Module 1:
Quantitative Approaches for Modeling COVID-19 and Other Pandemics
Coronavirus Could Overwhelm U.S. Without Urgent Action, Estimates Say

By James Glanz, Lauren Leatherby, Matthew Bloch, Mitch Smith, Larry Bachman, Jin Wu and Nicholas Bogel-Burroughs  March 20, 2020

Live updates: New York coronavirus deaths surpass 3,500; confirmed cases nationwide top 300,000

How the Virus Got Out

By Jin Wu, WeiJie Cai, Derek Perkins and James Glanz  March 21, 2020

The most extensive travel restrictions to stop an outbreak in human history haven’t been enough. We analyzed the movements of hundreds of millions of people to show why.

White House turns to statistical models for coronavirus forecast

BY RICARDO ALONSO-ZALDIVAR, LAURAN NEERGAARD, AND THE ASSOCIATED PRESS

March 31, 2020 8:30 AM EST
Outline

1) Some universal dynamical features of outbreaks

2) What is the basic reproduction number $R_0$ and why is it so important?

3) What stops disease outbreaks?

4) Essential core of models of disease dynamics: the Susceptible-Infectious-Recovered (SIR) model
Pandemics in History

**Black Plague, 1347-1351**
- Killed 30-60% of Europe’s population
- Fleas acted as the vector

**Plague of Athens, 430-427 BC**
- Killed ~1/4 of the city’s population

**Native American epidemics, 1518-1830’s**
- Smallpox and other diseases
- May have killed 90% of native Americans

**Spanish flu (H1N1), 1918-1920**
- Infected 1/3 of world population

**Recent outbreaks**
- Ebola, Zika, SARS, MERS

*Without strong control measures, can affect large fractions of population*
Pandemics differ in many ways...

- **Transmission rates**: fast or slow
- **Transmission routes**: human vs. non-human carrier, airborne vs. contact
- **Mortality & recovery rates**

But also share many core features...

These core shared features have allowed models of previous pandemics to be re-purposed for predicting the course of COVID-19
Key Concept: The “basic reproduction number” $R_0$

$R_0 = \text{average } \# \text{ of people that each infected person infects, in a nearly entirely susceptible population}$

Why is this number so important?
Consider $R_0 = 2$:

Initially 1 infected person ($I_0$)

Generation:

$I(n) = \#\text{ new infections in generation } n = (1)(2)^n = I_0 R_0^n$

Grows exponentially fast – outbreak!!!
Consider $R_0 = \frac{1}{2}$:

Initially 8 infected people ($I_0$)

$I(n) = \text{(# new infections in n\textsuperscript{th} generation)} = (8)(\frac{1}{2})^n = I_0 R_0^n$

Decays exponentially fast – rapid recovery!
Consider $R_0 = 1$:

Initially 8 infected people ($I_0$)

$$(n) \quad 0 \quad 1 \quad 2 \quad 3$$

No change: borderline between growth and decay

$$I(n) = (# \text{ new infections in } n^{th} \text{ generation}) = (8)(1)^n = I_0 R_0^n$$
R0: Summary So Far

R0 = average # of people that each infected person infects, in a nearly entirely susceptible population

R0 > 1: exponential growth (outbreak)
R0 < 1: no outbreak

\[ I(n) = I_0 R_0^n \]

Issue: Real data is plotted versus time, not generation
Generation Time

**Generation time** $g$: mean time interval between infection of one person, and infections of the people that this person infects.

- 3 days
- 5 days
- 7 days

$g = 5$ days

\[ I(n) = I_0 R_0^n \]

The number of new infections ($I(n)$) can be calculated using the formula $I(n) = I_0 R_0^n$, where $I_0$ is the initial number of infections, $R_0$ is the basic reproduction number, and $n$ is the number of generations. The time in days is calculated as $(\# \ of \ generations) \times (time \ per \ generation)$.

\[ t = n \times g \quad \text{or} \quad n = \frac{t}{g} \]

Graph showing the number of new infections ($I(n)$) as a function of the number of generations ($n$) and time in days ($t$).
**Generation Time**

Generation time $g$: mean time interval between infection of one person, and infections of the people that this person infects.

$g = 5$ days

$I(t) = I_0 R_0^{(t/g)}$

- $t = n \times g$ or $n = t / g$
- time in days = ($\#$ of generations) * (time per generation)

Graph showing the relationship between the number of generations and the time in days.
Question: How do we determine $R_0$ from real data?

\[ I(t) = I_0 R_0^{(t/g)} \]

Issue: Even if we know $g = 5$ days, still very hard to see what $R_0$ is from this graph.
**Answer:** Instead plot on semi-log plot!

Math: \[ I(t) = I_0 R(t/g) \]
**Answer:** Instead plot on semi-log plot!

### Original plot (linear y-axis)

\[ I(t) = I_0 R_0 \left( \frac{t}{g} \right) \]

### Semi-log plot (logarithmic y-axis)

Math: \[ \log_{10}(I(t)) = \log_{10}(I_0 R_0 \left( \frac{t}{g} \right)) \]

\[ \log_{10}(I(t)) = \log_{10}(R_0 \left( \frac{t}{g} \right)) + \log_{10}(I_0) \]

Recall:
- \( \log(a \cdot b) = \log(a) + \log(b) \)
- \( \log(a^t) = t \cdot \log(a) \)
**Answer:** Instead plot on semi-log plot!

\[
I(t) = I_0 R_0^{(t/g)}
\]

---

**Math:**

\[
\begin{align*}
\log_{10}(I(t)) &= \log_{10}(I_0 R_0^{(t/g)}) \\
\log_{10}(I(t)) &= \log_{10}(R_0^{(t/g)}) + \log_{10}(I_0)
\end{align*}
\]

**Recall:**

\[
\begin{align*}
\log(a \cdot b) &= \log(a) + \log(b) \\
\log(a^t) &= t \cdot \log(a)
\end{align*}
\]
**Answer:** Instead plot on semi-log plot!

Original plot (linear y-axis)

- \( I(t) = I_0 R_0(t/g) \)

Semi-log plot (logarithmic y-axis)

Math:

\[
\log_{10}(I(t)) = \log_{10}(I_0 R_0(t/g))
\]

\[
\log_{10}(I(t)) = \frac{t}{g} \log_{10}(R_0(t/g)) + \log_{10}(I_0)
\]

\[
\log_{10}(I(t)) = \frac{1}{g} \log_{10}(R_0) t + \log_{10}(I_0)
\]

Recall:

- \( \log(a \cdot b) = \log(a) + \log(b) \)
- \( \log(a^t) = t \cdot \log(a) \)
Answer: Instead plot on semi-log plot!

Math:

\[ \log_{10}(I(t)) = \log_{10}(I_0 R_0^{(t/g)}) \]
\[ \log_{10}(I(t)) = \frac{t}{g} \log_{10}(R_0^{(t/g)}) + \log_{10}(I_0) \]
\[ \log_{10}(I(t)) = \left(\frac{1}{g}\right)\log_{10}(R_0) t + \log_{10}(I_0) \]

"y" = "m x" + "b"

slope 
intercept

Recall:

\[ \log(a \times b) = \log(a) + \log(b) \]
\[ \log(a^t) = t \times \log(a) \]
Determine $R_0$ from slope of semi-log plot!

\[ \log_{10}(I(t)) = \left( \frac{1}{g} \right) \log_{10}(R_0) t + \log_{10}(I_0) \]

\[ \text{(1/g) log}_{10}(R_0) = \text{slope} \times g \]

\[ R_0 = 10^{(\text{slope} \times g)} = 10^{(2/33.2 \text{ days}) \times (5 \text{ days})} = 2 \]
Next Question:
What mechanistic factors influence R0?

R0 = average # of people that each infected person infects, in a nearly entirely susceptible population

R0 = (# people infected)

= (# people infected/day) * (# days person is contagious)

= (probability a contact causes infection) * (# contacted/day) * D

R0 = p * C * D
Thought Question:
Which factor does social distancing affect?

1) Probability of transmission p
2) # of contacts C
3) # of days infected D
4) All of the above

Answer: (2) # of contacts C

Explanation:
By keeping people away from each other, it reduces the # of contacts we make each day. It does not reduce the # of days we are infected. If you make a contact, it does not reduce the probability p of transmitting the disease.
Basic reproduction number $R_0$:
$R_0$ = average # of people that each infected person infects, in a nearly entirely susceptible population

Effective reproduction number $R_{eff}$:
$R_{eff}$ = average # of people that each infected person infects, when some of the population is immune (= $R_0$ at start of epidemic)

1999 Measles outbreak in the Netherlands

- $R_{eff} > 1$: infections rise
- $R_{eff} < 1$: infections decline
- peak: $R_{eff} = 1$
Effect of Vaccination

**Question:** What fraction of the population must be vaccinated to avoid outbreaks for a disease with basic reproduction number R0?

**Answer:** Let v = fraction of people vaccinated

\[ 1 - v = \text{fraction of susceptible people} \]

e.g. \( R_0 = 2, \ v = \frac{1}{2} \):

![Diagram showing the relationship between infected, immune, and susceptible individuals.]

**Math:** To avoid an outbreak, need \( R_{eff} < 1 \):

\[ R_{eff} = (\text{# of infections if all susceptible}) \times (\text{fraction susceptible}) < 1 \]

\[ R_0 \times (1 - v) < 1 \]

\[ 1 - v < \frac{1}{R_0} \]

\[ v > 1 - \frac{1}{R_0} \]

"Herd immunity": Only need to vaccinate a fraction of the herd (i.e. population) to eradicate a disease.

No increase in # of infected!

\[ R_{eff} = 1 \]
Part 2: Disease Dynamics & the SIR Model

Flattening the curve:

Key Question: How can we understand where the shapes of these curves come from?
The SIR Model: Susceptible-Infected-(Recovered or Removed)

Rule 1:
- Susceptible population ($S$)
- Infected population ($I$)

$\beta$ ...are infected by infected people at rate $\beta$

Rule 2:
- Infected population ($I$)
- Recovered or Removed population ($R$)

$\gamma$ recover (or die) at a rate $\gamma$

Assume total population size $N$ fixed:

$$S + I + R = N$$

Assumed to be immune (includes vaccinated as well)
Let’s turn this into a quantitative model:

Susceptible population \((S)\)

...are infected by infected people at rate \(\beta\)

Infected population \((I)\)

recover (or die) at a rate \(\gamma\)

Recovered/Removed population \((R)\)

In a time interval \(\Delta t\):

\[
\Delta S = \text{Change in } # \text{ susceptibles} = - (\# \text{ who get infected}) = - \beta S I \Delta t
\]

\[
\Delta R = \text{Change in } # \text{ recovered/removed} = \# \text{ infected who recover} = \gamma I \Delta t
\]

\[
\Delta I = \text{Change in } # \text{ infected} = (\# \text{ who get infected}) - (\# \text{ who recover})
\]

\[
\Delta I = \beta S I \Delta t - \gamma I \Delta t
\]

For small \(\Delta t\), we divide by \(\Delta t\) and write:

\[
\frac{dS}{dt} = -\beta S I
\]
SIR model summary

Susceptible (S) → \( \beta \) → Infected (I) → \( \gamma \) → Recovered/Removed (R)

\[
\frac{dS}{dt} = -\beta SI
\]
(Susceptibles get infected by infected)

\[
\frac{dI}{dt} = \beta SI - \gamma I
\]
(Susceptibles get infected & infecteds recover/get removed)

\[
\frac{dR}{dt} = \gamma I
\]
(Infecteds recover or are removed)
When Does an Outbreak Occur?

Outbreak occurs when # of infected grows:
\[ \frac{dI}{dt} = (\text{new infection rate}) - (\text{recovery/removal rate}) > 0 \]

\[ \beta S I > \gamma I \]

\[ \beta S \left( \frac{1}{\gamma} \right) > 1 \]

this is \( R_{\text{eff}} \) for the SIR model!

When whole population is susceptible, \( S = N \)

\[ R_0 = \beta N \left( \frac{1}{\gamma} \right) \] for the SIR model
Comparison between our 2 formulas for R0 formula

**Simple model: no dynamics**

\[ R_0 = (pC)D \]

- \( D = \text{mean # of days contagious} \)

**SIR dynamical model**

\[ R_0 = \beta N \left( \frac{1}{\gamma} \right) \]

- \( \gamma = \text{recovery rate} = \frac{1}{D} = \frac{1}{(\text{mean # days contagious})} \)

\( \beta = \frac{pC}{N} = \frac{\text{# people infect/day}}{\text{total # of people}} \)

- \( \beta = \text{fraction of population infected per day by each infected person} \)

\[ \beta N = pC \]

\[ 1/\gamma = D \]
Extensions of the SIR Model

The SIR model contains the essence of nearly all disease dynamics models.

There are many possible extensions, including...

...more states & connections between states:
  • A **latent** period after Exposure but before becoming infectious (‘SEIR model’)
  • An **asymptomatic** period after becoming infectious, but before showing symptoms.
  • Recovered people losing their immunity & returning to susceptible (e.g. flu)
  • Different susceptible or infected populations (e.g. young vs. old)

...spatial dynamics:
  • Travel, spatial spread between different regions
  • Different people contact more or fewer other people

...randomness + time-varying parameters (e.g. seasonal changes):

You will have a chance to play with a few of these in this week’s and next week’s laboratories...and can play with more on your own!