
| $\hat{\beta}_1 = \hat{R}_1(\hat{\gamma} + d_1)$ | 0.0229 | | |
| $\hat{\beta}_2 = \hat{R}_2(\hat{\gamma} + d_2)$ | 0.102 | | |

*Source: Deo and Grover (2021)*




*Source: Deo and Grover (2021)*

**Figure A.1: Predictions of number of infected and number of deaths in California under the base case/ intervention of targeted testing. The blue shaded ribbon is the region of 95% credible intervals.**




*Source: Deo and Grover (2021)*

**Figure A.2: Predictions of number of infected and number of deaths in Florida under the base case / intervention of targeted testing. The blue shaded ribbon is the region of 95% credible intervals.**