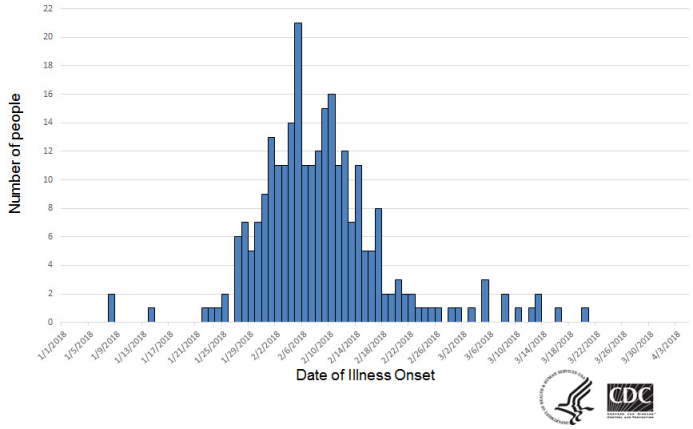
Simple Mathematical Models

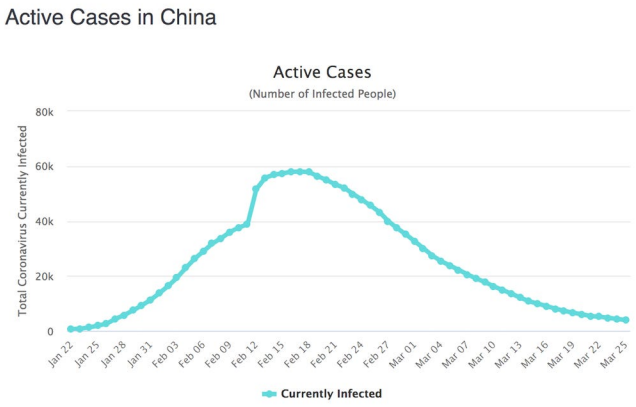
Clayton Mizgerd
Max Enis
Hemal Patel
Alexandra Bonat

November 2020

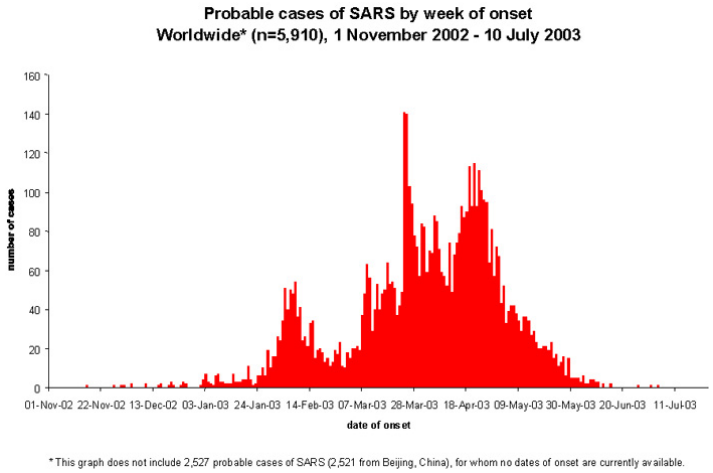
Do outbreak curves have a characteristic shape?



Salmonella outbreak



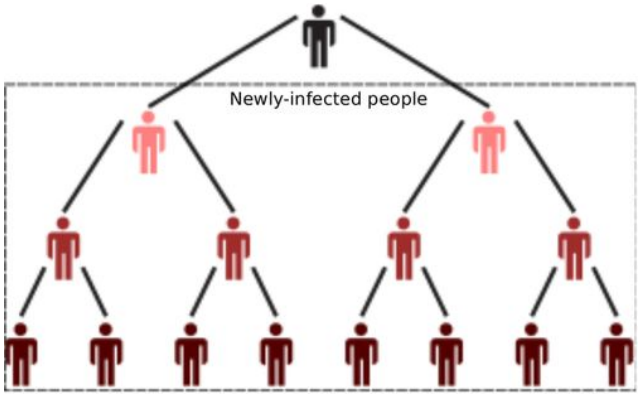
COVID-19 outbreak China

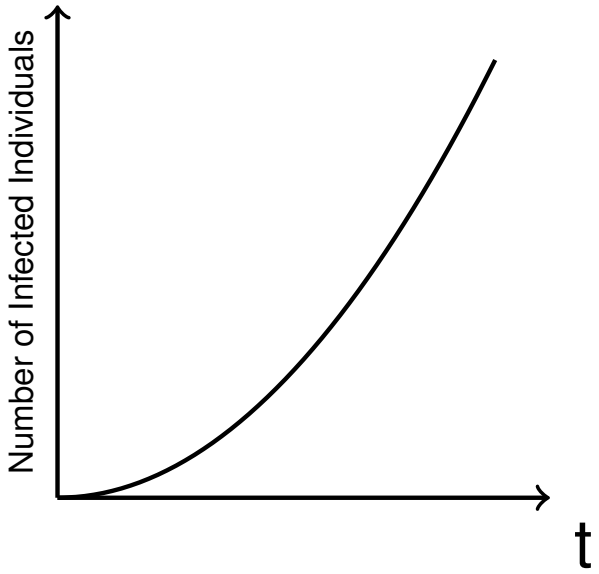


SARS outbreak

A simple model

Consider an infectious disease outbreak where each infected individual infects two other individuals





What's wrong with this model?

Compartmental Modelling

The disease will eventually run out of people to infect.

We can split population into "compartments":

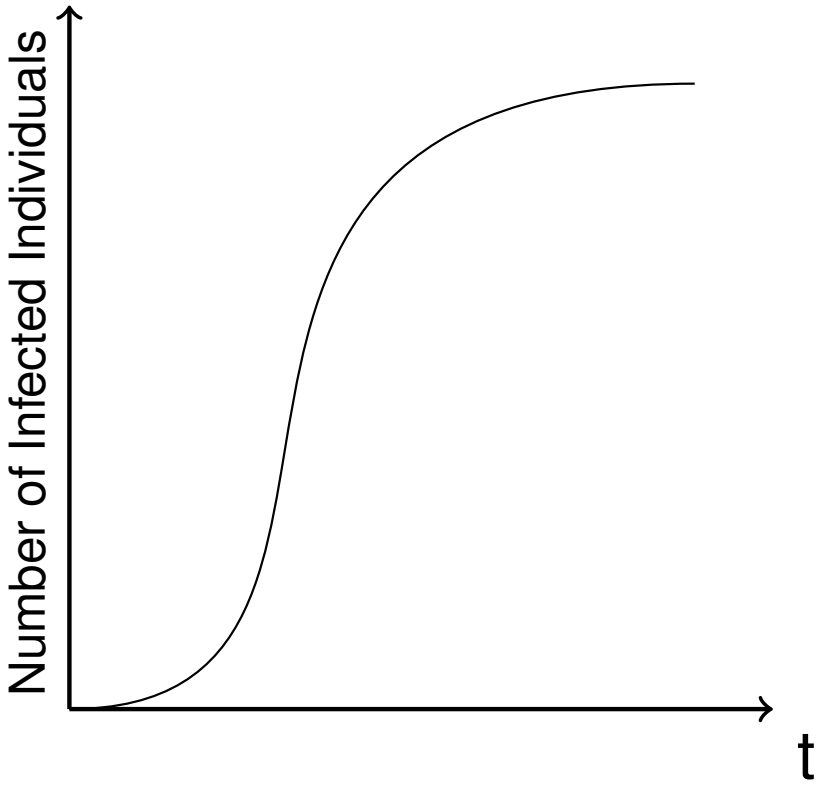
- Susceptible, S
- Infected, I

$$S \xrightarrow{\beta \mathrm{SI}} I$$

Observe the following

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI$$

Compartmental Modelling



We no longer infect infinitely, however all of the population remains infected.

Thus, we can introduce a new compartment, R, to indicate the recovered population.

Refined Model

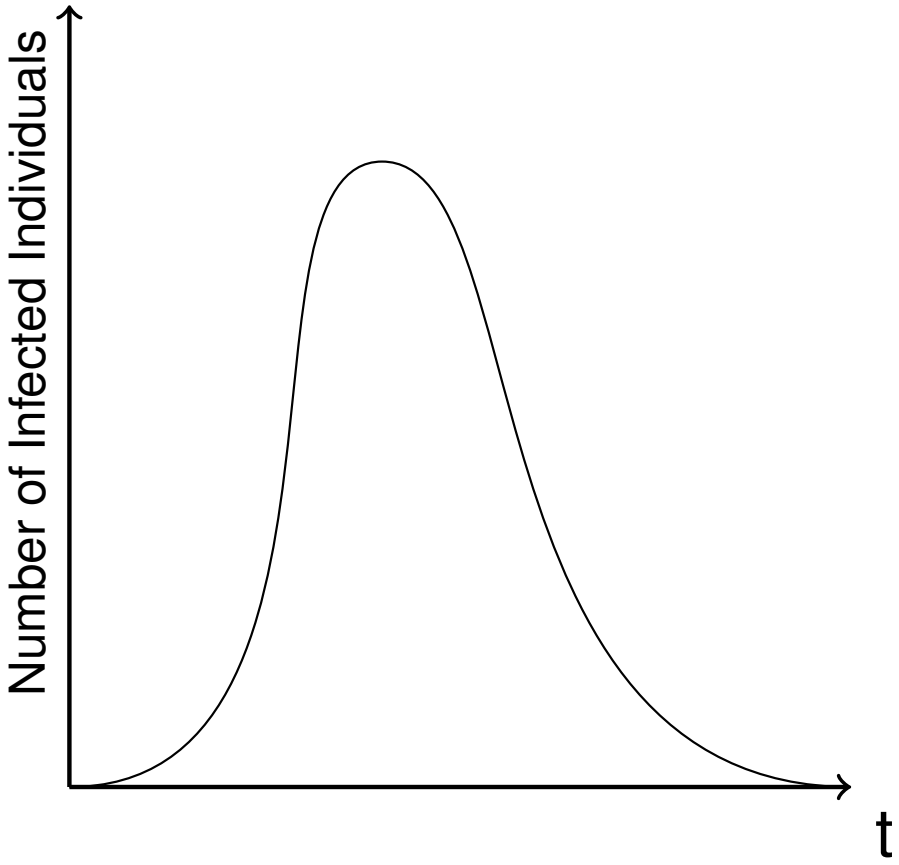
Our new model now looks like this:

$$S \xrightarrow{\beta \mathrm{SI}} I \xrightarrow{\mu \mathrm{I}} R$$

We now get

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \mu I \\ \frac{dR}{dt} &= \mu I \end{aligned}$$

SIR Model



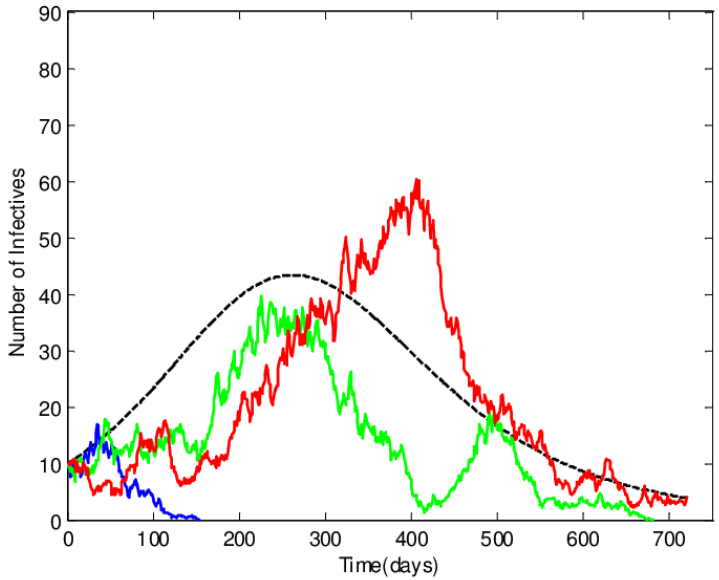
We now have a simple model which only implements two parameters β and μ.

$$S \xrightarrow{\beta \mathrm{SI}} I \xrightarrow{\mu \mathrm{I}} R$$

A Basic Extension

Infectious disease outbreaks are not smooth curves.

We can use stochastic modelling to capture the random nature of outbreaks.



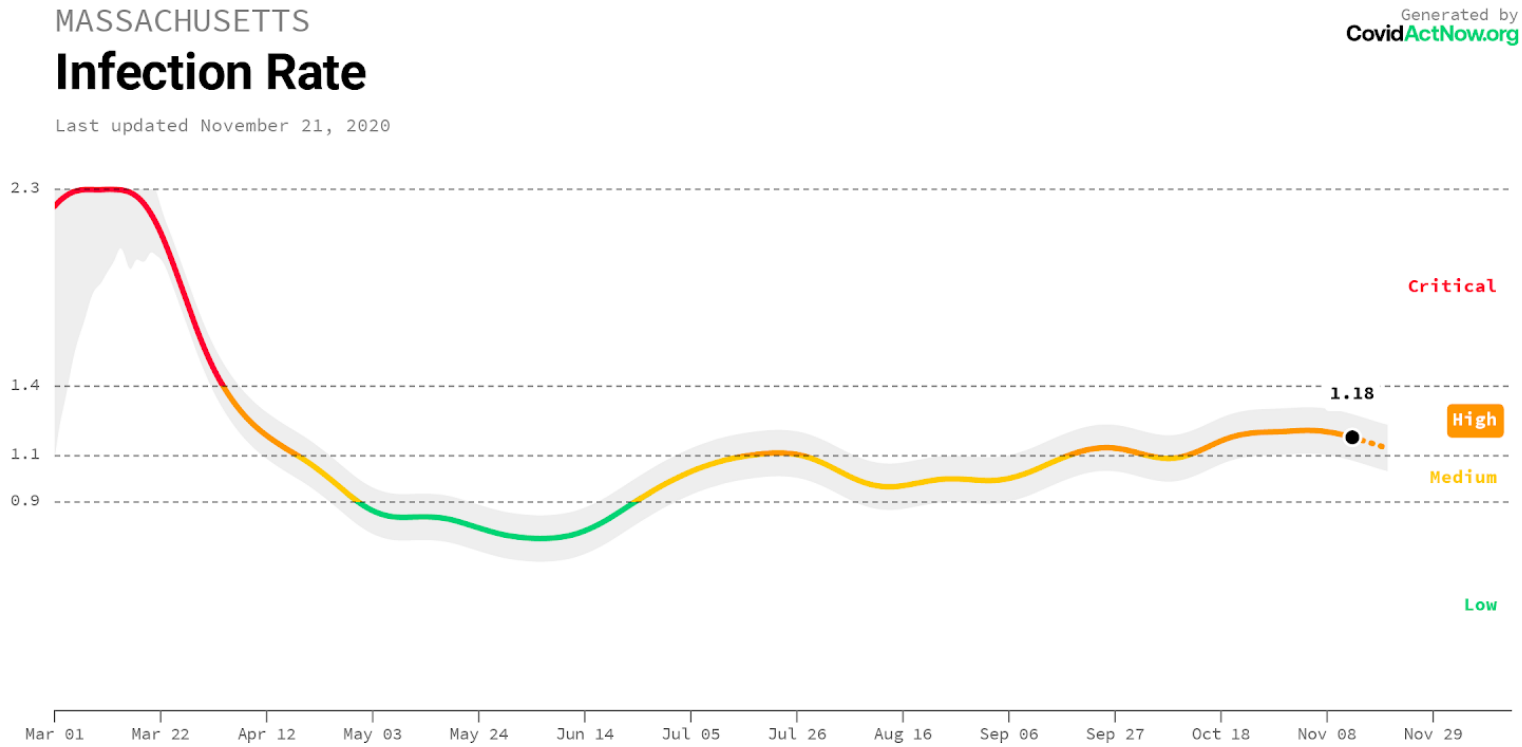
Model Parameters

- Effective Reproductive Number (R_t)
- Fatality Rate
- Length of Infection

Reproductive Number

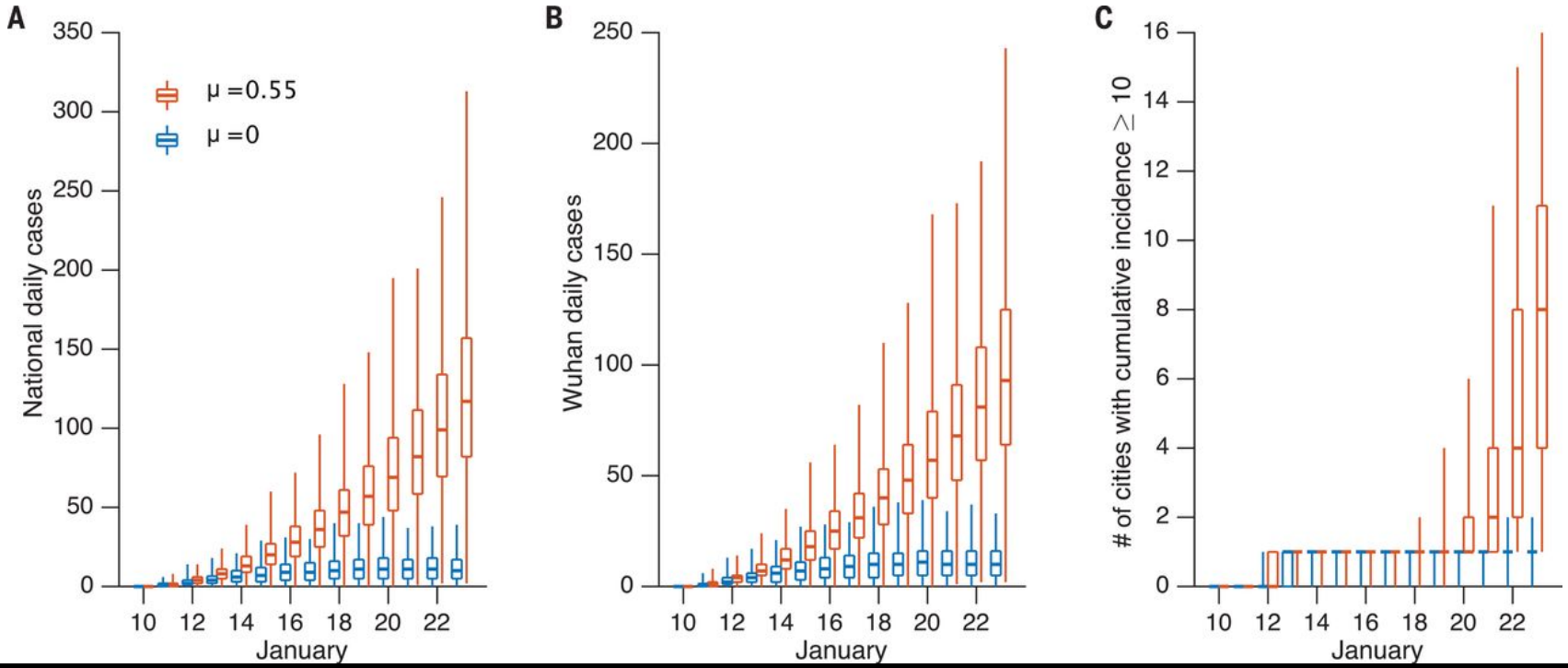
- R_0: Basic Reproductive Number
  - COVID-19: Likely around 2.5, estimates range from 1.4-4.0 (Majumder and Mandl, 2020)
  - Depends on location
- R_t: Effective Reproductive Number
  - Infection rate with public health measures initiated
  - Influenced by many factors

Massachusetts Infection Rate



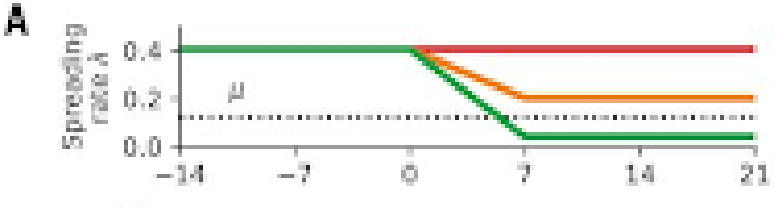
Factors Influencing R_t: Undocumented Cases

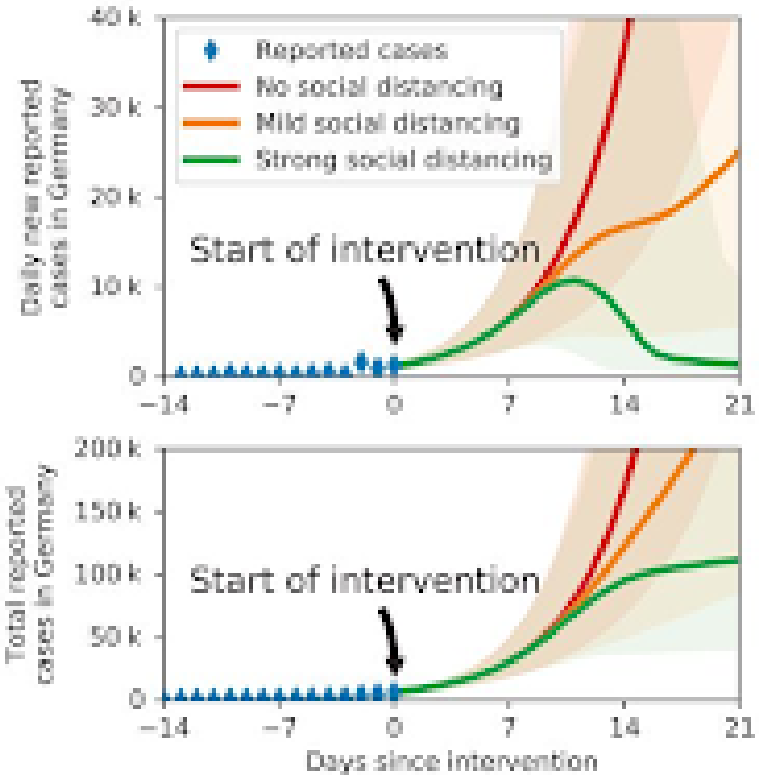
- 40% of cases are asymptomatic (CDC)
- Additional mild cases
- Generally less infectious than confirmed cases (75% for asymptomatic)
- Ruiyun et al., May 2020



Factors Influencing R_t: Non-Pharmaceutical Interventions

- Social distancing
  - Reducing each person's number of contacts
  - Dehning et al., July 2020

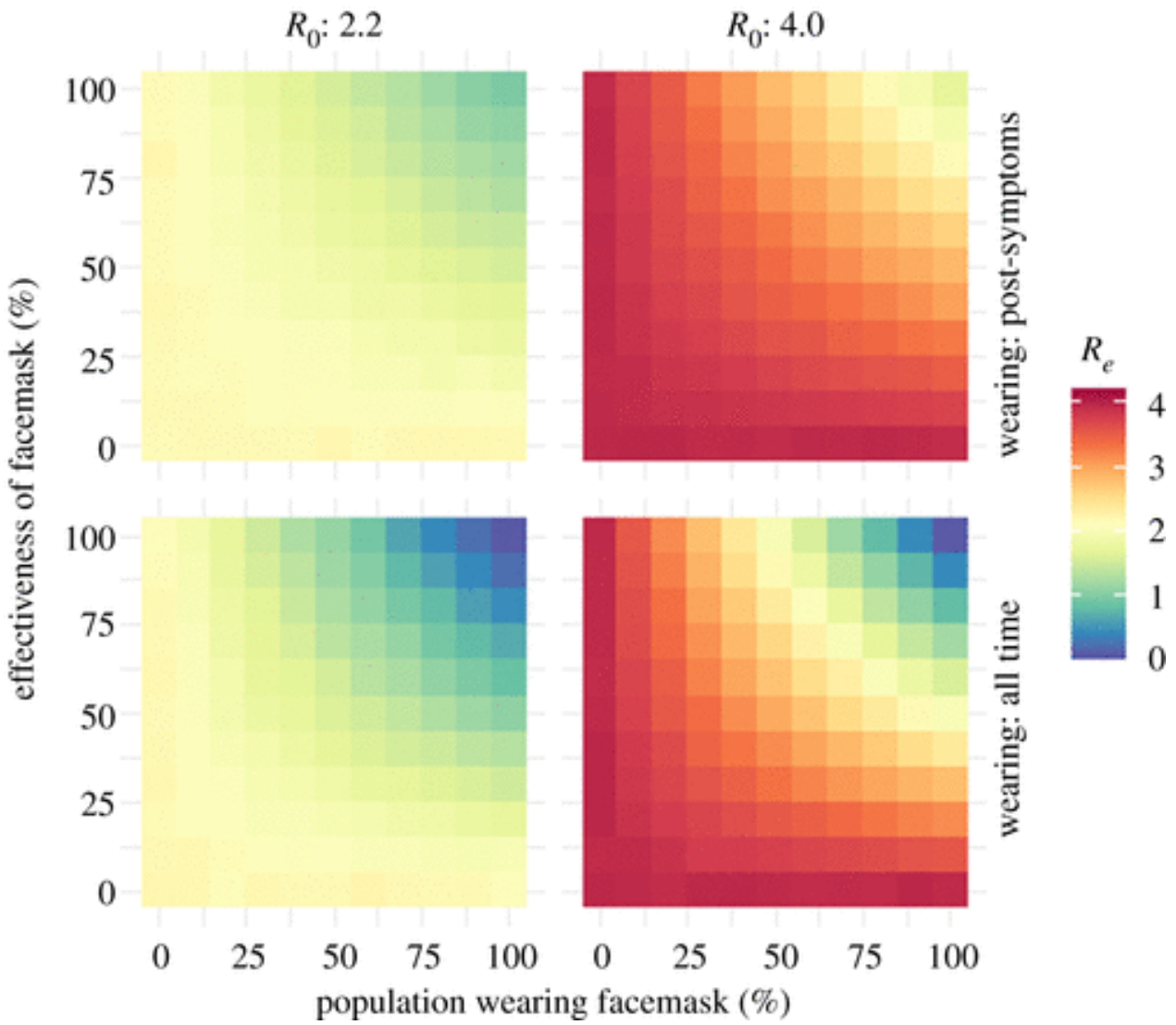




Factors Influencing R_t: Masking

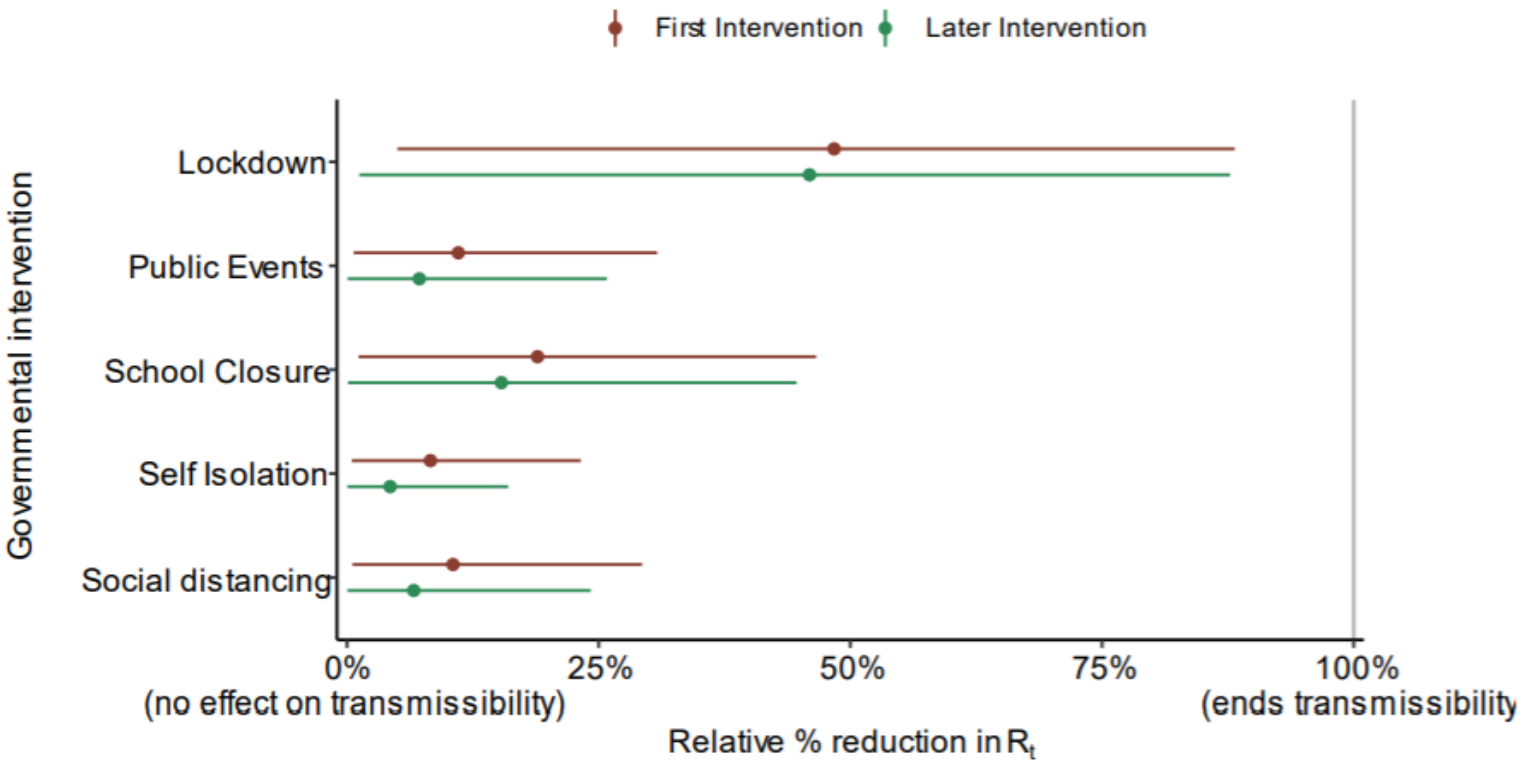
- Rader et al. (2020) found that an increase in mask wearing is associated with increased odds of transmission control ($R_t < 1$)
- Stutt et al. (2020) found that mask wearing can decrease R_t even if masks are as little as 50% effective

Factors Influencing R_t: Facemask Effectiveness



Factors Influencing R_t: Non-Pharmaceutical Interventions

- Imperial College Report



Fatality Rate

- Determines % of those moved from Infected to Removed who died/recovered
- Case Fatality Rate vs. Infection Fatality Rate
- Case Fatality (US): 2.2%
- Infection Fatality Rate:
  - Estimates range from 0.00% to 1.54% (median 0.27%) (Ioannidis, 2020)
  - Varies based on age demographics (CDC):
    - 0-19 years: 0.003%
    - 20-49 years: 0.02%
    - 50-69 years: 0.5%
    - 70+ years: 5.4%

Length of Infection

- 4-5 days average incubation period
- 10 days after onset of symptoms for mild to moderate cases (CDC)
  - Length of required isolation for asymptomatic cases
- For severe cases, likely no longer than 20 days
  - 88% no longer shedding replication-competent virus after 10 days
  - 95% after 15 days (van Kampen et al., 2020)

Matrix SIR Model

Definition (Matrix SIR Model)

The matrix SIR model looks at the population as a vector $\vec{v}_t = \{S_t, I_t, R_t\}$ and follows the equation $\vec{v}_{t+1} = A\vec{v}_t$ for some matrix A.

The simplest way to implement the SIR model is using a matrix.

Matrix SIR Model

Definition (Matrix SIR Model)

The matrix SIR model looks at the population as a vector $\vec{v}_t = \{S_t, I_t, R_t\}$ and follows the equation $\vec{v}_{t+1} = A\vec{v}_t$ for some matrix A.

Example

Consider starting population $\vec{v} = \{100, 10, 0\}$, and matrix

$$A = \begin{bmatrix} 0.9 & -0.1 & 0 \\ 0.1 & 0.9 & 0 \\ 0 & 0.2 & 1 \end{bmatrix}.$$

This is equivalent to the system of linear equations $S_{t+1} = 0.9S_t - 0.1I_t$, $I_{t+1} = 0.1S_t + 0.9I_t$, $R_{t+1} = R_t + 0.2I_t$.

Issues

Consider our last example. While the total population will remain constant, the susceptible population will be negative by time $t = 14$.
Adjusting parameters to prevent negative populations results in a simple exponential curve without the characteristic shape of an outbreak.

Remedy

How can we fix this problem?

Example

A matrix model might give the linear system
$S_{t+1} = S_t - 0.3I_t, \quad I_{t+1} = 1.1I_t, \quad R_{t+1} = R_t + 0.2I_t.$
Consider instead
$S_{t+1} = S_t - 0.3I_t\boldsymbol{S_t}/N,$
$I_{t+1} = 0.8I_t + 0.3I_t\boldsymbol{S_t}/N,$
$R_{t+1} = R_t + 0.2I_t,$
where N is the total population.
We now have $S_t, I_t, R_t \geq 0$ for any nonnegative input. The data will have the characteristic shape of an outbreak.

Nonlinear SIR Model

Definition (Nonlinear SIR Model)

The nonlinear SIR model uses a system of nonlinear equations to model the spread of disease.

Advantages:

- no population can go negative
- more accurate
- characteristic shape

Nonlinear SIR Model

Definition (Nonlinear SIR Model)

The nonlinear SIR model uses a system of nonlinear equations to model the spread of disease.

Disadvantages:

- cannot generally give closed-form expression

Differential SIR Model

Definition (Differential SIR Model)

The differential SIR model uses a system of differential equations to model disease.

Example

$$\frac{dS}{dt} = -\beta S(t)I(t)$$
$$\frac{dI}{dt} = \beta S(t)I(t) - \mu I(t)$$
$$\frac{dR}{dt} = \mu I$$

Differential SIR Model

Definition (Differential SIR Model)

The differential SIR model uses a system of differential equations to model disease.

Advantages:

- most realistic
- continuous

Differential SIR Model

Definition (Differential SIR Model)

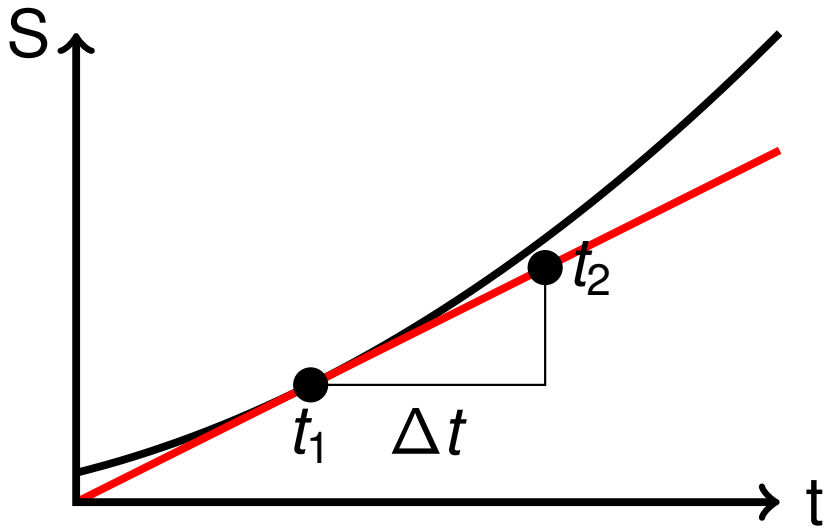
The differential SIR model uses a system of differential equations to model disease.

Disadvantages:

- can be more difficult to explain & understand
- cannot be explicitly solved

Euler's Method

Euler's method is one way mathematicians model differential equations that cannot be solved.
Euler's method treats each step of a differential equation as a linear equation.



$$S(t_2) = S(t_1) + \Delta t \frac{dS}{dt}(t_1).$$

Runge-Kutta

Runge-Kutta is a more accurate tool for estimating differential equations.

Runge-Kutta

Runge-Kutta is a more accurate tool for estimating differential equations.
Given the system $dS/dt, dI/dt, dR/dt$, calculate

$$k_1 = \Delta t \frac{dS}{dt}(S(t), I(t)) \qquad k_2 = \Delta t \frac{dS}{dt}\left(S(t) + \frac{k_1}{2}, I(t) + \frac{m_1}{2}\right) \qquad k_3 = \Delta t \frac{dS}{dt}\left(S(t) + \frac{k_2}{2}, I(t) + \frac{m_2}{2}\right)$$

$$m_1 = \Delta t \frac{dI}{dt}(S(t), I(t)) \qquad m_2 = \Delta t \frac{dI}{dt}\left(S(t) + \frac{k_1}{2}, I(t) + \frac{m_1}{2}\right) \qquad m_3 = \Delta t \frac{dI}{dt}\left(S(t) + \frac{k_2}{2}, I(t) + \frac{m_2}{2}\right)$$

$$n_1 = \Delta t \frac{dR}{dt}(S(t), I(t)) \qquad n_2 = \Delta t \frac{dR}{dt}\left(S(t) + \frac{k_1}{2}, I(t) + \frac{m_1}{2}\right) \qquad n_3 = \Delta t \frac{dR}{dt}\left(S(t) + \frac{k_2}{2}, I(t) + \frac{m_2}{2}\right)$$

$$k_4 = \Delta t \frac{dS}{dt}(S(t) + k_3, I(t) + m_3) \qquad k = \frac{k_1 + 2k_2 + 2k_3 + k_4}{6} \qquad S(t + \Delta t) = S(t) + k$$

$$m_4 = \Delta t \frac{dI}{dt}(S(t) + k_3, I(t) + m_3) \qquad m = \frac{m_1 + 2m_2 + 2m_3 + m_4}{6} \qquad I(t + \Delta t) = I(t) + m$$

$$n_4 = \Delta t \frac{dR}{dt}(S(t) + k_3, I(t) + m_3) \qquad n = \frac{n_1 + 2n_2 + 2n_3 + n_4}{6} \qquad R(t + \Delta t) = R(t) + n$$

Our Approach

We decided to use the nonlinear system of equations model.
Euler's method is effectively the same as the recursive method with small step size.

SIS Model (Susceptible-Infectious-Susceptible)

The SIS model is a model for diseases that can be contracted multiple times by the same person.

Definition (SIS Model)

The SIS model is a variant of the SIR model with only two compartments, Susceptible and Infectious. Persons who would be removed are returned to susceptible.

SIS Model (Susceptible-Infectious-Susceptible)

The SIS model is a model for diseases that can be contracted multiple times by the same person.

Definition (SIS Model)

The SIS model is a variant of the SIR model with only two compartments, Susceptible and Infectious. Persons who would be removed are returned to susceptible.

- Evidence so far suggests that COVID-19 reinfections are uncommon
- Nobody dies and is removed

SEIR Model (Susceptible-Exposed-Infectious-Removed)

The SEIR model introduces the Exposed compartment. This groups people who are in the incubation period many viruses have.

Definition (SEIR Model)

The SEIR model is a variant of the SIR model with four compartments. The exposed category marks people who cannot spread the disease, but have been exposed to it and will become infectious.

SEIR Model (Susceptible-Exposed-Infectious-Removed)

The SEIR model introduces the Exposed compartment. This groups people who are in the incubation period many viruses have.

Definition (SEIR Model)

The SEIR model is a variant of the SIR model with four compartments. The exposed category marks people who cannot spread the disease, but have been exposed to it and will become infectious.

- SARS-CoV-2 is infectious during the incubation period

SIRD Model (Susceptible-Infectious-Recovered-Deceased)

The SIRD model allows easy tracking of how many people have died of the disease.

Definition (SIRD Model)

The SIRD model is a variant of the SIR model with four compartments. The Recovered compartment marks only those who recovered, and Deceased marks those who died of the disease.

SIRD Model (Susceptible-Infectious-Recovered-Deceased)

The SIRD model allows easy tracking of how many people have died of the disease.

Definition (SIRD Model)

The SIRD model is a variant of the SIR model with four compartments. The Recovered compartment marks only those who recovered, and Deceased marks those who died of the disease.

- Useful for displaying value of flattening the curve
- Deaths can also be calculated later as proportion of Removed

Two-Group Model

$$\begin{cases} S_{u,n+1} &= S_{u,n} - \dfrac{I_{u,n}S_{u,n}k_u}{N} - \dfrac{I_{e,n}S_{u,n}k_e}{N} - jS_{u,n} \\ S_{e,n+1} &= S_{e,n} - \dfrac{I_{u,n}S_{e,n}k_u}{N} - \dfrac{I_{e,n}S_{e,n}k_e}{N} + jS_{u,n} \\ I_{u,n+1} &= I_{u,n} + \dfrac{I_{u,n}S_{u,n}k_u}{N} + \dfrac{I_{e,n}S_{u,n}k_e}{N} - I_{u,n}m \\ I_{e,n+1} &= I_{e,n} + \dfrac{I_{u,n}S_{e,n}k_u}{N} + \dfrac{I_{e,n}S_{e,n}k_e}{N} - I_{e,n}m \\ R_{n+1} &= R_n + I_{u,n}m + I_{e,n}m \end{cases}$$

where $S_{u,n} + S_{e,n} + I_{u,n} + I_{e,n} + R_n = N$ and
$S_{u,0} = N - 1, S_{e,0} = 0, I_{u,0} = 1, I_{e,0} = 0, R_0 = 0.$

Parameters

- Two-group SIR model has four parameters: k_u, k_e, j, and m
- Note: Disease progresses from nth stage to $(n+1)$st stage after one "time period"
- k_u represents the quantity of people that an uneducated, infected individual interacts with times the probability that a susceptible individual who interacts with this person gets sick
- k_e represents the same thing, but for educated, infected individuals

Excel Demonstration

Parameters

- m represents 1 divided by the number of time periods for the average person to recover
- j represents the proportion of uneducated who become educated after each time period
- Under this model, the R_0 value for early stages of the disease to be represented as k_u/m for uneducated individuals and k_e/m for educated individuals

Limitations

- Does not allow for the handling of multiple NPIs
- Does not allow for separate treatment of individuals by age
- Parameters cannot be determined to sufficiently great accuracy
- Model cannot be used to assess the spread of the virus to particular locations within the USA
- Model does not admit a closed form (most likely)

Assumptions

The model assumes that

- educated individuals are equally likely to be infected from the virus as uneducated individuals
- educated infected individuals are less likely to spread the virus than uneducated infected individuals
- the USA is a closed and localized system (there is no travel to and from the USA)
- within the USA, location has no impact
- the only artificial inhibitor to the spread of the virus is "education"
- age does not affect the spread of the virus
- . . . and more

General Recursive Model

Suppose that the population is split into p distinct groups, so that we have groups $S_1, S_2, \ldots, S_p, I_1, I_2, \ldots, I_p$, and $R_1, R_2, \ldots, R_p$. Let A be the $m \times m$ matrix of *interaction coefficients* $k_{1,1}, \ k_{1,2}, \ldots, k_{2,1}, \ldots, k_{p,p}$ representing the number of people belonging to a group that a person from another group may infect. Then,

$$A = \begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,p} \\ k_{2,1} & k_{2,2} & \cdots & k_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ k_{p,1} & k_{p,2} & \cdots & k_{p,p} \end{pmatrix}.$$

General Recursive Model

Furthermore, define X_n to be the $p \times p$ diagonal matrix with diagonal elements equal to $S_{1,n}, S_{2,n}, \ldots S_{p,n}$, so that $X_n = \begin{pmatrix} S_{1,n} & 0 & \cdots & 0 \\ 0 & S_{2,n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & S_{p,n} \end{pmatrix}$. Finally, denote $\mathbf{S}_n$ to be the N-dimensional vector $\mathbf{S}_n = \begin{pmatrix} S_{1,n} \\ S_{2,n} \\ \vdots \\ S_{m,n} \end{pmatrix}$, and define $\mathbf{I}_n$ and $\mathbf{R}_n$ similarly.

General Recursive Model

Then, the general recursive SIR model for a population of N groups is

$$\begin{cases} \mathbf{S}_{n+1} = \mathbf{S}_n - \dfrac{X_n A \mathbf{I}_n}{N} \\ \mathbf{I}_{n+1} = \mathbf{I}_n + \dfrac{X_n A \mathbf{I}_n}{N} - m\mathbf{I}_n \\ \mathbf{R}_{n+1} = \mathbf{R}_n + m\mathbf{I}_n \end{cases}$$

where m represents the frequency of recovery of the virus, and N is the total population.

Note that if the p groups are split into uneducated and educated groups, an additional "education" term may be added to the first equation which takes some proportion of the people in the uneducated group and moves them into their respective educated group.

Parameters

- Interaction coefficients $k_{a,b}$ determine the rate at which infected people from group b will infect susceptible people from group a
- The m parameter is the same between the two-group model and the general model: represents the frequency of which infected people are removed from the population, or, in other words, $1/(\text{length of infection})$

Excel Demonstration

Benefits over Two-Group Model

- The model can account for the drastically varying effects of COVID based on age
- The model can predict COVID effects in a partially vaccinated population by creating a new population which rarely infected
- The model can also predict certain NPIs other than "education" by creating a new population
- The model achieves more generality while maintaining simple calculations

Limitations

- The quantity of variables is $p^2 + 2$, which is quadratic in the number of population groups
- The interaction coefficients are practically infeasible to calculate so the model likely has limited application to reality
- Model still cannot account for location effects

Assumptions

The model assumes that

- as with the two-group model, the USA is a closed and localized system (there is no travel to and from the USA)
- within the USA, location has no impact
- there are no societal effects that exist temporally across multiple populations (for example, legal mandates)
- the interaction coefficients between any two populations remain constant over the lifetime of the virus

Thank You