COVID-19 DECISION-ANALYSIS TOOL
Interactive Simulation Tool For Evaluating Combination Interventions for Phased Reopening: Application to State of NY
Working Paper

Sonza Singh, Hanisha Tatapudi, Xinmeng Zhao, Shifali Bansal, Seyedeh Nazanin Khatami, and Chaitra Gopalappa
Disease Modeling Lab, University of Massachusetts Amherst
Corresponding author: Chaitra Gopalappa, chaitrag@umass.edu

Contents

1. Overview: How to use online tool

Link to tool . This is an interactive simulation tool that can be used to evaluate alternative response decisions for the COVID-19 epidemic. Users can enter decisions for three interventions over a 52-week period from the beginning of May

Input decision choices for the following three interventions
1. **Intervention 1**: Proportion reduction in contacts per person through social distancing measures (Enter values between 0 to 1)
   a. **Interpretation**: A value of 0 will use a upper bound for number of contacts per person (replication a normal pre-COVID scenario), a value of 1 will use a lower bound for number of contacts per person (replicating maximum stay-at-home order scenario), and values between 0 and 1 will select from within an upper and lower bound (lower and upper bounds for contacts per persons were estimated using methods in Section 5 and 6)
   b. This can be used for evaluating variants of social distancing measures such as modified work schedules,
      Example 1: To evaluate a rotating schedule that ensures that at any time only 50% of personnel work on site, decision entered should be: 1, 52, 0.5
      Example 2: To evaluate a rotating schedule that ensures that at any time only 50% of personnel work on site for the first 2 weeks, and it increases to 75% (i.e., 25% reduction in contacts) for weeks 3 to 52, enter: 1, 2, 0.5, 3, 52, 0.25

2. **Intervention 2:** Testing capacity -Number of tests administered per day through universal testing (Enter positive numbers)
   a. **Interpretation:** We assume testing is available to the full population and thus there is an equal chance for a positive and a negative case to get tested (see Sections 6 below for methodology). We assume persons testing positive would be immediately quarantined and cannot infect others, so the value entered for this intervention should be the number of tests successfully completed.

3. **Intervention 3:** Testing capacity -Number of tests administered per day through contact tracing and testing (Enter positive numbers)
   a. **Interpretation:** We assume that a certain proportion of the persons tested (equal to secondary attack rate) would be positive (see Sections 6 below for methodology). We assume persons testing positive would be immediately quarantined and cannot infect others, so the value entered for this intervention should be the number of tests successfully completed.

Timeline: Flexibility to enter decisions for up to weekly time-units

2. Simulation methodology

We developed a compartmental model for simulating epidemic projections overtime. The epidemic flow diagram is depicted in Figure 1. Each box is an epidemic state, and each arrow represents a transition from one state to another. Note, each compartment is further split by age and gender, but for clarity of notations, we do not include it in the equations below.
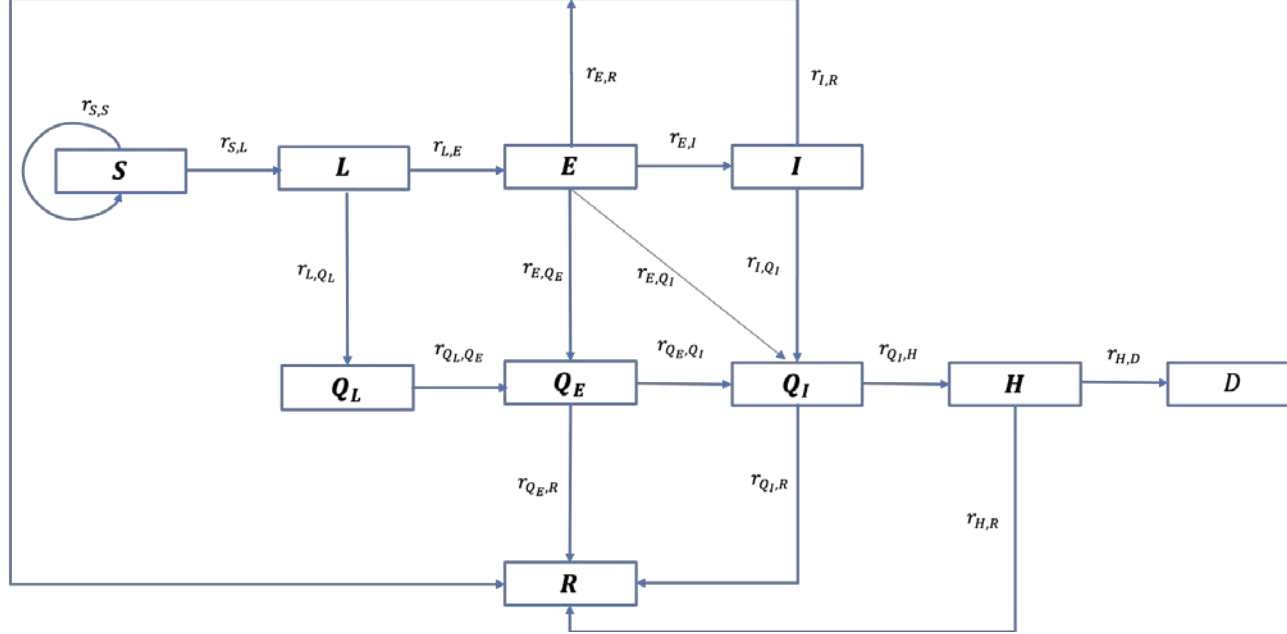

Figure 1: Extended SEIR compartmental model

Let $\pi_{t+1} = [S, L, E, I, Q_L, Q_E, Q_I, H, R, D]$ be a vector, with each element representing the number of people in a compartment at time $t + 1$, where,

S = the number **Susceptible**,
L = the number in **Latent stage** (not infectious and asymptomatic),
E = the number in **Exposed** (asymptomatic or pre-symptomatic and infectious),
I = the number **Infected** (symptomatic and infectious),
Q_L = the number in **Latent and Quarantined** (tested and diagnosed),
Q_E = the number in **Exposed and Quarantined** (tested and diagnosed),
Q_I = the number **Infected and Quarantined** (tested and diagnosed),
H = the number **Hospitalized**,

R = the number **Recovered**, and

D = the number **Deaths.**

.

Let,

N = total population who are alive,

a_B = symptom-based testing rate,

$a_{SD,t}$ = proportion contact reduction compared to pre-COVID scenario at time t,

$a_{C,t}$ = rate of testing through contact tracing at time t,

$a_{U,t}$ = rate of testing through universal testing at time t,

β_{max} = transmission rate under no social distancing,

β_{min} = transmission rate under maximum social distancing,

$days_L$ = duration in latent period,

$days_{incub}$ = duration in incubation period,

$days_{IR}$ = time from onset of symptoms to recovery,

$days_{Q_IR}$ = time from diagnosis to recovery,

$days_{Q_IH}$ = time from diagnosis to hospitalization,

$days_{HR}$ = time from hospitalization to recovery,

$days_{HD}$ = time from hospitalization to death,

$prop_{asymp}$ = proportion asymptomatic,

$prop_{hosp}$ = proportion hospitalized, and

$prop_{severe}$ = proportion hospitalized

Then, we can write the equations for transition rates (arrows in Figure 1) as follows:

$$r_{S,L} = \frac{[\beta_{min} + (1 - a_{SD,t})(\beta_{max} - \beta_{min})](E + I)}{N}$$, which assumes that only infected persons in E and I can transmit, that persons in Q_E and Q_I self-quarantine, and persons in L and Q_L are not infectious.

$$r_{L,E} = \frac{1}{days_L}$$

$r_{E,Q_I} = \dfrac{(prop_{severe})(1-prop_{asymp})}{days_{incub}-days_L}$, which assumes that a certain proportion of cases that are severe $(prop_{severe})$ get diagnosed immediately because of symptoms; the denominator is based on the assumption that the duration of exposed phase is equal to the duration of the incubation period minus the duration of the latent period.

$r_{E,I} = \dfrac{(1-prop_{severe})(1-prop_{asymp})}{days_{incub}-days_L}$, which follows from above.

$r_{I,Q_I} = a_B$, which assumes that, under symptom-based testing, only persons who show moderate to severe symptoms get diagnosed and those who show mild symptoms do not.

$r_{Q_I,H} = \dfrac{prop_{hosp}}{days_{Q_I H}}$, for $prop_{hosp}$ it uses the proportion of persons hospitalized among those diagnosed through symptom-based testing.

$r_{L,Q_L} = a_{U,t} + (1 - a_{U,t})a_{C,t}$, which assumes that when both universal testing and contact tracing and testing are administered, persons diagnosed through universal testing will not be again tested through contact tracing.

$r_{E,Q_E} = a_{U,t} + (1 - a_{U,t})a_{C,t}$, which is similar to above.

$r_{E,R} = \dfrac{prop_{asymp}}{days_{incub}-days_L}$, which assumes that a certain proportion of persons $(prop_{asymp})$ never show symptoms and thus directly go from exposed to recovered.

$r_{I,R} = a_{U,t} + (1 - a_{U,t})a_{C,t} + \dfrac{1}{days_{IR}}$, which assumes that mild cases that did not get diagnosed through symptoms-based testing have a chance of getting tested through additional testing options, and upon diagnoses, self-quarantine.

$r_{Q_L,Q_E} = \dfrac{1}{days_L}$

$$r_{Q_E,Q_I} = \frac{[a_B(1 - prop_{severe}) + (prop_{severe})](1 - prop_{asymp})}{days_{incub} - days_L}$$, theoretically, r_{Q_E,Q_I} should be the same as $r_{E,I}$, however, as the rate of transitioning from Q_I to H is fixed to proportion hospitalized under symptom-based testing, if extensive testing is conducted, the number of persons in Q_I would increase thus incorrectly inflating the number of persons who are hospitalized; therefore, to avoid this, we modified the equation to consider that the number of persons flowing into Q_I would be equal to the proportion flowing from I to Q_I under symptom-based testing.

$$r_{Q_E,R} = \frac{(1 - [a_B(1 - prop_{severe}) + (prop_{hosp})])(1 - prop_{asymp})}{days_{incub} - days_L}$$, which follows from the above equation.

$$r_{Q_I,R} = \frac{1 - prop_{hosp}}{days_{Q_IR}}$$

$$r_{H,R} = \frac{prop_{Recover}}{days_{HR}}$$

$$r_{H,D} = \frac{(1 - prop_{Recover})}{days_{HD}}$$

We simulate the epidemic over time using the following system of differential equations

- $\pi_{t+1} = \pi_t + \pi_t Q_t \Delta t$

where,

- Q_t = a matrix of transition rates between states (arrows in Figure 1), and
- Δt = time-step

We use a time-unit of per day for the transition rates in Q_t and set $\Delta t = \frac{1}{10}$, and thus, the model simulates every 10th of a day.

3. Input data assumptions and sources for simulation model

We use data estimates from other studies in the literature. The description of the data, its sources, and values (with ranges and medians where applicable) for all parameters described in Section 2 are found on GitHub.

4. Calibration of unknown data inputs

As population distribution and density varies by location (we modeled state-level for the U.S.), it is expected that the number of contacts per person also varies, and thus the transmission rate β, as $\beta = pc$, where p is the probability of transmission per susceptible-infected contact, and c = number of contacts per person. We estimated β by calibrating to the number of diagnosed cases in the State. (Data and sources on GitHub)

a. **Transmission rate under absence of social distancing (Upper bound for transmission rate)** (β_{max})
   As implementation of social distancing measures influences c, we estimated β by calibrating to the diagnosed cases reported prior to implementation of any social distancing measures in the State modeled. We call this value of transmission rate β_{max} and assume this is the upper bound of transmission rate for that State.

b. **Transmission rate under maximum social distancing (lower bound for transmission rate)** (β_{min})
   As social distancing reduces the number of contacts per person, transmission rate would be lower. Therefore, we calibrate a second value as the transmission rate under social distancing by fitting the model to the diagnosed cases at the time maximum social distancing was implemented. We call this β_{min} and assume this is the lower bound of transmission rate for that State.

5. Methodology for modeling interventions

We setup the model to take the following parameters from the user, we discuss how these inputs would be used in the simulation. **Note:** all three decision choices a_{SD}, T_C, T_U have a subscript t for time to indicate that decisions can change over time For clarity of notations we exclude t in the equations.

a. **Proportion reduction in contacts (a_{SD}) (takes values between 0 to 1)**:
   From previous section,
   $c_{max} = \beta_{max}/p$ = the upper bound for the number of contacts per person (equal to contacts if normal operations similar to a pre-COVID situation were to resume), and

$c_{min} = \beta_{min}/p$ = the lower bound for the number of contacts per person (equal to the lowest that was achieved during maximum social distancing implementation in the State),

Thus,

if $a_{SD} = 0$ it will use c_{max} i.e., there will be no reduction in contacts

if $a_{SD} = 1$ it will use c_{min}, i.e., the lowest that can be achieved for the population modeled, and

for a_{SD} between 0 and 1, it would take a value between c_{max} and c_{min} through interpolation.

Proportion reduction in contacts (a_{SD}) would influence the transmission rate, and thus, we calculate the corresponding transmission rate β_{SD} as

$$(1 - a_{SD}) = \frac{\left(\frac{\beta_{SD}}{p} - \frac{\beta_{min}}{p}\right)}{\frac{\beta_{max}}{p} - \frac{\beta_{min}}{p}} = \frac{(\beta_{SD} - \beta_{min})}{\beta_{max} - \beta_{min}}$$

And the corresponding transmission rate (β_{SD}) can be written as

$$\beta_{SD} = \beta_{min} + (1 - a_{SD})(\beta_{max} - \beta_{min})$$

b. **Testing capacity for universal testing as number of tests per day (T_U) and unit cost of testing (u_U)**

The value to input for the number of tests per day should consider resource availabilities, such as trained personnel, equipment, and infrastructure. We assume that all persons would be provided the test on a first-come-first basis.

We assume that infected persons once diagnosed would be quarantined and thus would not transmit the infection to others

Let

    T_U = the number of persons who can be tested on any given day,

    u_U = unit cost of testing per person through universal testing, and

    a_U = the proportion of undiagnosed infected cases that get detected through universal testing

Then, for every time-unit (daily), we estimate

    the proportion of undiagnosed infected cases that get detected as $= a_U = \frac{T_U}{S+E+L+I}$

    the number of infected cases that get newly diagnosed as $= a_U(L + E + I)$, and

    the cost of universal testing as $= u_U T_U$.

c. **Testing capacity for contact tracing and testing as number of tests per day (T_C) and unit cost of testing (u_c)(USD)**

The value to input for the number of tests per day should consider resource availabilities, such as trained personnel, equipment, and infrastructure. We assume that infected persons once diagnosed would be quarantined and thus would not transmit the infection to others

Let

T_C = the number of tests per day, and

δ = secondary attack rate,

u_c= unit cost of tracing and testing per person

a_C = the proportion of undiagnosed infected cases that get detected

Then, for every time-unit (day), we estimate,

the maximum number of positive cases that can be detected as = T_C δ,

the proportion of undiagnosed infected cases that get detected as = $a_C = \min\left\{1, \frac{T_C\delta}{(1-a_U)(L+E+I)}\right\}$,

where the denominator of the second term in the bracket is persons who did not get tested through universal testing,

the number of infected cases that are newly diagnosed through contact tracing and testing as= $a_C(1 - a_U)(L + E + I)$, and

the cost of testing as = $u_c a_C(1 - a_U)(L + E + I)/\delta$.

6. **Output measures from simulation**
   a. Number of new cases diagnosed per day by type of testing (tracked in the simulation- see methodology)
   b. Cost of testing by type of test
   c. Number of new deaths (tracked in the simulation- see methodology)

   Economic impact* (See data and sources on GitHub)
   d. Value of statistical life years (VSL) loss
      i. Calculated as age-specific VSL times the number of new deaths at that time point
   e. Unemployment rate
      ii. We make some assumptions for employment rate (see below).

f. Wage loss
    iii. Calculated as unemployment rate due to COVID-19 times labor work force under normal situations times the number of people in the population times the average daily wage of jobs lost due to COVID
    iv. We assume 40 hours per week per person

Interpretation*
Economic costs help understand the impact of decisions. VSL is the economic impact from deaths due to COVID-19 if appropriate interventions are not implemented. Wage loss is the economic impact from implementing social distancing. These numbers can be used for informing decision-making, e.g., by reducing deaths, the costs that would be saved (VSL) could be used for covering lost wage through relief packages.

Unemployment rate assumptions:
We use a simple estimation for changes in unemployment as a function of social distancing.

$$K_t = \max(a_{SD,t}K, y_t) \, ; \, A_t = \max\left(A, \min(a_{SD,t}K, y_t)\right)$$

$$y_{t+1} = \begin{cases} y_t - \mu_- \frac{(K_t - A_t)}{K-A} & \text{if } y_t = K_t \\ y_t + \mu_+ \frac{(K_t - A_t)}{K-A} & \text{otherwise} \end{cases}$$

$$\mu_+ \sim \frac{K-A}{(d_K - d_A)}$$
$$\mu_- \sim 0.5 \frac{K-A}{(d_K - d_A)}$$

$$y_{t+1} = \begin{cases} y_t - \frac{K-A}{(d_K - d_A)} \frac{(K_t - A_t)}{K-A} & \text{if } y_t = K_t \\ y_t + 0.5 \frac{K-A}{(d_K - d_A)} \frac{(K_t - A_t)}{K-A} & \text{otherwise} \end{cases}$$

$$y_{t+1} = \begin{cases} y_t - \frac{1}{(d_K - d_A)} \frac{(K_t - A_t)}{1} & \text{if } y_t = K_t \\ y_t + 0.5 \frac{1}{(d_K - d_A)} \frac{(K_t - A_t)}{1} & \text{otherwise} \end{cases}$$

Where,

- y_t is the unemployment rate at time t (day) used as input in the simulation,
- A is the minimum unemployment rate in the population simulated (pre-COVID value),
- K is the maximum unemployment rate in the population simulated,
- $a_{sd,t}$ is the social distancing intervention at time t,
- $d_K - d_A$ is the duration (in days) that it took for unemployment rate to go from A to K,
- μ_+ = rate of increase in the unemployment, and
- μ_- = rate of decrease in unemployment

Disclaimer: This equation is only meant as an approximation for estimation of Wage loss and should NOT be used as a projection. The equation is set to account for varying levels of social distancing measures.
We assume that rate of decline in employment would be slower than rate of increase by half.

7. Calibration of simulation model to State of New York

We simulated the model to the State of New York. We compare the model estimated number of hospitalization and deaths with that reported and verified that the predictions match well, as seen in Figures 2, 3, and 4. In this base-case figures presented below, we assumed that social distancing would be retained for the remaining duration of the simulation, and thus represents the best case scenario.
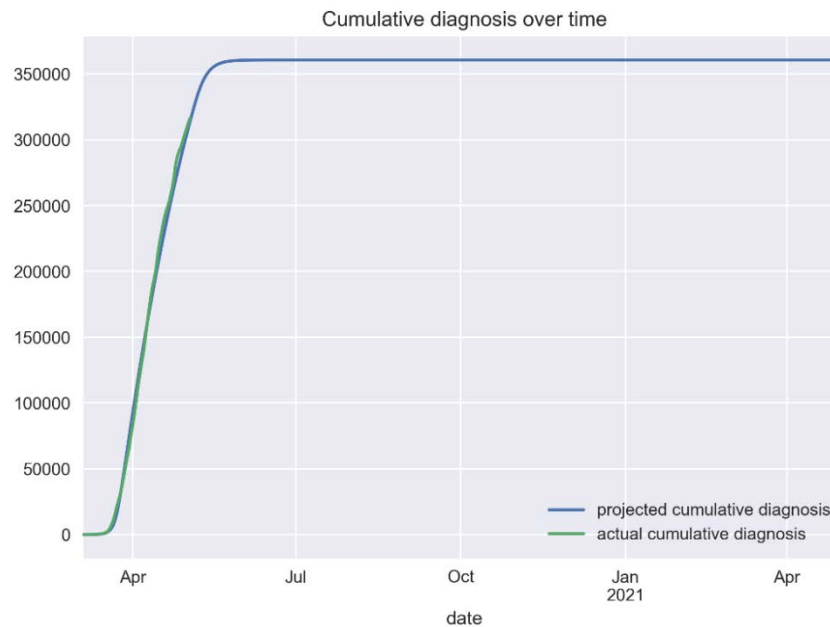

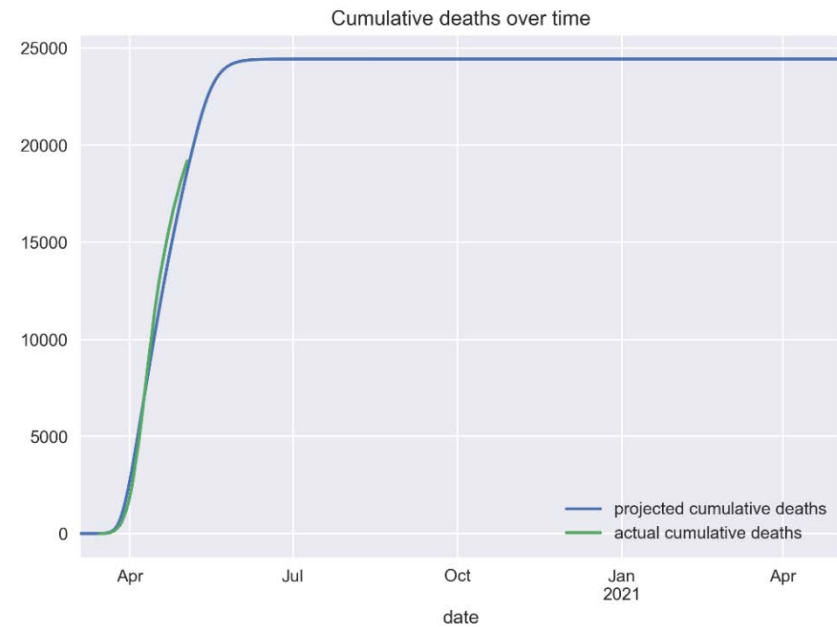Figure 2 Cumulative number of diagnosis
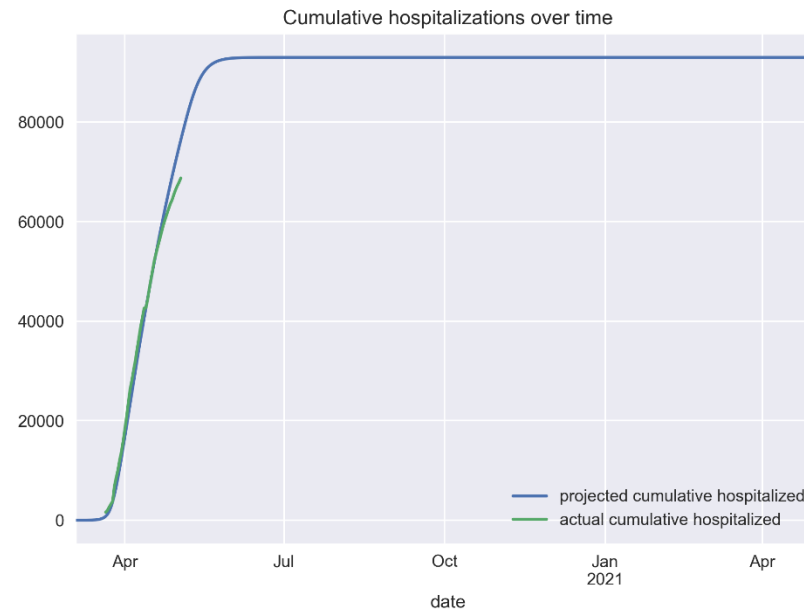
Figure 3 Cumulative number of deaths

Cumulative hospitalizations over time

Figure 4 Cumulative number of hospitalizations

8.   Model and references
- The model can be found on GitHub COVID-19-decision-analysis-model.
- All data inputs used in the model can be found on GitHub in Data folder.
- Summary of sources for data inputs to model can be found on GitHub SummaryofSource.xlsx .