Modeling transmission dynamics with SIR models

Molecular Epidemiology of Infectious Diseases

Lecture 9

March 18th, 2024

The road ahead

This week: Modeling epidemic dynamics with SIR models

Next week: Stochastic models for simulation and inference

The week after: putting everything together with phylodynamic modeling

The general goal of infectious disease modeling is to better understand and explain epidemic dynamics.

Understanding epidemic dynamics

Why do epidemics rapidly grow and then decline?

Why do some outbreaks take off and not others?

Why do some diseases persist continuously and become endemic?

Why do some pathogens periodically re-emerge?

Why do some pathogens go extinct? How do we eliminate others?

These are all questions about infectious disease *dynamics* — changes over space and time.

Modeling dynamical systems

We describe the world around us using dynamical systems — a system that evolves over time.

We try to explain changes mechanistically in terms of processes we can empirically observe and measure (e.g. height of an apple).

Changes in important variables are described mathematically using ***differential equations***.

Understanding differential equations with almost no calculus***

***It all turns out to be a bunch of calculus in the end

The bathtub model

Let's consider a bathtub where *W(t)* denotes the amount of water in the tub at time *t*.

The amount of water that **flows *in*** over time *Δt*

$$\text{Inflow rate} \times \Delta t$$

The amount of water that **flows *out*** is:

$$\text{Outflow rate} \times \Delta t$$

The bathtub model

So if we start with *W(t)*, the amount of water after *Δt* is:

$$W(t+\Delta t) = W(t) + \text{Inflow rate} \times \Delta t - \text{Outflow rate} \times \Delta t$$

Using *I* and *O* to denote the flow rates:

$$W(t+\Delta t) = W(t) + I\Delta t - O\Delta t$$

The bathtub model

The difference between two time points gives the total amount of change over *Δt:*

$$\begin{aligned} W(t+\Delta t) - W(t) &= W(t) + I\Delta t - O\Delta t - W(t) \\ &= I\Delta t - O\Delta t \end{aligned}$$

We can then consider the *instantaneous* rate of change over a small interval of time *Δt*:

$$\frac{W(t+\Delta t) - W(t)}{\Delta t} = \frac{I\Delta t - O\Delta t}{\Delta t} = I - O$$

We can write this as the derivative of *W* with respect to time, which gives us the differential equation:

$$\frac{dW}{dt} = I - O$$

Differential equations

Differential equations track the change in a variable (or the amount of “stuff”) *per unit time*.

Often, this is simple as accounting for the rate at which stuff flows in minus the rate at which flow stuff out, e.g.:

$$\frac{dW}{dt} = I - O$$

Compartmental models

Compartmental models track the rate at which “stuff” flows between different compartments or states.

Compartments can represent different tubs, populations, ect.

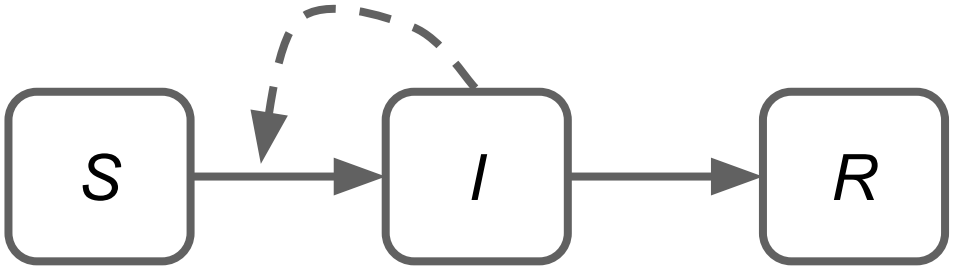
Flow between compartments can be tracked using a coupled system of differential equations.

SIR models

SIR Models

Susceptible-**I**nfected-**R**ecovered (SIR) models are compartmental models that focus on the infection status of individual hosts

The infection status of individual hosts changes over time. This can be represented as a flow diagram:



Modeling transmission

Transmission is normally modeled assuming random mixing. Incidence, or the rate of new infections, depends on the rate susceptible hosts contact infected hosts.

$$\text{Incidence} = \beta SI$$

The transmission rate β reflects the contact rate between hosts multiplied by the probability of infection per contact (i.e. the rate of infectious contacts)

These models assume *mass-action*: individuals contact one another randomly so incidence only depends on the density of S and I hosts. Hosts within each compartment are assumed to be exchangeable.

Contact rates

Contact rates between hosts can either be ***frequency*** or ***density-dependent***.

Frequency-dependent: contact rate per infected host is independent of total population size N.

$$\text{Incidence} = \beta \frac{S}{N} I$$

Density-dependent: contact rate depends on the population densities.

$$\text{Incidence} = \beta S I$$

Generally, human infectious diseases are modeled as frequency-dependent whereas animal/plant diseases are modeled as density-dependent.

Modeling recovery

Infected individuals are assumed to recover at a constant rate γ over the course of infection. The total rate of recoveries in the infected population is:

$$\text{Total recoveries} = \gamma I$$

The average infectious period is the reciprocal of the recovery rate:

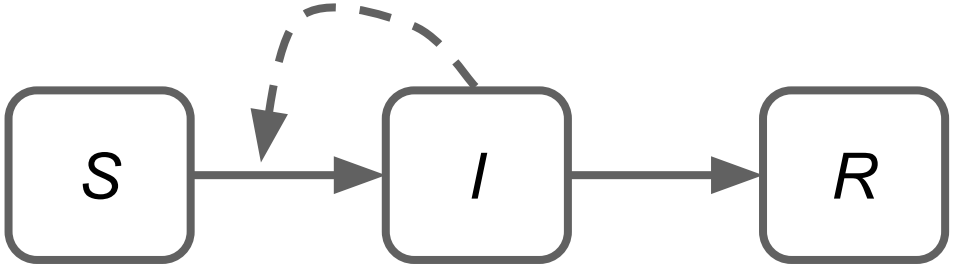
$$\text{Mean infectious period} = \frac{1}{\gamma}$$

A constant recovery rate assumes that individuals remain infected for an exponentially distributed amount of time.

The basic SIR model

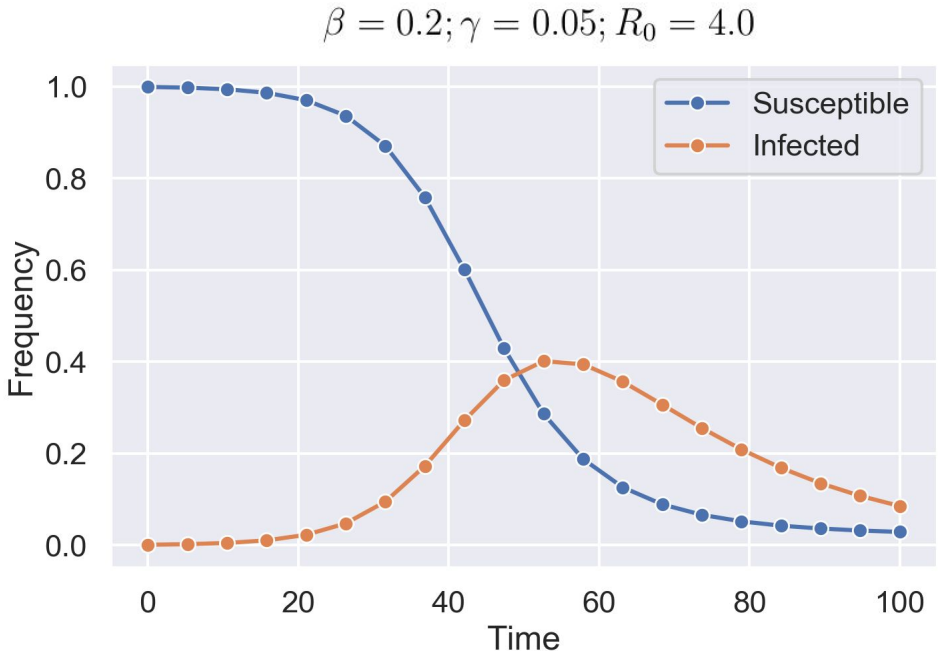
The basic SIR model can be represented as a system of differential equations:

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



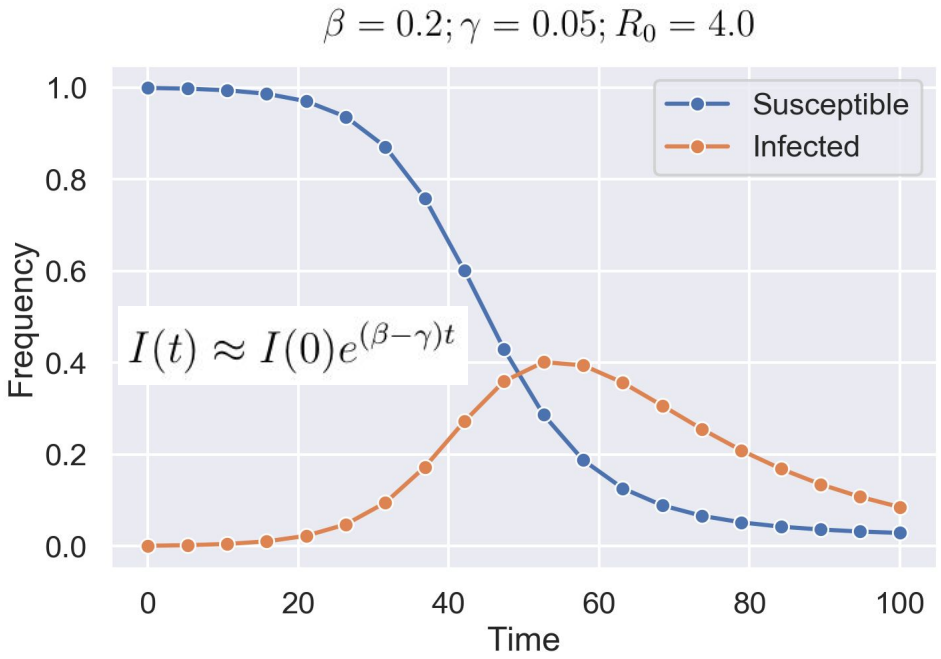
Epidemic dynamics under the SIR model

In the early stages of an epidemic, the number of infections (prevalence) grows exponentially as the number of infections increases.



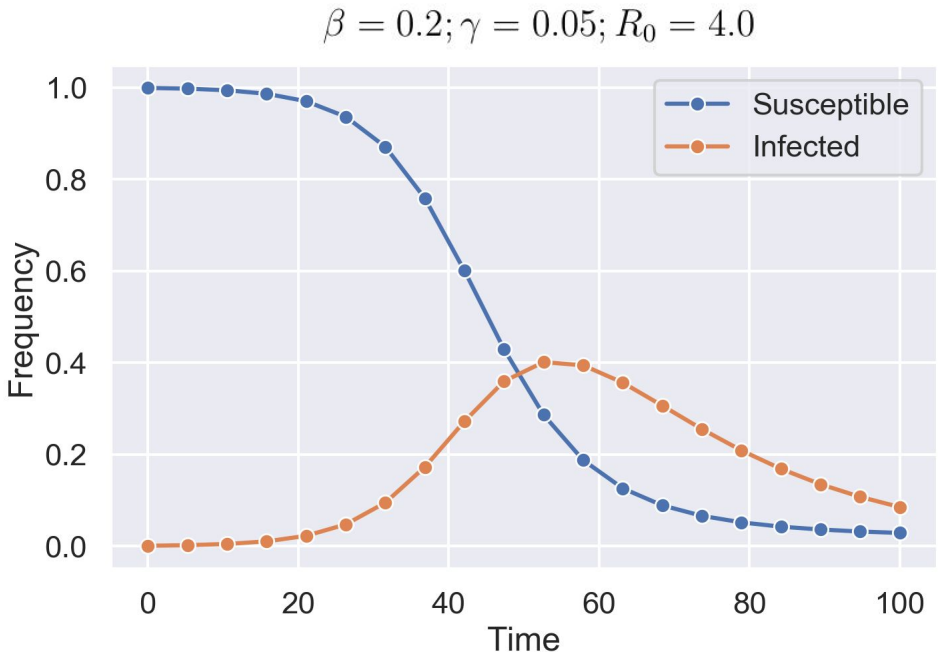
Epidemic dynamics under the SIR model

In the early stages of an epidemic, the number of infections (prevalence) grows exponentially as the number of infections increases.



Epidemic dynamics under the SIR model

Eventually the epidemic "*burns out*" as incidence declines due to a depletion of susceptible hosts.



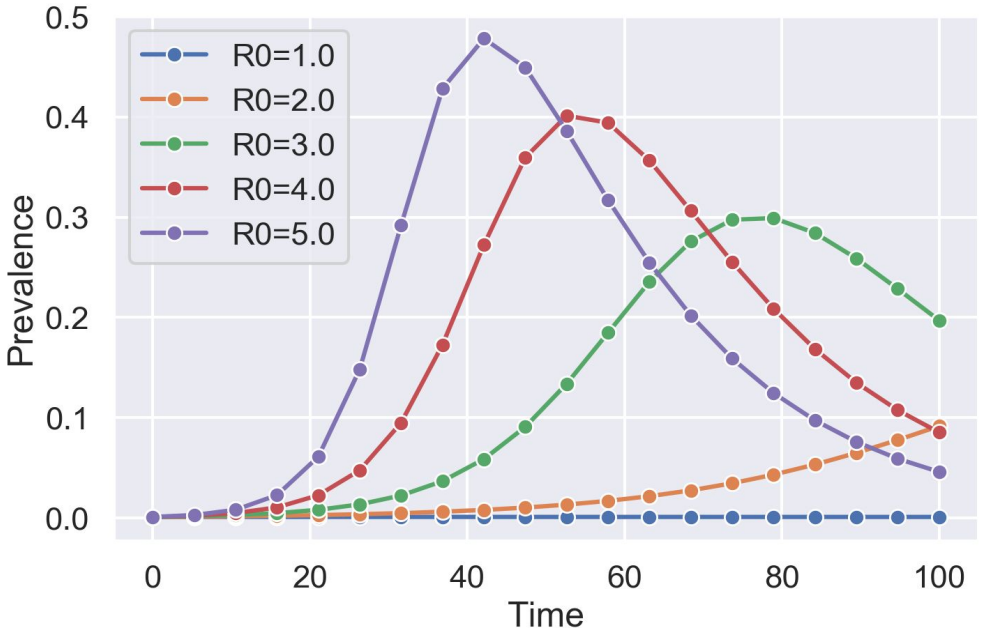
The epidemic threshold: R_0

The **basic reproductive ratio** R_0 is the average number of secondary cases arising from a single infection in an entirely susceptible population:

$$R_0 = \frac{\beta}{\gamma} = \beta \times \text{Average infectious period}$$

R_0 must be greater than one for an epidemic to take off.

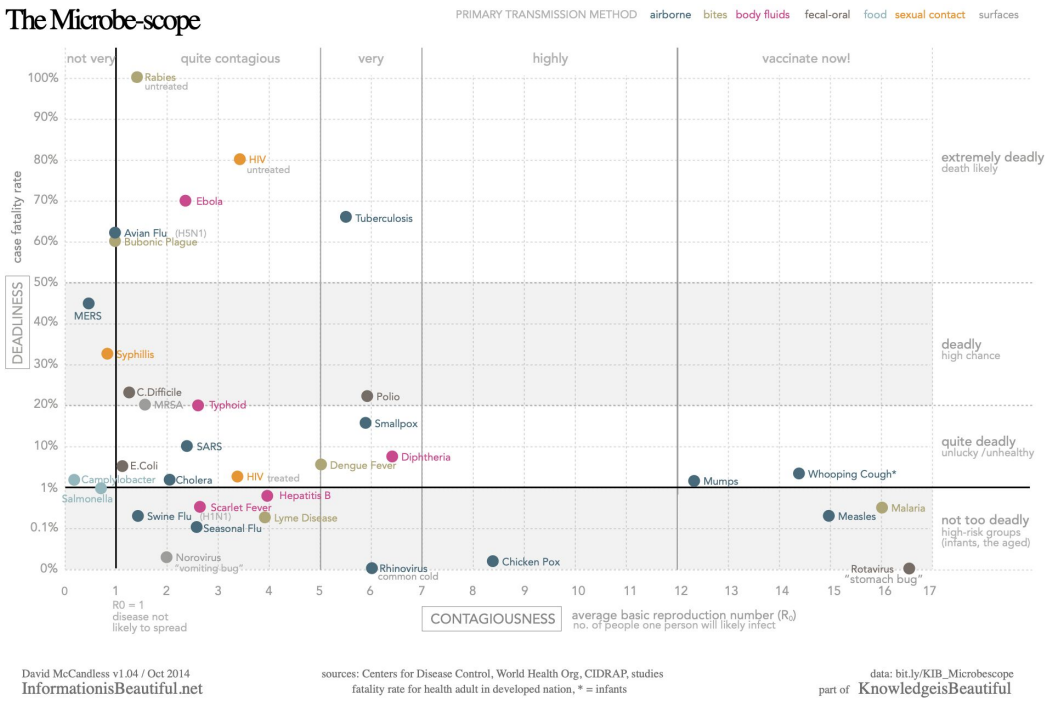
Epidemic dynamics under the SIR model



R_0 estimates for different pathogens

R_0 varies significantly between pathogens.

Because R_0 depends on β, it is a function of both a pathogen's infectiousness and host contact rates.



Just hypothetically...
say you're a mad scientist with the goal of designing the perfect pathogen to eradicate humanity?

To design the perfect pathogen, what would you maximize?
(This is a hard one!)

A) Transmission rate β

B) Growth rate r

C) Reproductive number R_0

D) Virulence (pathogenicity)

To design the perfect pathogen, what would you maximize? (This is a hard one!)

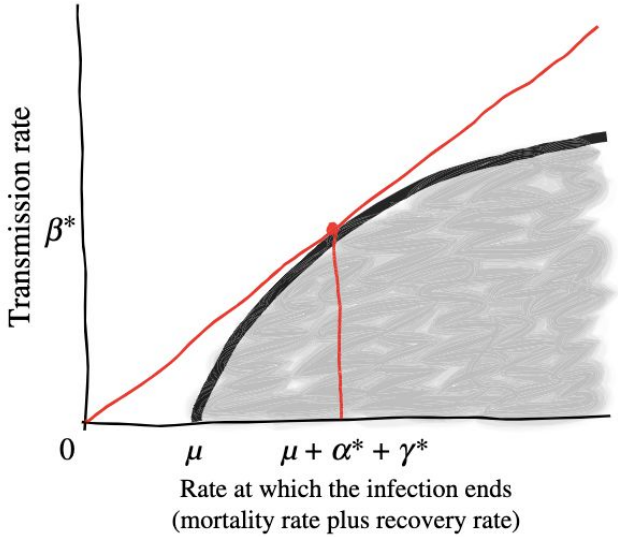
A) Transmission rate β

B) Growth rate r

C) Reproductive number R_0

D) Virulence (pathogenicity)

The transmission-virulence trade-off
(Alison et al., 2008)



To design the perfect pathogen, what would you maximize? (This is a hard one!)

A) Transmission rate β

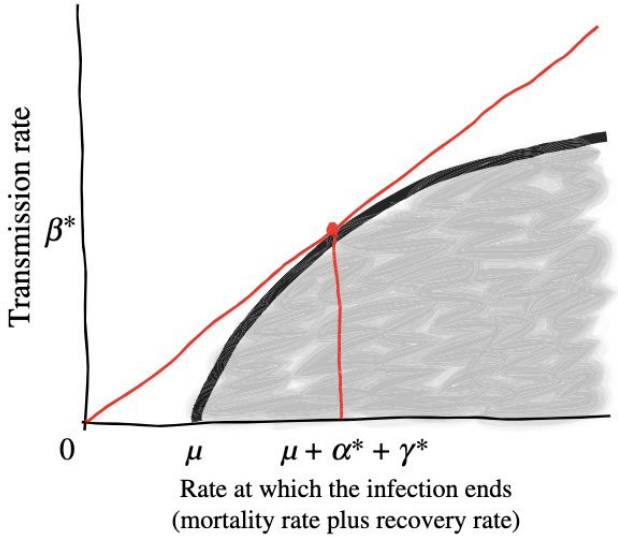
B) Growth rate r

C) Reproductive number R_0

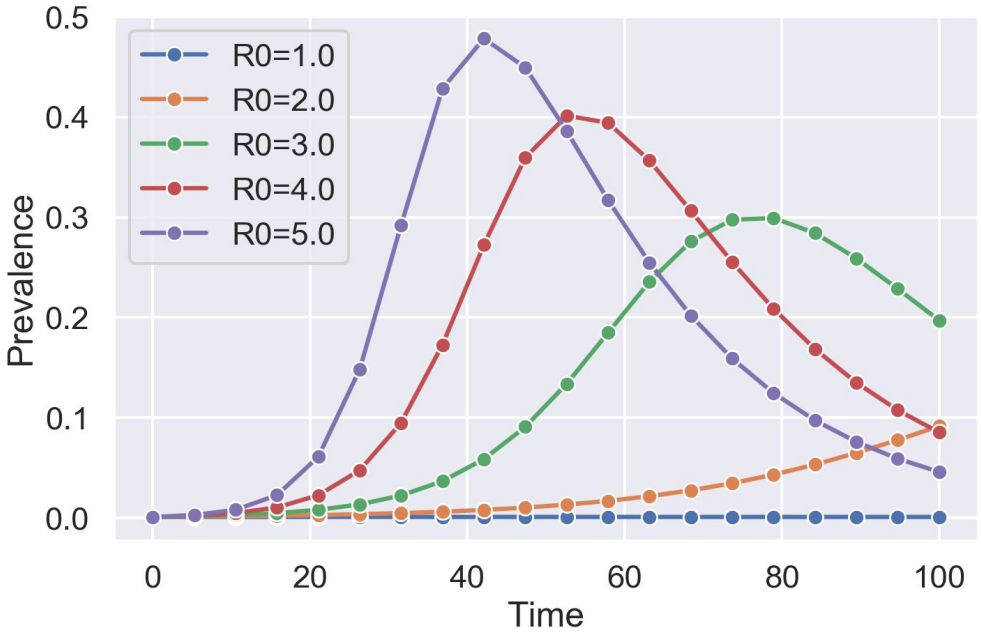
D) Virulence (pathogenicity)

Maximizing R_0 maximizes the cumulative number of infections over time and therefore the long-term fitness of a pathogen.

The transmission-virulence trade-off
(Alison et al., 2008)



The basic SIR model is pretty boring...



Incorporating host demography

Incorporating host demography allows us to consider long-term persistence and the endemic dynamics of a pathogen. Births replenish the susceptible population.

Births are generally assumed to occur at a constant rate ν and deaths at a constant rate μ. The average lifespan of an individual is therefore $1/\mu$.

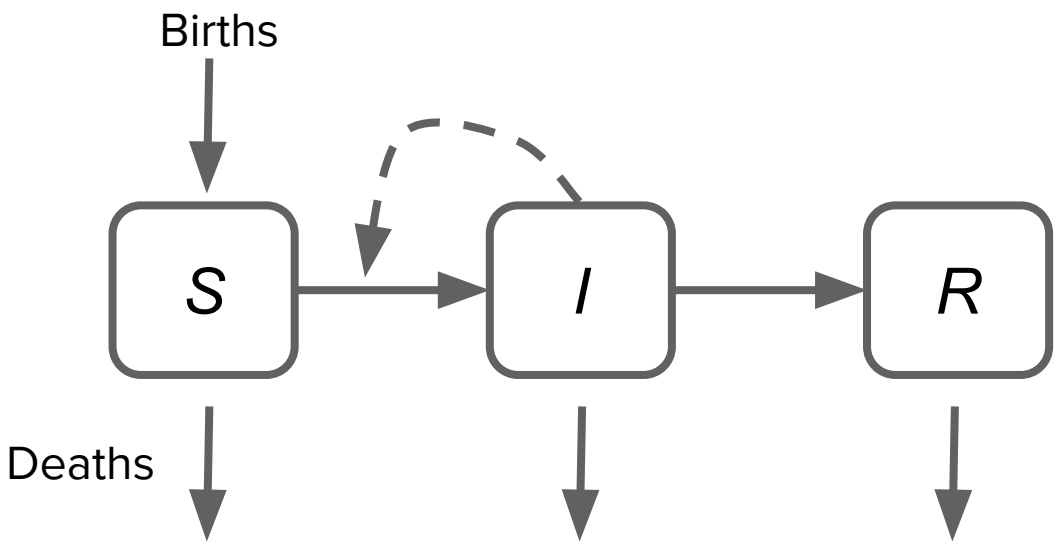
Note: it is common to assume that $\mu = \nu$ such that the host population is in demographic equilibrium and *N* remains constant through time.

The SIR model with host demography

The SIR with demography can be represented as a system of differential equations:

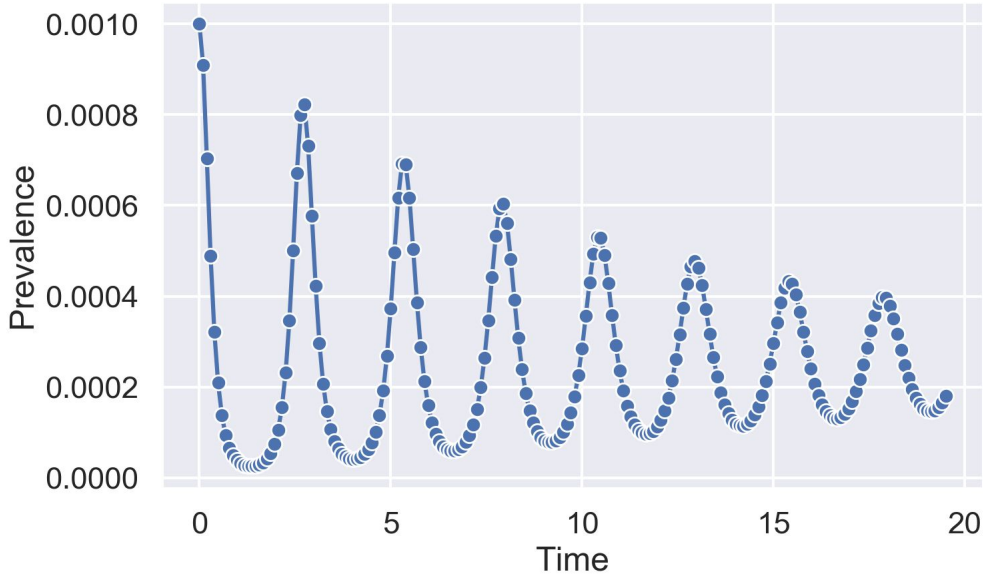
$$\begin{aligned}
\frac{dS}{dt} &= \nu - \beta SI - \mu S \\
\frac{dI}{dt} &= \beta SI - (\gamma + \mu) I \\
\frac{dR}{dt} &= \gamma I - \mu R
\end{aligned}$$

$$R_0 = \frac{\beta}{\gamma + \mu}$$



Epidemic dynamics with demography

Adding births allows the susceptible population to be replenished, which can give rise to complex oscillatory dynamics.

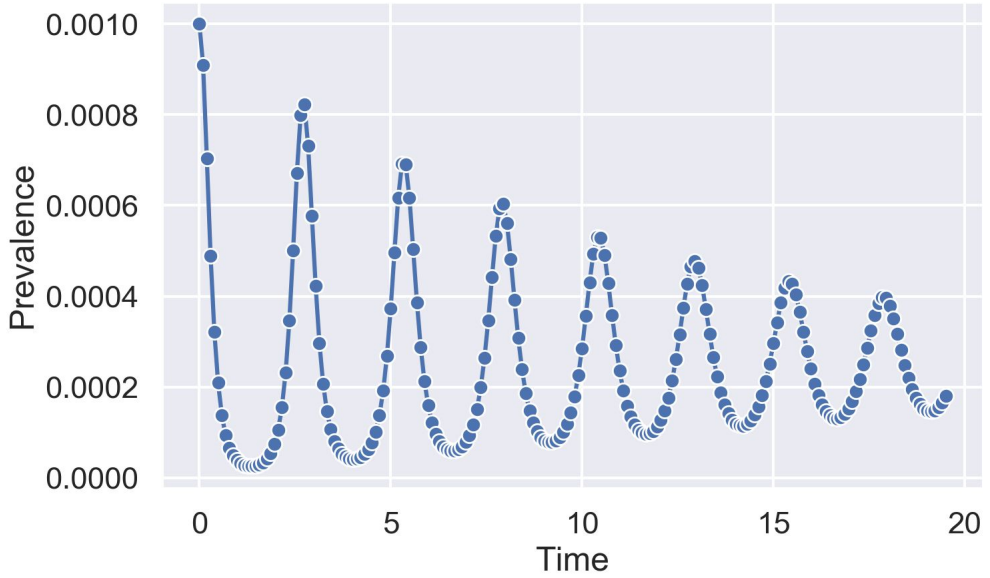


Epidemic dynamics with demography

Adding births allows the susceptible population to be replenished, which can give rise to complex oscillatory dynamics.

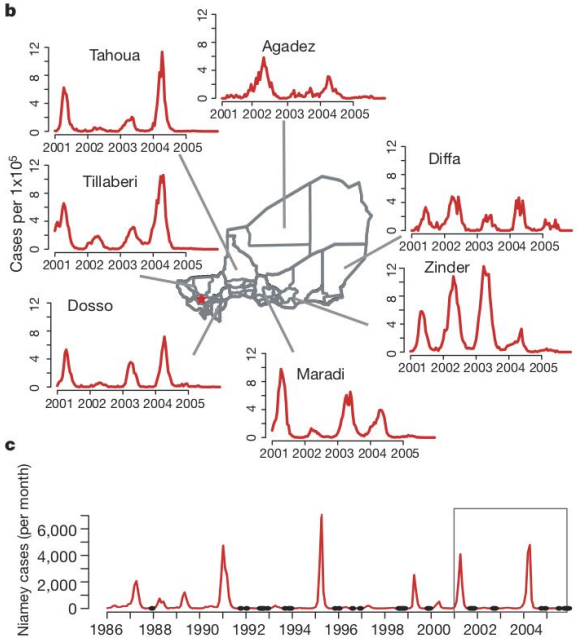
Eventually the system will reach *endemic equilibrium*:

$$S^* = \frac{\gamma+\mu}{\beta} = \frac{1}{R_0}$$
$$I^* = \frac{\mu}{\beta}(R_0 - 1)$$



Dynamical transitions in measles

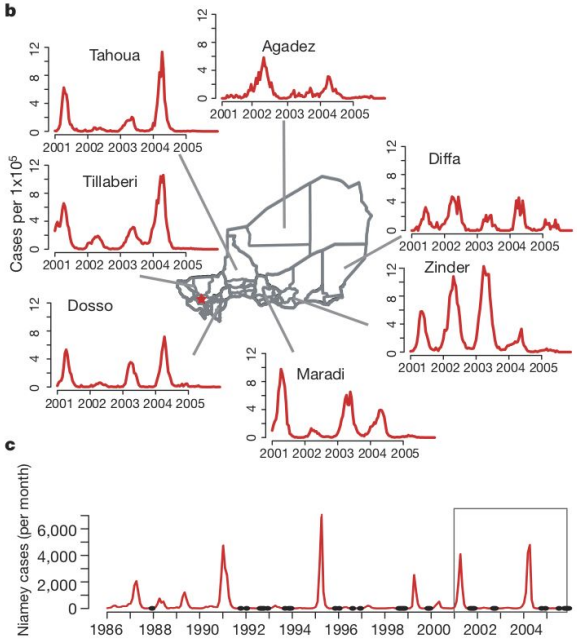
Birth rates drive transitions in the periodicity of measles epidemics

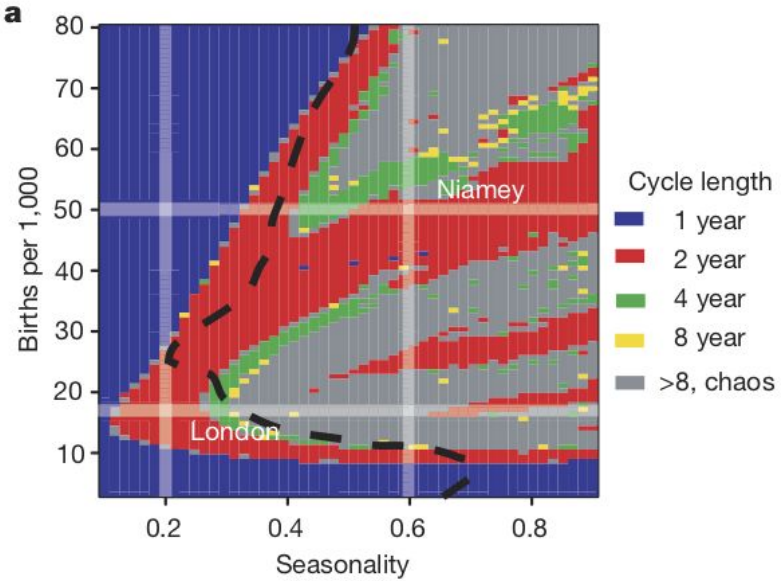


Ferrari *et al.* (Nature, 2008)

Dynamical transitions in measles

Birth rates drive transitions in the periodicity of measles epidemics





Ferrari *et al.* (Nature, 2008)

The vaccination threshold

We can consider a modified version of the SIR model with demography where a fraction *p* of newborns are vaccinated:

$$\begin{aligned}
\frac{dS}{dt} &= \nu(1-p) - \beta SI - \mu S \\
\frac{dI}{dt} &= \beta SI - (\gamma + \mu)I \\
\frac{dR}{dt} &= \gamma I + \nu p - \mu R
\end{aligned}$$

This system turns out to be dynamically identical to a system with a basic reproductive ratio R_0' reduced by the fraction *1 - p*:

$$R_0' = (1-p)R_0$$

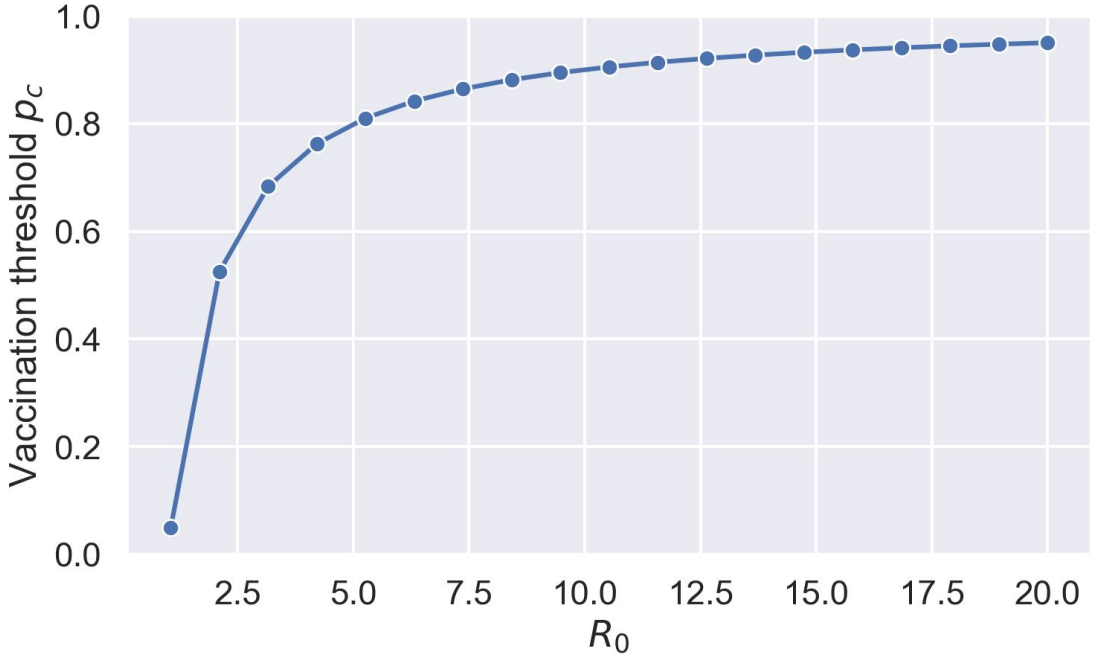
The vaccination threshold

In order to eradicate a pathogen through long-term vaccination, we need:

$$R_0' = (1-p)R_0 < 1$$

We therefore need to vaccinate a critical proportion of newborns p_c:

$$p_c = 1 - \frac{1}{R_0}$$



The vaccination threshold

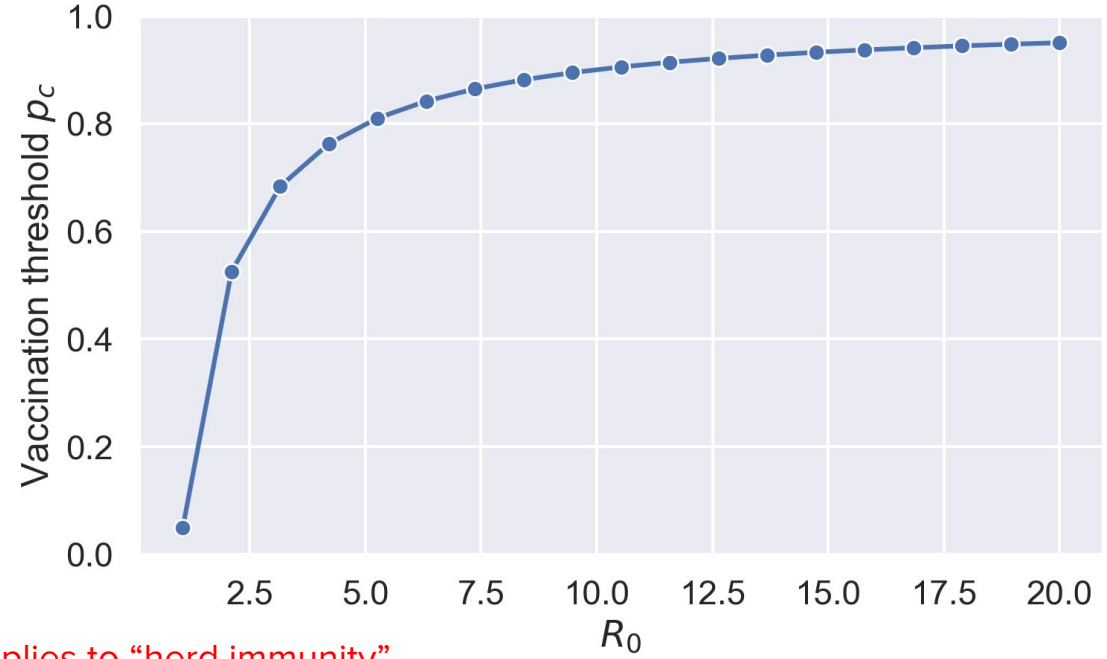
In order to eradicate a pathogen through long-term vaccination, we need:

$$R_0' = (1-p)R_0 < 1$$

We therefore need to vaccinate a critical proportion of newborns p_c:

$$p_c = 1 - \frac{1}{R_0}$$

Also applies to “herd immunity”



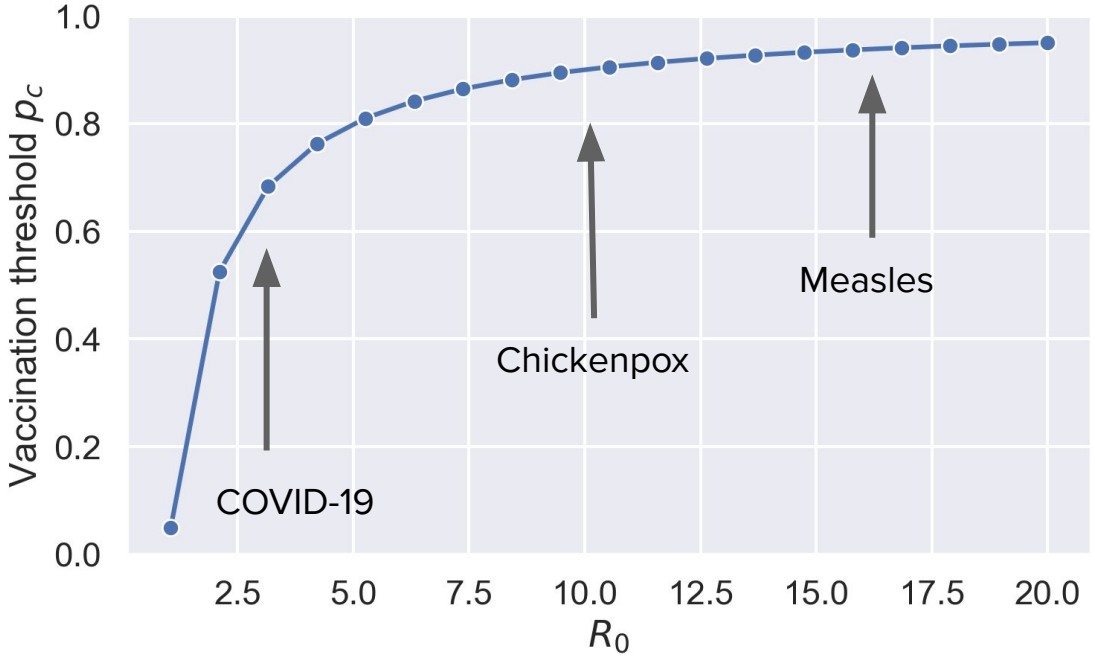
The vaccination threshold

In order to eradicate a pathogen through long-term vaccination, we need:

$$R_0' = (1-p)R_0 < 1$$

We therefore need to vaccinate a critical proportion of newborns p_c:

$$p_c = 1 - \frac{1}{R_0}$$



There are *many* common variants of the SIR model

Common SIR model variants

SI models — infections with no recovery (e.g. HIV before antiretrovirals)

SIS models — infections with no long term immunity (e.g. gonorrhea)

SEIR models — infections with an “exposed” but not yet infectious incubation period (e.g. many plant pathogens)

SIRS — immunizing infections with waning immunity (e.g. pertussis)

More complex SIR model extensions

Multi-pathogen systems with competing pathogen strains (e.g. influenza)

Multi-host systems with host heterogeneity

Host population structure

Adding heterogeneity in the host population allows us to consider multiple forms of population structure:

- Age structure
- Gender structure
- Risk/contact structure
- Spatial structure

SIR models with multiple host classes

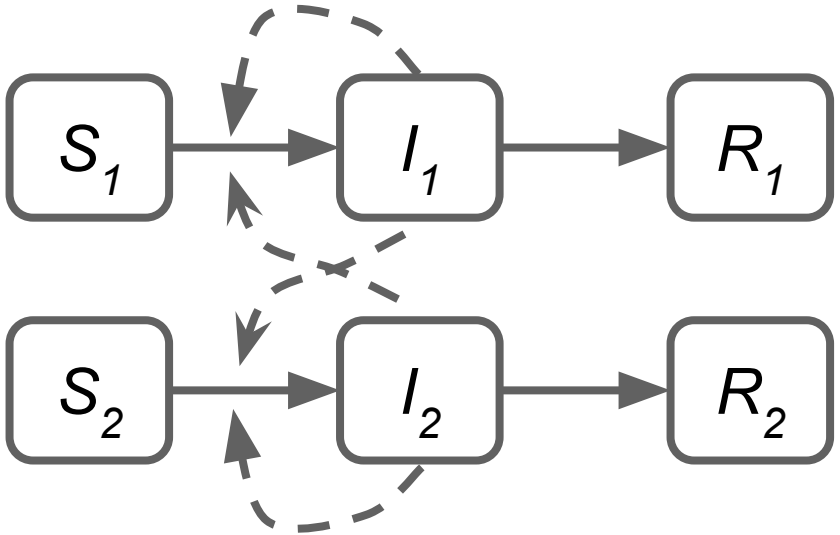
Each host as an assigned infection status and host class (e.g. S_1, S_2, I_1, I_2)

Hosts can transmit pathogens between different classes. Transmission between classes is parameterized as a transmission rate matrix:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{bmatrix}$$

SIR model with two classes

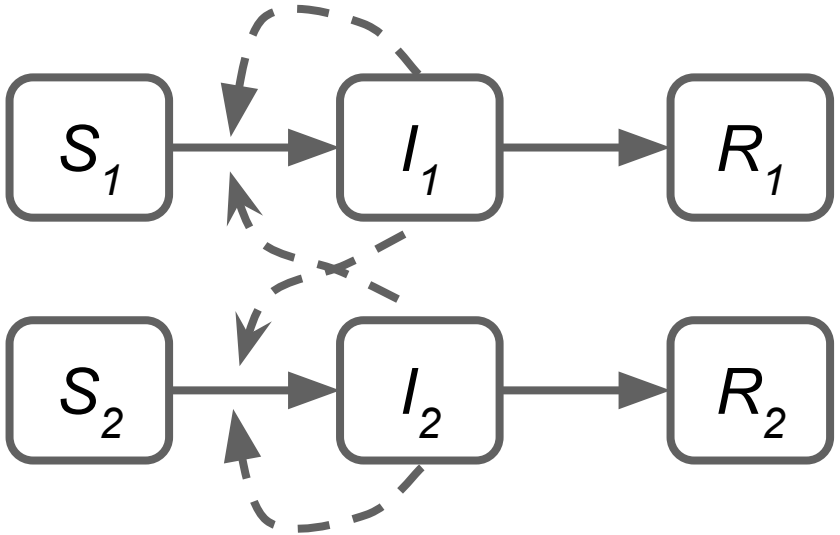
$$\frac{dI_1}{dt} = \beta_{11} S_1 I_1 + \beta_{21} S_1 I_2 - \gamma I_1$$
$$\frac{dI_2}{dt} = \beta_{22} S_2 I_2 + \beta_{12} S_2 I_1 - \gamma I_2$$



SIR model with *n* classes

This model can easily be generalized to consider any number of classes:

$$\frac{dI_i}{dt} = \sum_j \beta_{ji} S_i I_j - \gamma I_i$$



Modeling risk structure

Many host populations exhibit risk structure where there are individuals at high and low risk of becoming infected and/or transmitting.

This is one instance of the 80-20 rule where the majority of new infections (80%) result from just a small fraction of the host population (20%).

A SIS model with high/low risk groups

We'll consider a two-class SIS model as appropriate for a non-immunizing STD:

$$
\begin{aligned}
\frac{dS_H}{dt} &= -\beta_{HH} S_H I_H - \beta_{LH} S_H I_L + \gamma I_H \\
\frac{dS_L}{dt} &= -\beta_{LL} S_L I_L - \beta_{HL} S_L I_H + \gamma I_L \\
\frac{dI_H}{dt} &= \beta_{HH} S_H I_H + \beta_{LH} S_H I_L - \gamma I_H \\
\frac{dI_L}{dt} &= \beta_{LL} S_L I_L - \beta_{HL} S_L I_H - \gamma I_L
\end{aligned}
$$

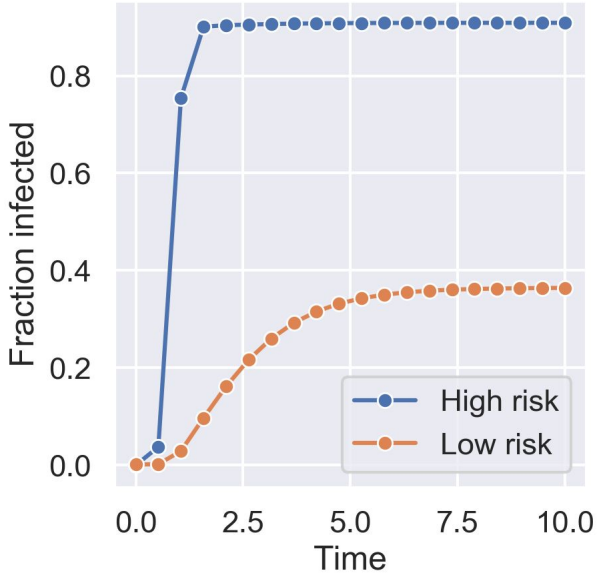
$$N_H = 200; N_L = 800$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{HH} & \beta_{HL} \\ \beta_{LH} & \beta_{LL} \end{bmatrix} = \begin{bmatrix} 10.0 & 0.6 \\ 0.6 & 1.2 \end{bmatrix}$$

An SIS model with high/low risk groups

We'll consider a two-class SIS model as appropriate for a non-immunizing STD:

$$
\begin{aligned}
\frac{dS_H}{dt} &= -\beta_{HH} S_H I_H - \beta_{LH} S_H I_L + \gamma I_H \\
\frac{dS_L}{dt} &= -\beta_{LL} S_L I_L - \beta_{HL} S_L I_H + \gamma I_L \\
\frac{dI_H}{dt} &= \beta_{HH} S_H I_H + \beta_{LH} S_H I_L - \gamma I_H \\
\frac{dI_L}{dt} &= \beta_{LL} S_L I_L - \beta_{HL} S_L I_H - \gamma I_L
\end{aligned}
$$



Targeted control with risk-structure

As before, R_0 must be less than one to prevent the epidemic from spreading and eradicate the pathogen. But how much do we need to vaccinate in each group?

Computing R_0 is now more complex, but we can compute it using a next-generation matrix that gives the expected number of new infections in each risk group arising from a single infection in every other risk group:

Targeted control with risk-structure

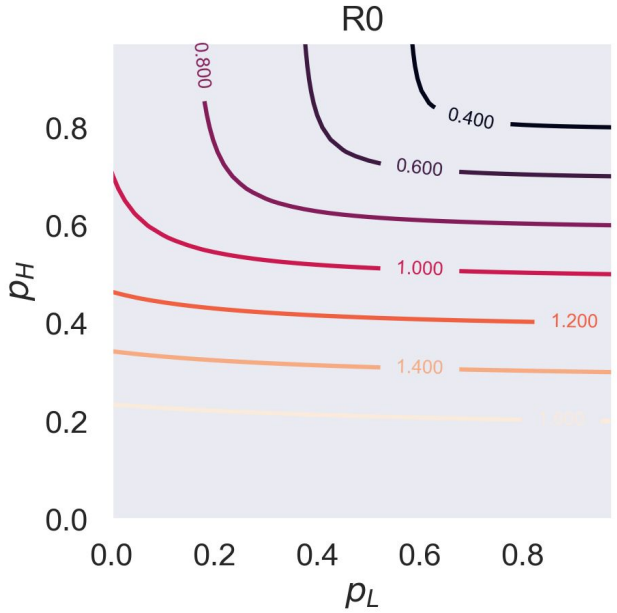
As before, R_0 must be less than one to prevent the epidemic from spreading and eradicate the pathogen. But how much do we need to vaccinate in each group?

Computing R_0 is now more complex, but we can compute it using a next-generation matrix that gives the expected number of new infections in each risk group arising from a single infection in every other risk group:

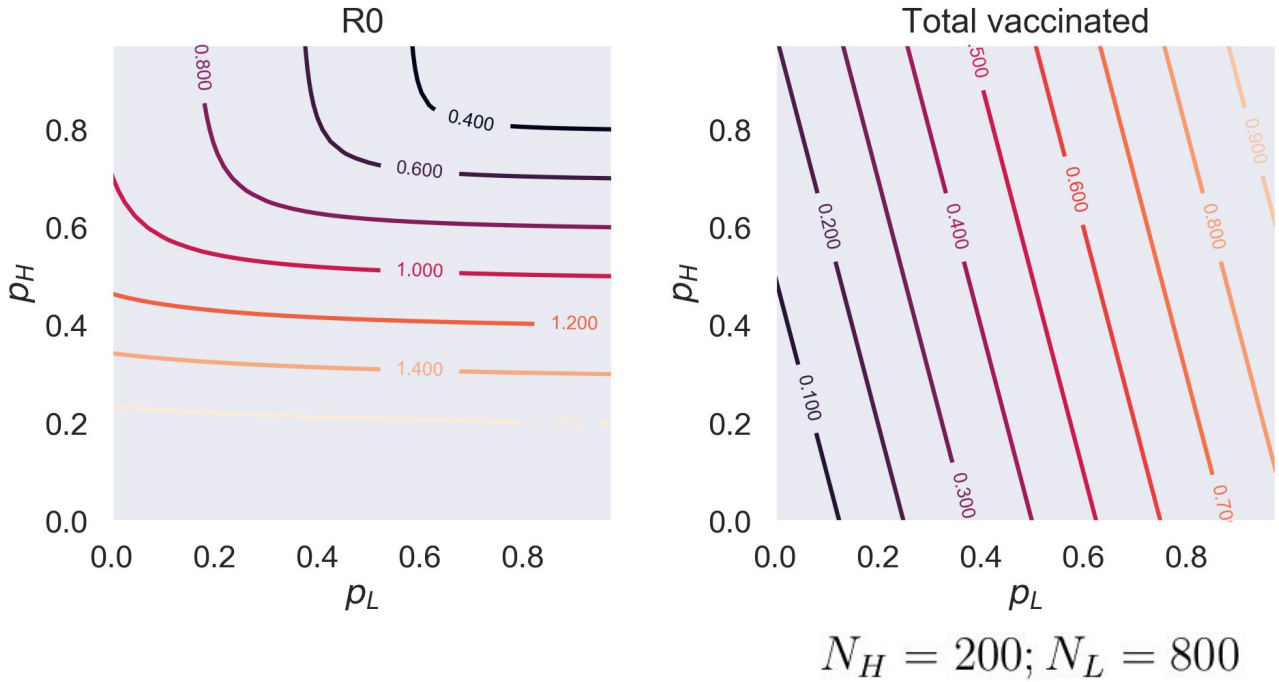
$$\boldsymbol{G} = \begin{bmatrix} (1-p_H)\frac{\beta_{HH}}{\gamma}n_H & (1-p_H)\frac{\beta_{LH}}{\gamma}n_H \\ (1-p_L)\frac{\beta_{HL}}{\gamma}n_L & (1-p_L)\frac{\beta_{LL}}{\gamma}n_L \end{bmatrix}$$

Technical note: R_0 is the leading eigenvalue of this next-generation matrix.

Targeted control with risk structure



Targeted control with risk structure



Some final thoughts

Understanding epidemic dynamics

Why do epidemics rapidly grow and then decline?

Why do some outbreaks take off and not others?

Why are some diseases endemic and others epidemic?

Why do some pathogens go extinct? How do we eliminate others?

Final comments

The flexibility of SIR-type models lets us easily model the population dynamics of many different types of pathogens with almost any type of host population structure.

Multi-host models can easily be extended to include many different subpopulations. This becomes especially useful for modeling the spatio-temporal dynamics of metapopulations.

Challenge: Parameterizing these models becomes very difficult without disaggregated data on the number and source of infections in each population. ***This is where pathogen genetic data can help us!***