Module 2

Lecture 2

Epidemics & The SIR model

Sarah Groves, Ph.D.

Shayn Peirce-Cottler, Ph.D.



New information about our mystery virus

CONFIDENTIAL OUTBREAK INVESTIGATION
DAY 2 BRIEF

OUTBREAK UPDATE — UNDERSTANDING EPIDEMIC SPREAD
TO: Industry modeling consultants hired by UVA
FROM: UVA Student Health & Wellness
DATE: February 24, 2026

NEW INFORMATION
Surveillance teams note that early transmission clusters were associated with large indoor gatherings and shared dining facilities. Environmental sampling has detected viral RNA on high-touch surfaces, though viability outside the host appears limited. Routine vaccination records have not demonstrated clear protective effects. Investigations remain ongoing and has refined the clinical timeline.

UPDATED DISEASE PARAMETERS
- **Symptomatic period:** approximately **5-9 days**
  - **Infectious period**: **~2 days** before symptom onset
- **Mild respiratory symptoms** predominate early in illness
- Some students report a **rash** and **low-grade fever.** A minority of cases describe **sore throat** and **fatigue** preceding fever.
- **Gastrointestinal** symptoms have been reported in ~8% of cases.

DATA AVAILABLE
No new data is available at this time. Investigations are ongoing, and new data is expected in the next briefing.

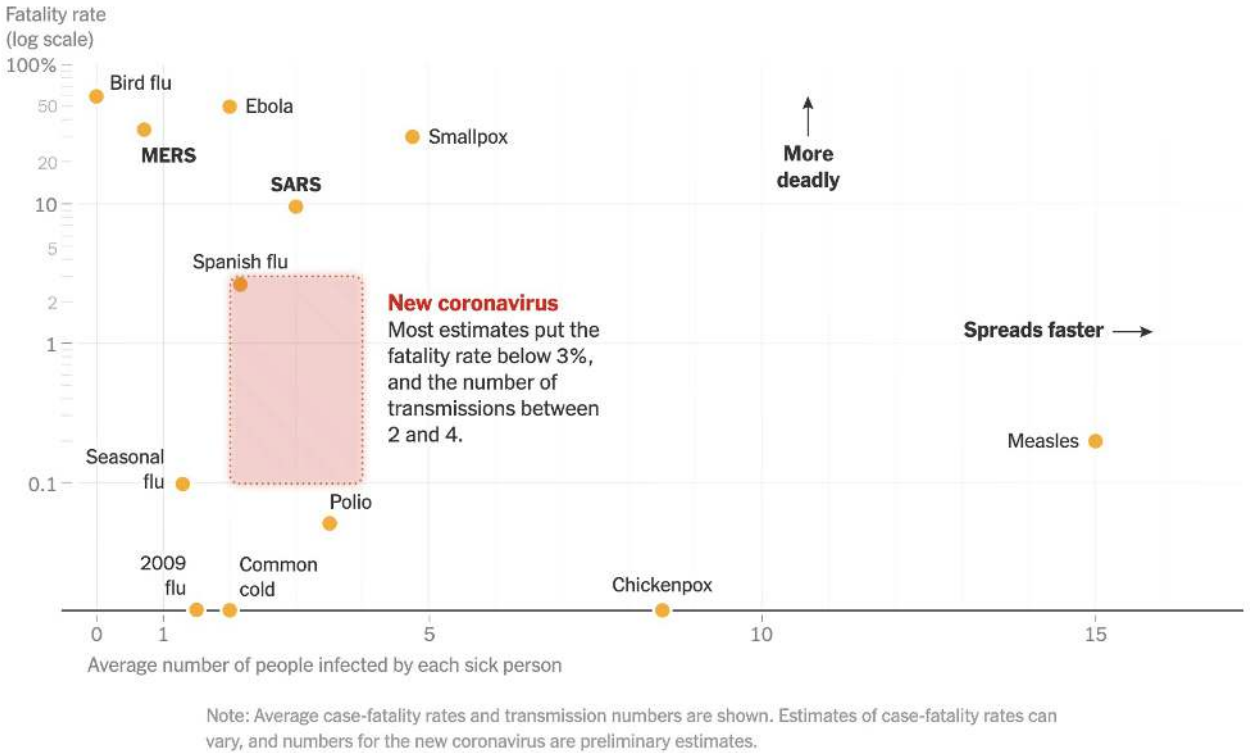
CURRENT TASKS
- Understand the relationship between **disease characteristics** and **epidemic characteristics**
- Estimate the **initial growth rate** from early outbreak data

- Symptomatic period: 5-9 days, with 2 days infectious before symptoms appear
- Symptoms can help us better understand what type of virus it is
  - Is it in the same family as a known virus?

Using Biology to Understand Outbreaks

- **Key parameters we can estimate from biological characteristics:**
  - **R_0 (Basic reproduction number):** How many people will one case infect?
  - **Incubation period:** How long after someone is exposed can they infect others?
  - **Infectious period:** How long can someone transmit the disease?
  - **Recovery rate:** How quickly does population become immune?

Why do some viruses spread faster than others?



See more with viruses.html in your GitHub repo

- **Transmission route** (airborne? Respiratory droplets?)
- **Infectious dose**: min # of virus particles needed to start an infection
- **Viral shedding**: high shedding + long infectious period = higher R0
- **Host susceptibility** (cross-immunity, genetic factors)

February 2020

https://www.nytimes.com/interactive/2020/world/asia/china-coronavirus-contain.html

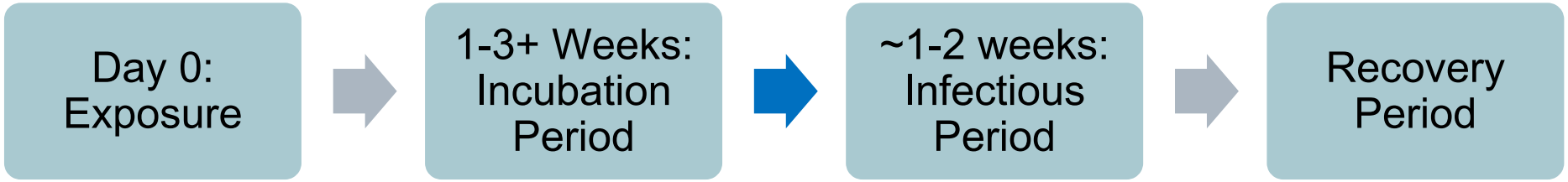
Infectiousness: β= transmission rate



- How fast infectious people transmit the disease to susceptible people
- β = “beta” = “transmission rate” = contact rate X transmission probability per contact
- Hard to estimate directly from data

- If transmission probability is 0.5 and contact rate is 5 per day, β = 2.5 people/day

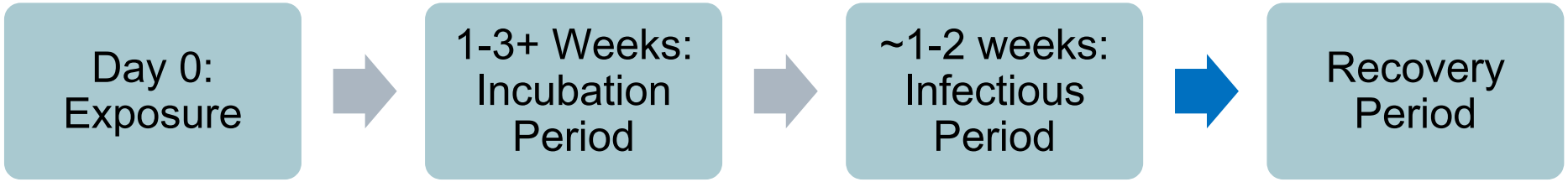
Latency: σ = Incubation rate



- How fast exposed people become infectious
- σ = “sigma” = 1/latent period
- Remember this depends on biological factors like replication speed of the virus, initial dose received, and speed of immune system activation
- Sometimes can estimate based on virus biology & testing vs symptoms

- If latent period is 10 days, σ is 0.1/day, meaning “on average it takes ten days for an exposed person to be become infectious”

Infectious period: γ = Recovery rate

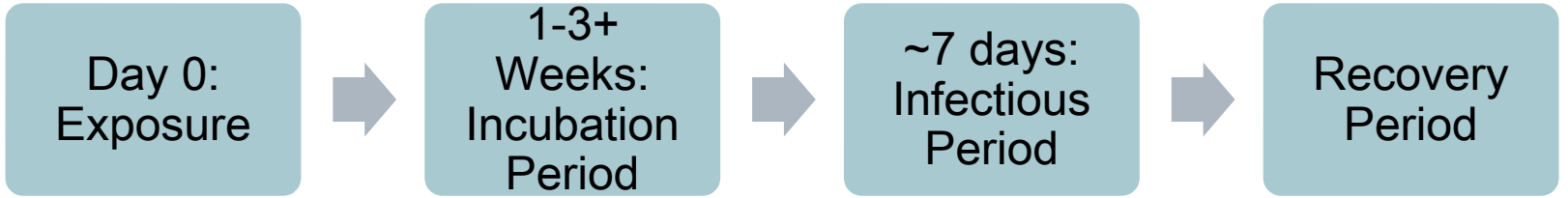


- How fast infectious people recover
- γ = “gamma” = 1 / infectious period
- Usually have an estimate of γ based on symptoms

- If the infectious period is 7 days (until recovery), γ is 1/7, meaning on average 1 out of 7 infected people will recover each day

R0: Basic reproduction number

- R_0 = “Average number of people one infected person will infect in a completely susceptible population”
- R0 = transmission probability per contact X contact rate X infectious period
- Can also define this as R0 = β / γ
- Early in the epidemic, we can use the exponential growth in infections to predict R0



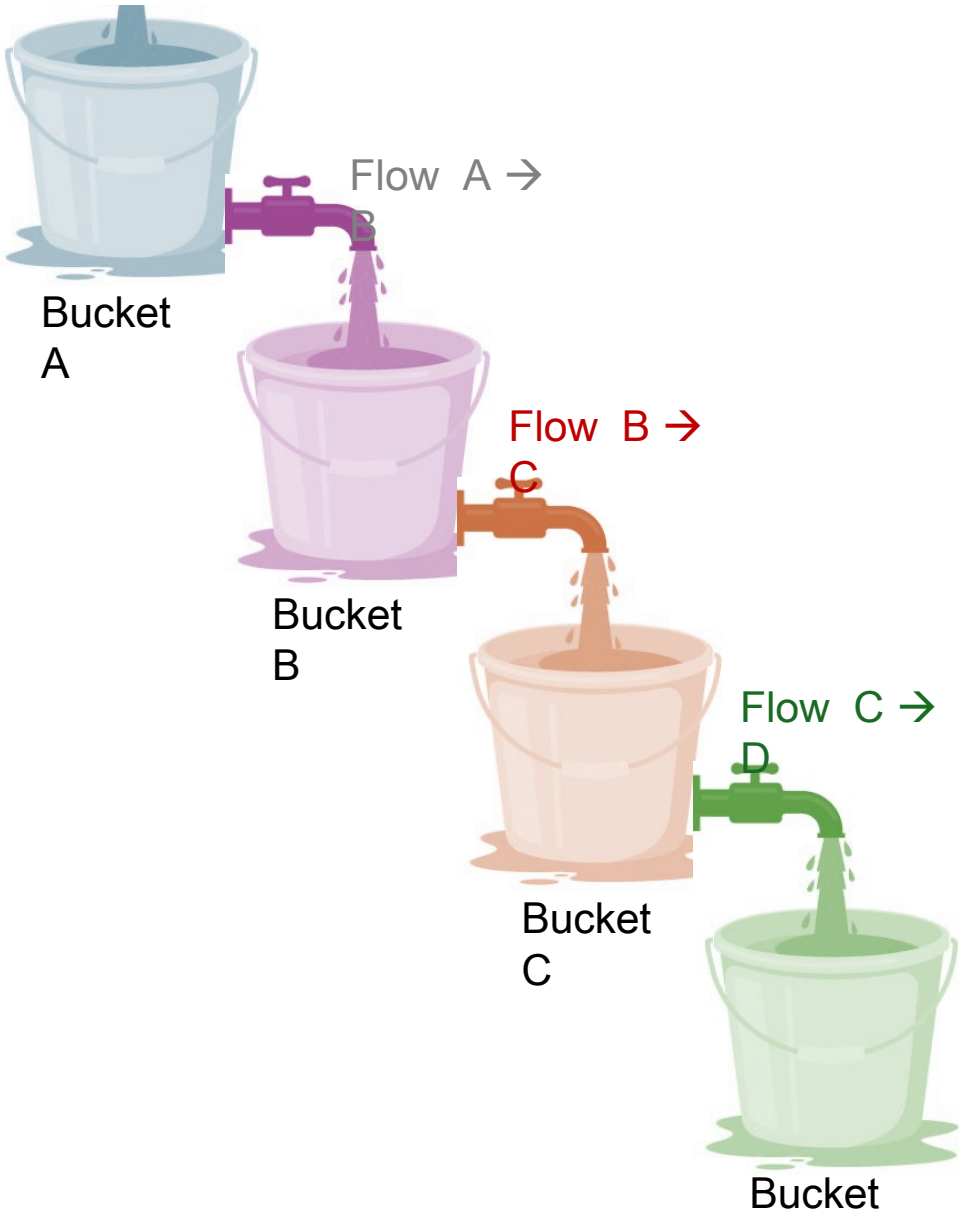
Transmission probability = 0.2
Contact rate = 5 per day
β = 1 person /day
Infectious period = 7 days
R0 = 7

From individual to population

- Imagine a population of 10,000 people:
  - Day 1: 1 person exposed
  - Day 4: Original person now infectious, plus 2 new people exposed
  - Day 7: Original person + 2 new people now infectious, plus 4 more exposed
  - Day 10: 7 people infectious, plus 8 more exposed...
- How many people will be infected by next week?
- When will the epidemic peak?

We need to track groups of people, not just individuals

Compartmental Models



(Finite) difference equations

Bucket A (t+1) = Bucket A (t) – flow A→B
Bucket B (t+1) = Bucket B(t) + flow A→B - flow B→C
Bucket C (t+1) = Bucket C(t) + flow B→C - flow C→D
Bucket D (t+1) = Bucket D(t) + flow C→D

Differential equations

$$\frac{dA}{dt} = -aA$$

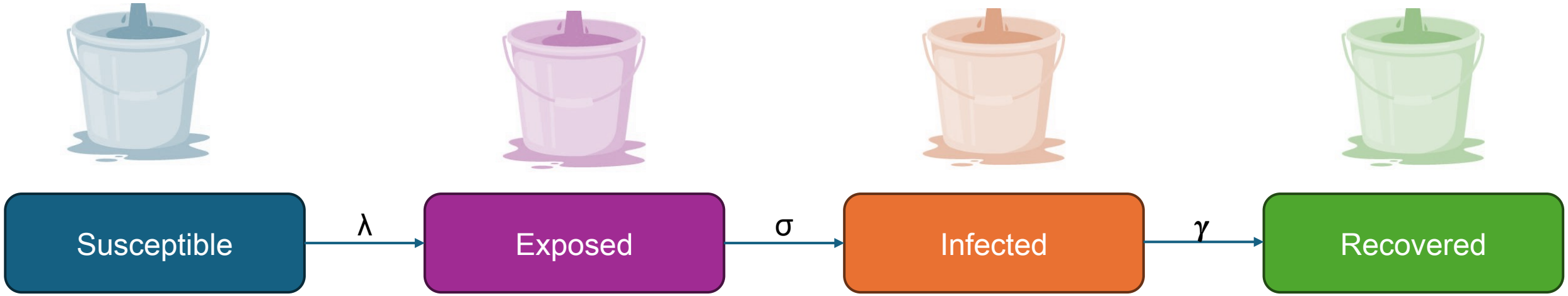
$$\frac{dB}{dt} = aA - bB$$

$$\frac{dC}{dt} = bB - cC$$

$$\frac{dD}{dt} = cC$$

Epidemiological compartment models: Group the population into compartments depending on disease status

The SIR model



How would you write the differential equations given the flowchart above?

Epidemic activity

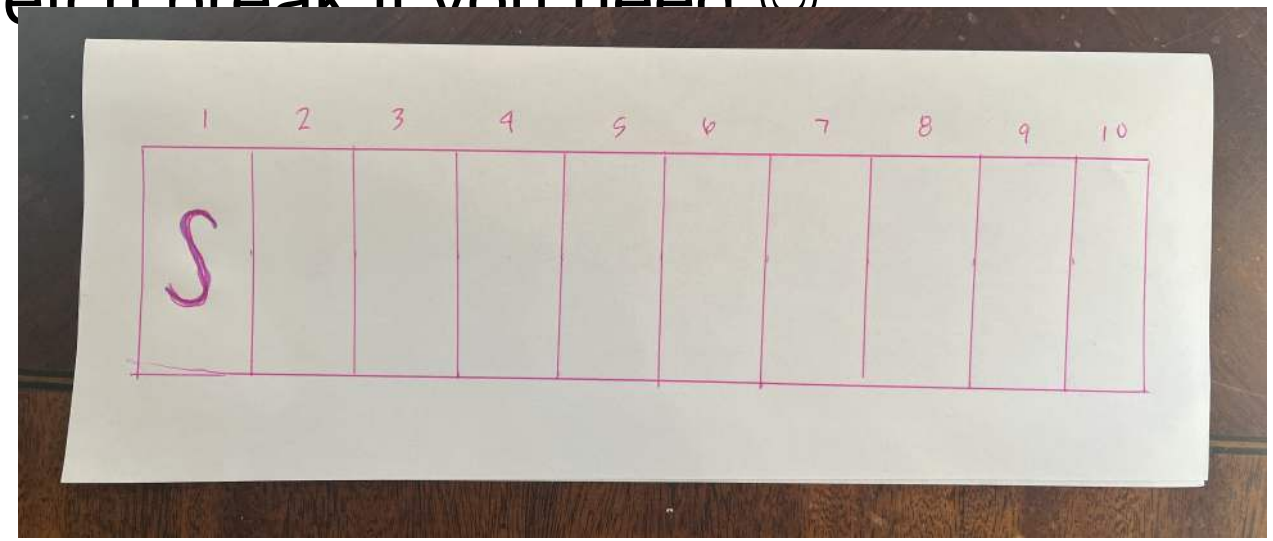
- **Susceptible** population: the entire class! (minus one…)
- **Exposed** population: ignore for this game (we are assuming exposure quickly leads to infection)
- **Infectious** population : 1 (Patient zero!)
- **Recovered** population: 0 (beginning of epidemic)
- **Transmission probability** = 0.5 (“finger matching game”)
- **Length of infection** = 3 days (recovery rate γ = 1/3 days)

Game play

- During the **day**, infections can spread if infected & susceptible individuals choose the same number
  - Infected individuals challenge susceptible neighbors; can't challenge the same person twice in a row
  - "1 2 3 shoot"
  - When you are infected, wait until the next day (round) to start infecting others
- During the **night**, we will keep track of how many people are in each compartment & figure out each person's status for the upcoming day.
  - If you were not infected, write "S": you are still susceptible
  - If you were infected during the day, write "I" for the upcoming day (i.e. day 1 of infection)
  - Once you've written "I" three times, you are recovered ("R")

Game set up and predicting SIR data: 5 min break

1. Add 10 squares to the inside of your name-tag and number them.
   - Turn it inside-out and place in front of you so others can see your status
   - Unless you are patient zero, write “S” in the first box under Day 1.
2. Go up to the board and draw a guess for what the data will look like for ONE of the compartments: S, I, or R. Label your guess with your name.
3. Take a little stretch break if you need ☺



Calculating R0 from parameters

- **Transmission probability** = 0.5 (“finger matching game”)
- **Contact rate** = ~3 neighbors per day
- **Length of infection** = 3 days (recovery rate = 1/3 days)
- $\beta = 0.5 * 3 = 1.5$ contacts per day
- $\gamma = 1/3$ days
- **R0** = $\beta / \gamma = 4.5$ people, on average, are infected by each infectious person in total

Predicting R0 from game data

- R0 = "Average number of people one infected person will infect in a completely susceptible population"
  1. Measure growth of infected population from one round to the next (multiplicative growth factor per round, g)
  2. Early epidemic growth looks like $I(t) = I_0 g^t$ (discrete growth)
  3. Over an infectious period of D, $R_0 \approx g^D$
- Continuous growth (like our mystery virus):

$$g = e^r \text{ so } I(t) = I_0 e^{rt}$$

  - Over an infectious period D, $R_0 \approx e^{rD} \approx 1 + rD$ (if rD is small)
  - If we know D (infectious period) and find r (by fitting an exponential to our data), we can estimate R0!

Looking ahead & Homework

- In a .py file (that will be copied to your final notebook): Estimate R0 for the mystery virus data using a fit to the exponential growth in I.
  - We don't know the transmission rate or recovery rate for the mystery virus, like we do in the game → we will estimate these parameters based on the data (working backwards from what we did in the game)
  - Plot the estimated exponential curve over your scatterplot from last class.
- In your final notebook:
  - What viruses have a similar R0? Use the viruses.html file to find a virus or 2 with a similar R0 and give a 1-2 sentence background of the diseases.
  - How accurate do you think your R0 estimate is?

(This will all be part of your Github check in #1 due Friday 2/27)