Introduction to Mathematical Epidemiology

Gerardo Chowell, PhD
Professor of Epidemiology and Biostatistics

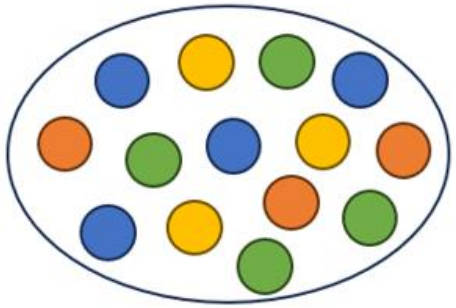
SISMID 2024

Georgia State University | SCHOOL OF PUBLIC HEALTH

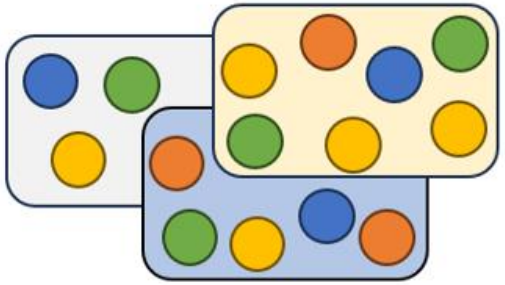
Mathematical Models

From toy models to high-fidelity models..

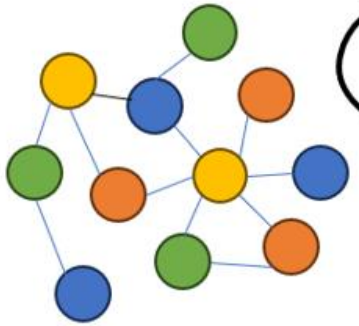
Homogeneous Mixing



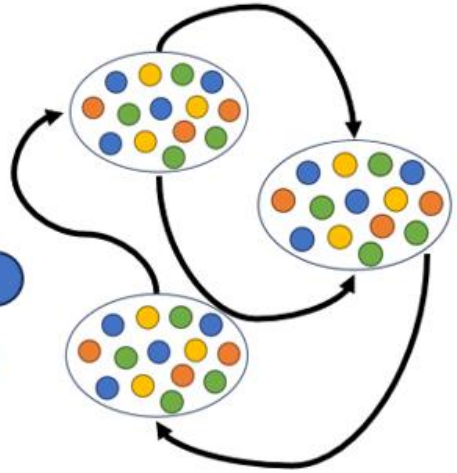
Social Structured Individuals classified by age, gender, etc.



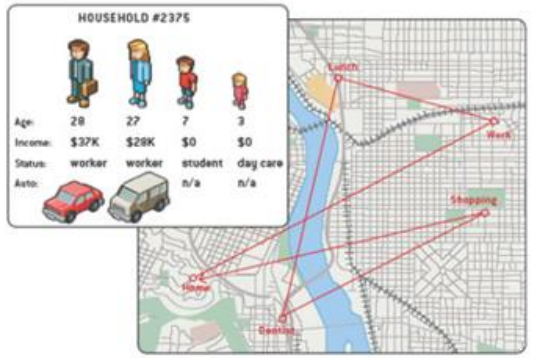
Contact Network



Multi-Scale (patch) Model



Agent-based simulation (Barrett et al. 2005, Sci. American)



- Homogeneous mixing
- Minimum number of disease stages
- No social/demography structure

COMPLEXITY

- Very high degree of heterogeneity
- Large number of disease stages
- Sophisticated social network

Why Use Mathematical Models to Study Disease Spread?

Mathematical models provide a consistent framework for analysis, prediction, understanding, and providing feasible forecasts for:

Predicting the emergence of new pathogens

Estimating the degree to which the disease will spread among different populations

Understanding of the early history of an outbreak and identifying the initial infection

Optimizing the impact of prevention strategies on disease transmission;

Assessing the impact of proposed medical interventions

Assessing the effectiveness of partial protection to infection by vaccination or protective clothing.

Improving our **understanding** of the interaction of risk factors as they affect transmission (example: schools, and work environment); and

Possible Research Topics

Investigate epidemiology and transmission parameters

Evaluate the contribution of different transmission mechanisms (asymptomatic vs symptomatic transmission)

Optimal allocation of social distancing and pharmaceutical interventions (vaccines, therapeutics)

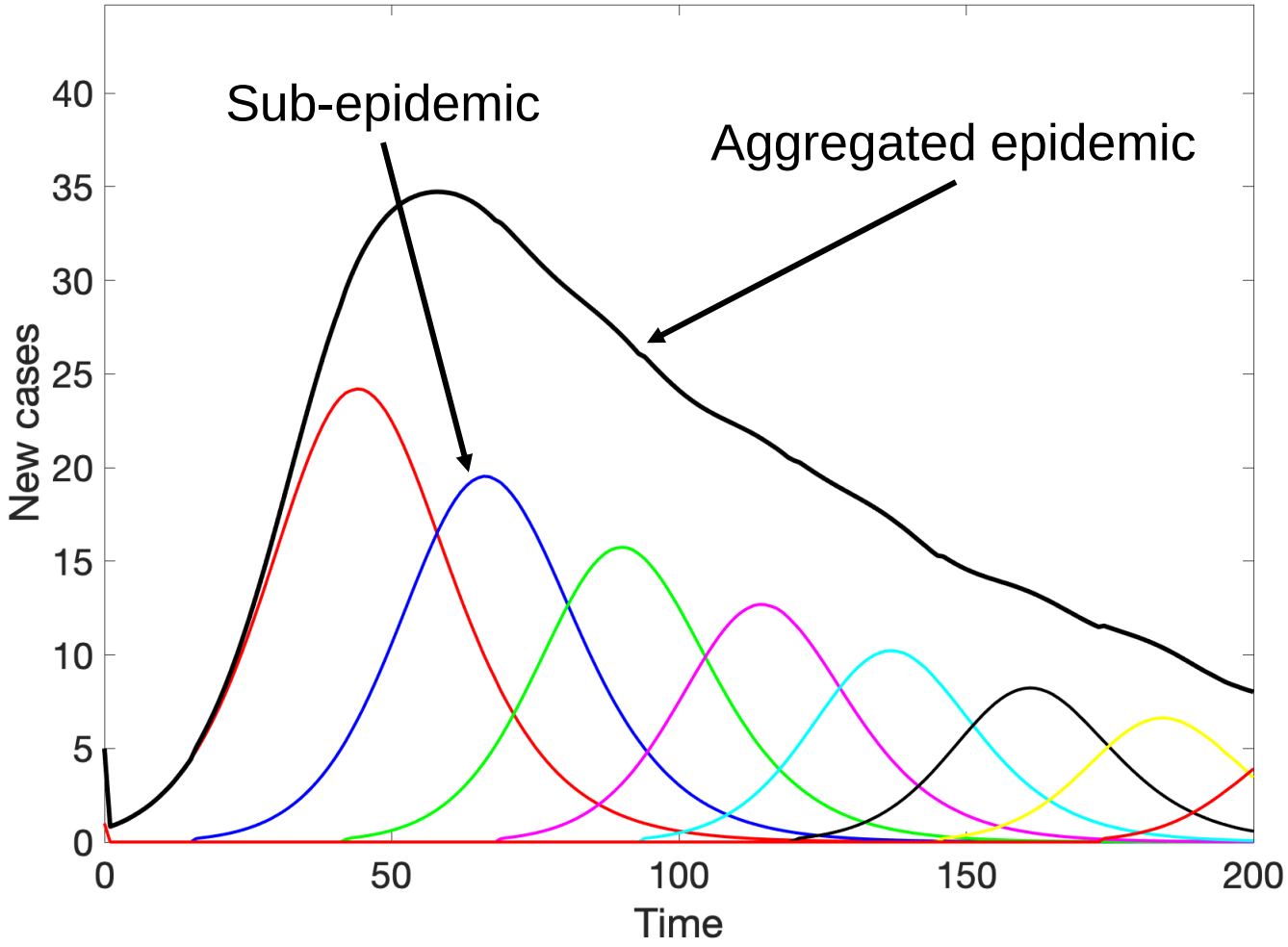
Forecast the course of the epidemic in terms of cases, hospitalizations, deaths, etc.

Complexities of epidemics

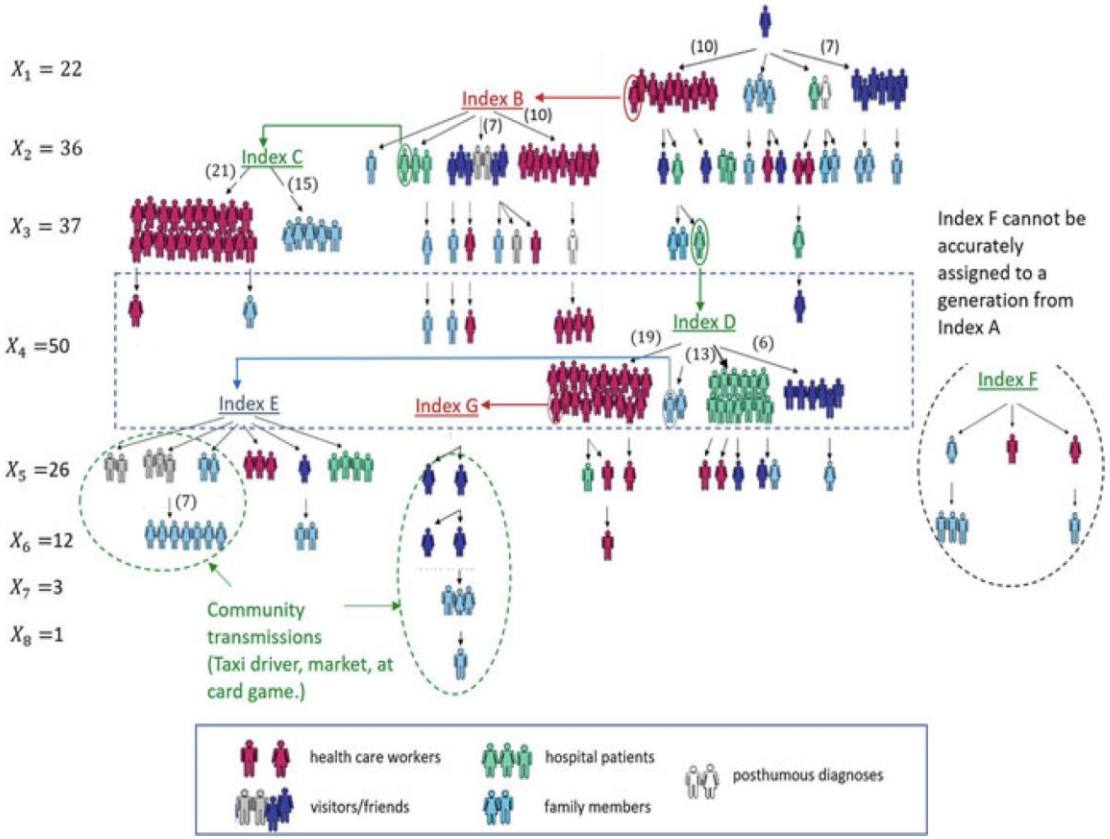
The complexity of epidemics: Sub-Epidemics

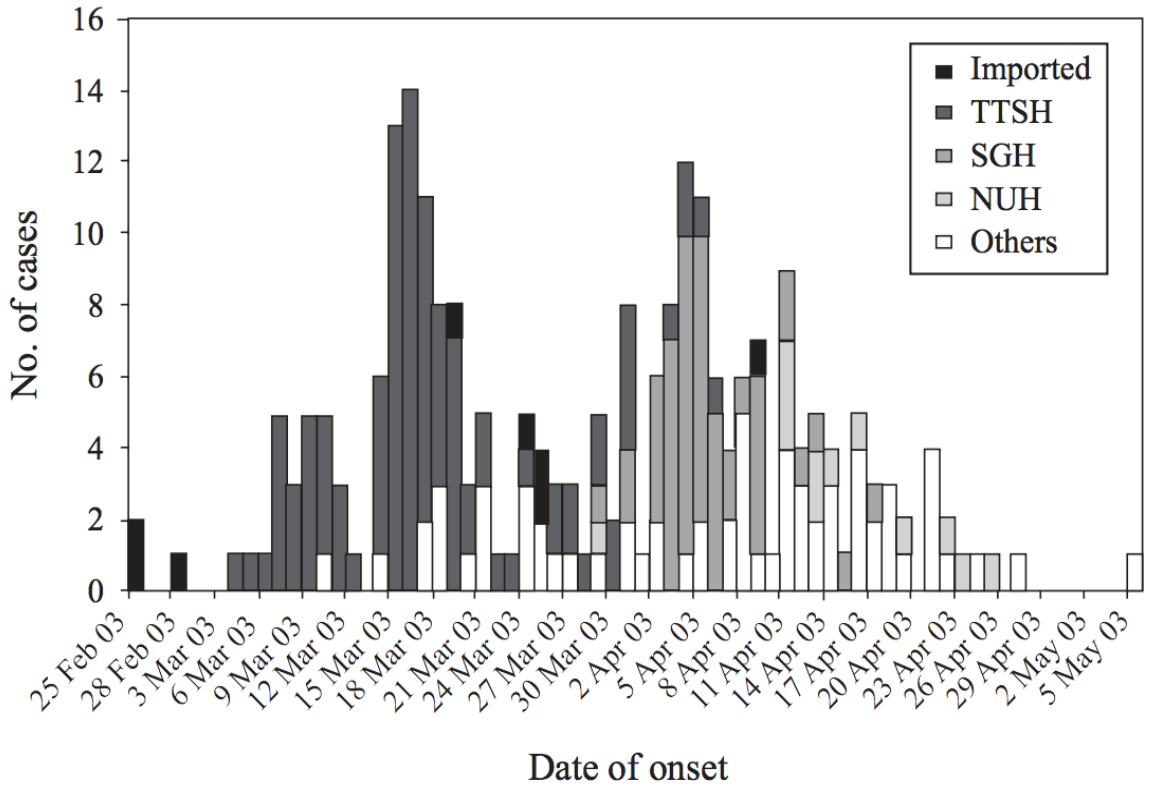
- Epidemics are usually presented as single event (i.e., *Aggregated epidemic*), however they often consist of multiple interacting, overlapping "sub-epidemics"

Sub-Epidemic
Outbreaks with unique characteristics (i.e., transmission rates, affected populations, responses to interventions) underlying an epidemic



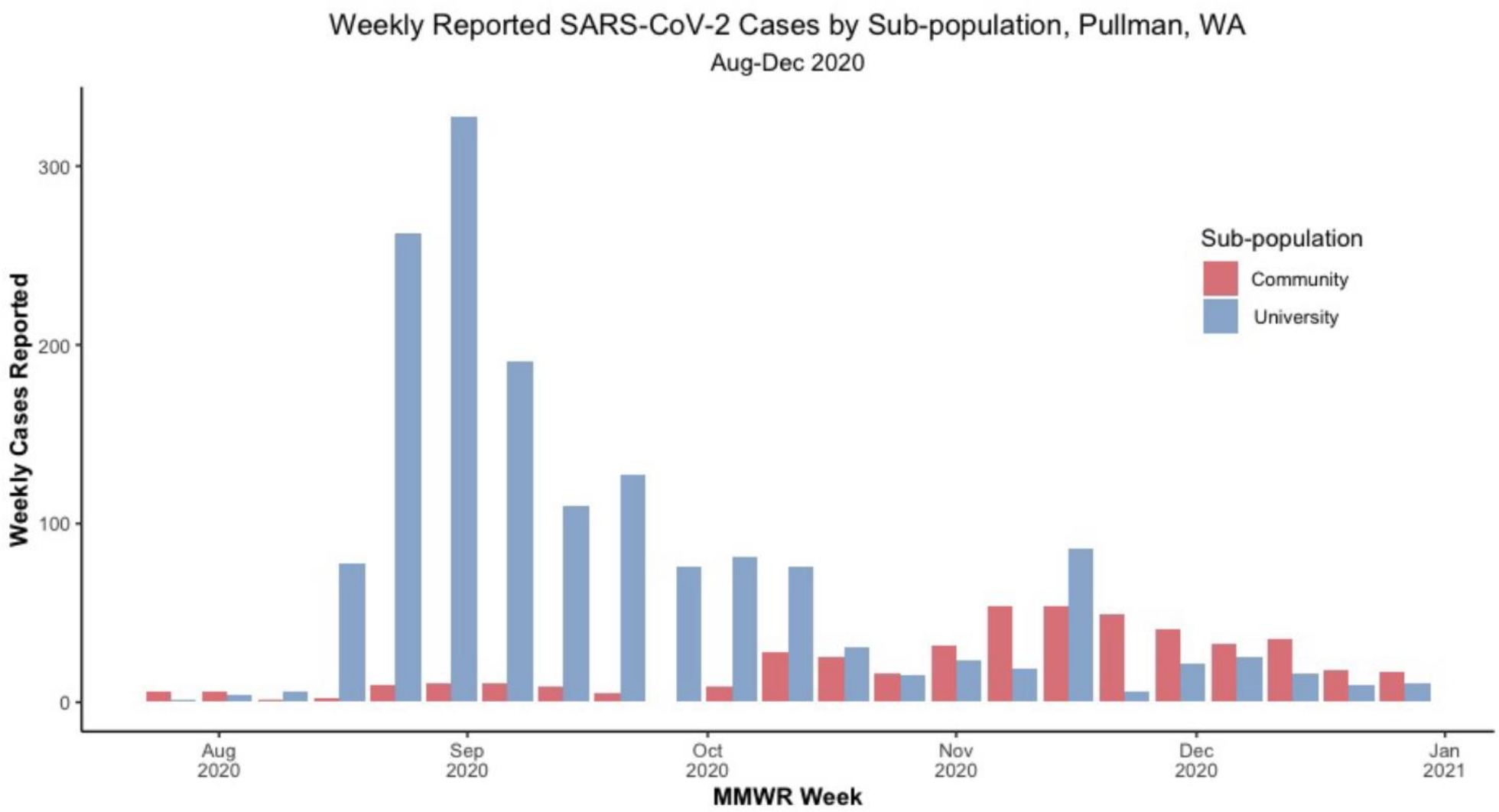
Example: The complexity of epidemics





Adapted from Goh et al. 2006

Example: The complexity of epidemics

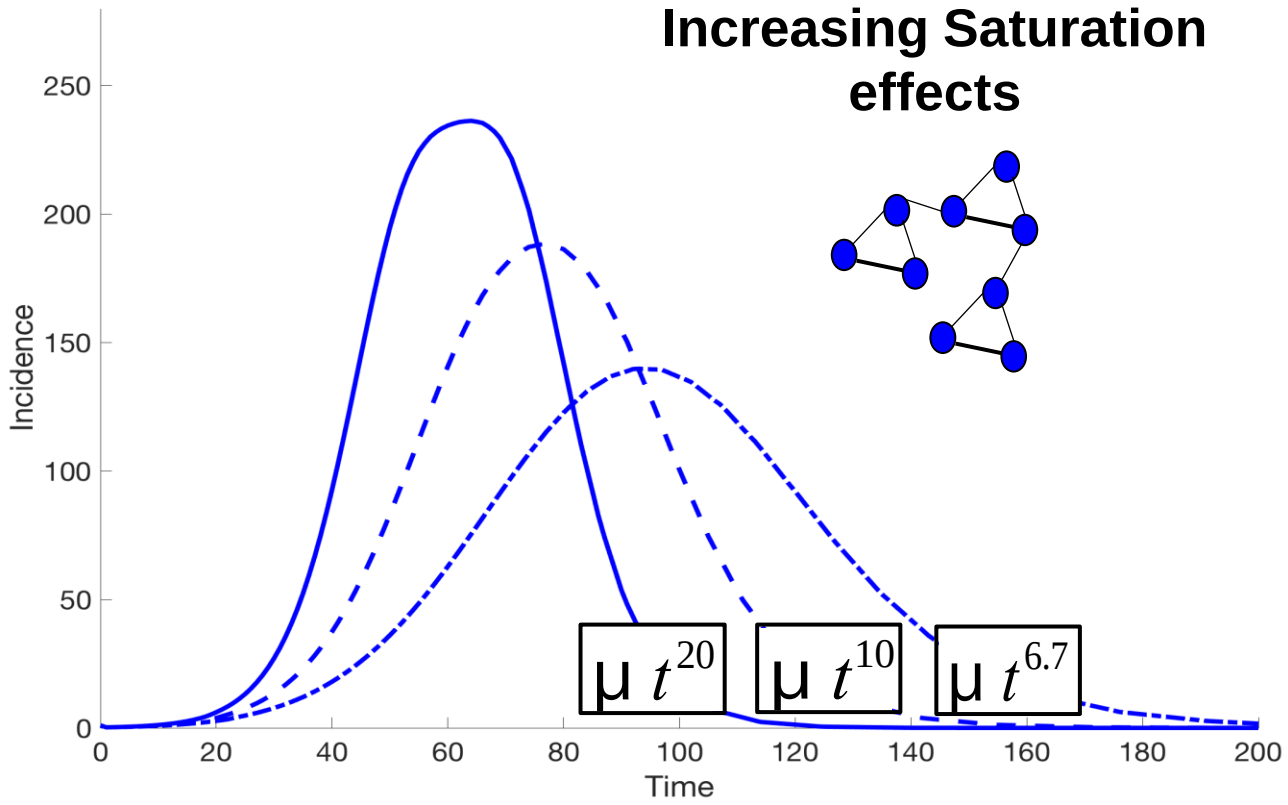


Mietchen et al. medrxiv (2024)

The complexity of epidemics: Variable epidemic growth scaling

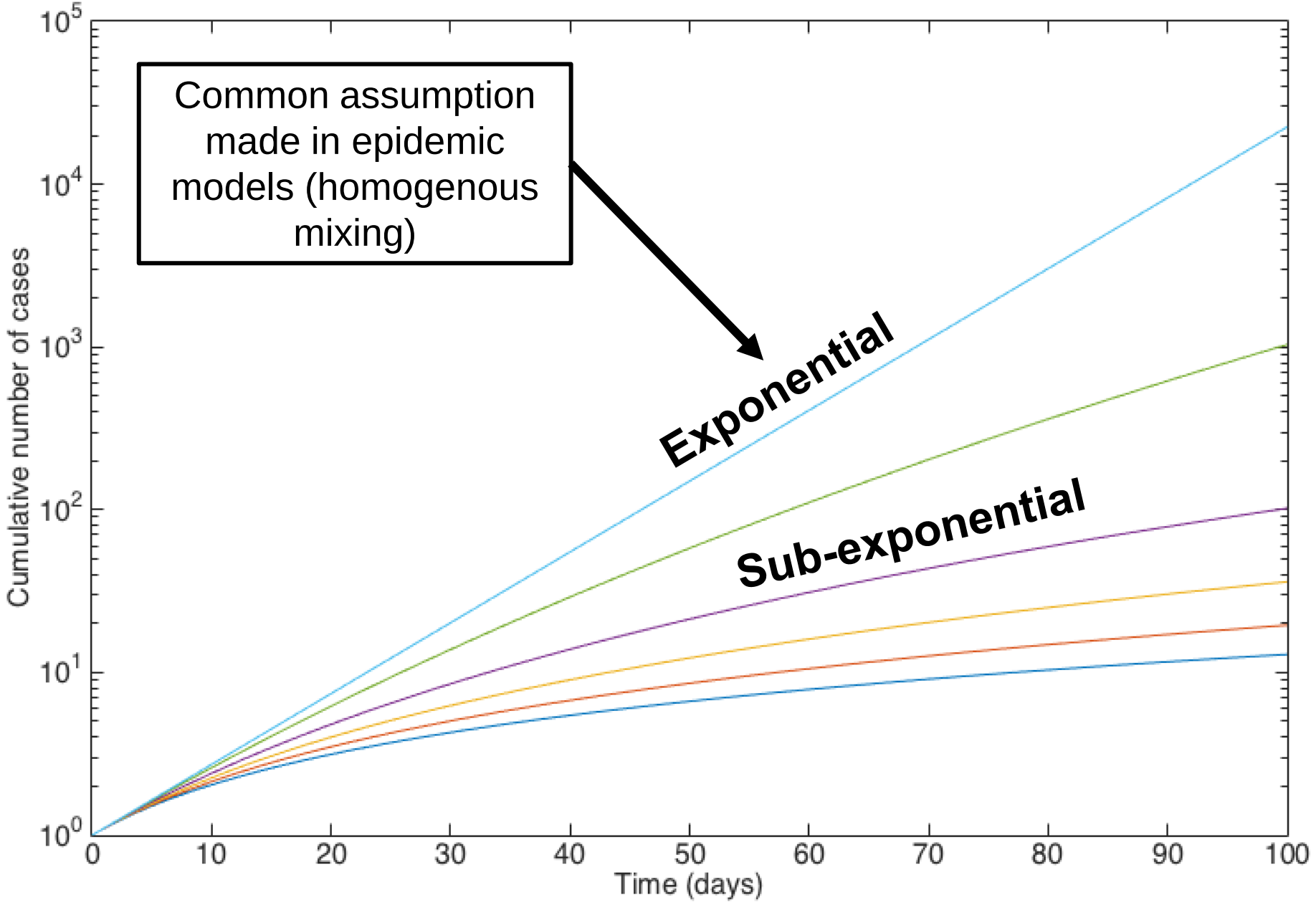
Possible reasons for different epidemic growth scaling:

- Mode of transmission
- Reactive behavior changes
- Spatial effects (clustering)
- Other individual-level heterogeneities in susceptibility and infectivity

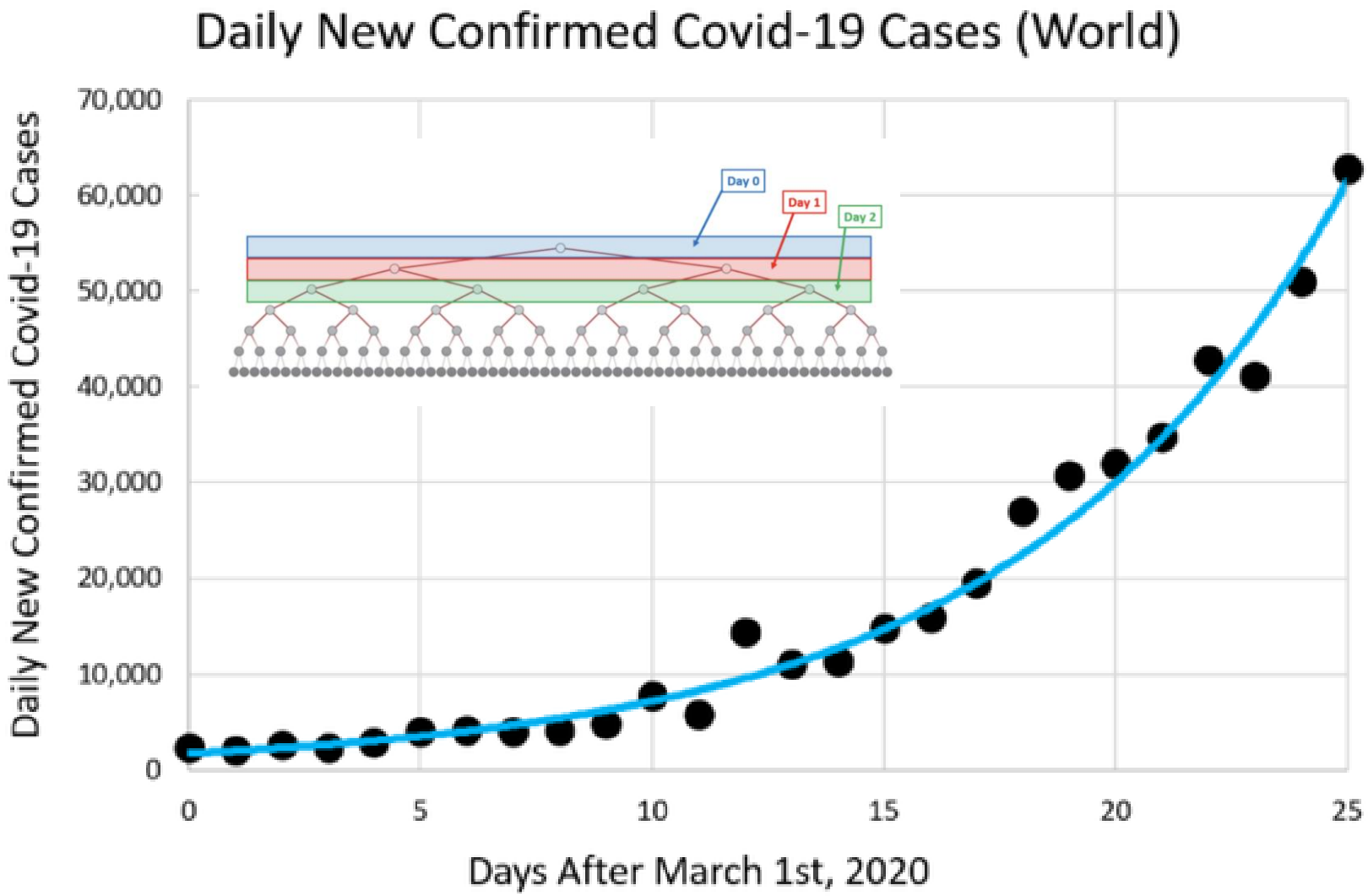


Viboud et al. Epidemics (2016); Chowell et al. JR Interface (2016)

Exponential vs. sub-exponential epidemic growth: Implications for epidemic modeling


Common assumption made in epidemic models (homogenous mixing)
Exponential
Sub-exponential
Cumulative number of cases
Time (days)
10^0
10^1
10^2
10^3
10^4
10^5
0
10
20
30
40
50
60
70
80
90
100

Example: SARS-CoV-2 Pandemic



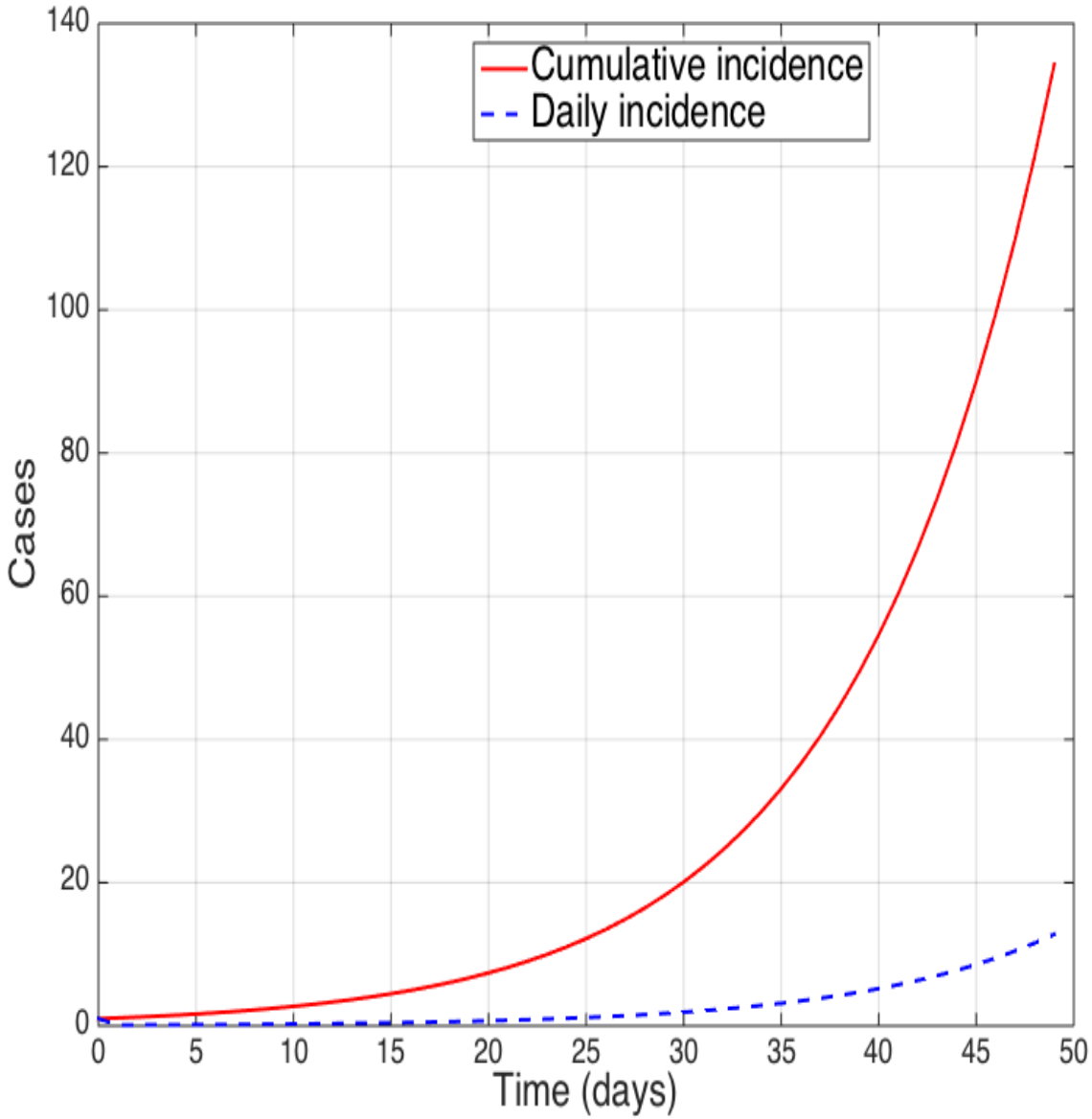
The Math of Ending the Pandemic: Exponential Growth and Decay. The New York Times. *By Dashiell Young-Saver 2021.*

Exponential growth (EXP)

$$\frac{dC(t)}{dt} = C'(t) = rC(t)$$

- $\boldsymbol{C'(t)}$: Incidence curve over time t
- $\boldsymbol{r}$: Positive parameter denoting the growth rate

Exponential growth curves



Doubling Time

Definition
The time it takes for the number of cases to double.

If the epidemic is growing exponentially fast, the doubling time (t_d) is given by:

$$t_d = \frac{\log(2)}{r}$$

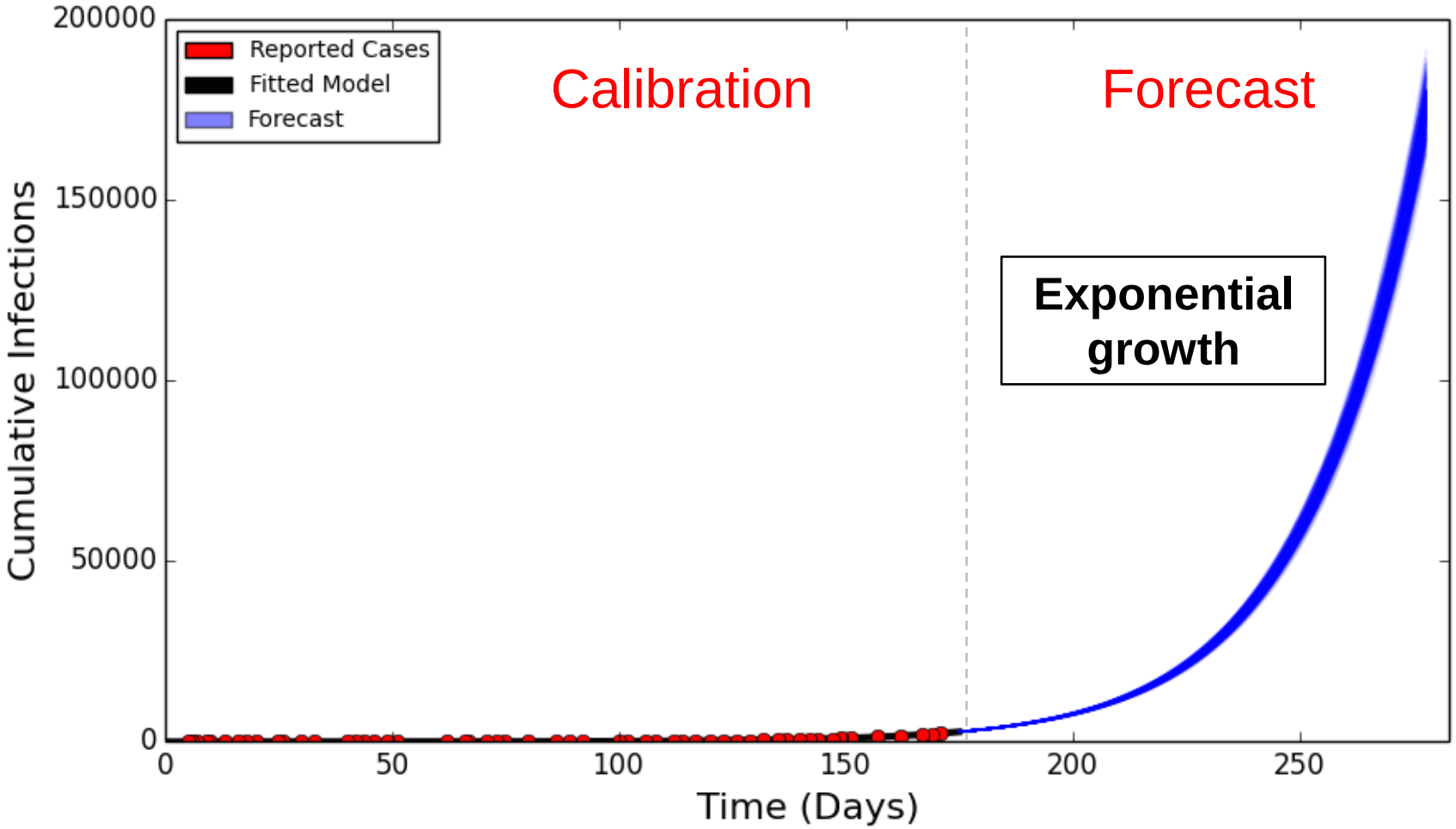
Example: Given a growth rate of $r = 0.1$ per day, what is the doubling time?

$$t_d = \frac{\log(2)}{r} = \frac{\log(2)}{0.1} \sim 7\ days$$

We conclude that the epidemic doubles in size approximately every 7 days during the exponential growth phase.

Example: “Worst-case” Scenarios

No changes in behavior or increased medical interventions.



Generalized growth model (GGM)

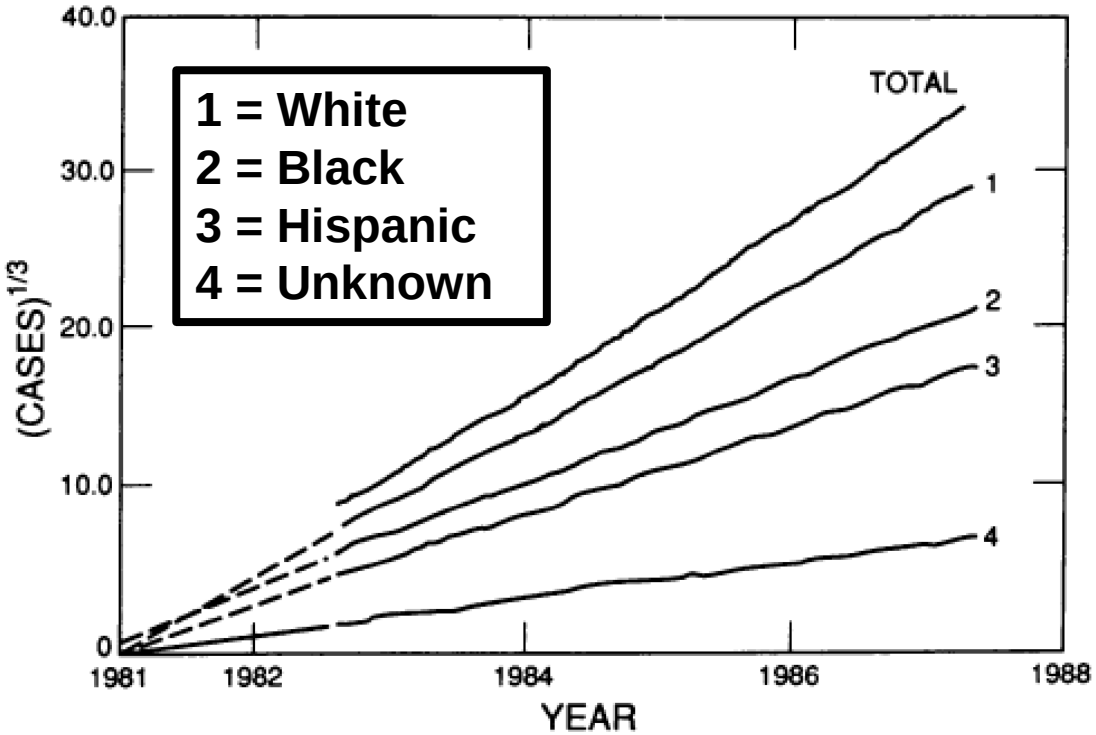
$$\frac{dC(t)}{dt} = C'(t) = rC(t)^p$$

- $\boldsymbol{C'(t)}$: Incidence curve over time t
- $\boldsymbol{r}$: Positive parameter denoting the growth rate
- $\boldsymbol{p} \in [\mathbf{0}, \mathbf{1}]$: Scaling of growth parameter

Chowell & Viboud. Infectious Disease Modeling (2016); Viboud, Simonsen, Chowell. Epidemics (2016)

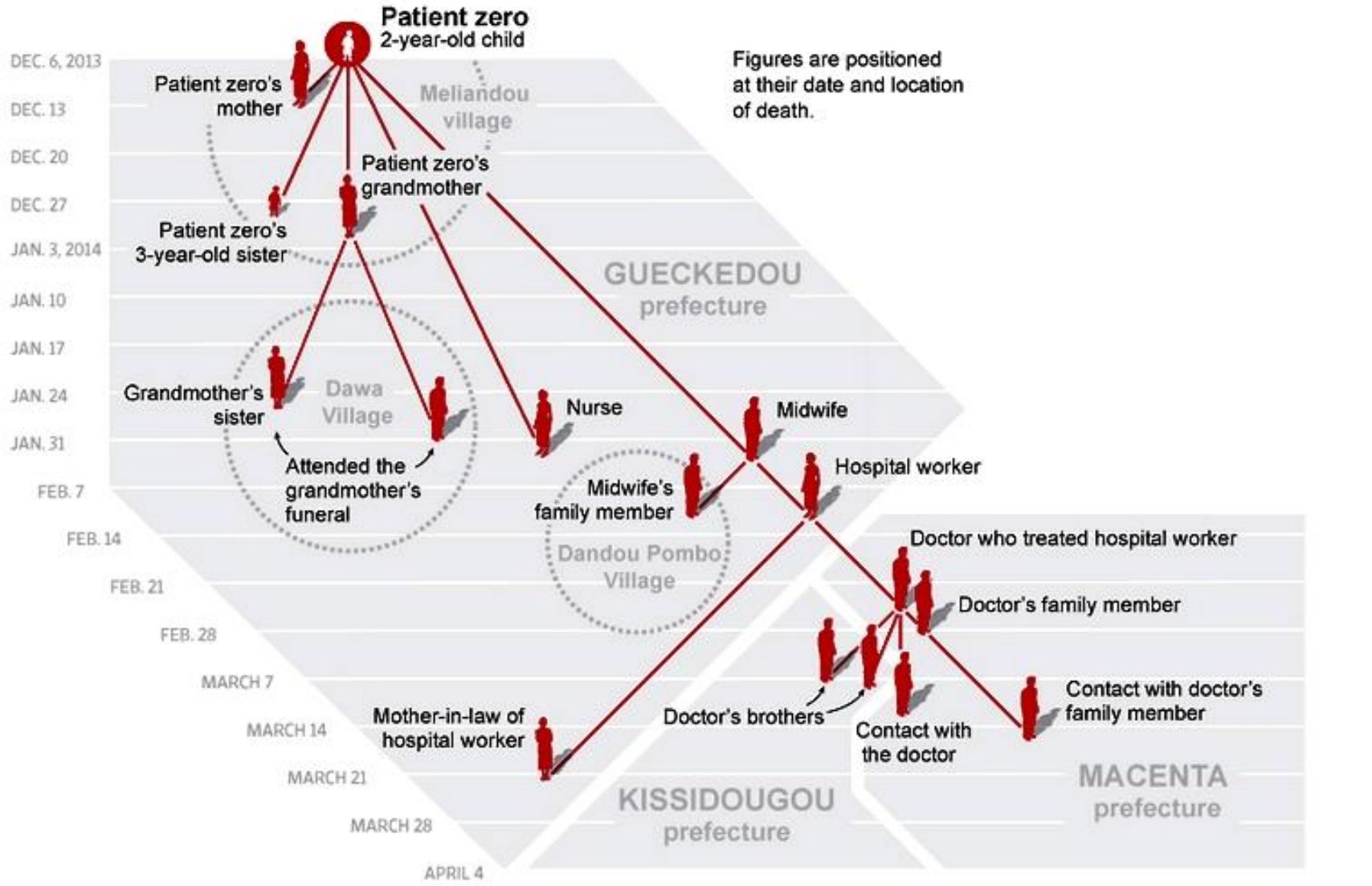
Example: The initial phase of the HIV/AIDS epidemic in the United States

https://www.pbs.org/newshour/science/america-hiv-outbreak-origins-nyc-gaetan-dugas



Colgate et al. PNAS (1989)

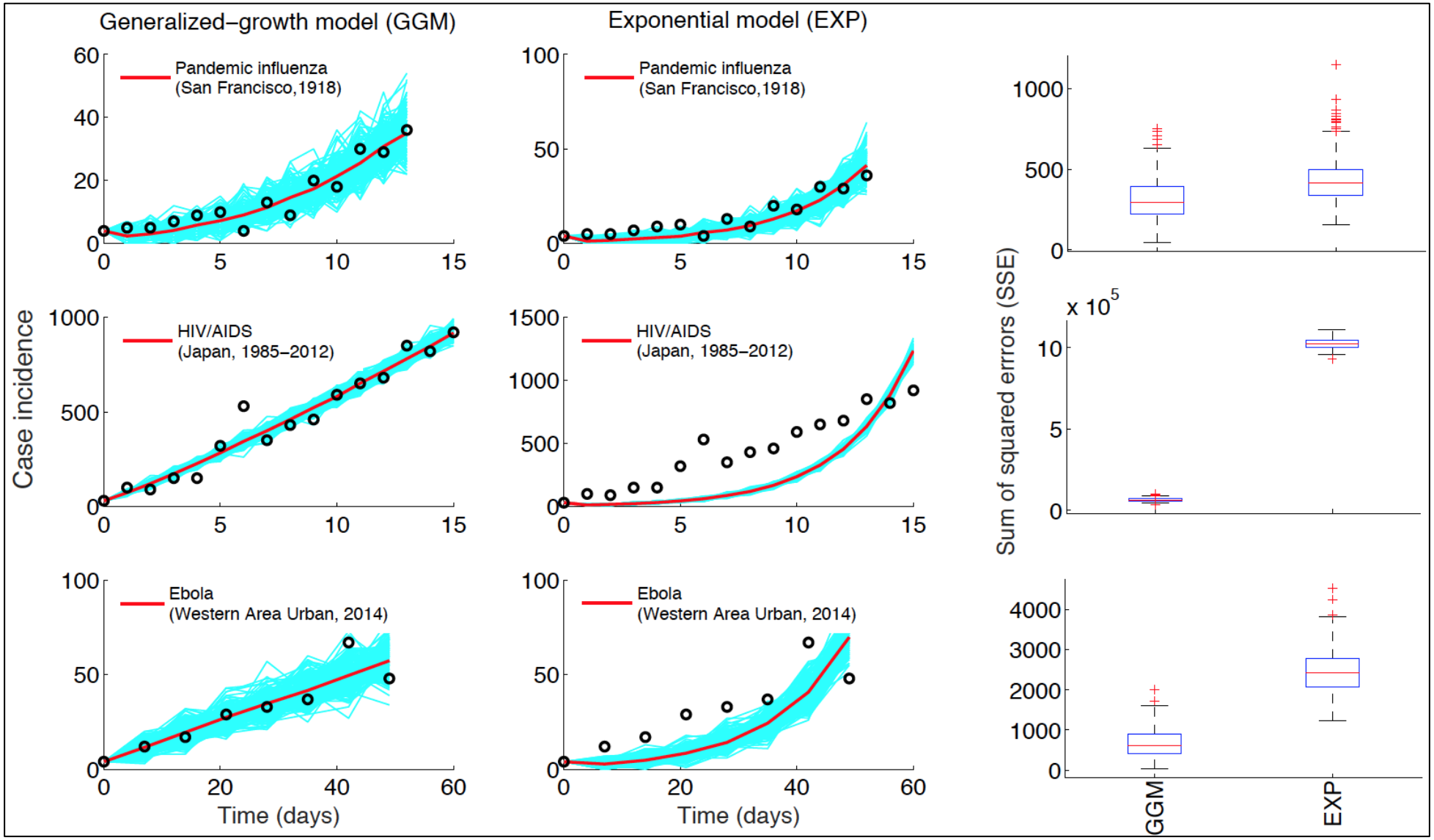
Example: Early Ebola transmission tree



Sources: New England Journal of Medicine

The Wall Street Journal

Example: Exponential vs. sub-exponential growth

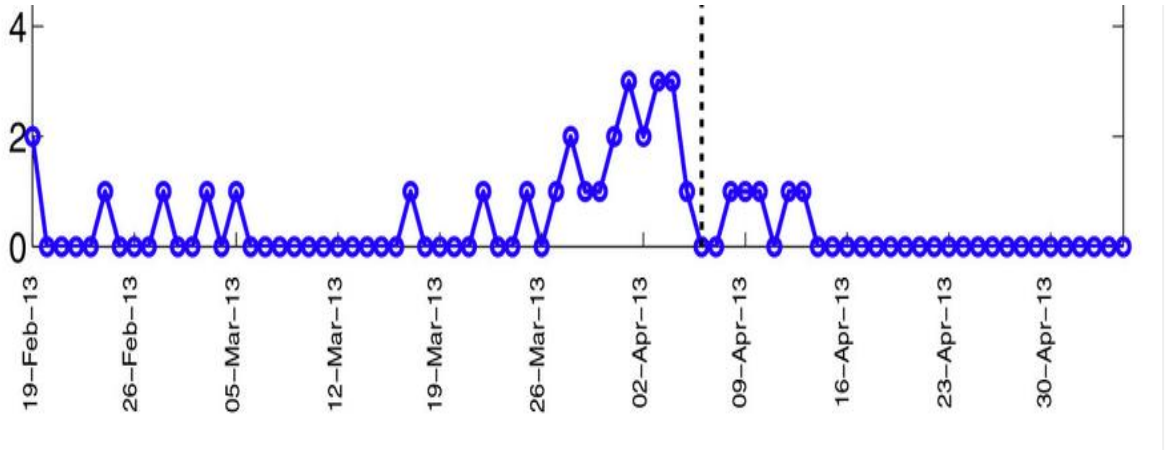


The basic reproduction number, R_0

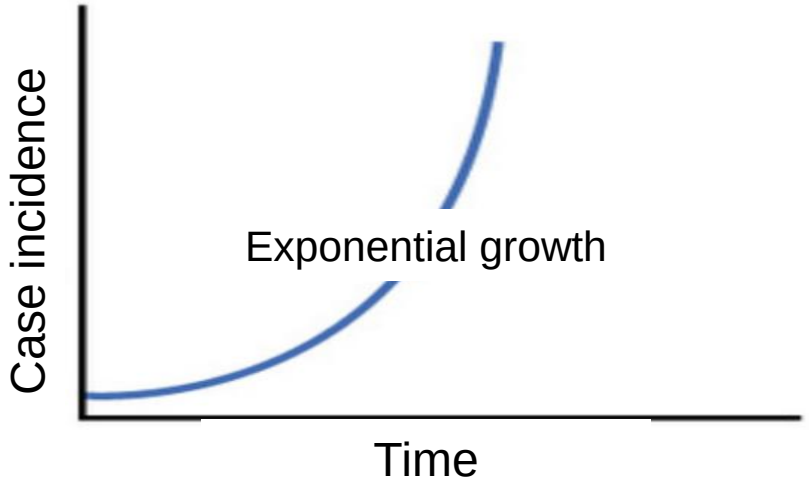
Definition

The number of secondary cases generated by a primary infectious case during its period of infectiousness in an entirely susceptible population. It is an important indicator to guide intervention strategies (i.e., higher R_0 mean stronger interventions required).

Stuttering chains (subcritical), **R_0<1**



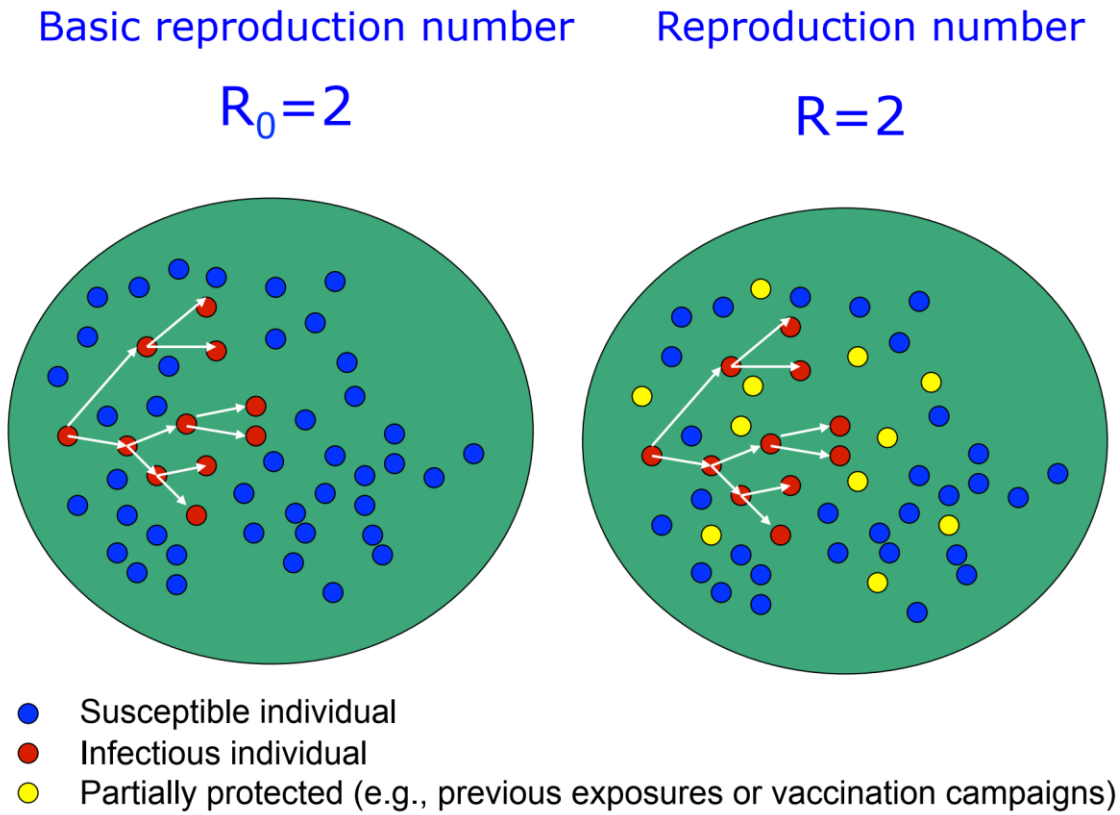
Epidemics (supercritical), **R_0>1**



Reproduction numbers

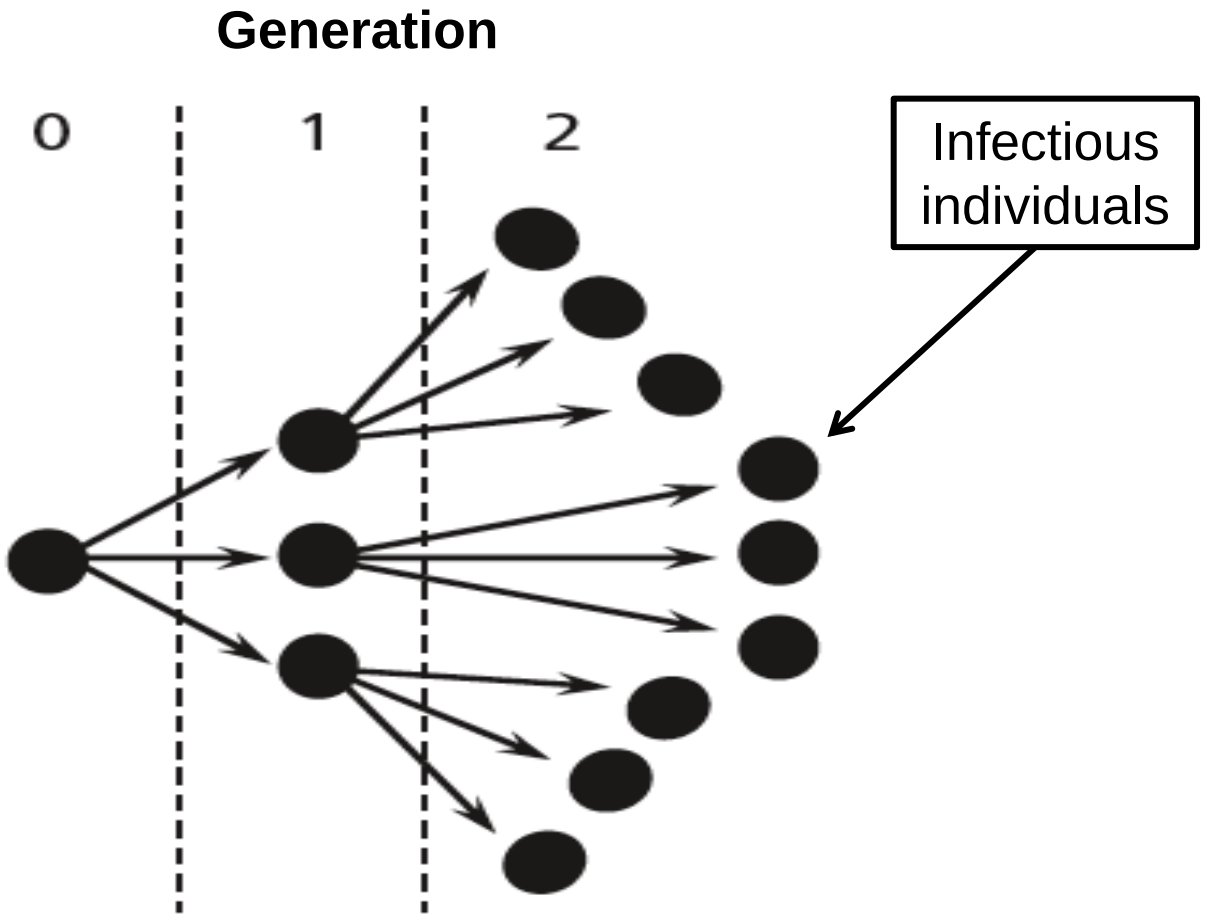
Measures transmissibility in a partially immune population because of previous exposures to infectious disease agents or vaccination campaigns.

- More practical quantity
- In a well mixed population, $R = (1-p)R_0$ where p is the fraction of population that is effectively protected against infection
  - $R \leq R_0$



Generation Time

The time between the infection of a primary case and one of its secondary cases.



Initial phase of epidemic ($R_0 = 3$)

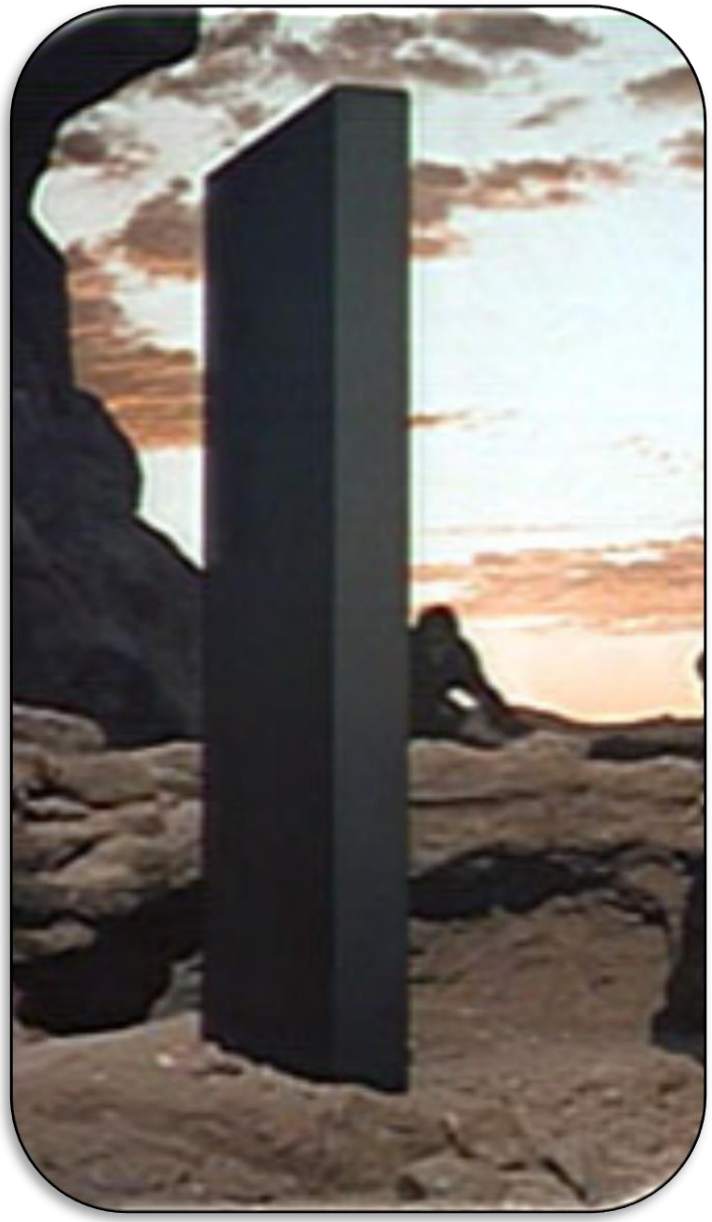
Common Epidemic models

- Epidemic models can help us characterize an epidemic process including:
  - Identify the relevant epidemiological states of the individuals and other actors (e.g., mosquitoes, rats) involved in the transmission dynamics of infection.
  - Devise equations that describe transmission mechanisms (e.g., close contact, sexual contact, airborne transmission, indirect transmission).
  - Evaluate the impact of different mitigation strategies.

Sir Ronald Ross

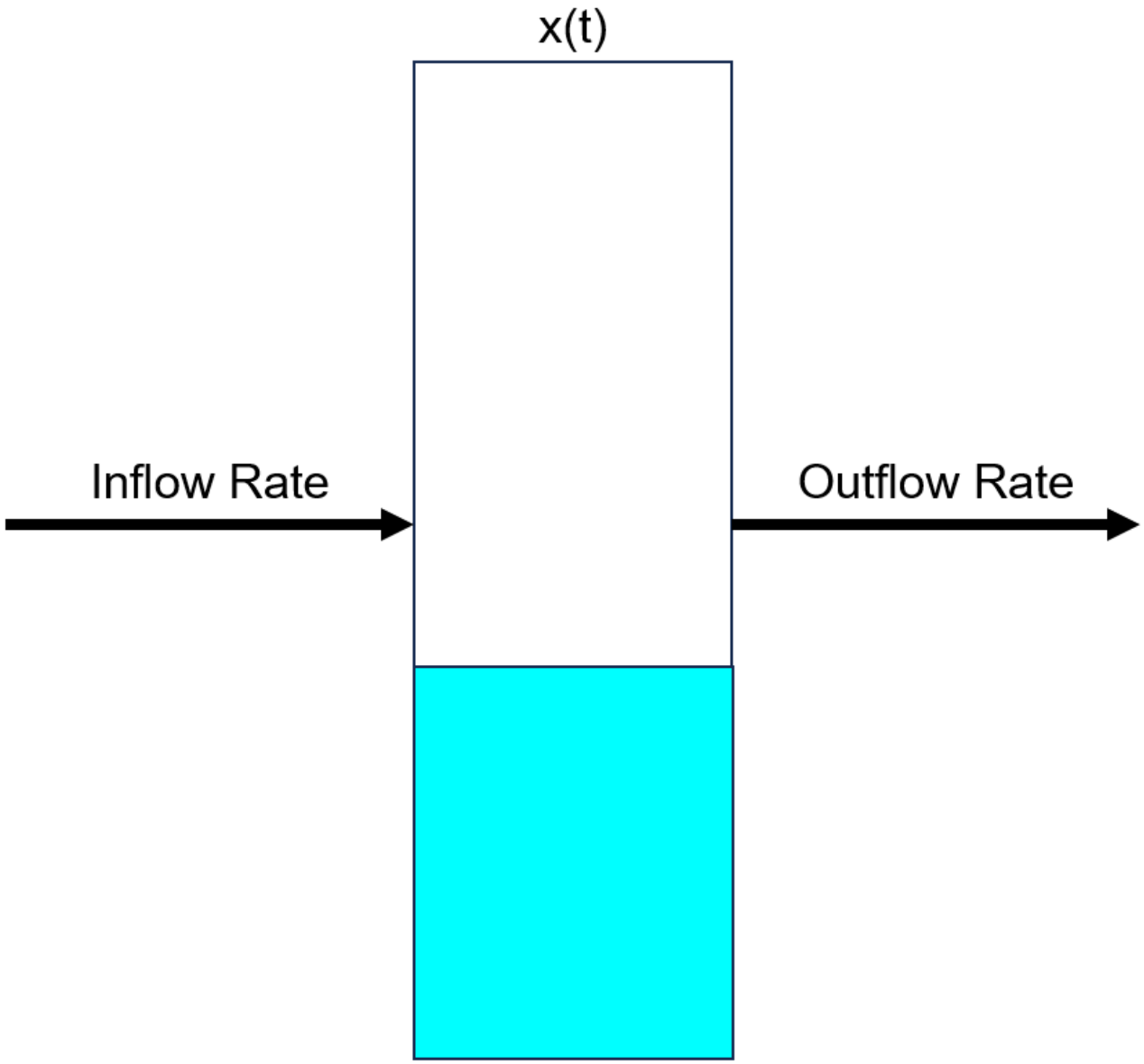
What is NOT needed:

A black box

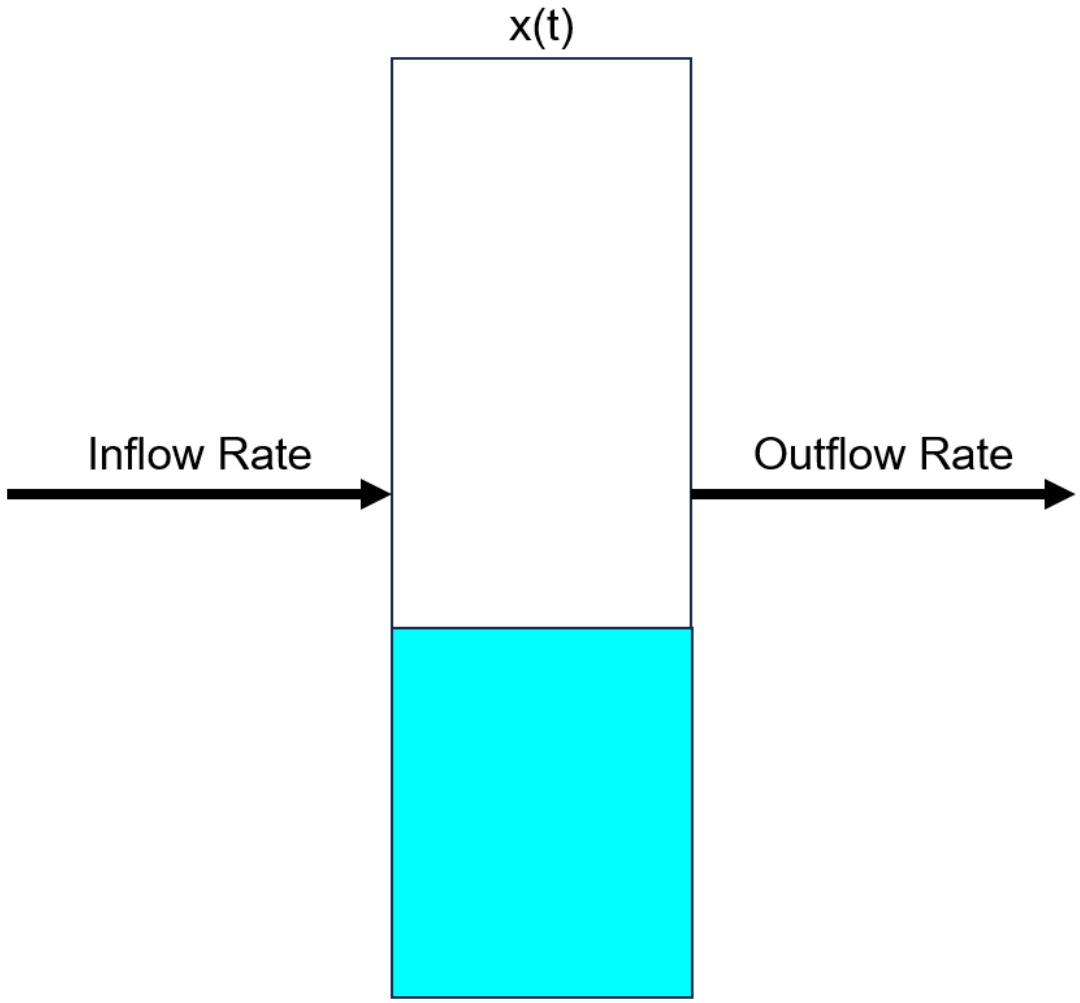


Model Building Blocks: Differential Equations

What is a differential equation?

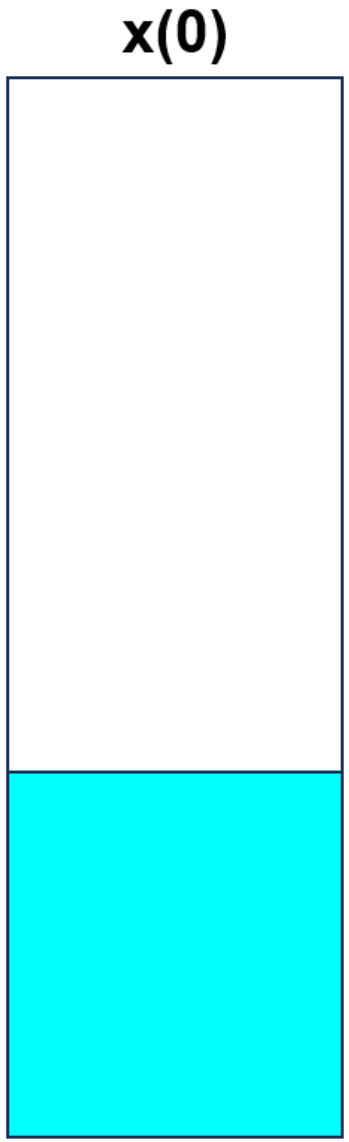


A differential equation describes the state of a variable as a function of time

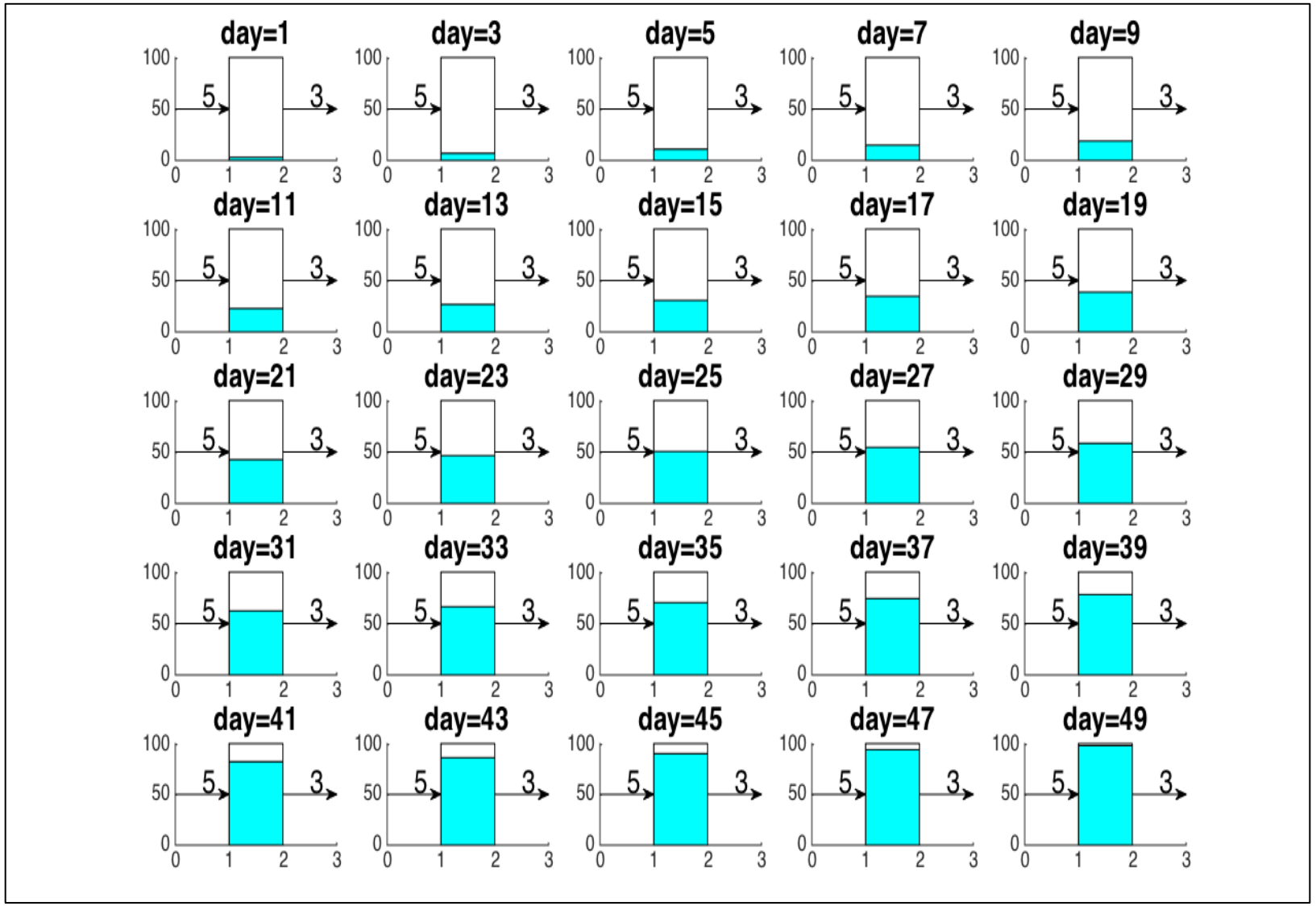


$$\frac{dx}{dt} = Rate\ of\ Change\ in\ x = Inflow\ rate(t) - Outflow\ rate(t)$$

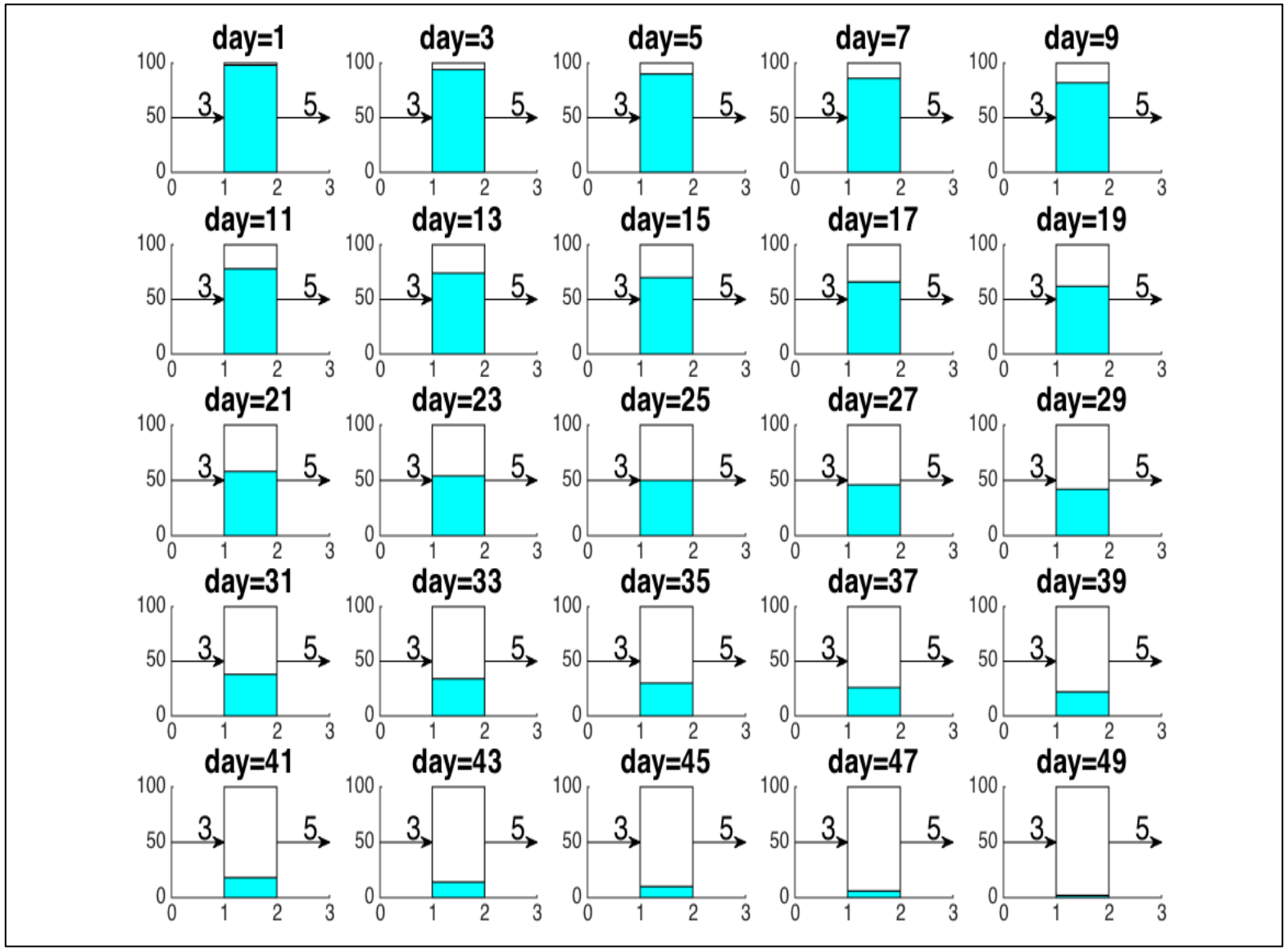
Initial state / initial conditions



Example: One differential equation with constant rates of inflow and outflow (quantity grows)



Example: One differential equation with constant rates of inflow and outflow (quantity declines)



Example: Exponential Growth

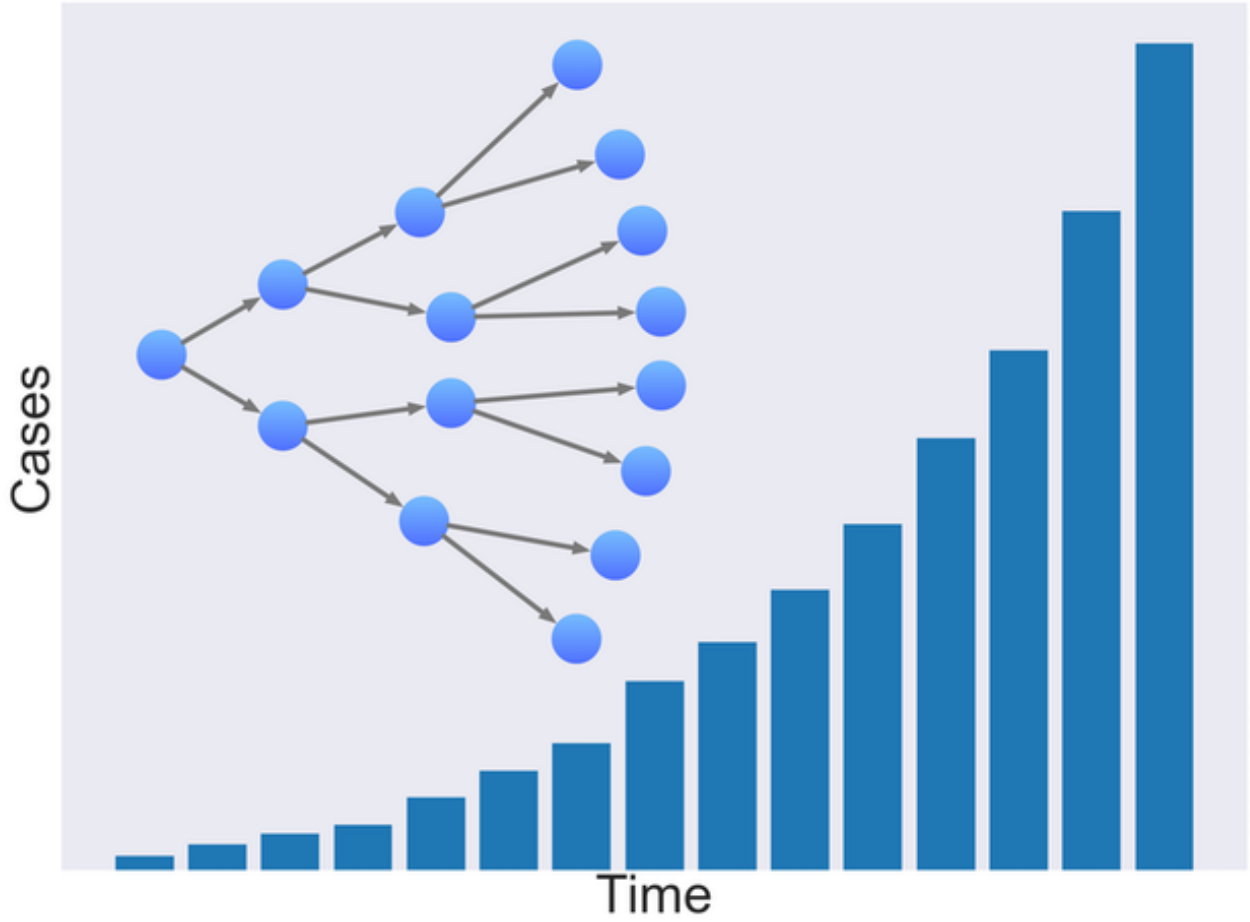
$$C(t) = C_0\, e^{rt}$$

Differential Equation

$$\frac{dC(t)}{dt} = C'(t) = rC(t)$$

$$r = Growth\ rate\left(\frac{1}{time}\right)$$

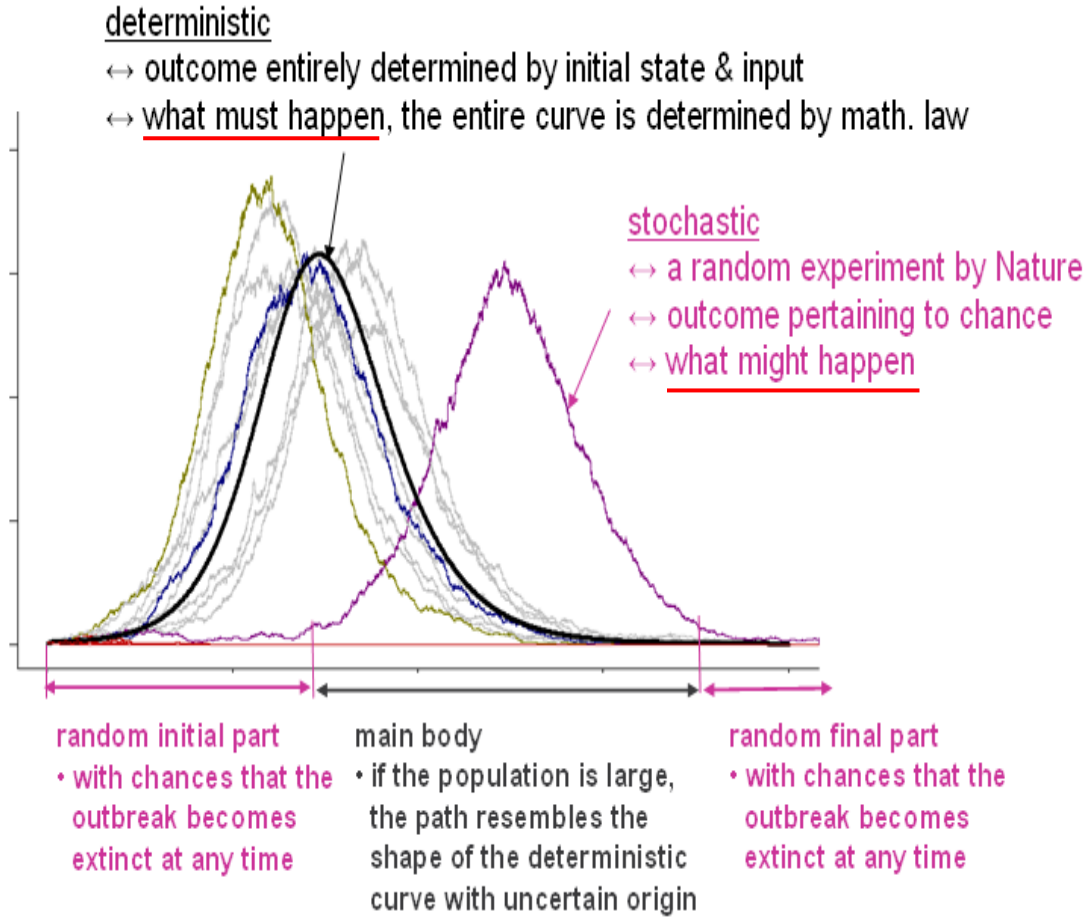
- $\boldsymbol{C'(t)}$: Incidence curve over time t
- $\boldsymbol{C_0}$: Initial number of cases



Model Types: Deterministic vs. Stochastic

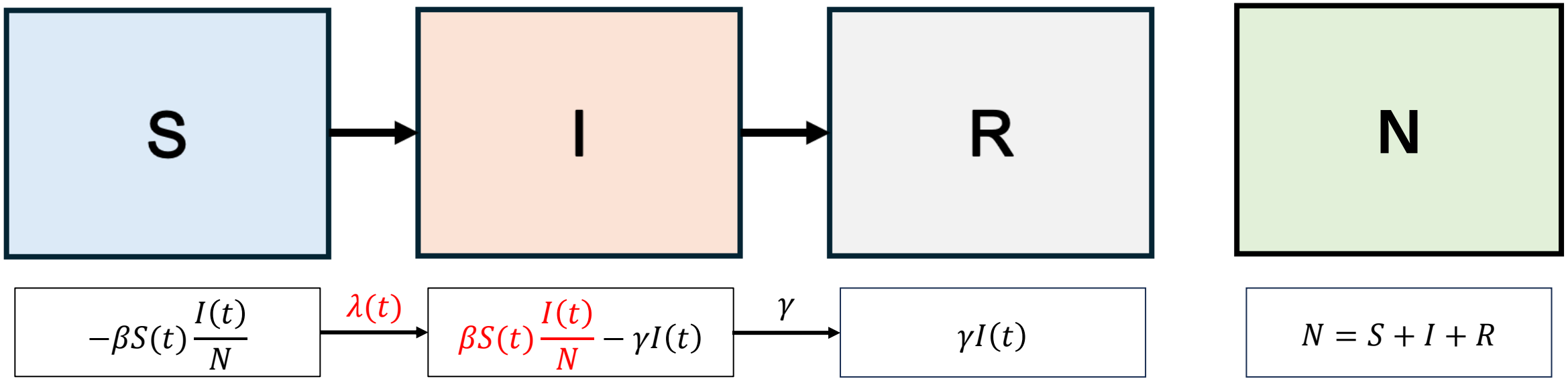
Deterministic vs. stochastic models

- Deterministic models produce the same exact results for a particular set of parameters and initial conditions.
  - Rely on the average rates of the process
  - Applicable to large populations
- Stochastic models account for certain levels of unpredictability or randomness.
  - More appropriate to deal with small populations
  - Each realization of the stochastic model is different



Example Model 1: Susceptible-Infectious-Recovered (SIR)

A simple epidemic model (SIR)



Parameters (Θ): β (transmission rate), N (population size), γ (removal/recovery rate)

State Variables ($\dot{x}_i$): $\boldsymbol{S}$usceptible, $\boldsymbol{I}$nfectious, $\boldsymbol{R}$ecovered

Note: Negative values indicate moving out of a compartment and positive values indicate inward movement

Assumptions: SIR model (Deterministic)

Homogenous Mixing

"For a typical susceptible individual, the force of infection is proportional to % of individuals 'currently infectious', $\frac{I(t)}{N}$, where N is the population size."

Complete Susceptibility

The entire population is equally susceptible to the infectious agent.

Timing of Infection

Individuals become infectious instantaneously after exposure to the infectious agent.

SIR model (Deterministic)

Homogenous Mixing

"For a typical susceptible individual, the force of infection is proportional to % of individuals 'currently infectious', $\frac{I(t)}{N}$, where N is the population size."

Incident cases at time *t* = Incidence rate ($\frac{\beta I(t)}{N}$) X Susceptible population (*S(t)*)

Change of prevalence: $$dI/dt = \beta \frac{SI}{N} - \gamma I$$

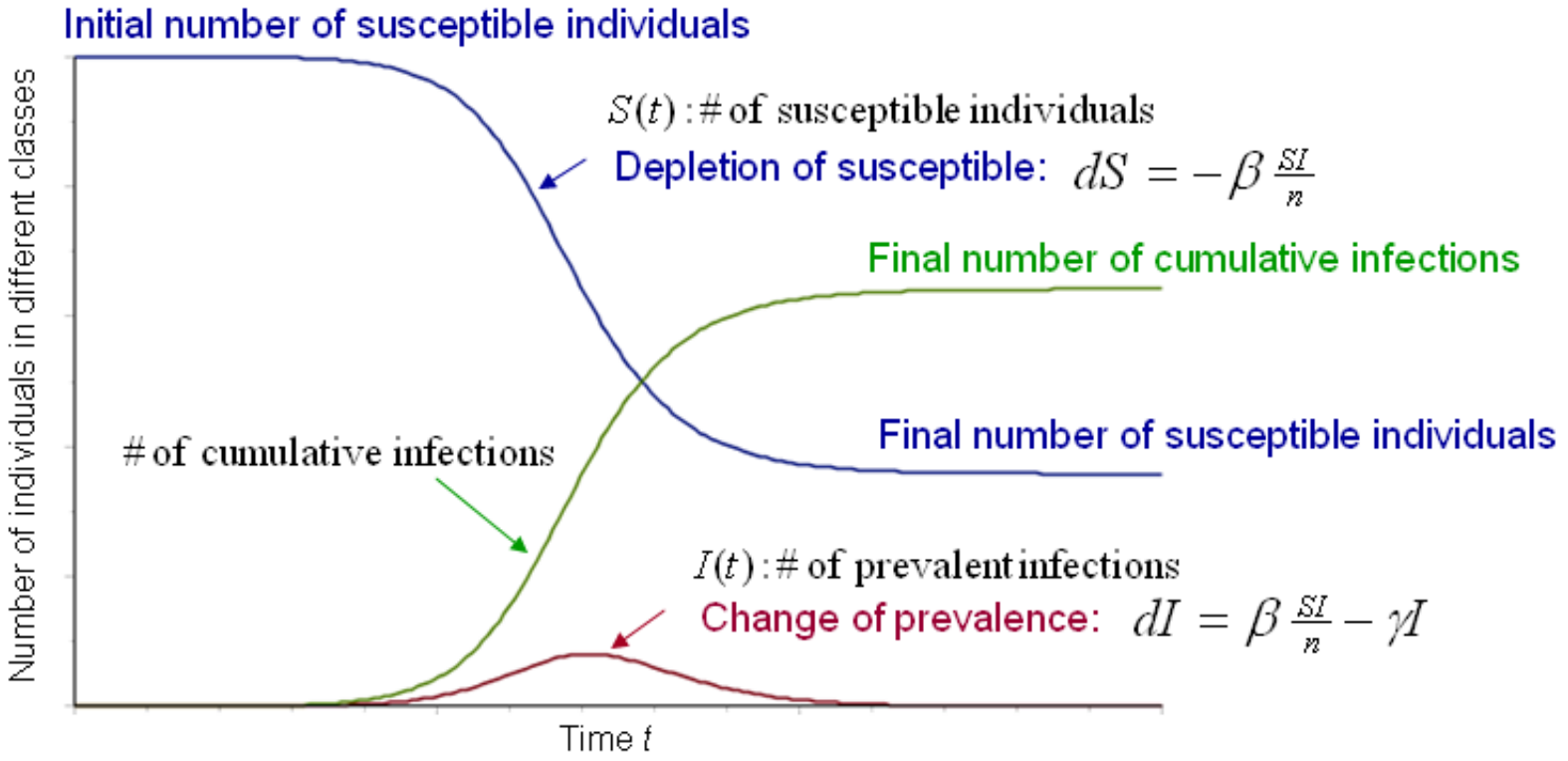
Depletion of susceptible: $$dS/dt = -\beta \frac{SI}{N}$$

Assuming constant *n*, at time *t*

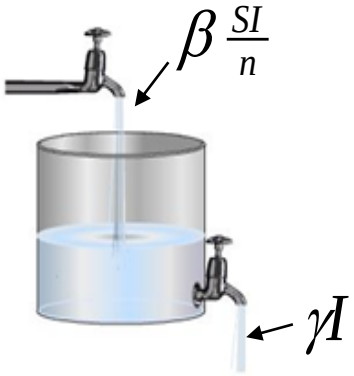
$$R(t) = N - S(t) - I(t)$$

individuals **recovered with immunity**.

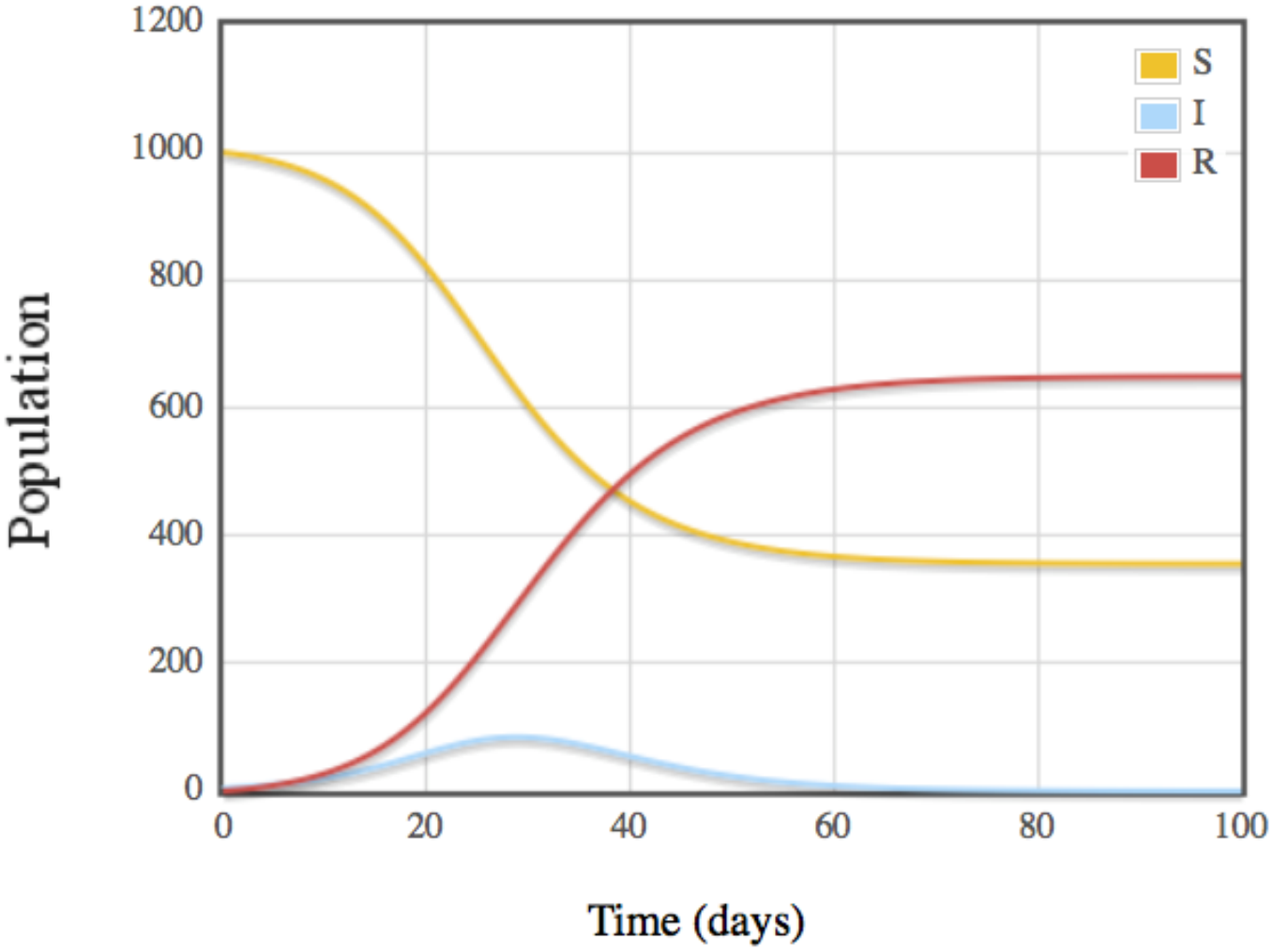
SIR model (Deterministic)



Implicitly implied:
Population size is constant

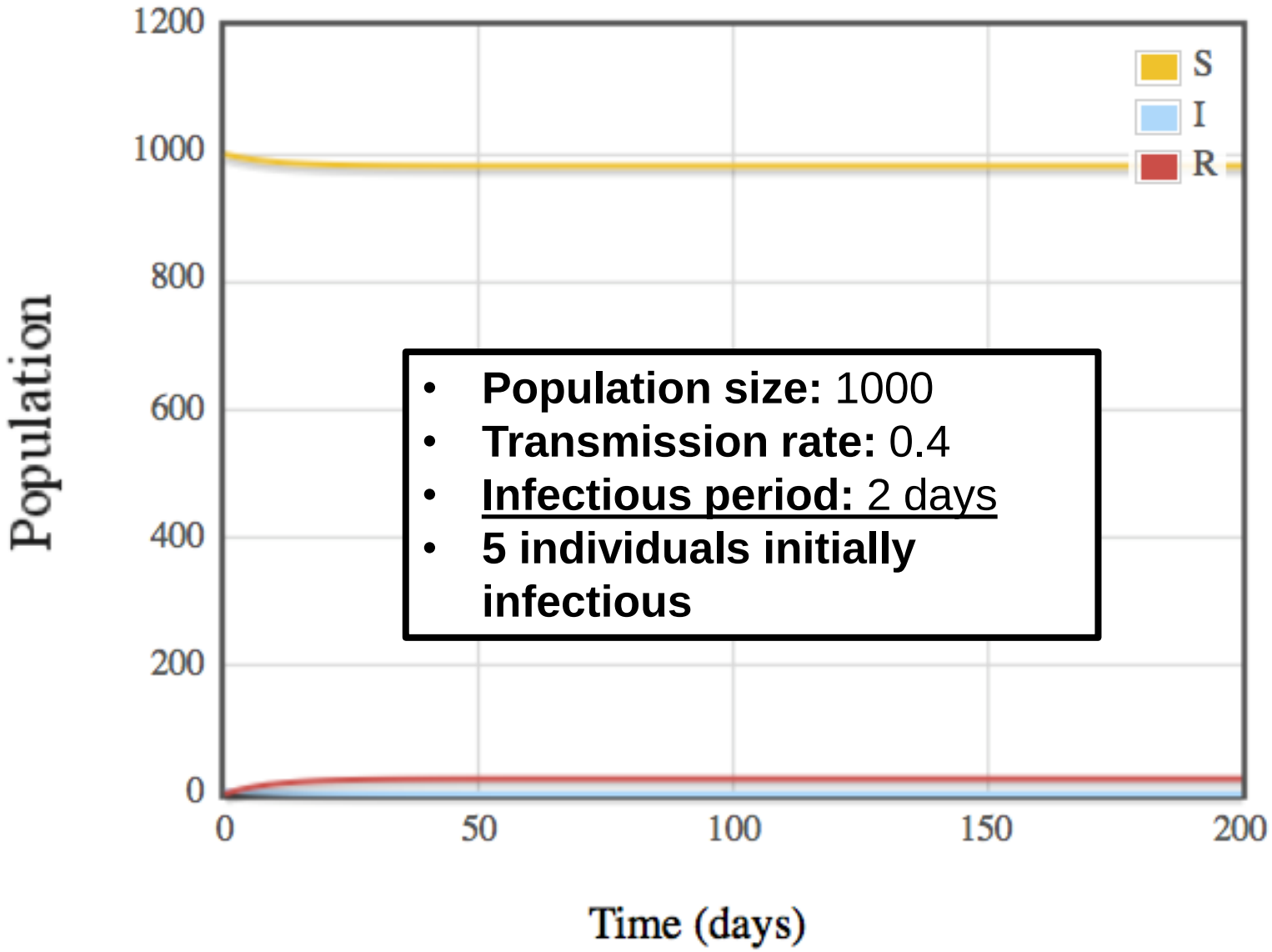


Example: SIR mathematical model results



- **Population size:**1000
- **Transmission rate:** 0.4
- **Infectious period:** 4 days
- **5 individuals initially infectious**

Example: SIR model (2-day infectious period)



- **Population size:** 1000
- **Transmission rate:** 0.4
- **Infectious period:** 2 days
- **5 individuals initially infectious**

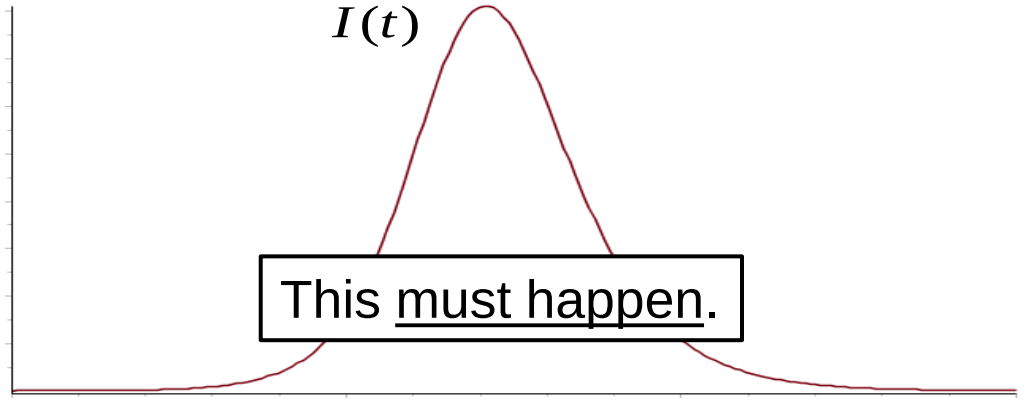
Beta: 0.4
Gamma: 0.5
Nu: 0

Initial
Susceptible: 1000
Infected: 5
Recovered: 0
Days: 200
Submit

Deterministic vs. Stochastic SIR Model

Per the previous deterministic SIR model:

if $R_0 = \frac{\beta}{\gamma} > 1$ ⟶

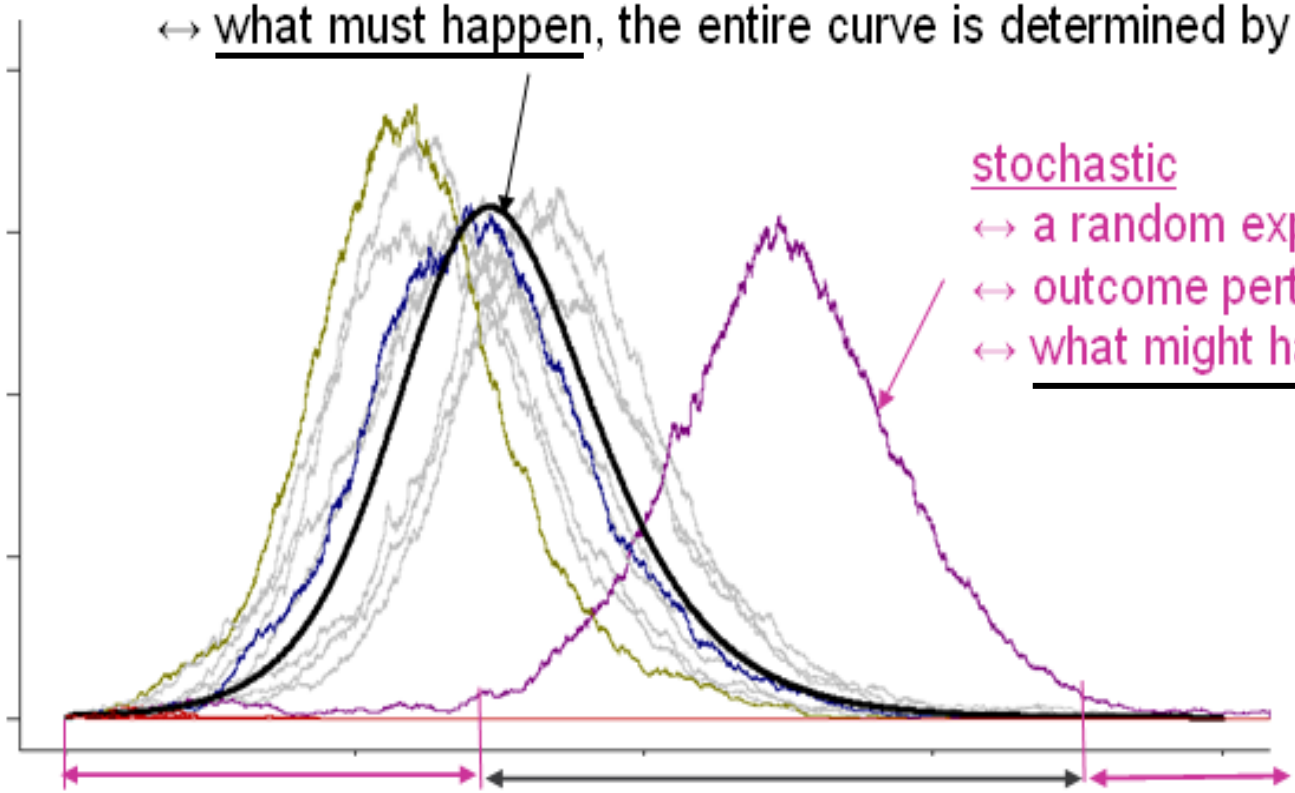


Computer simulated scenarios (Stochastic)	
One repeatedly seeds an initial infectious individual in a large and closed population, recovery occurs with and lifelong immunity, with parameters (β, γ, N) identical to the deterministic SIR model.	In many tries, there is no outbreak except for a very few initial transmissions. So, an outbreak is not longer a "must happen" situation no matter how contagious is the disease.

deterministic

↔ outcome entirely determined by initial state & input

↔ what must happen, the entire curve is determined by math. law



stochastic

↔ a random experiment by Nature

↔ outcome pertaining to chance

↔ what might happen

random initial part

- with chances that the outbreak becomes extinct at any time

main body

- if the population is large, the path resembles the shape of the deterministic curve with uncertain origin

random final part

- with chances that the outbreak becomes extinct at any time

Basic Reproduction Number (R_0) for Deterministic SIR

Definition

The average number of secondary infections a single infectious individual could produce when seeded in a very large susceptible population.

R_0 = (transmission rate) x (infectious period) $= \beta/\gamma$

- Reproduction number at the beginning of an outbreak
- The reproduction number during the outbreak is always smaller.
  - Even without control, the effective reproduction number ($R_t = \frac{R_0 S(t)}{N}$) is reduced through the depletion of the susceptible population

Example

For a disease like influenza, the mean infectious period is about 4 days and the mean transmission rate is about 0.4 new infections per day per infected person. Hence,

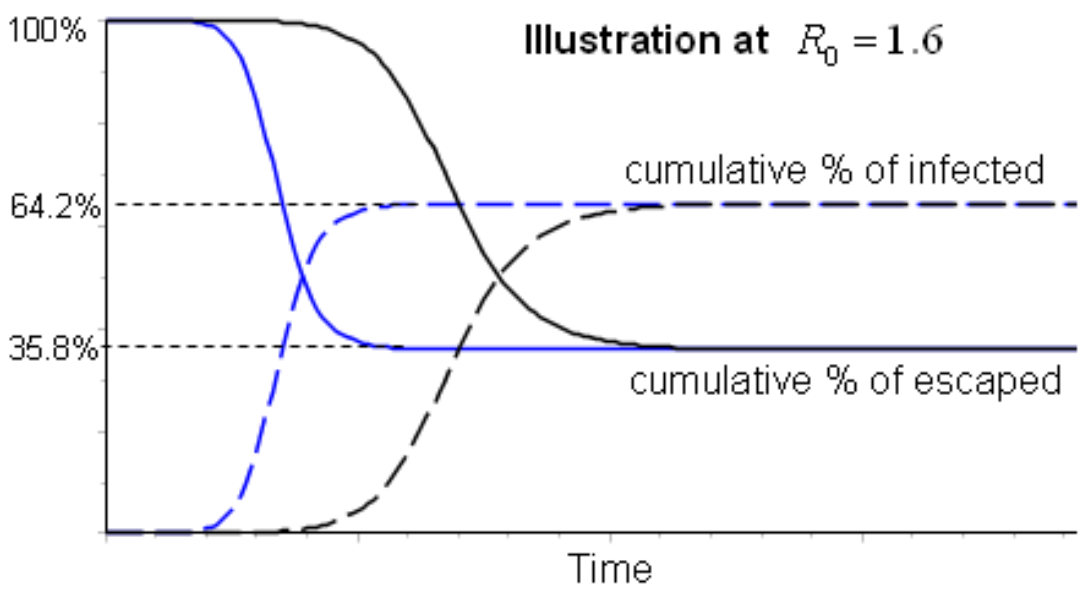
R_0 = 0.4 X 4 = 1.6

General conclusion: An outbreak is possible if $R_0 > 1$.

R_0 and final epidemic size (Deterministic SIR)

General conclusion: There is always a proportion "susceptibles" that escape infection (i.e. the final % of infected < 100%), and this proportion is predicted by R_0 before the outbreak.

$R_0 \leq 1.0$	final % infected $\rightarrow$ 0% as $n \rightarrow \infty$
$R_0 = 1.2$	final % infected = 31.4%
$R_0 = 1.4$	final % infected = 51.1%
$R_0 = 1.6$	final % infected = 64.2%
$R_0 = 2.0$	final % infected = 79.7%
$R_0 = 2.5$	final % infected = 89.3%
$R_0 = 3.0$	final % infected = 94.0%

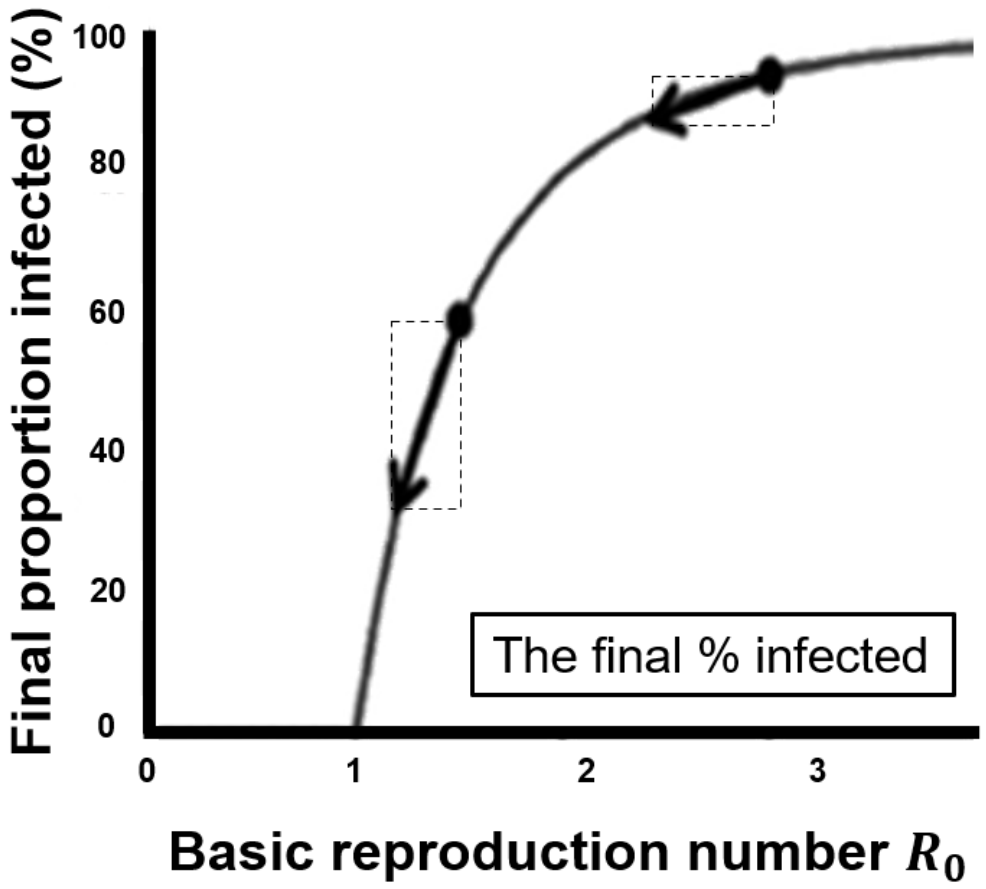


Final size formula:

$$z = 1 - \exp(-Ro * z)$$

R_0 and final epidemic size (Deterministic SIR)

$R_0 = \beta/\gamma$ is essential. Different values of β and γ only affect the time-course towards the final outcome.



R_0 = 3.0: 20% reduction of R_0
6% reduction of the final proportion of infected

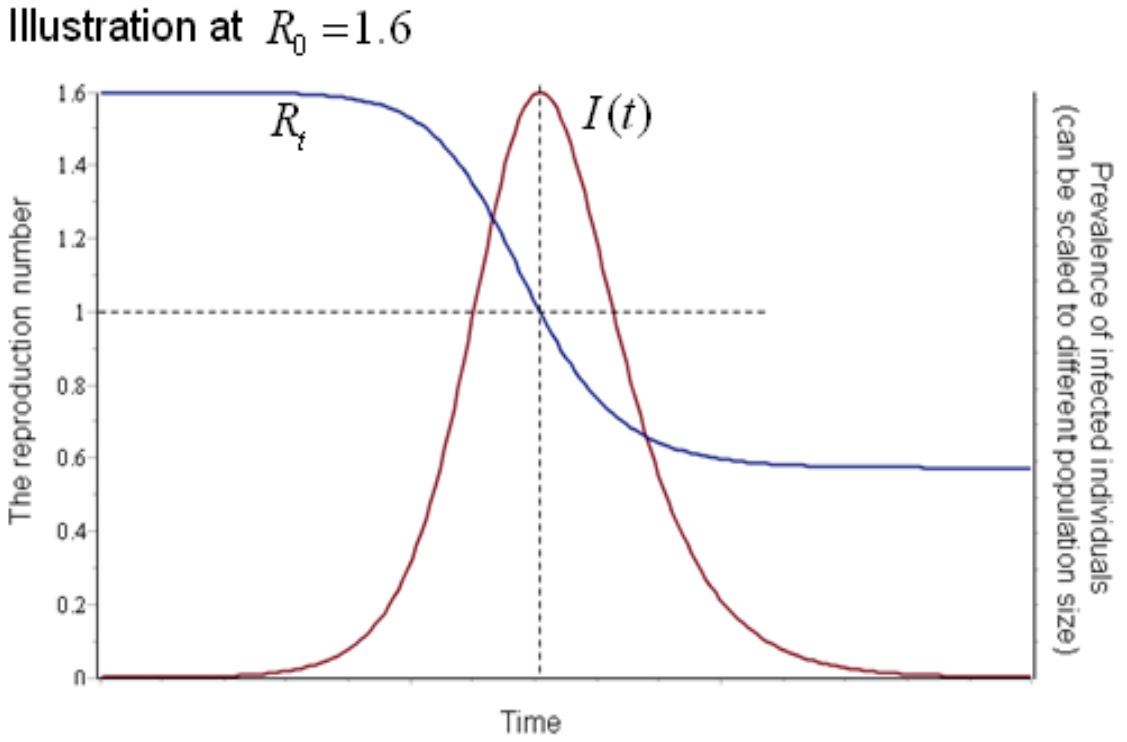
R_0 = 1.5: 20% reduction of R_0
27% reduction of the final proportion of infected

Vaccination coverage: Assume a vaccine is available and is less than perfect with effectiveness $\Phi \leq 1$, define a quantity $V_c = \frac{R_0 - 1}{R_0 \Phi}$. If vaccine coverage is $\geq V_c$, the reproduction number is controlled to < 1. We say then can say that the whole population has "herd immunity". (If Φ is too small, then $V_c \geq 1$. Herd immunity is not achievable).

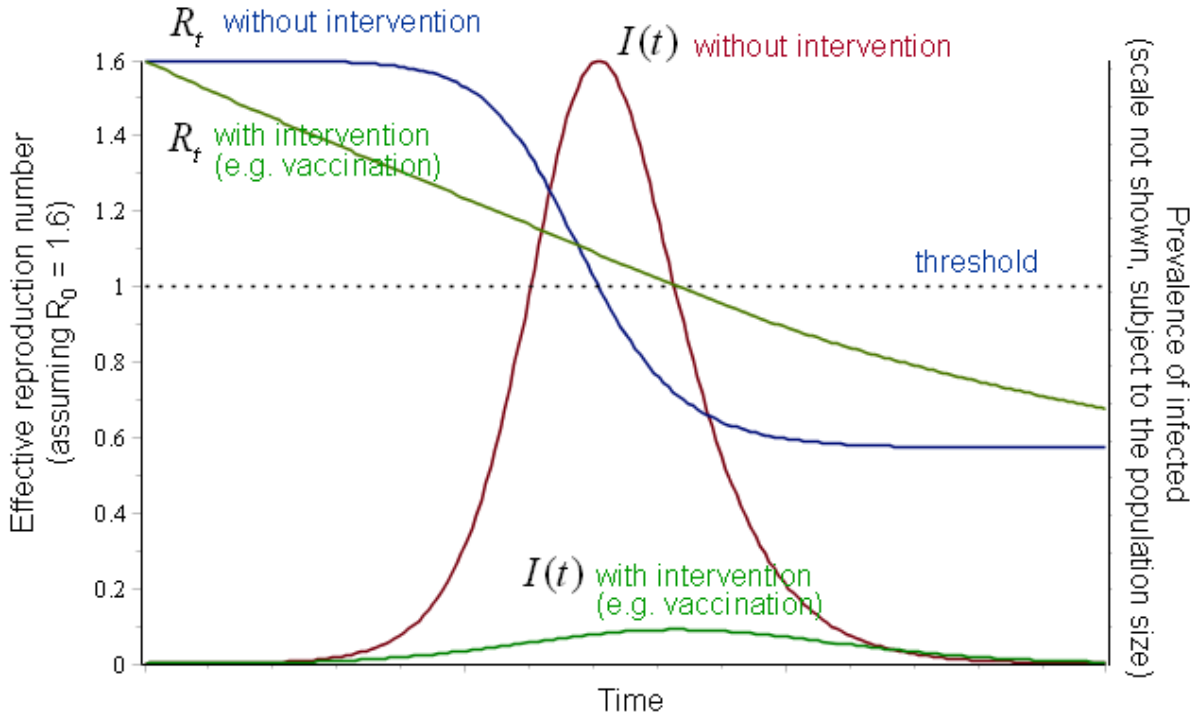
Impact of interventions

General conclusion: The prevalence of infected but not yet recovered individuals, $I(t)$, peaks when the effective reproduction number $R_t = \frac{R_0 S(t)}{N}$ is reduced to cross $R_t = 1$.

No Interventions



How interventions change the course of the outbreak



Impact of interventions

Qualitatively speaking, if a public health prevention action (e.g., vaccination) started early, it will:

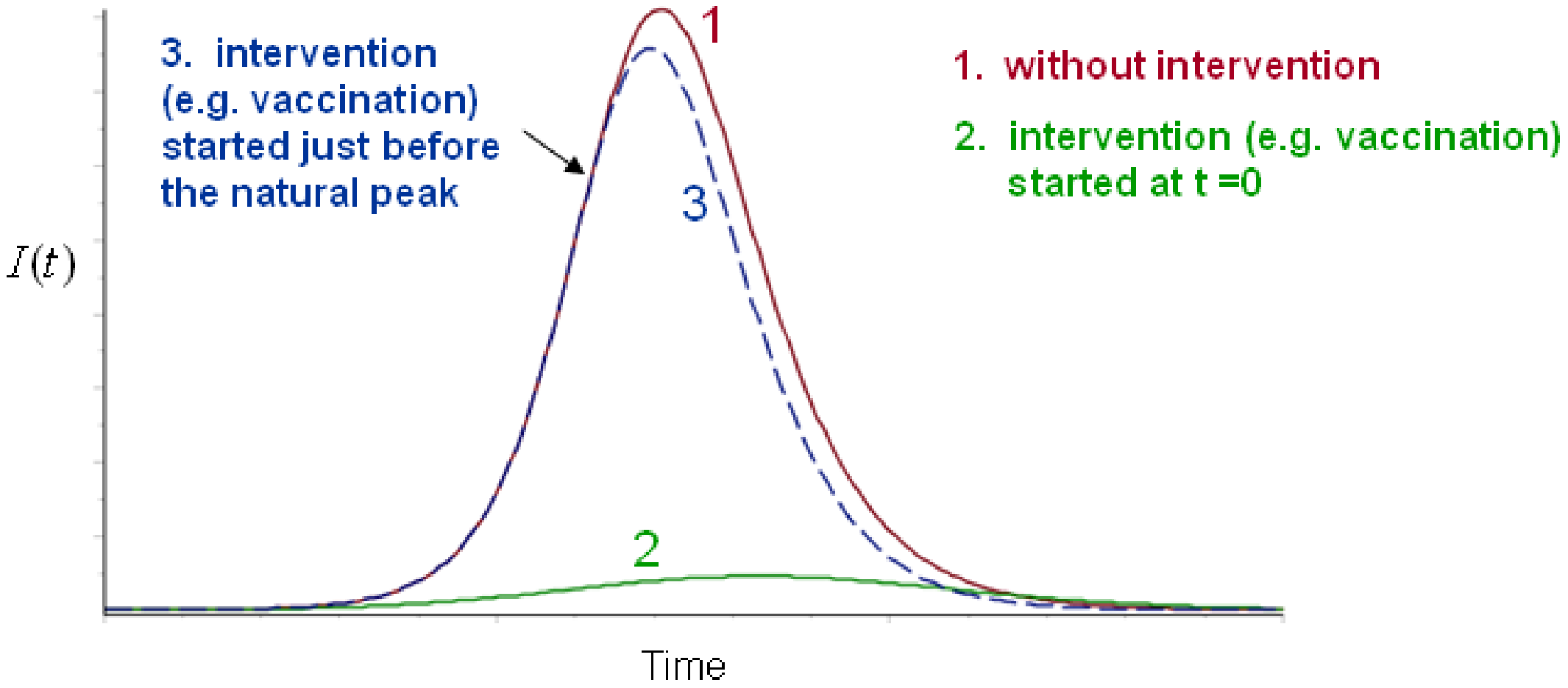
Slow down the initial growth of the “epidemic curve”, because of the further reduction of the effective reproduction number R_t

Reduce the peak level of the “epidemic curve” and the final number of totally infected

Delay the peak time of the outbreak or epidemic.

Timing of interventions is key

General conclusion: The timing of interventions can change the course of the outbreak



Timing of interventions is key

Qualitatively speaking,

1. Early intervention (i.e. before the “exponential growth” of an outbreak) may avoid or minimize the outbreak: Does successful intervention mean “shooting at own foot”?
   - If the outbreak is avoided, are we “crying wolves”? (as well as lawsuits for adverse effects …)
   - If not avoided but minimized, how do we prove that intervention was effective? Without comparison with the worse scenario (curve 1), what happened (curve 2) looks as if there is a bell-shaped epi-curve anyway !
2. Once exponential growth has been observed in the data (i.e., with delays in reporting etc.) the effectiveness of prevention is minimal … but …

Conclusions hold in more complex (deterministic) SIR-type models

$$dI/dt = \beta \frac{SI}{N} - \gamma I$$

$$dS/dt = -\beta \frac{SI}{N}$$

Assuming a constant n, at time t, there are

$$R(t) = N - S(t) - I(t)$$

individuals recovered with immunity.

seems over-simplified. But the conclusions shown in previous slides are general, and also true in:

- more complicated models such as incorporating a latent (infected but not infectious) period
- recovery rate not constant (thus not a simple γ as a parameter),
- large populations regardless the population size,
- different time scales,
- and to some extent, some non-homogeneous mixing patterns.

Conclusions hold in more complex (deterministic) SIR-type models

In the more complicated models, R_0 may be not expressed as $\frac{\beta}{\gamma}$ **but** as long as keeping the definition as

> "the expected number of secondary infections generated by a typical infectious individual during its entire infectious period, if seeded in a large population at the beginning",

then the general conclusions hold:

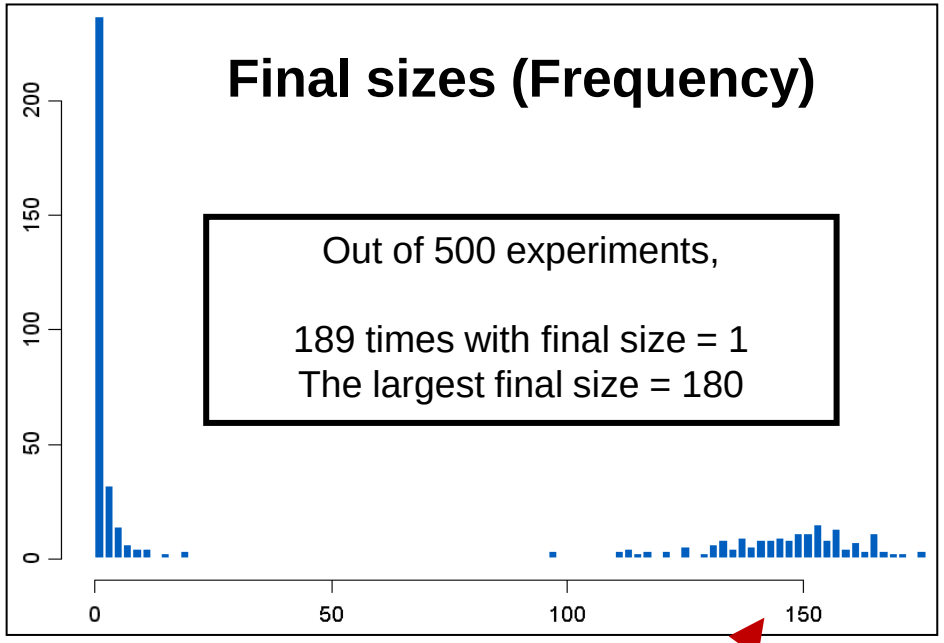
$R_0 \leq 1.0$	final % infected $\rightarrow$ 0% as $n \rightarrow \infty$
$R_0 = 1.2$	final % infected = 31.4%
$R_0 = 1.4$	final % infected = 51.1%
$R_0 = 1.6$	final % infected = 64.2%
$R_0 = 2.0$	final % infected = 79.7%
$R_0 = 2.5$	final % infected = 89.3%
$R_0 = 3.0$	final % infected = 94.0%

and the calculation of vaccination coverage.

What does the stochastic SIR model say about the final size?

A hypothetical susceptible population of size n = 200 (imagine a very large class, amphitheater, of students)

- 500 repeated experiments under identical initial conditions, seeding one infected individual, $R_0 = 1.8$
- Final size of the outbreaks: between 1 and 200
  - Final Size = 1: No transmission beside the seed itself
  - Final size = 200: The entire population is infected



Stochastic simulation

In approximately 56% of the repeated experiments, the final size accounts for a few cases (e.g. < 10).

Otherwise, it might happen that large outbreak occurs and should it happen, the final size is distributed around 146.5.

Deterministic SIR

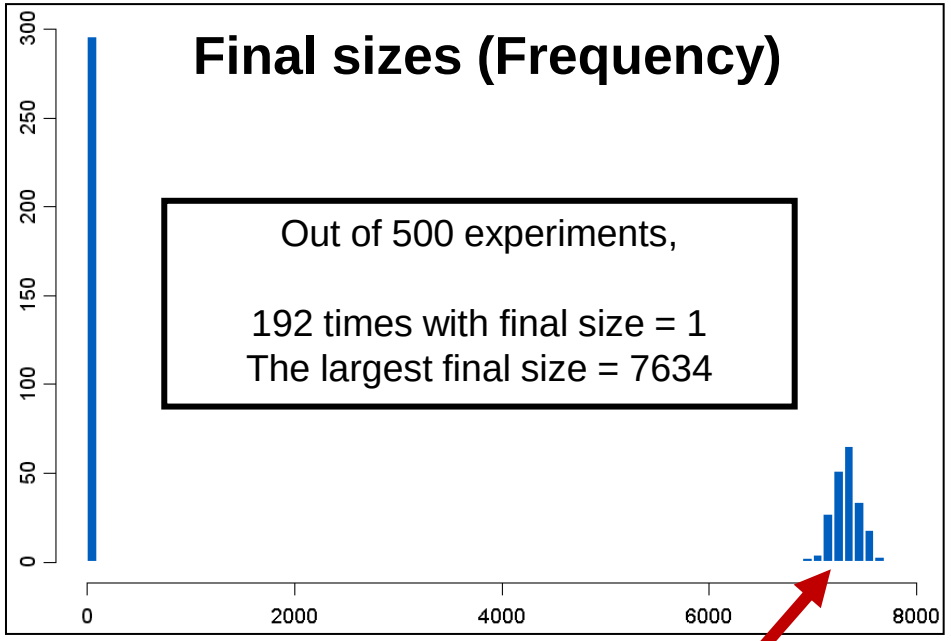
$R_0 = 1.8$ and Final % Infected $= 73.2\%$

$n = 200$, expect 146.5 as final number of infections and this must happen.

What does the stochastic SIR model say about the final size?

A hypothetical susceptible population of size n = 10000 (imagine a sporting event or similar large gathering)

- 500 repeated experiments under identical initial conditions, seeding one infected individual, $R_0 = 1.8$
- Final size of the outbreaks: between 1 and 10000
  - Final Size = 1: No transmission beside the seed itself
  - Final size = 10000: The entire population is infected



Stochastic simulation

In approximately 56% of the repeated experiments, the final size accounts for a few cases.

Otherwise, it might happen that large outbreak occurs and should it happen, the final size is distributed around 7324.

Deterministic SIR

$R_0 = 1.8$ and Final % Infected $= 73.2\%$

$n = 10000$, expect 7324 as final number of infections and this must happen.

What does the stochastic SIR model say about the final size?

Helps us to articulate a “small outbreak” vs. a “large outbreak”

Small Outbreak

After initial seeding, the final number of infected does not scale with the population size and usually remains at a few cases. If the population size goes infinitely large, the final % of infected is approximately zero.

$\boldsymbol{R_0 \leq 1}$: Small outbreak with certainty, large outbreak impossible

Large Outbreak

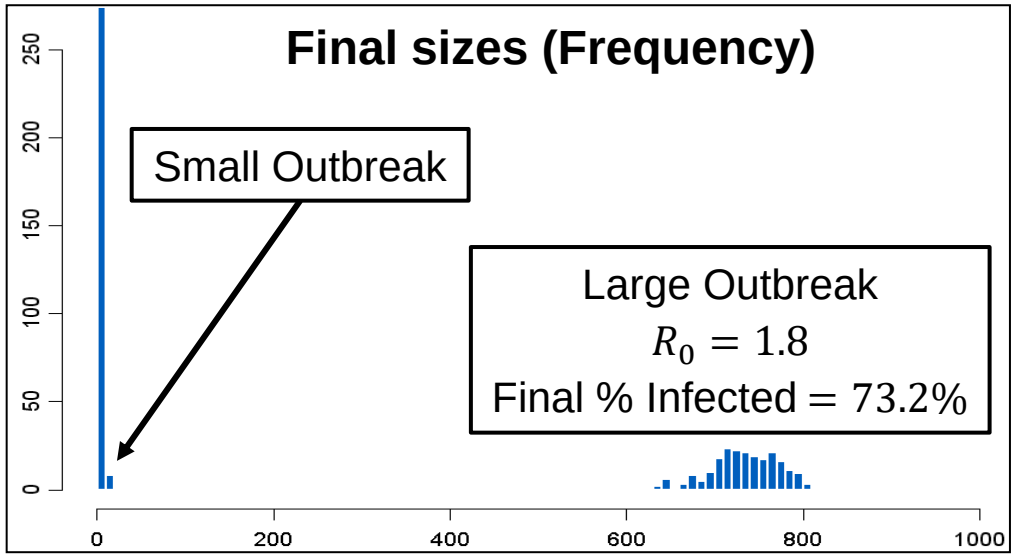
After initial seeding, the final number of infected scales with the population size as a positive proportion. This proportion, on average, agrees with that given by the deterministic SIR model but with some random variation.

$\boldsymbol{R_0 > 1}$: Large outbreak might happen, not “must” happen

Example: **A hypothetical susceptible population of size n = 1000 (imagine a high school)**

- 500 repeated experiment under identical initial conditions, seeding one infected individual

Stochastic model further predicts the probability (risk) of a large outbreak when $R_0 > 1$.



Deterministic vs. Stochastic: Parts of an Outbreak

An outbreak contains:

Initial Part + Main Part + Ending Part

The initial part (stochastic model to calculate):

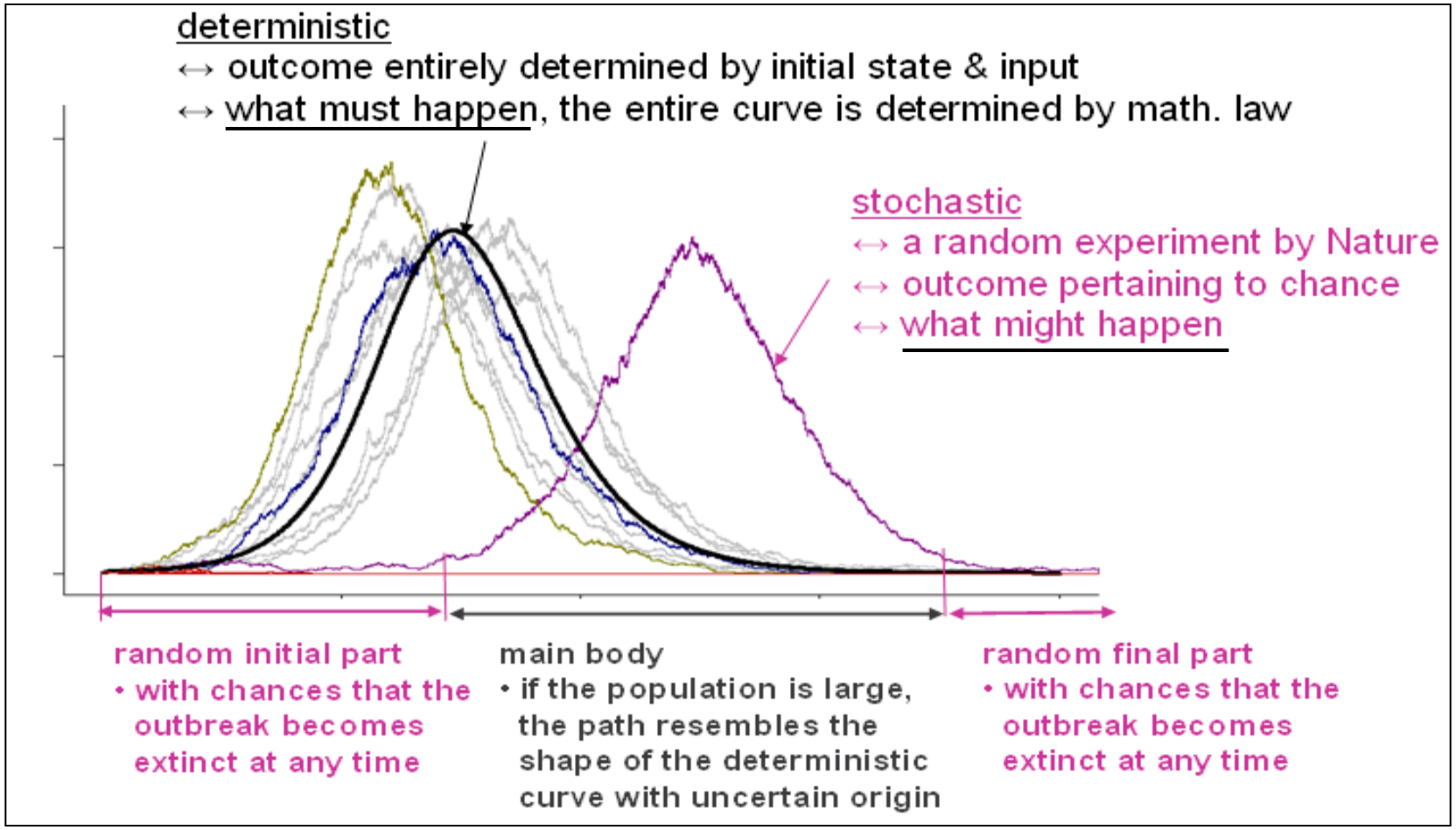
- What is the probability that an initial case will cause a large outbreak?
- How long is expected from initial seeding to a large outbreak (or to early extinction) ?

The main part if the population is large and relatively homogeneous (deterministic model):

- The deterministic model approximates the main part in large populations.
- Disease transmission is a stochastic, but a deterministic model is easier to implement.

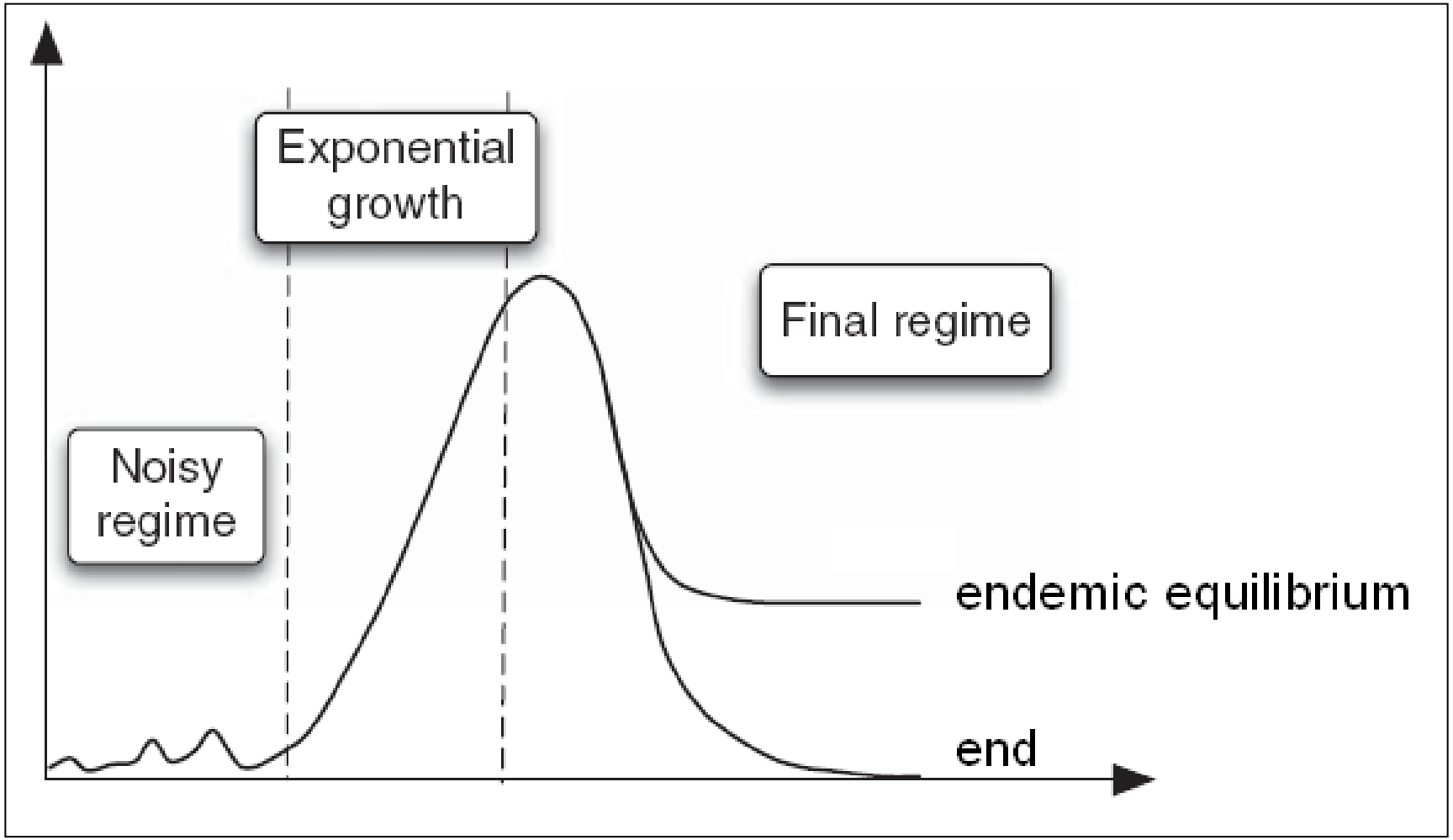
Deterministic vs. Stochastic

This diagram represents a typical acute outbreak in a population with relatively constant size and the outbreak “ends”.



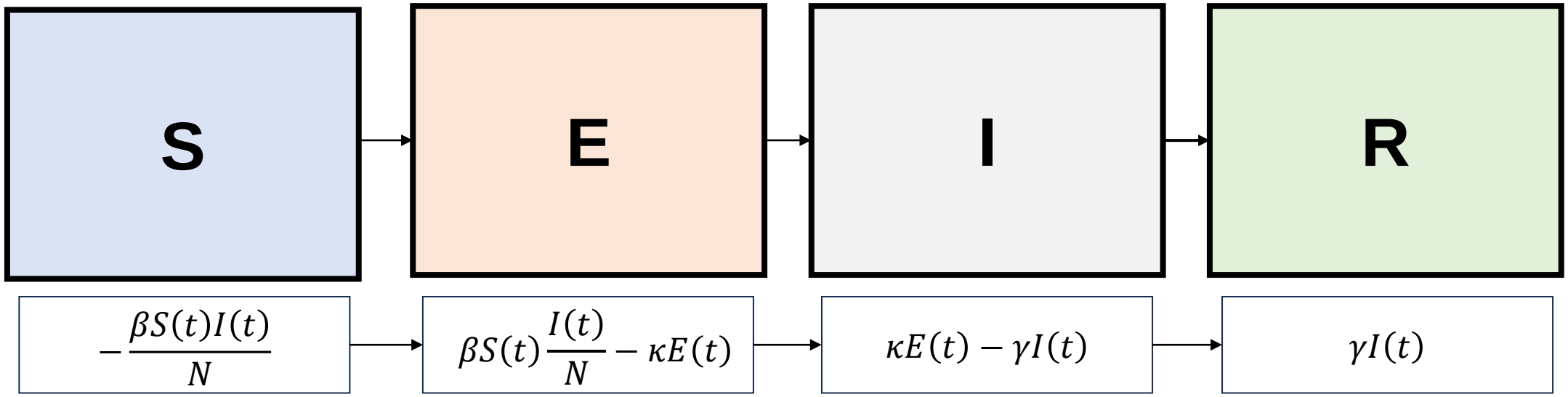
Deterministic vs. Stochastic

In general: Different deterministic models and their stochastic counterparts can be developed for other scenarios.



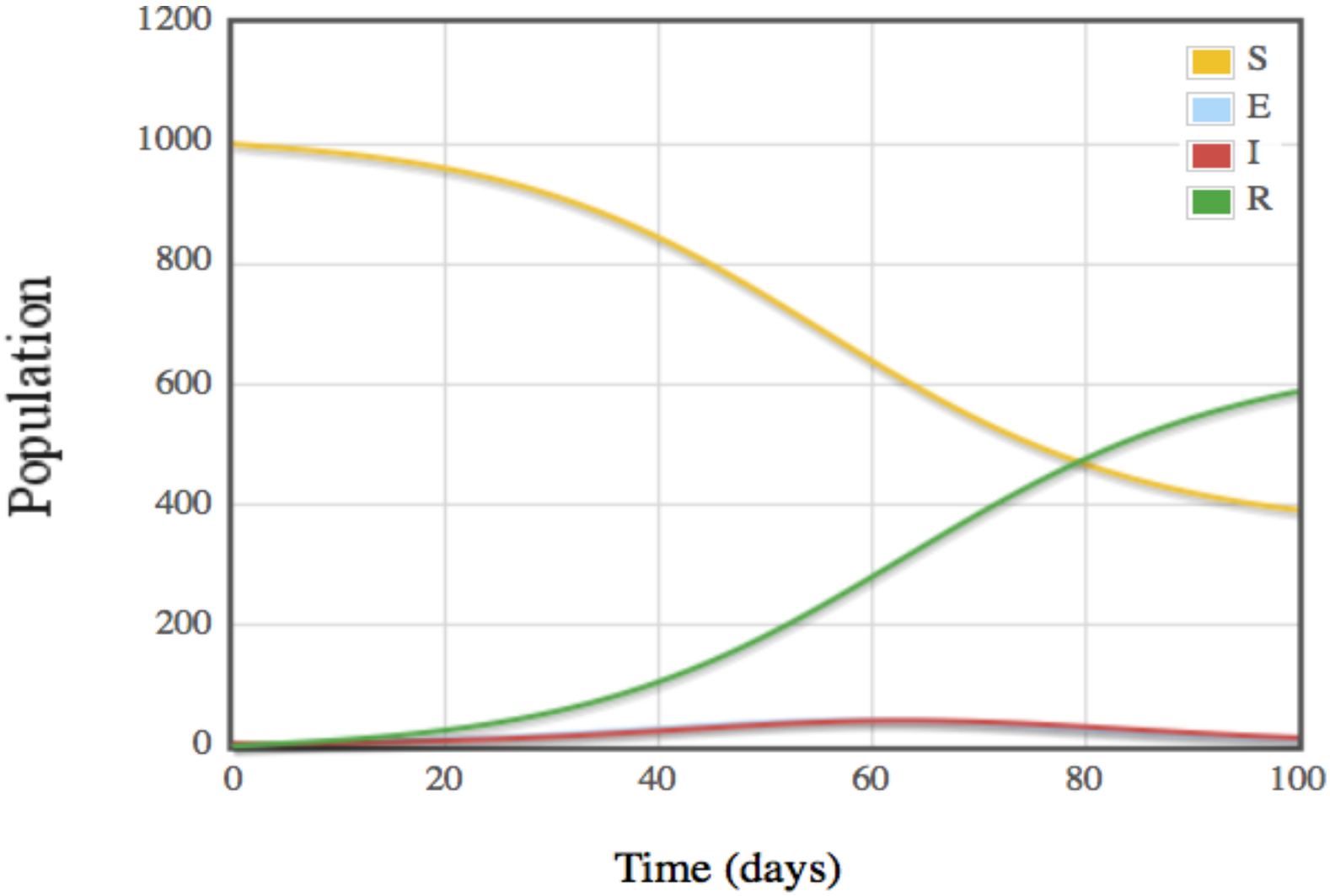
More Complex Compartmental Models

Susceptible-Exposed-Infectious-Recovered model



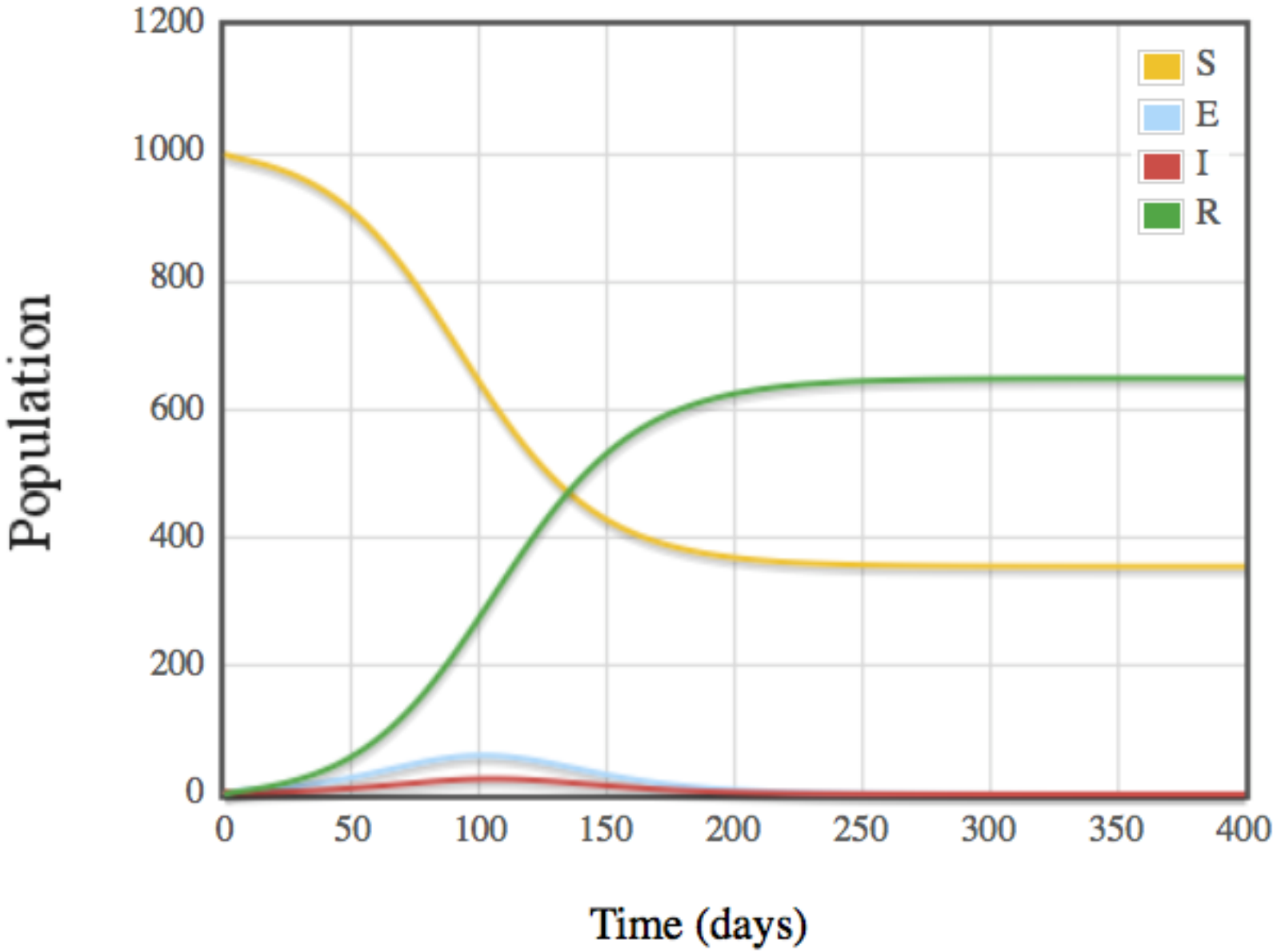
- Expands upon the SIR model with the inclusion of latency period via an exposed class.
- **Latency Period ($1/\kappa$):** Time elapsed from effective exposure to the infectious agent to infectiousness.
- R_0 can still be calculated as β/γ

Example: Susceptible-Exposed-Infectious-Recovered model



- **Population size:** 1000
- **Transmission rate:** 0.4
- **Latent period:** 4 days
- **Infectious period:** 4 days
- **5 individuals initially infectious**

Example: SEIR model (10-day latent period)

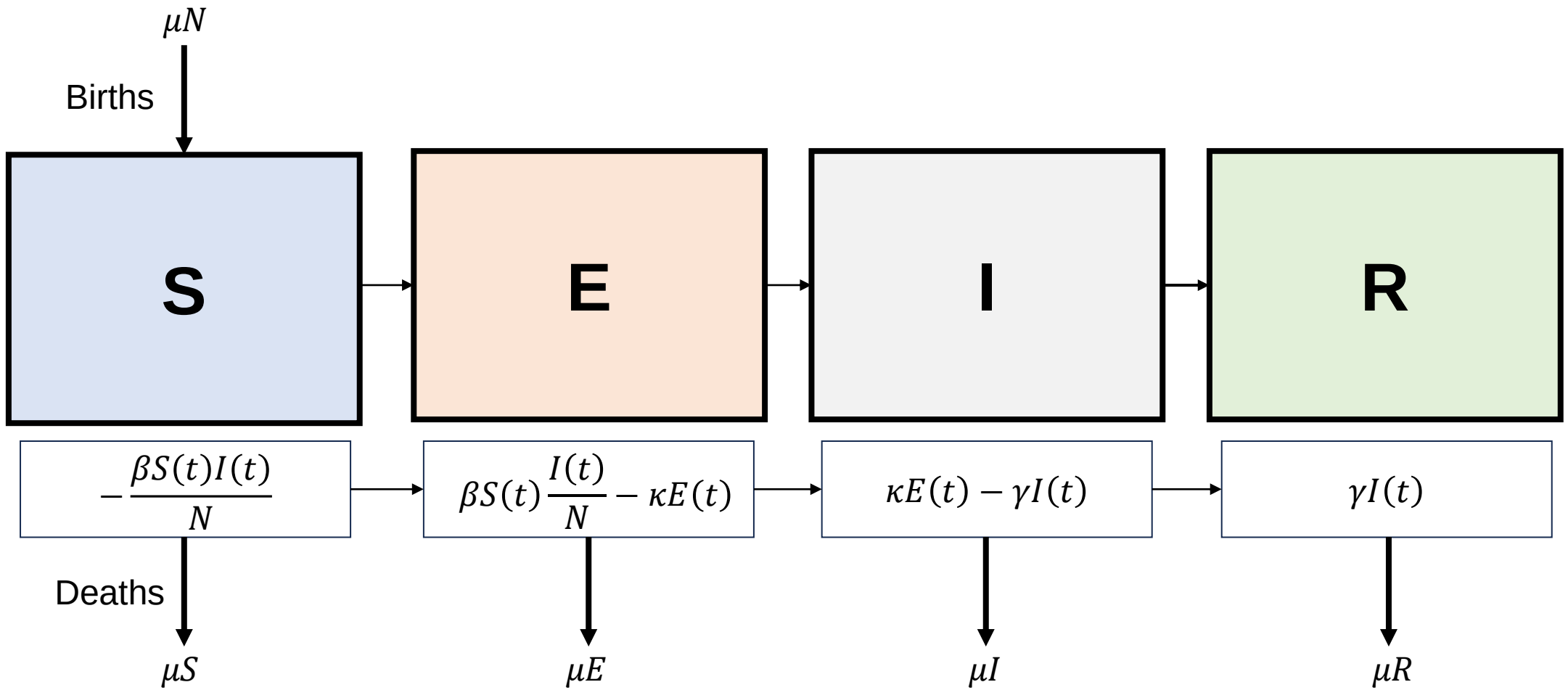


Beta: 0.4
Gamma: 0.25
Sigma: 0.1
Mu: 0
Nu: 0

Initial
Susceptible: 1000
Exposed: 0
Infected: 5
Recovered: 0

Days: 400
Submit

Accounting for demographic processes (Births and Deaths)

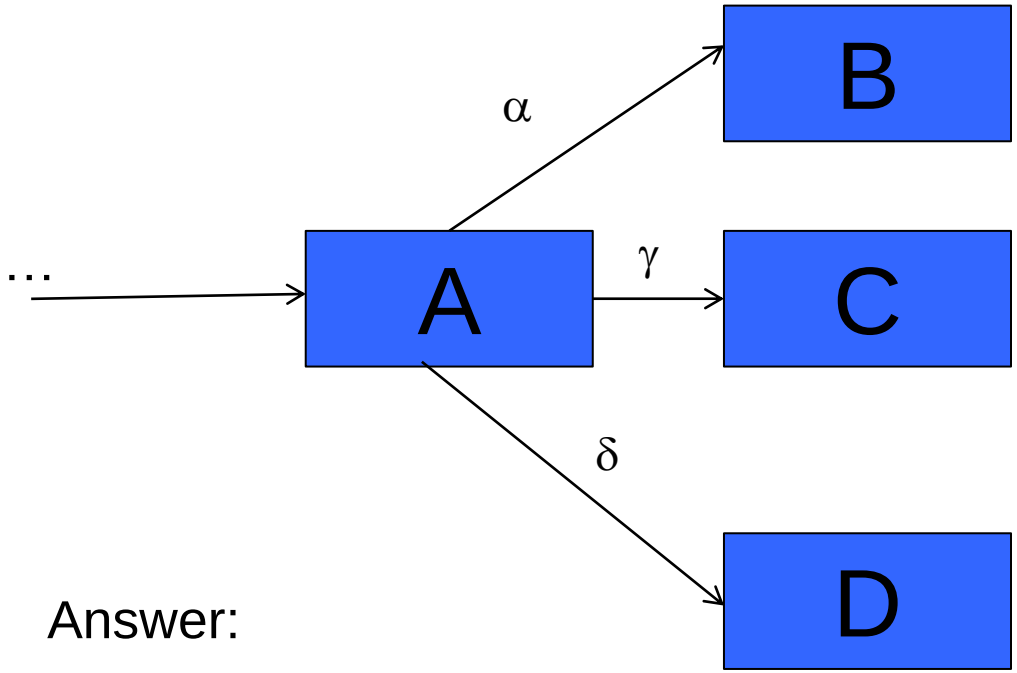


- $\boldsymbol{\mu}$: Migration Process
- How should birth and death rates be in order to achieve a constant population size?

R_0 for compartmental epidemic models

$$R_0 = (Transmisson\ Rate) * (Proportion\ Infected\ Reaching\ Infectious\ Compartment) * (Infectious\ Period)$$

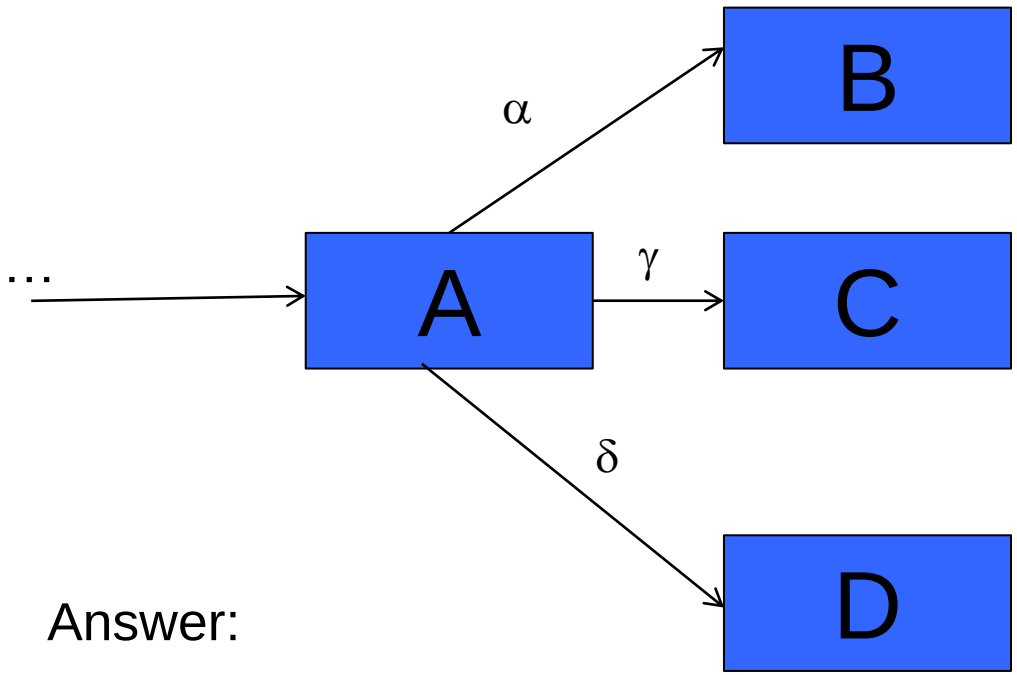
What is the average time that individuals spend in compartment A?



Answer:

$$1/(\alpha+\gamma+\delta)$$

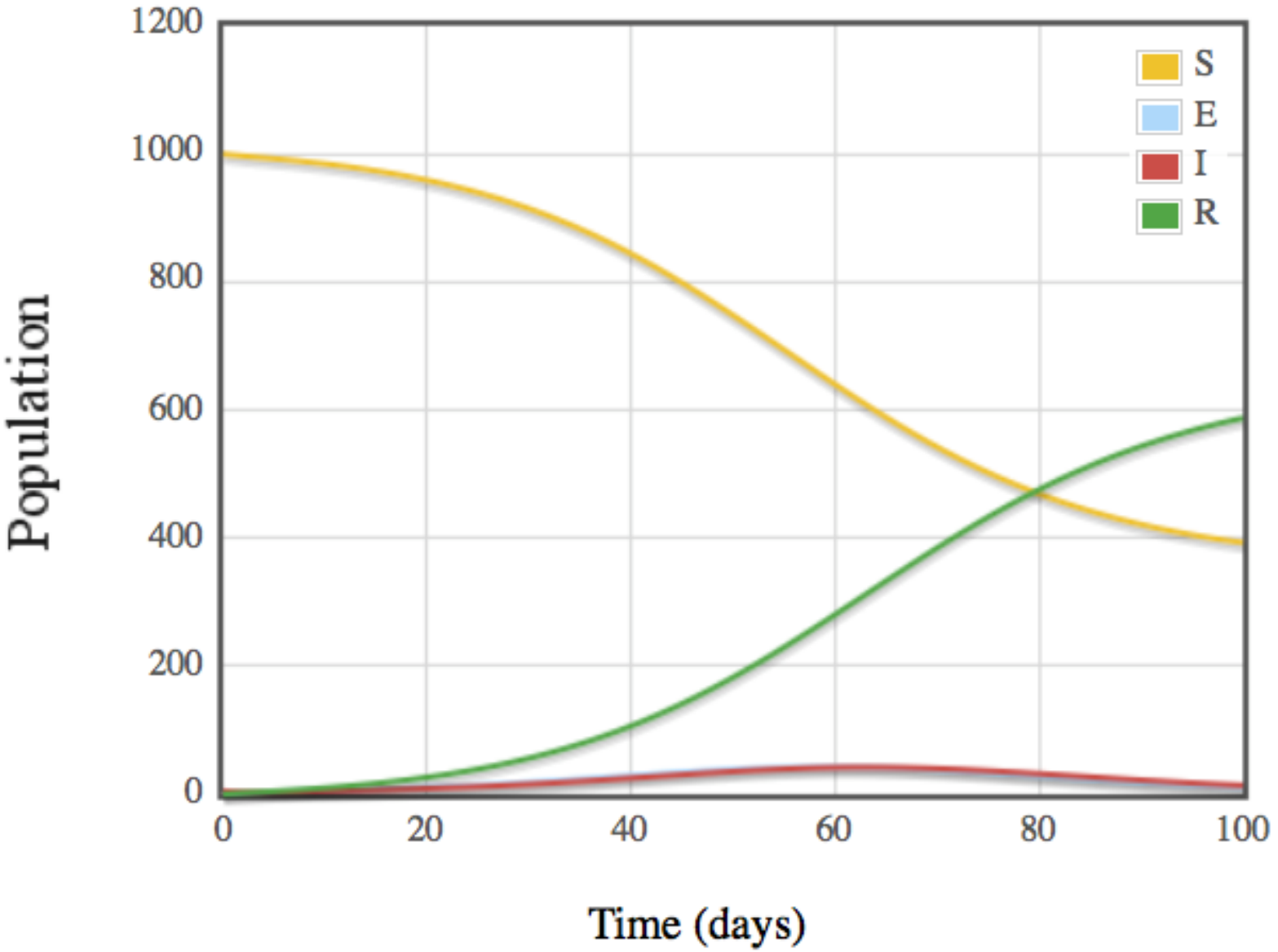
What fraction of individuals reach compartment C?



Answer:

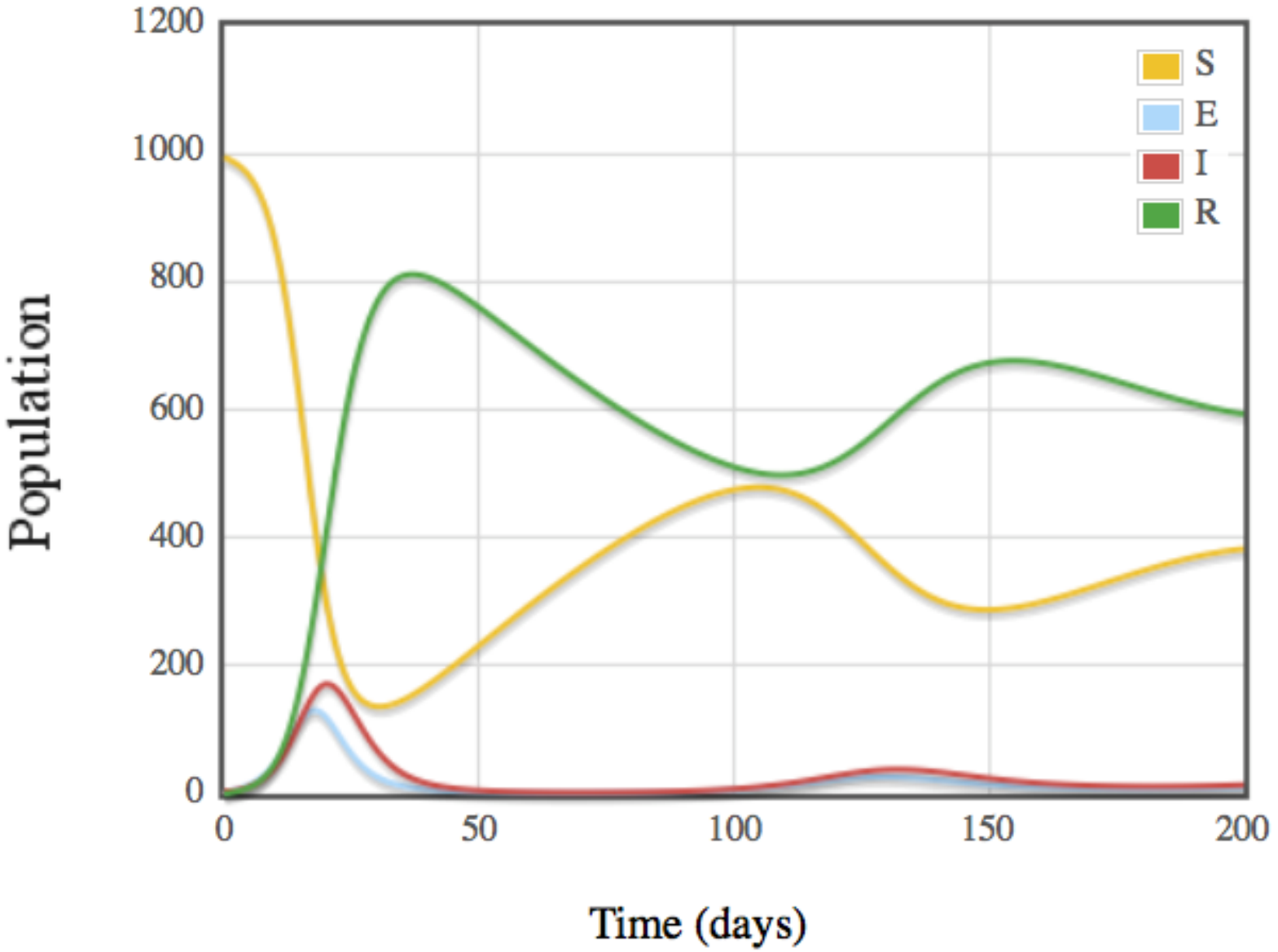
$$\gamma/(\alpha+\gamma+\delta)$$

Example: SEIR model with birth and natural death rates



- **Population size:** 1000
- **Transmission rate:** 0.4
- **Latent period:** 4 days
- **Birth & Death Rate:** 1/70 years
- **Infectious period:** 4 days
- **5 individuals initially infectious**

Example: SEIR model with migration process ($\mu = 1/100$ per day)



Beta: 1
Gamma: 0.3333
Sigma: 0.5
Mu: 0.01
Nu: 0

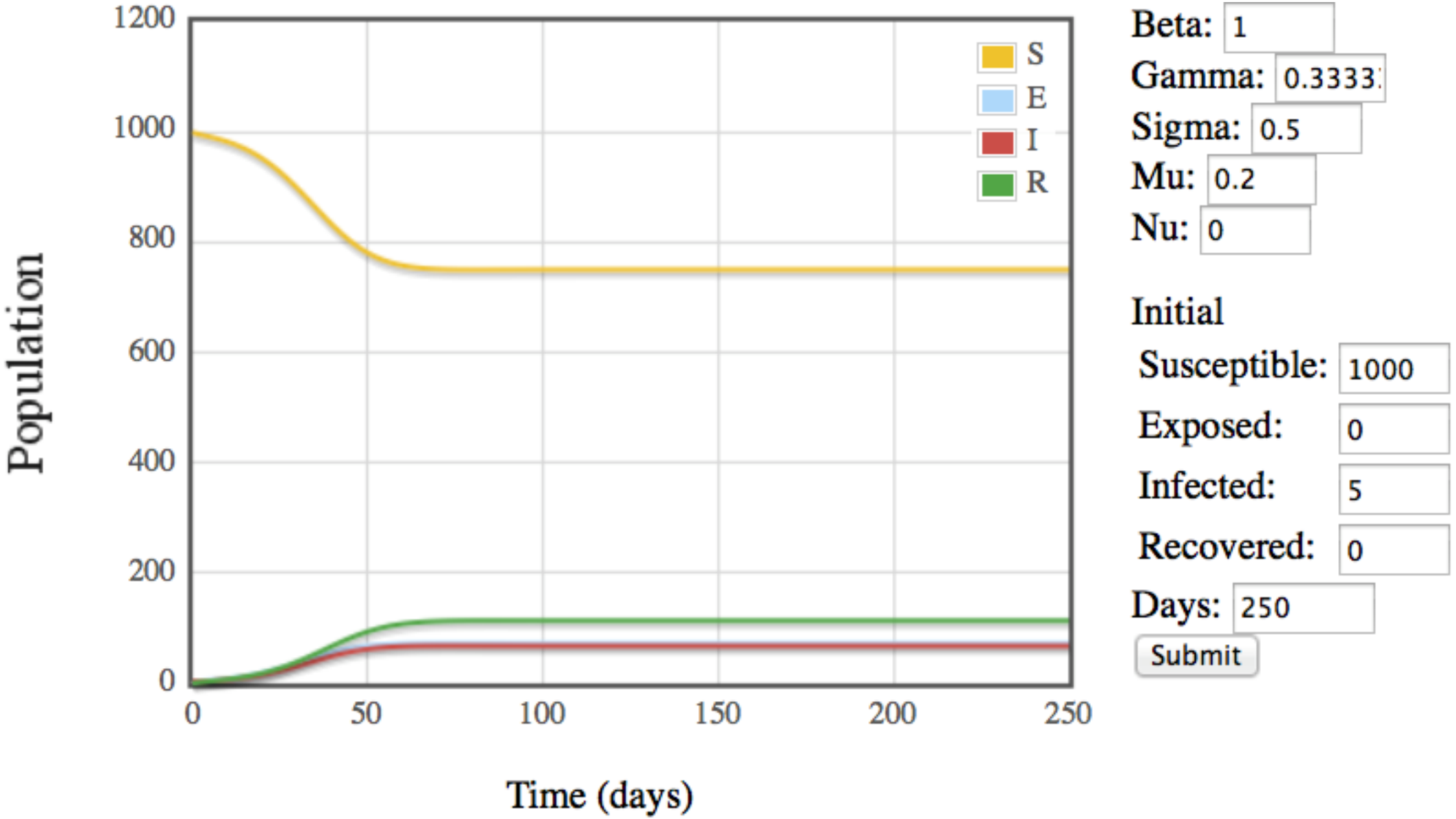
Initial

Susceptible: 995
Exposed: 0
Infected: 5
Recovered: 0

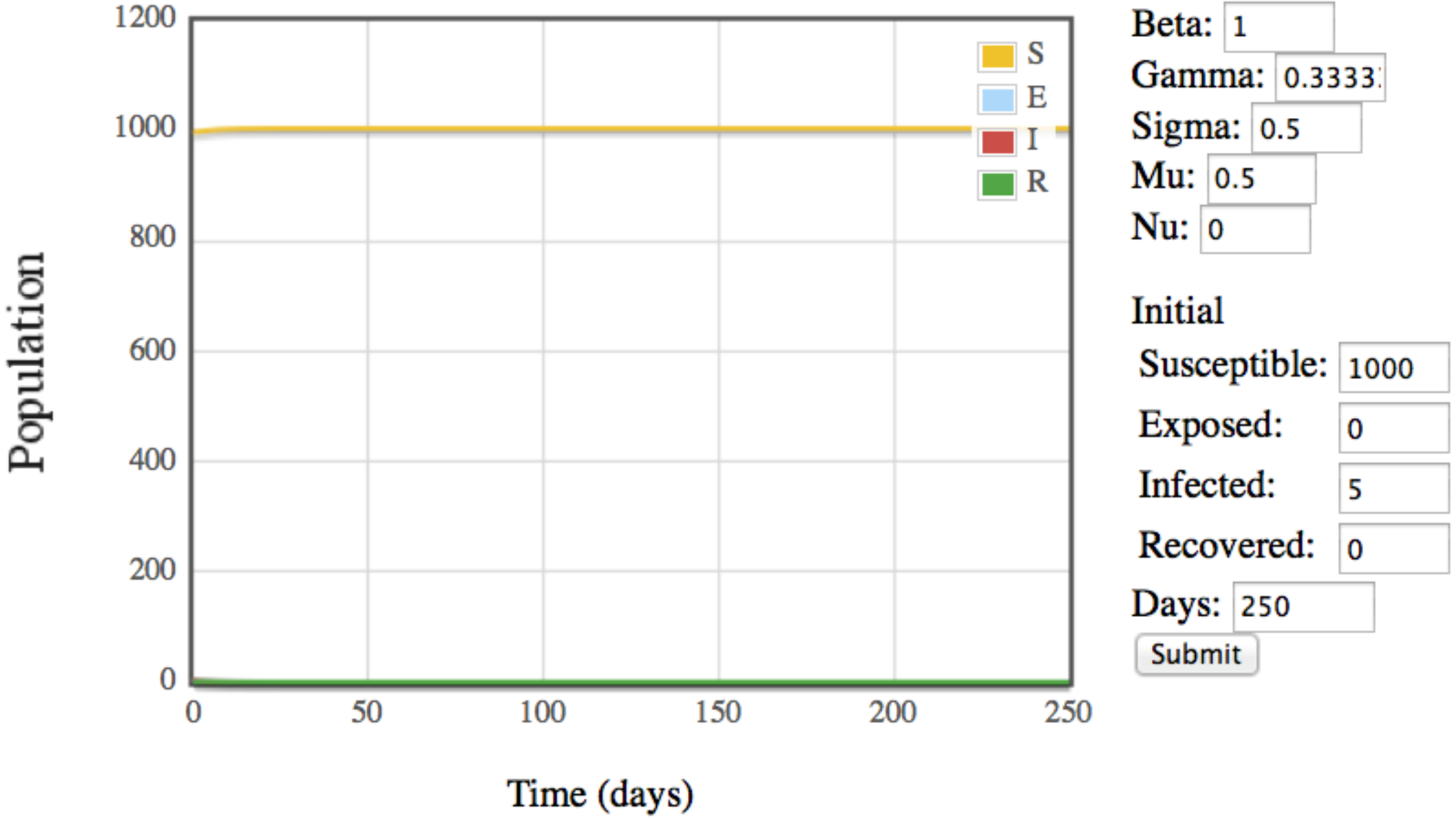
Days: 200

Submit

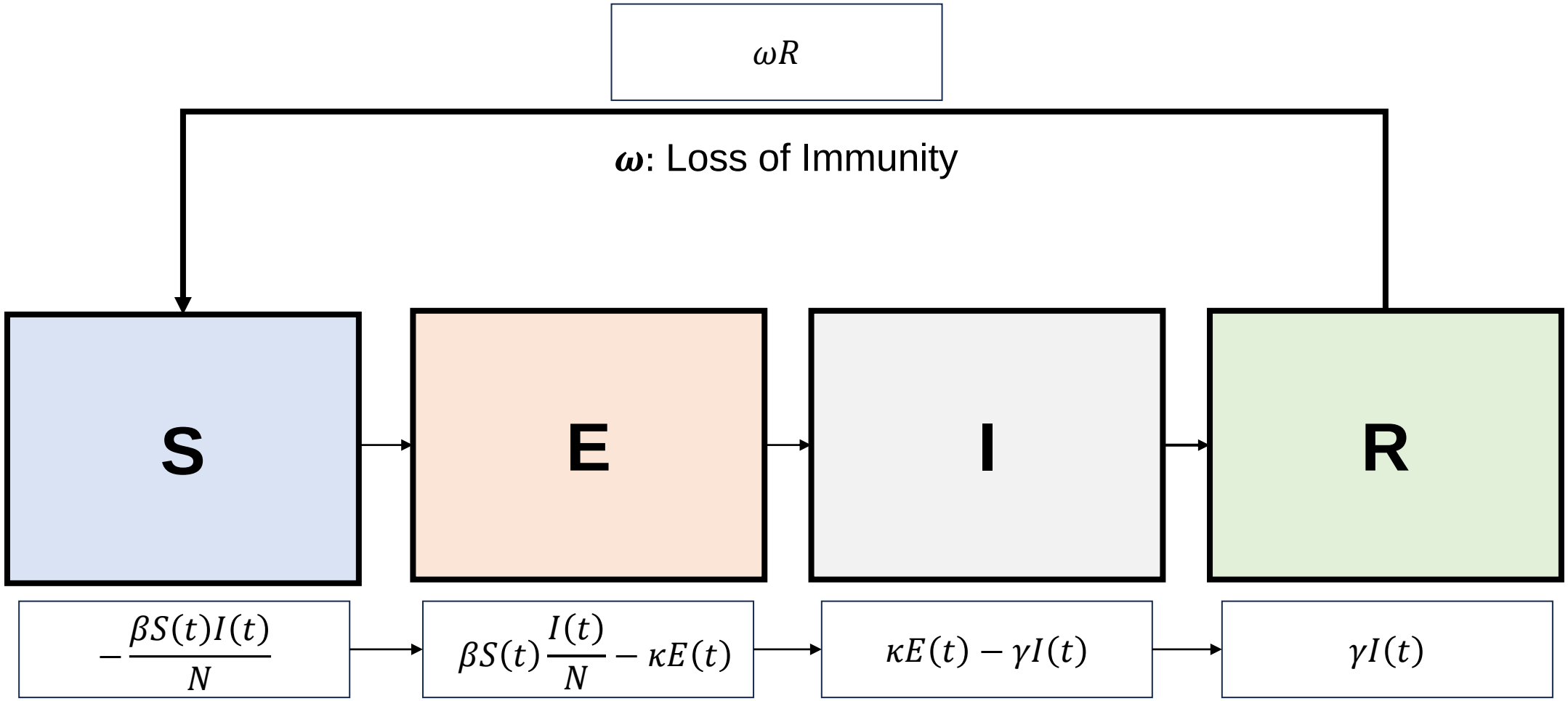
Example: Migration process ($\mu = 1/5$ per day)



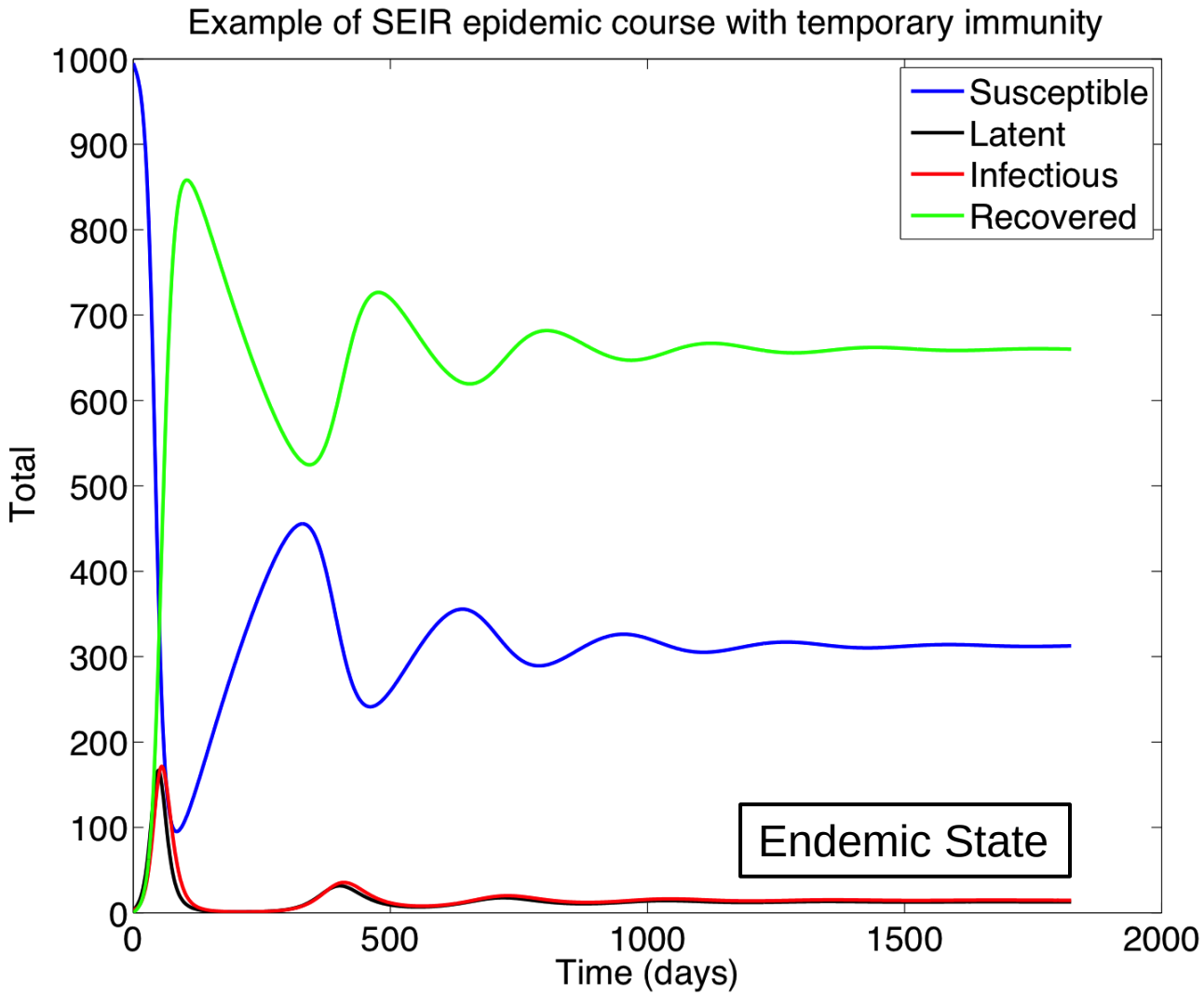
Example: Migration process ($\mu = 1/2$ per day)



Modeling temporary immunity for SEIR epidemics (SEIRS)

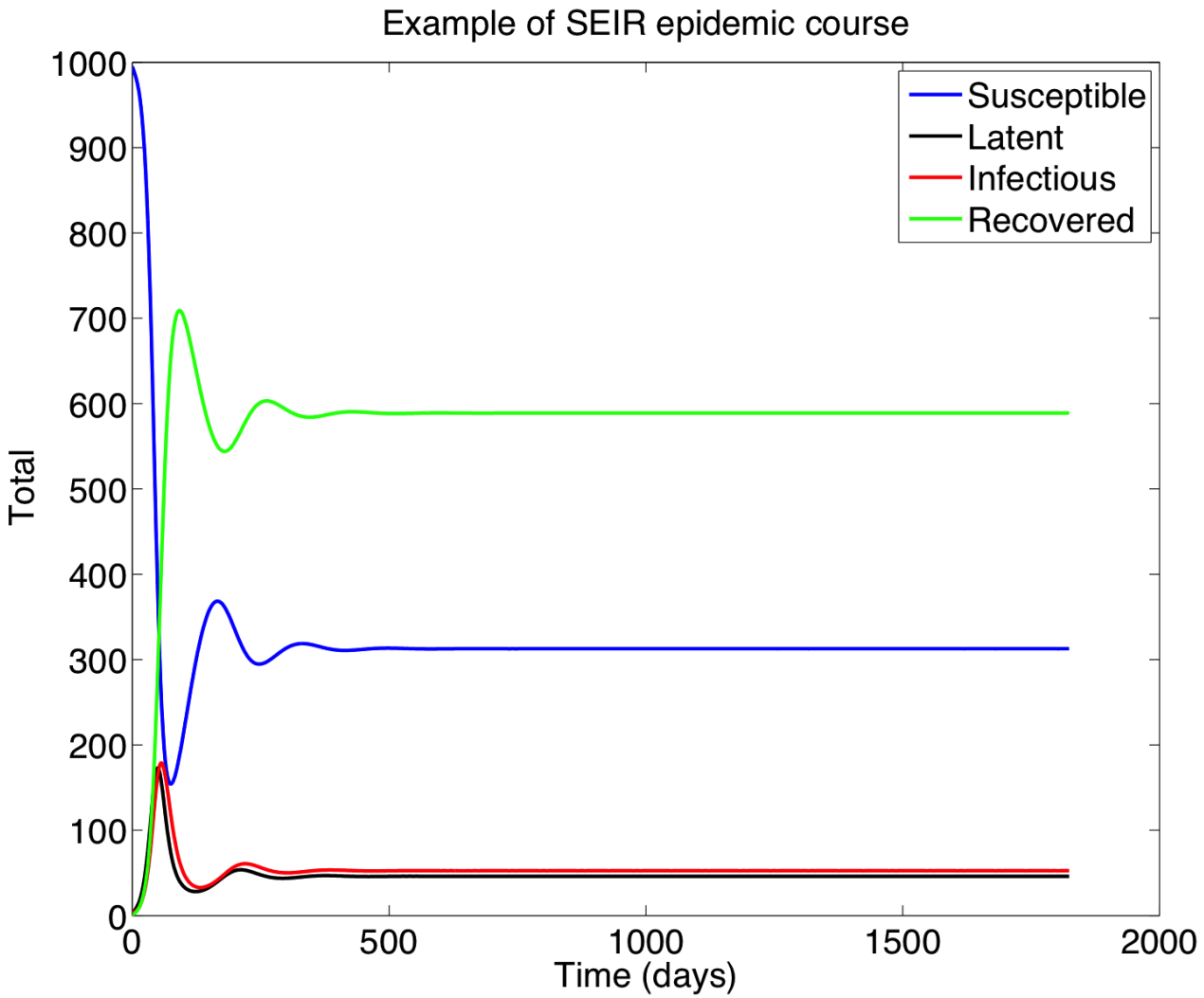


Example: 1-year protection after infection



- **Population size:**1000
- **Transmission rate:** 0.4
- **Latent period:** 7 days
- **Infectious period:** 8 days
- **5 individuals initially infectious**
- **5-year progression of disease**

Example: 3-month protection after infection



- **Population size:**1000
- **Transmission rate:** 0.4
- **Latent period:** 7 days
- **Infectious period:** 8 days
- **5 individuals initially infectious**
- **5-year progression of disease**