



# **Advanced infectious disease modeling**

March 17 - April 4

2-4 PM

AIMS-Rwanda, Kigali

Christian Selinger, Israel Ukawuba

# Christian's Agenda:

DATE	TITLE	CONTENT
Monday March 17, 2025	Malaria model parameters	Parameterizing and simulating ODE models of malaria in R
Tuesday March 18, 2025	R0	Basic reproduction number theory and examples
Wednesday March 19, 2025	R0	Sensitivity analysis for infectious disease models: gradients, Sobol index
Thursday March 20, 2025 (T,\$)	Sensitivity analysis	Sensitivity analysis for infectious disease models: gradients, Sobol index
Friday March 21, 2025 (*)	Model fitting	Fitting model output curves to data using optimisation algorithms
WEEKEND BREAK		
Monday March 24, 2025	Statistical inference	Concepts of statistical inference (frequentist, Bayesian, maximum likelihood)
Tuesday March 25, 2025	Statistical inference	Bayesian updating, Monte Carlo integration
Wednesday March 26, 2025	Statistical inference	Monte Carlo integration
Wednesday March 26, 2025	Resistance modeling	bioinformatics meets disease modeling

(\*) Quiz

(\$) Assignment

(T) Tutorial

# Israel's Agenda:

---

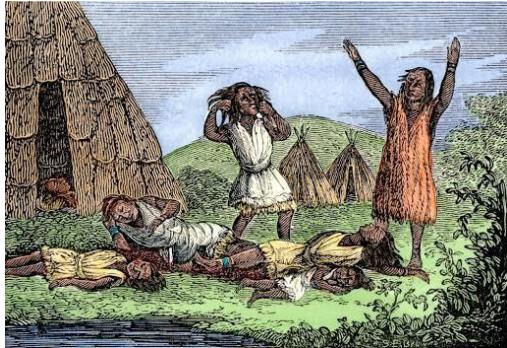
DATE	TITLE	CONTENT
Thursday March 27, 2025	Malaria ecology & climate	Lecture: intro to climate and climate regulation of malaria ecology
Friday March 28, 2025 (*)	Malaria ecology & climate	Exercise: climate and malaria ecology
WEEKEND BREAK		
Monday March 31, 2025	Climate-regulated modelling 1	Lecture: Implicit forms of climate malaria modeling
Tuesday April 1, 2025	Climate-regulated modelling 2	Lecture: Explicit forms of climate malaria modeling
Wednesday April 2, 2025	Climate modeling cont'd	Exercise: implicit and explicit models
Thursday April 3, 2025	Applied climate-driven models	Lecture: Climate models for intervention impact assessment, early warning and forecasting systems
Friday April 4, 2025 (+)	Applied climate-driven models	Exercise: climate models for intervention and forecasting

(\*) Quiz

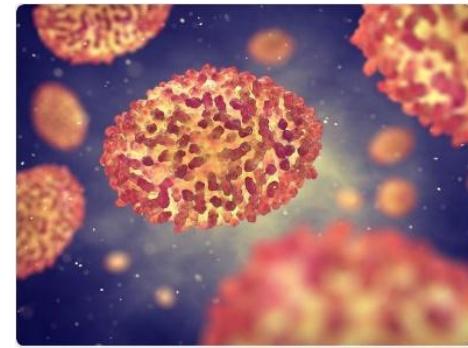
(+) group presentations

# Phenomenology – describing the invisible

Find the name of the pathogen!



Native populations  
in the Americas



**Acute** infectious disease that begins with a high **fever**, headache, and back pain and then proceeds to an **eruption on the skin** that leaves the face and limbs covered with cratered marks.

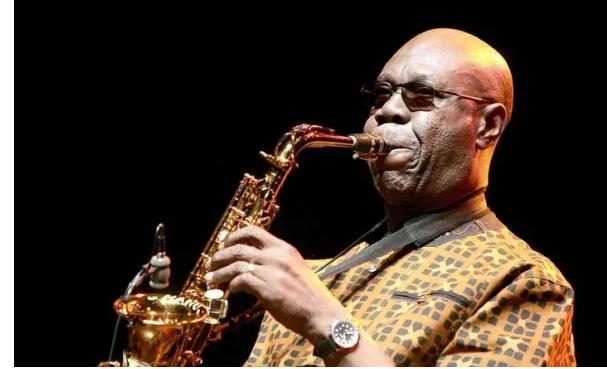
*Encyclopedia Britannica*

Smallpox virus

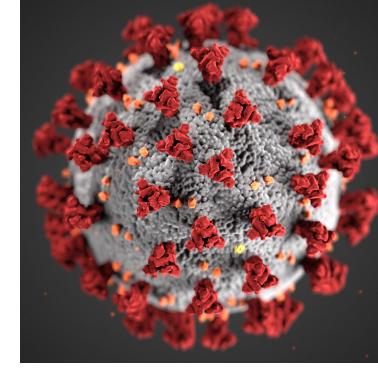
# Phenomenology – describing the invisible

Find the name of the pathogen!

**Acute** viral infection characterized primarily by fever and respiratory symptoms.



Manu Dibango



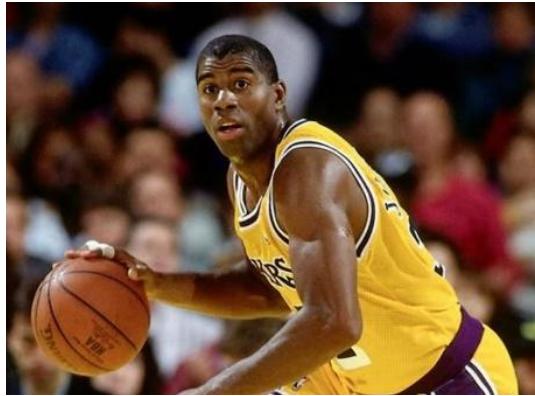
50-140 nm

*Encyclopedia Britannica*

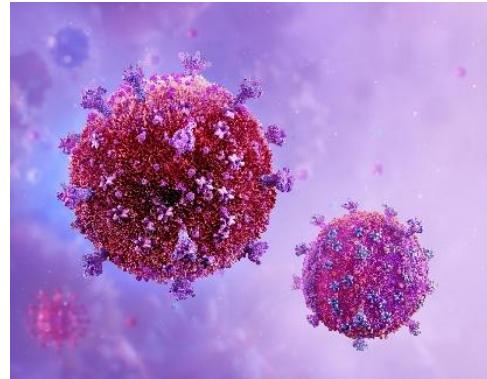
Severe Acute Respiratory Syndrome  
Coronavirus 2

# Phenomenology – describing the invisible

Find the name of the pathogen!



Ervin Johnson



120 nm

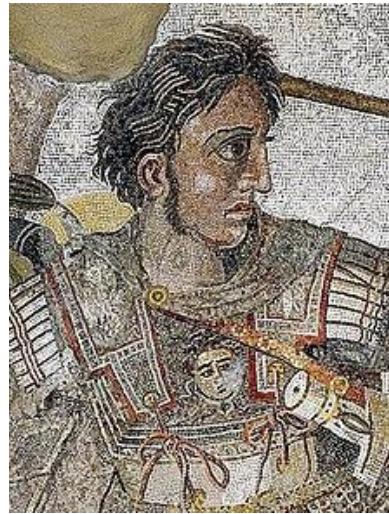
Transmissible disease of the **immune system** caused by a virus that **slowly** attacks and destroys the immune system, the body's defense against infection, leaving an **individual vulnerable** to a variety of other infections and certain malignancies that eventually cause death.

*Encyclopedia Britannica*

## Human Immuno-deficiency Virus

# Phenomenology – describing the invisible

Find the name of the pathogen!



Alexander the Great



10000-15000 nm

Many of the early Greeks thought the disease was contracted by drinking swamp water; later, the Romans attributed it to **breathing vapours**, arising from bodies of stagnant water.

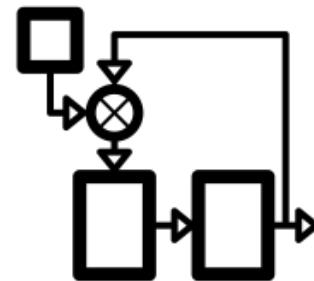
*Encyclopedia Britannica*

Plasmodium falciparum sporozoites  
(and Red Blood Cells)

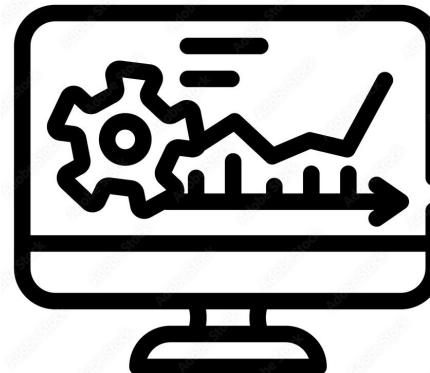
# Modeling = a process where people from different worlds meet on a common ground



**biologist/epidemiologist/malaria program:**  
*translate biological and population-level observations into flow diagrams*



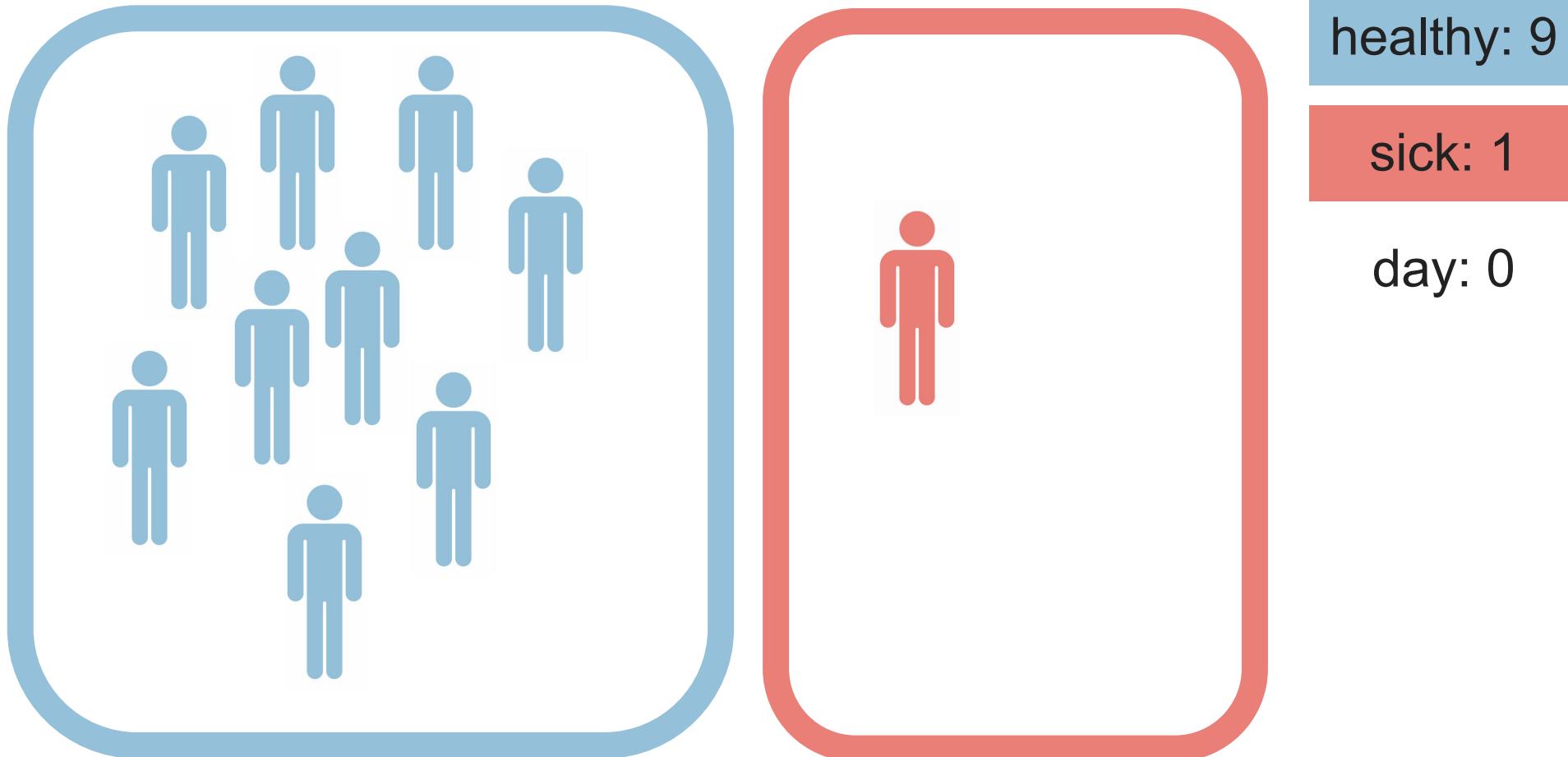
**mathematician:**  
*translates flow diagrams into equations*



**software engineer/computer science:**  
*translates equations into computer code*

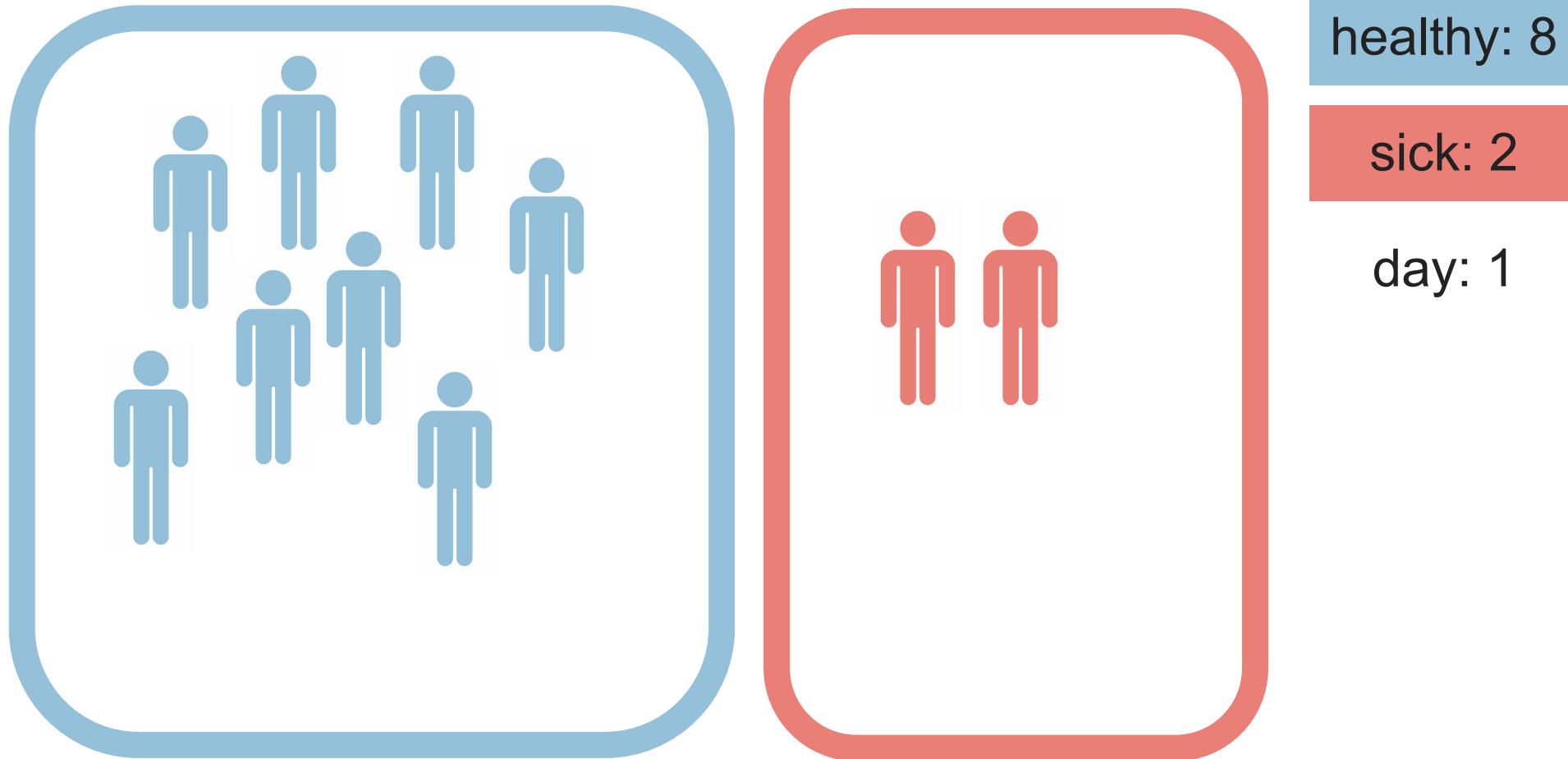
# Putting people into boxes - compartments

## Contagion process - simplified



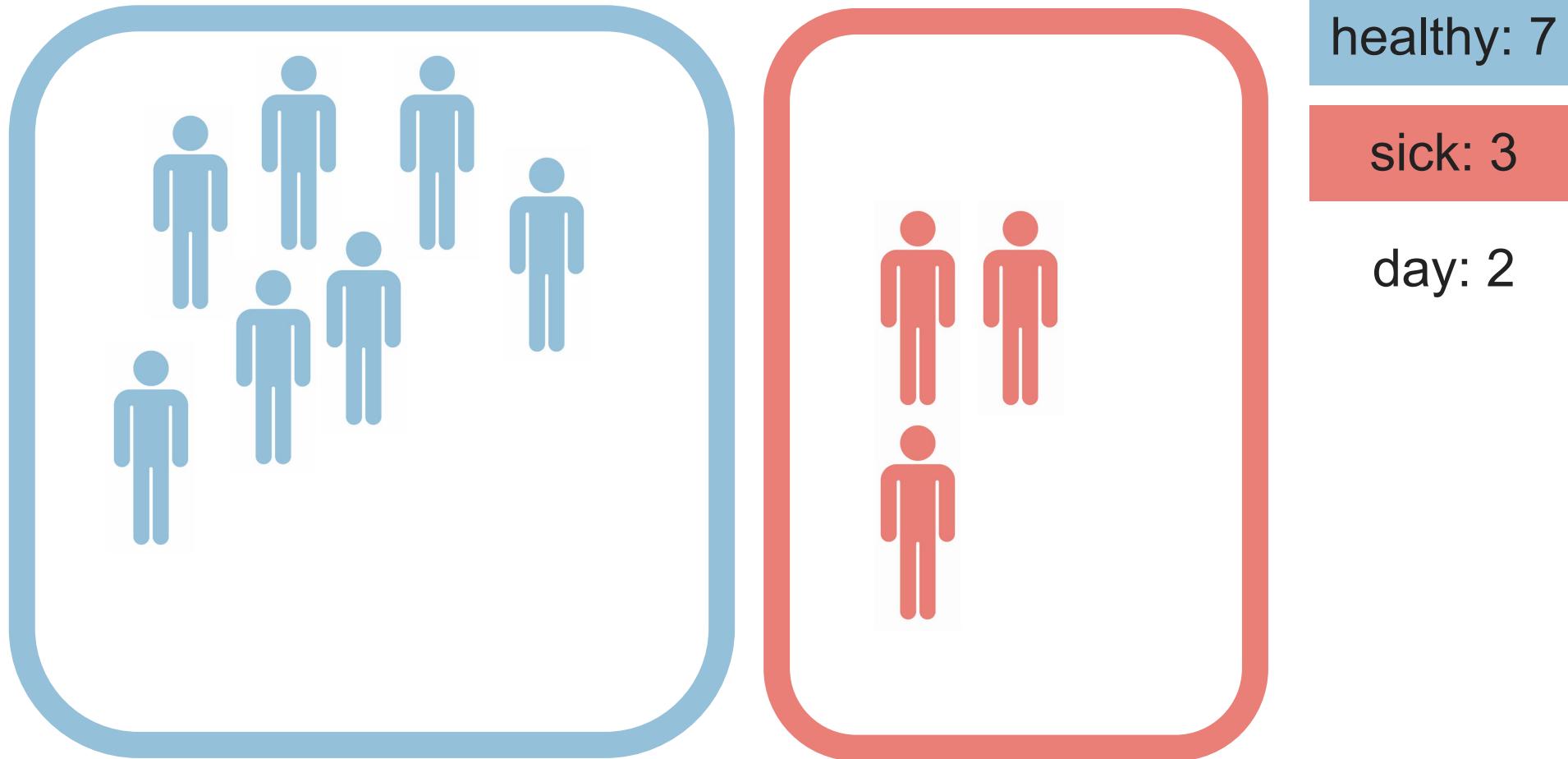
# Putting people into boxes - compartments

## Contagion process - simplified



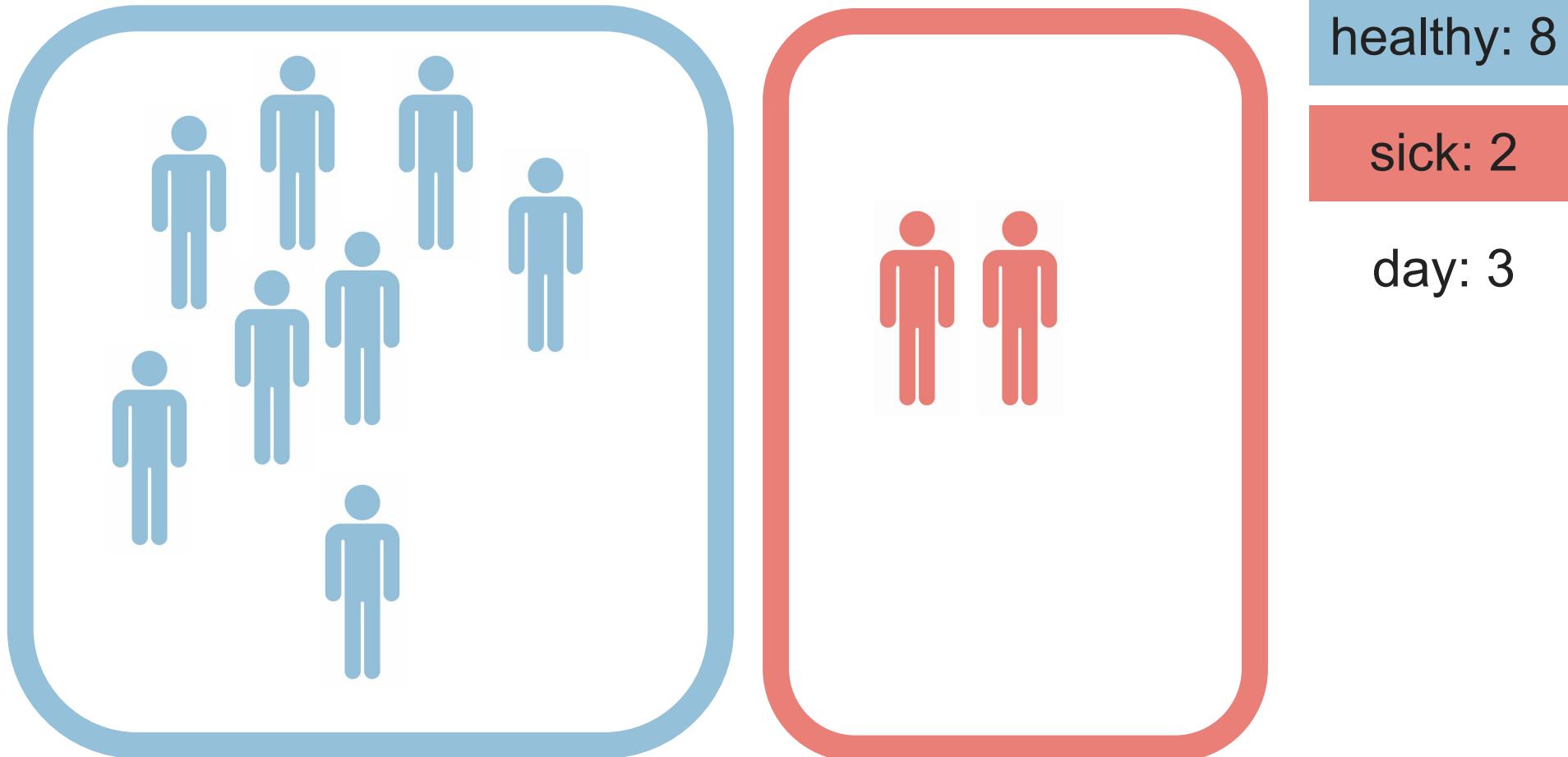
# Putting people into boxes - compartments

## Contagion process - simplified



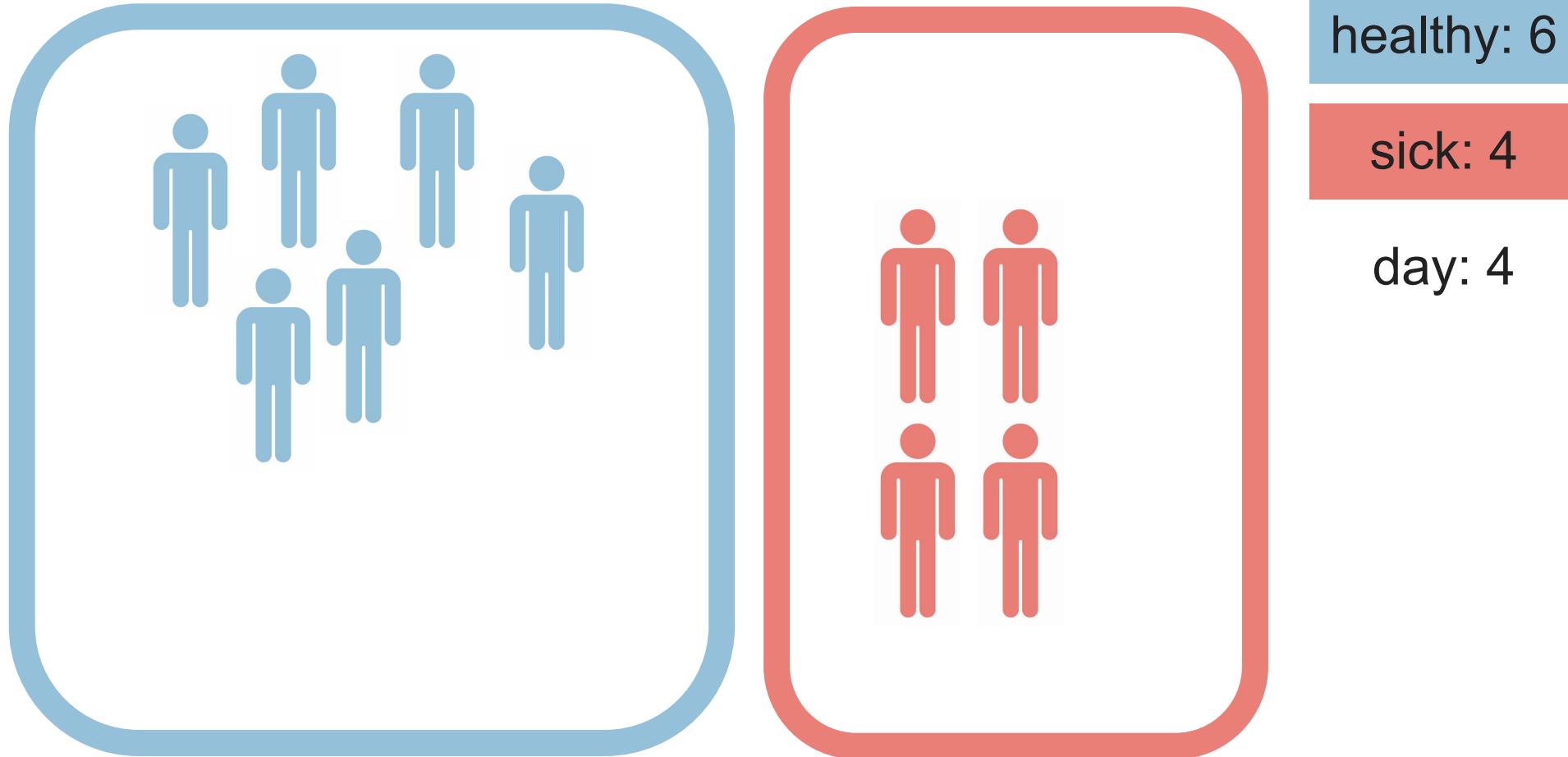
# Putting people into boxes - compartments

## Contagion process - simplified



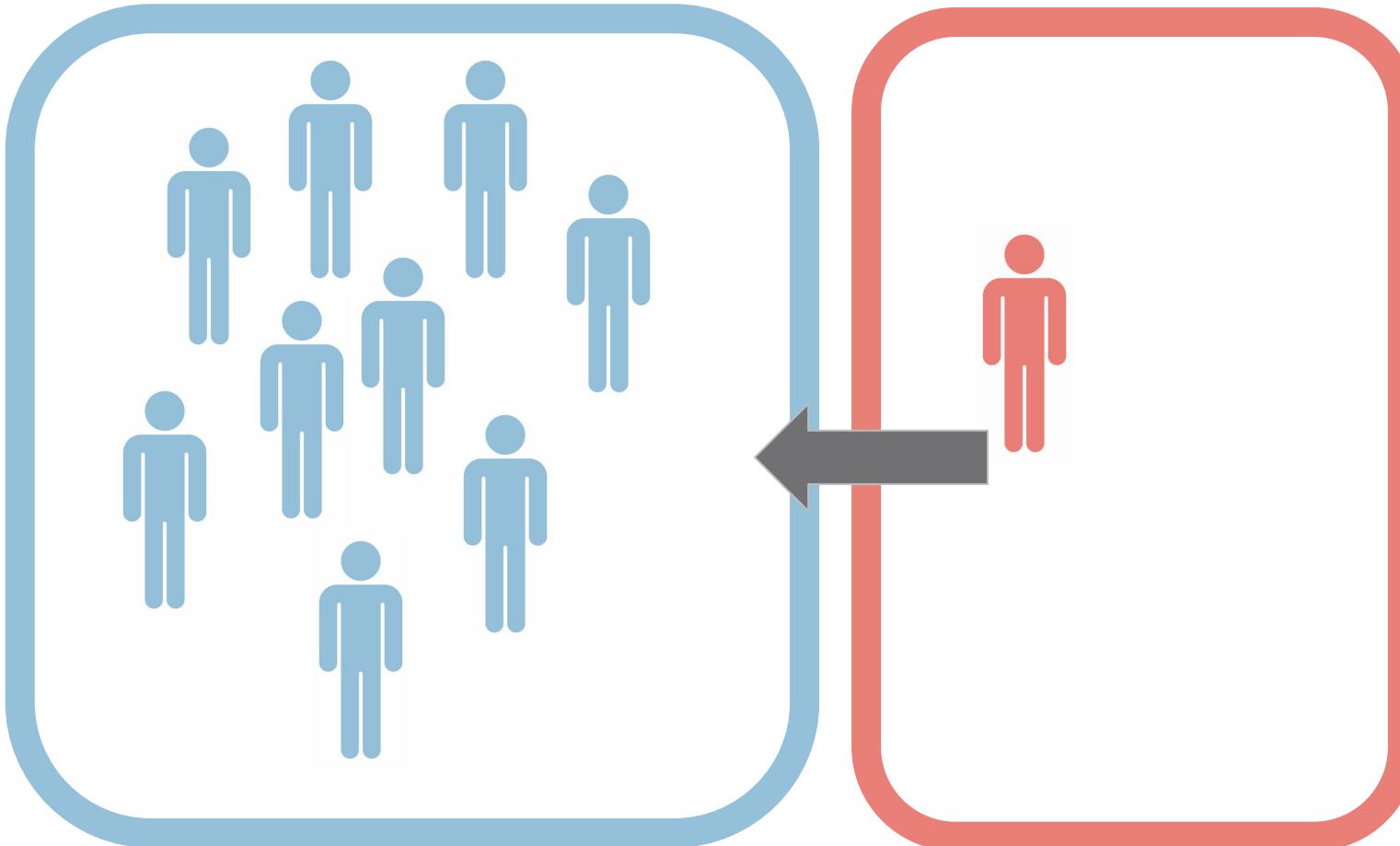
# Putting people into boxes - compartments

## Contagion process - simplified



# Transferring people between boxes - transitions

## Contagion process - simplified

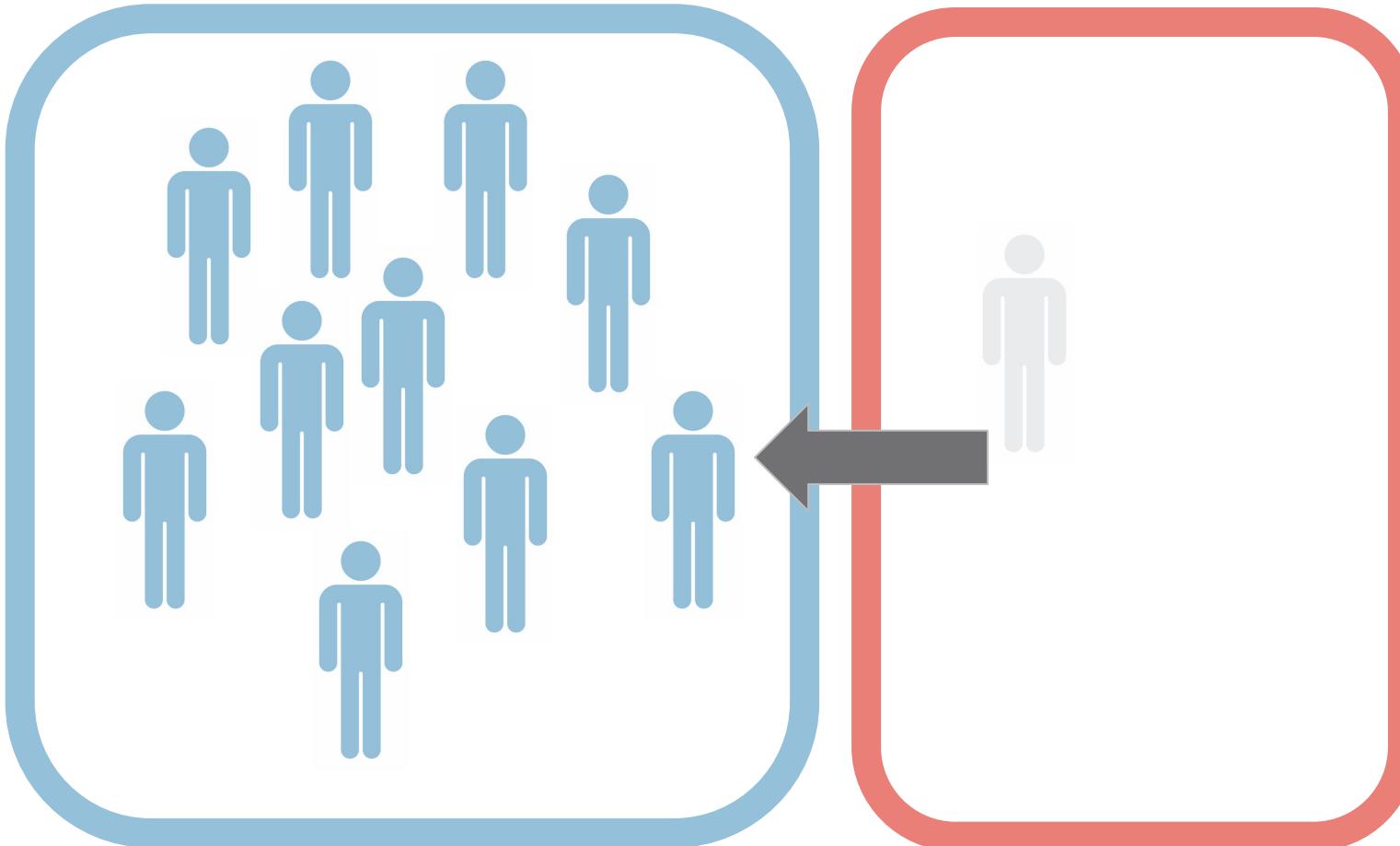


from sick to healthy  
**RECOVERY**

depends **only** on sick  
individual

# Transferring people between boxes - transitions

## Contagion process - simplified

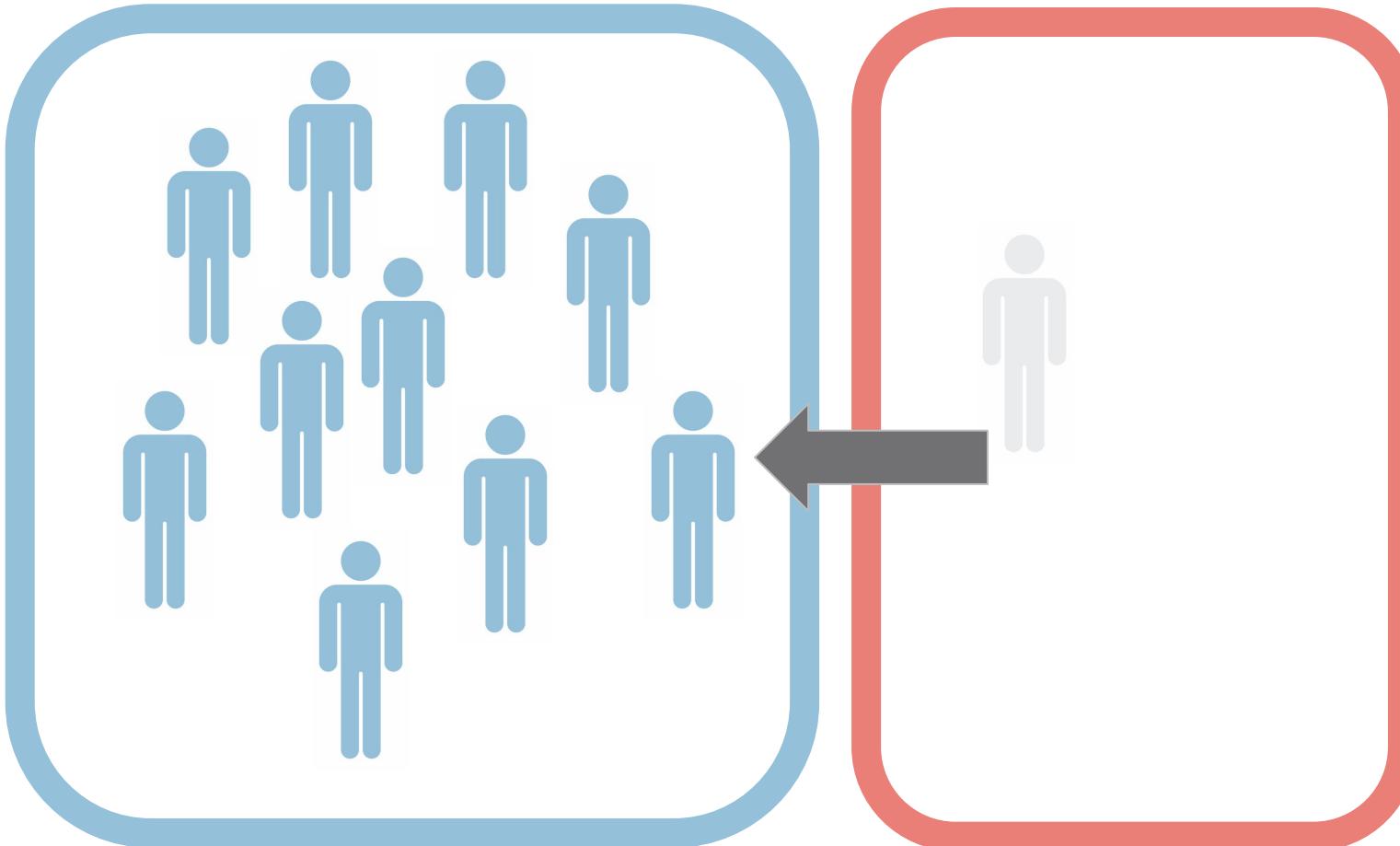


from sick to healthy  
**RECOVERY**

depends **only** on **sick**  
**individual**

# Transferring people between boxes - transitions

## Contagion process - simplified



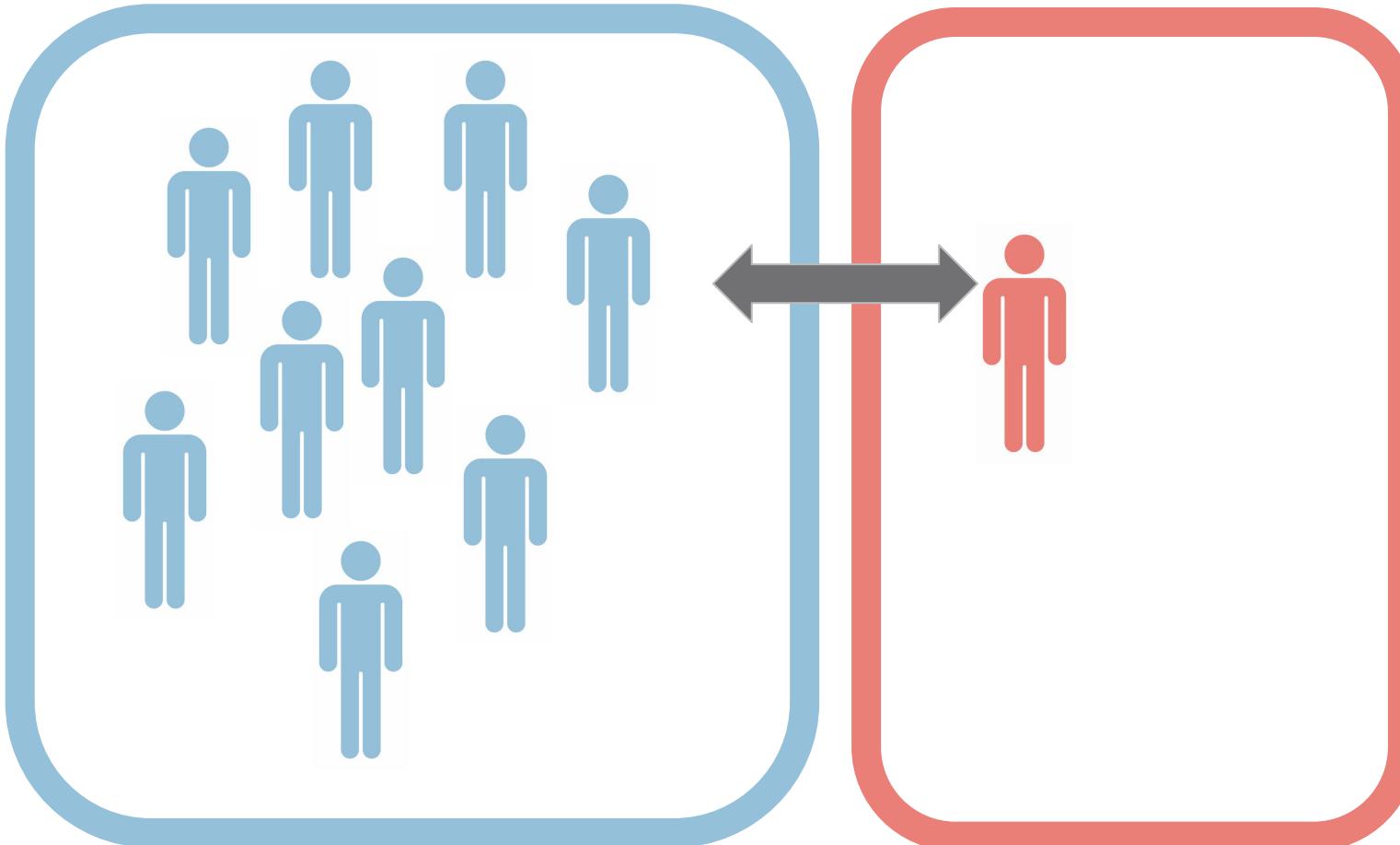
from sick to healthy  
**RECOVERY**

depends **only** on sick  
individual



# Transferring people between boxes - transitions

## Contagion process - simplified

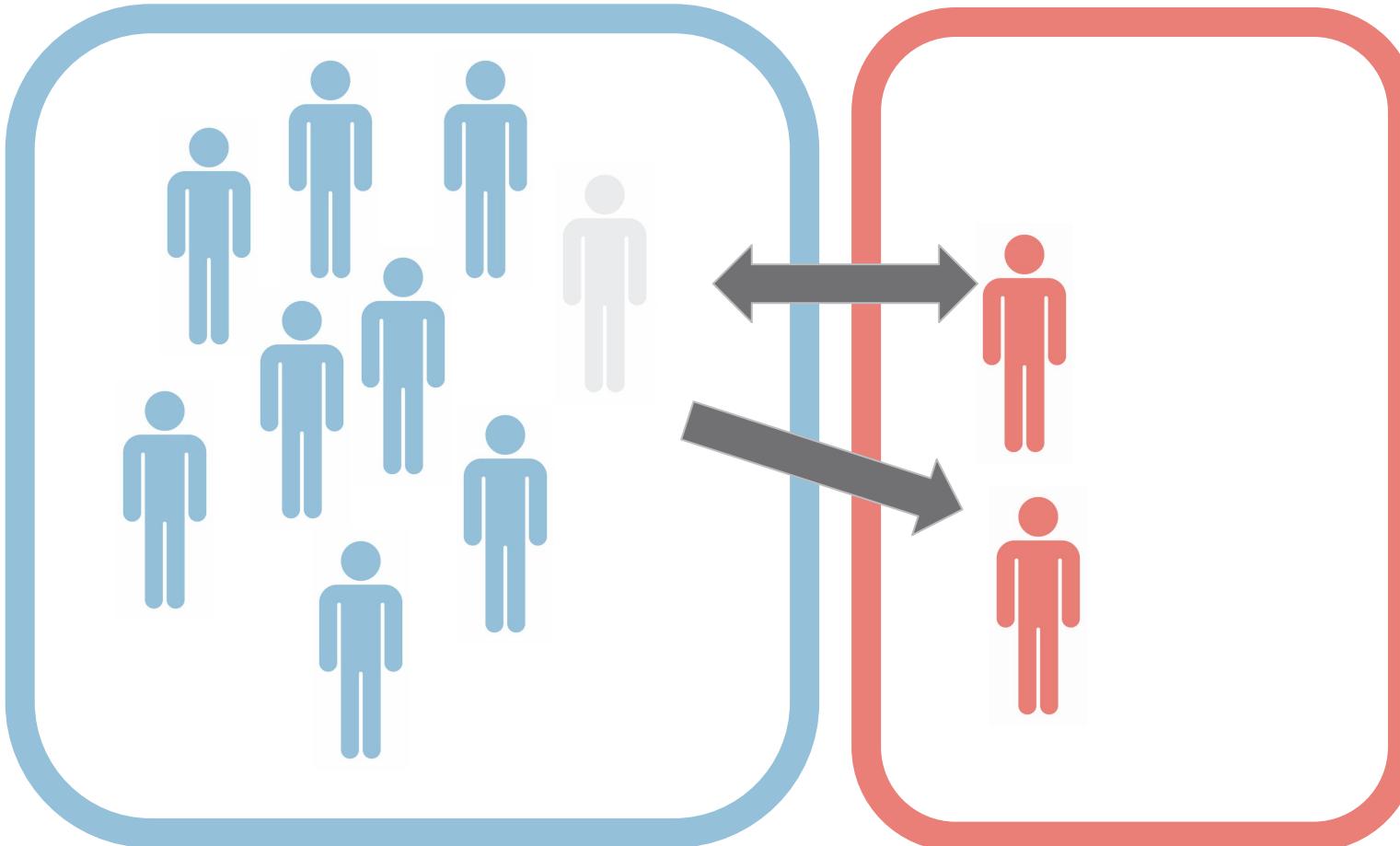


from healthy to sick  
**TRANSMISSION**

depends on **encounter**  
between sick and healthy  
individual

# Transferring people between boxes - transitions

## Contagion process - simplified

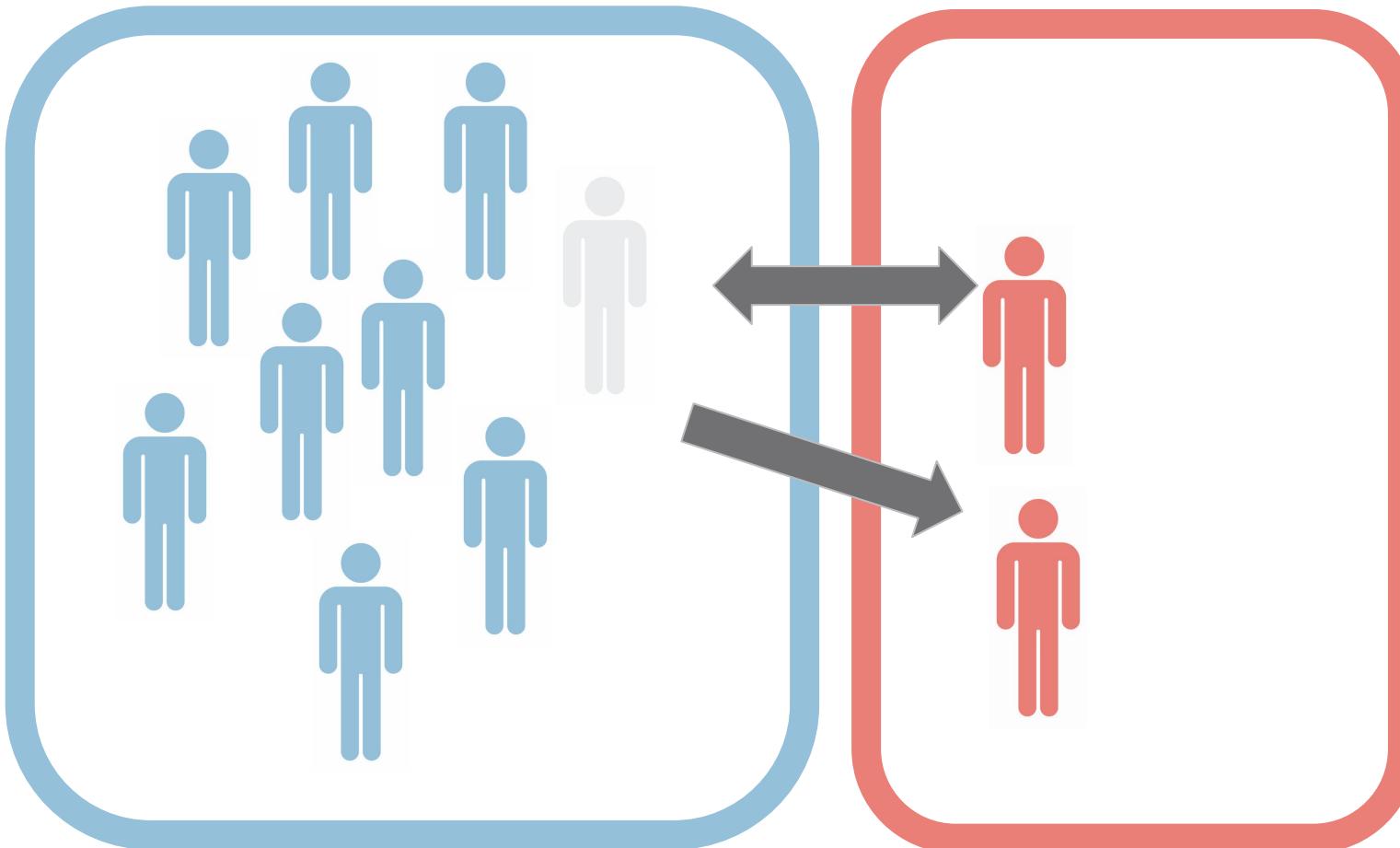


from healthy to sick  
**TRANSMISSION**

depends on **encounter**  
between sick and healthy  
individual

# Transferring people between boxes - transitions

## Contagion process - simplified



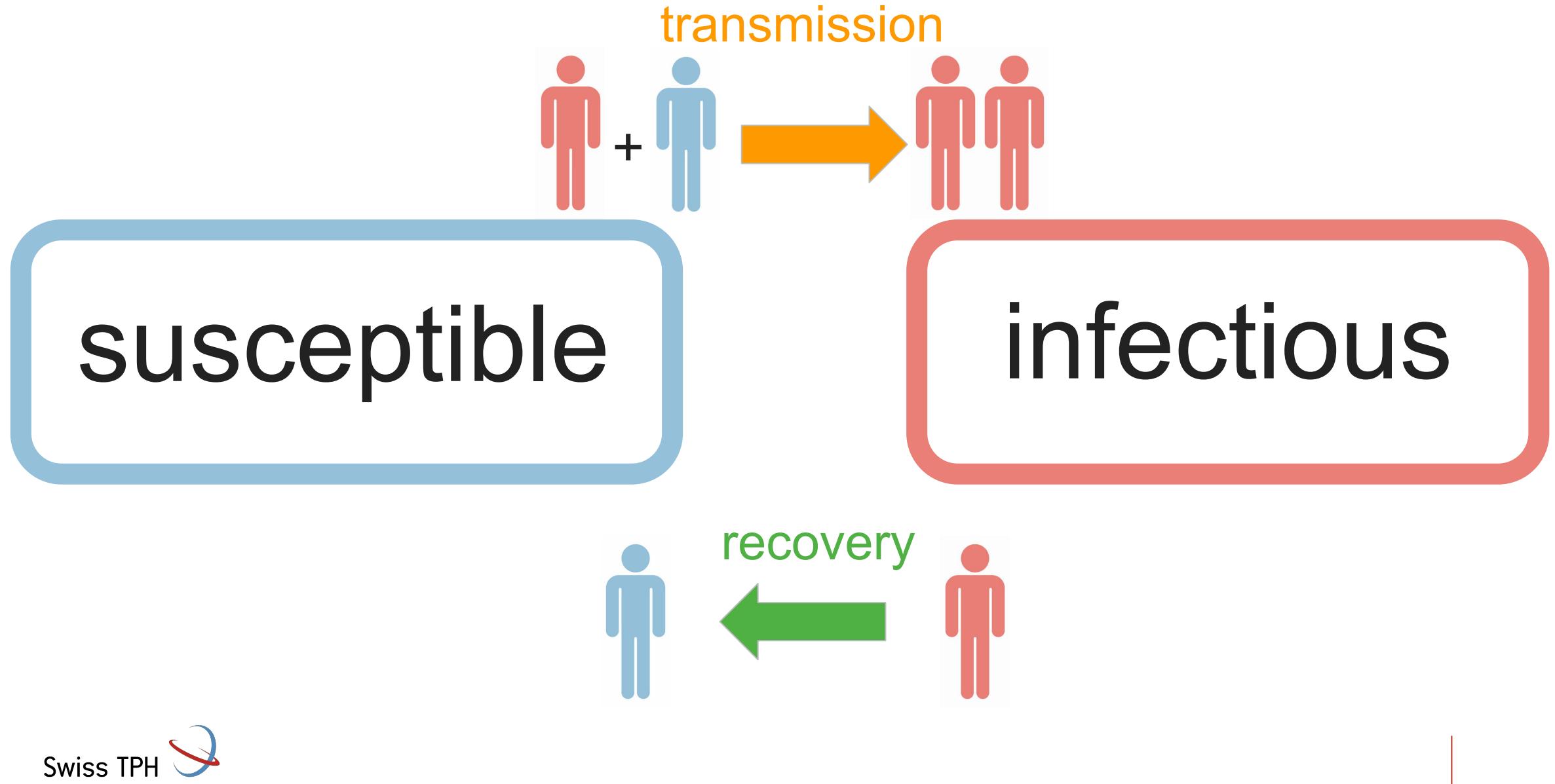
from healthy to sick  
**TRANSMISSION**

depends on **encounter**  
between sick and healthy  
individual

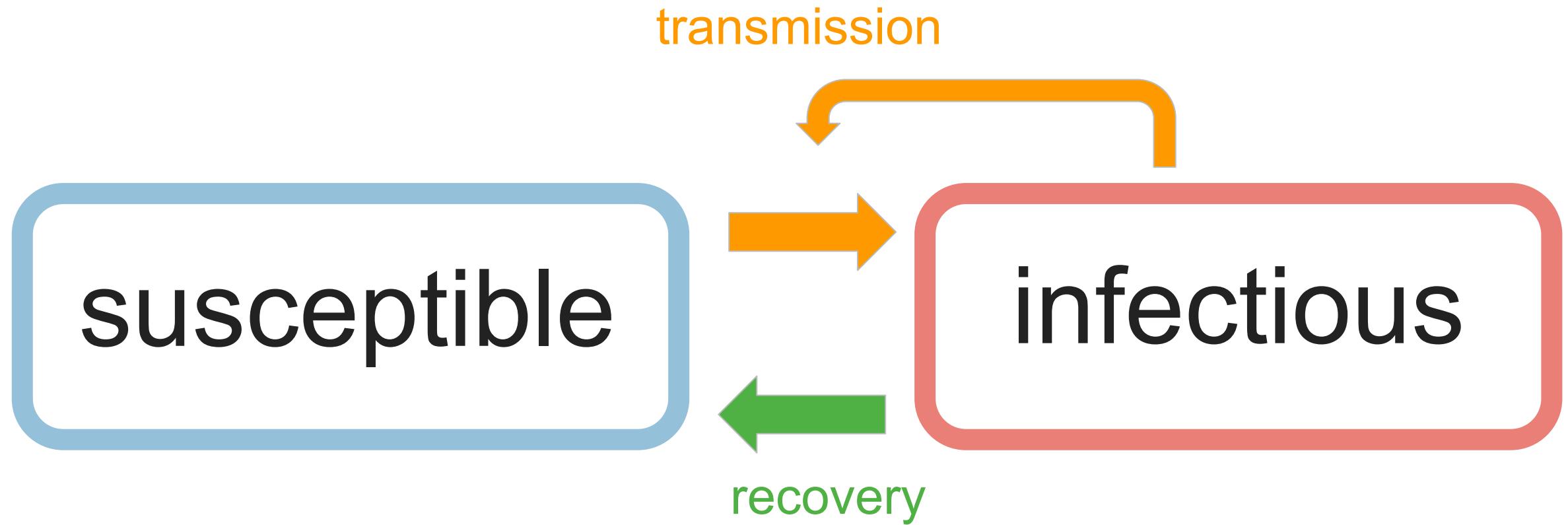


*“transmission kernel”*

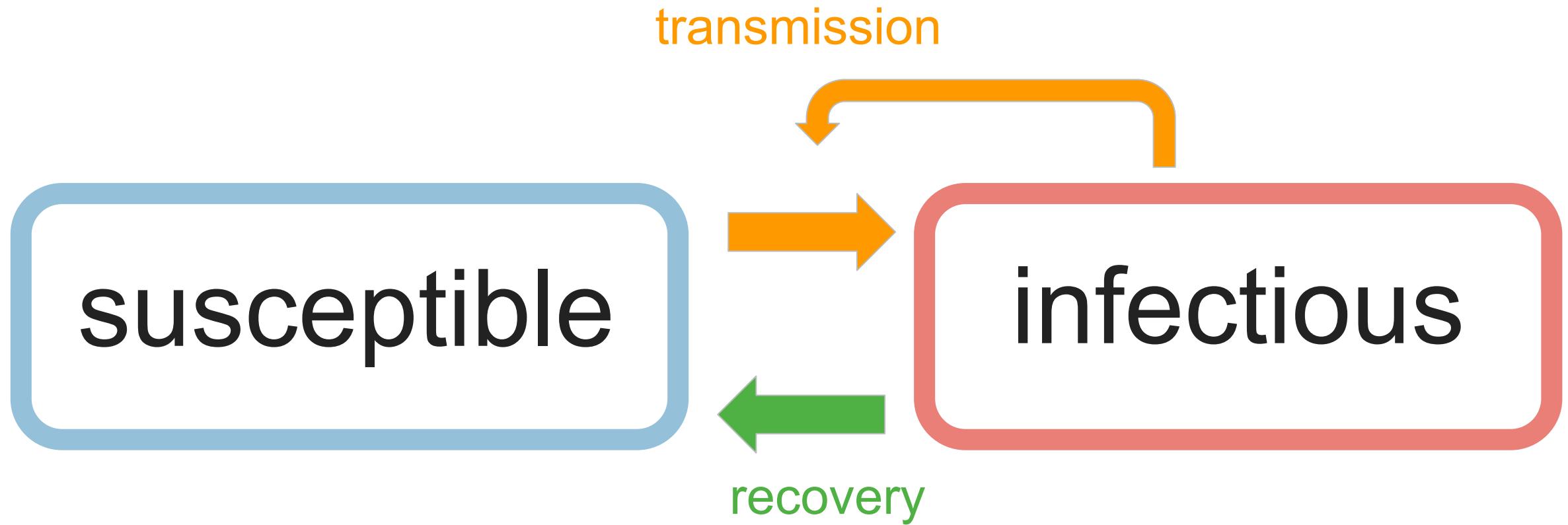
# Flow diagrams for disease transmission



# Flow diagrams for disease transmission



# Flow diagrams for disease transmission



*SIS model*

# Flow diagrams for disease transmission: assumptions!

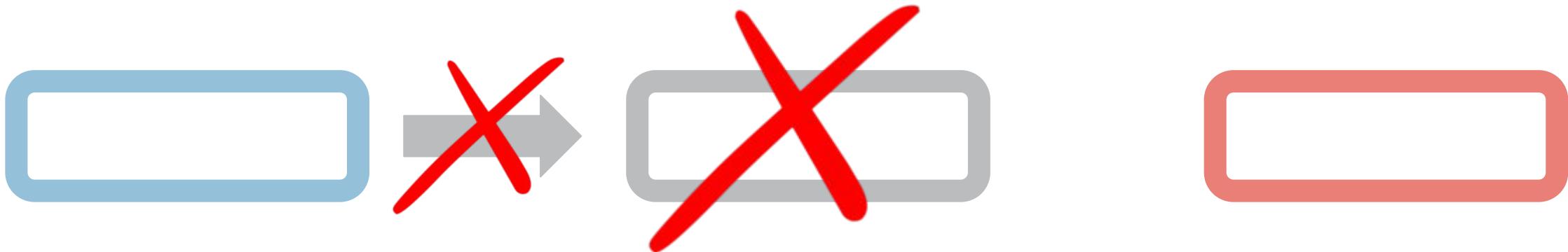
We need to be clear about the underlying assumptions:

**Hypothesis A:** Individuals become immediately also infectious upon infection

**Hypothesis B:** There are no births or deaths in the population, there is no migration.

**Hypothesis C:** Once an individual recovers, it will become immediately susceptible again.

**Hypothesis D:** All susceptible hosts are equally likely to meet an infected host.



# Flow diagrams for disease transmission: assumptions!

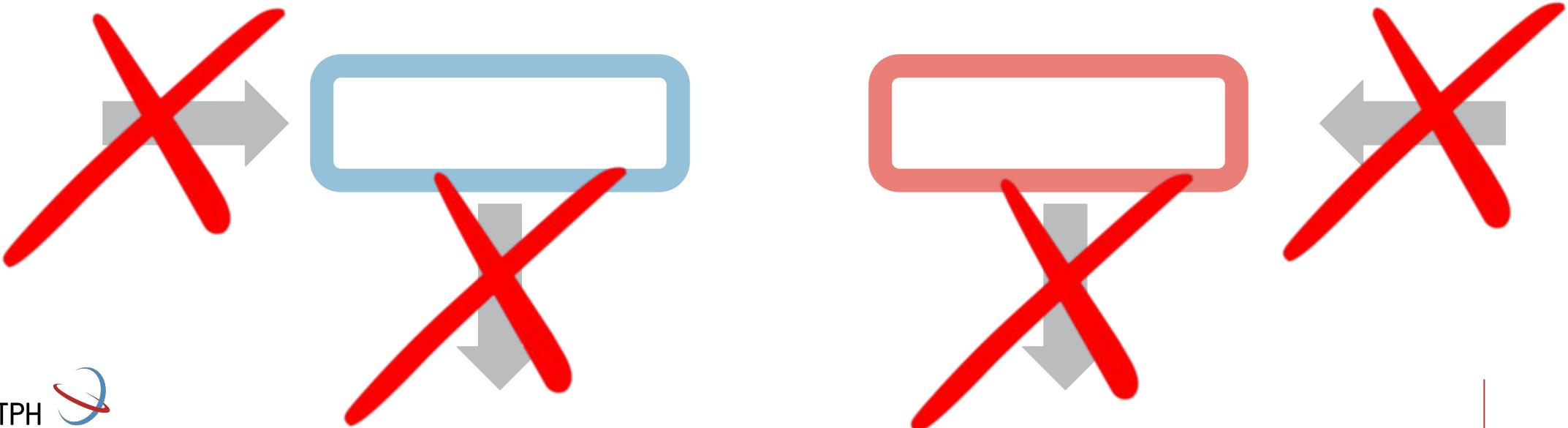
We need to be clear about the underlying assumptions:

**Hypothesis A:** Individuals become immediately also infectious upon infection

**Hypothesis B:** There are no births or deaths in the population, there is no migration.

**Hypothesis C:** Once an individual recovers, it will become immediately susceptible again.

**Hypothesis D:** All susceptible hosts are equally likely to meet an infected host.



# Flow diagrams for disease transmission: assumptions!

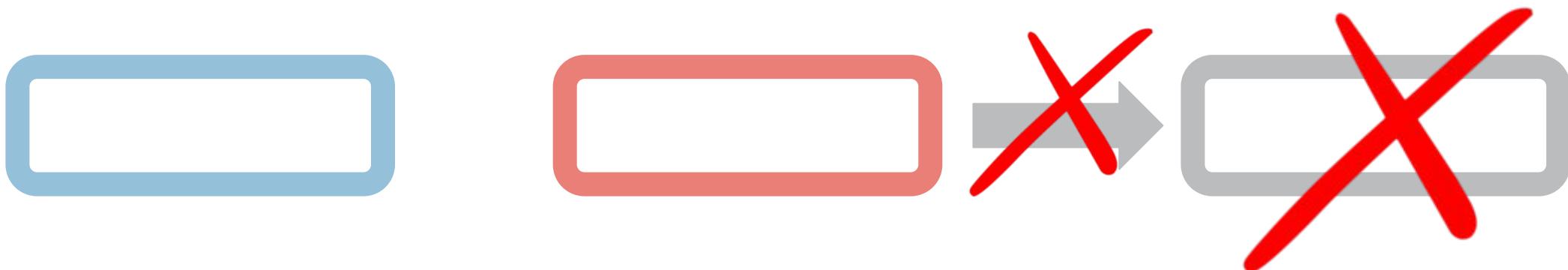
We need to be clear about the underlying assumptions:

**Hypothesis A:** Individuals become immediately also infectious upon infection

**Hypothesis B:** There are no births or deaths in the population, there is no migration.

**Hypothesis C:** Once an individual recovers, it will become immediately susceptible again.

**Hypothesis D:** All susceptible hosts are equally likely to meet an infected host.



# Flow diagrams for disease transmission: assumptions!

We need to be clear about the underlying assumptions:

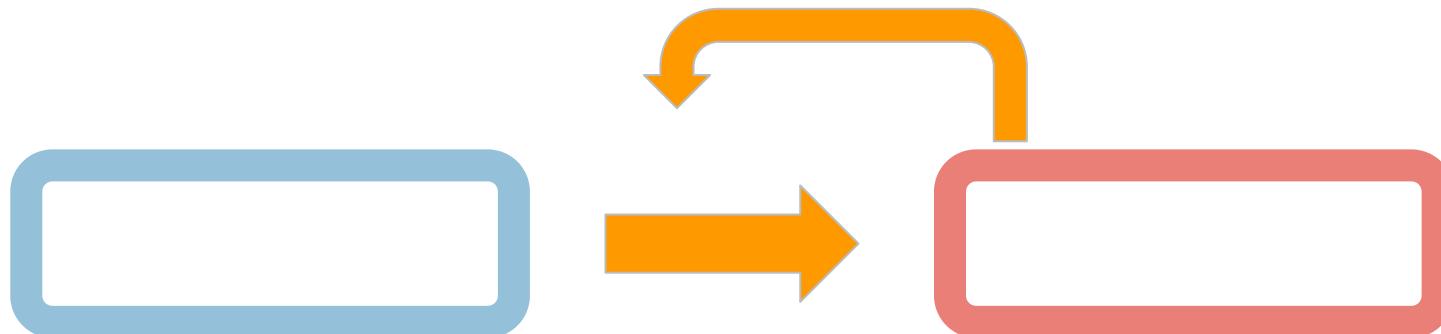
**Hypothesis A:** Individuals become immediately also infectious upon infection

**Hypothesis B:** There are no births or deaths in the population, there is no migration.

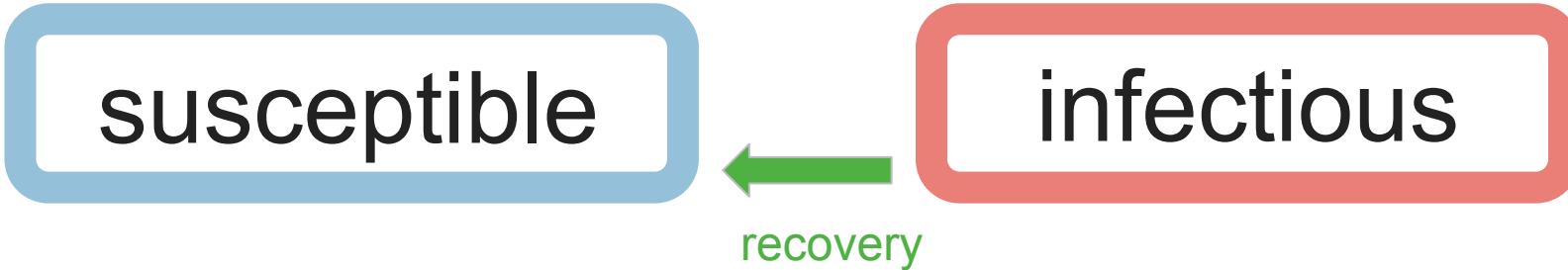
**Hypothesis C:** Once an individual recovers, it will become immediately susceptible again.

**Hypothesis D: All susceptible hosts are equally likely to meet an infected host.**

rate for new infection depends on relative proportion of infectious individuals



# Flow diagrams as differential equations



inflow **into** susceptible

$$S(t + 1) - S(t) = \gamma I(t)$$

recovery

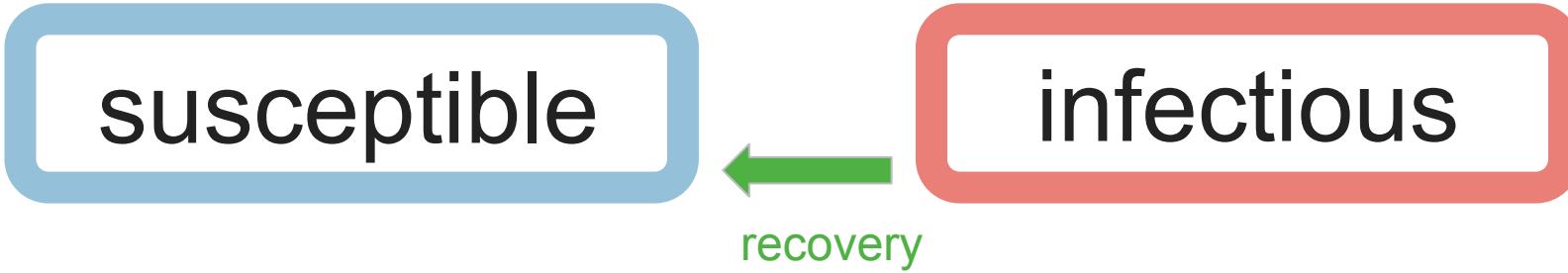


What is  $\gamma$  ?

**rate** =expected number of events per time unit



# Flow diagrams as differential equations



inflow **into** susceptible

$$S(t + 1) - S(t) = \gamma I(t)$$

outflow **from** infectious

$$I(t + 1) - I(t) = -\gamma I(t)$$

recovery



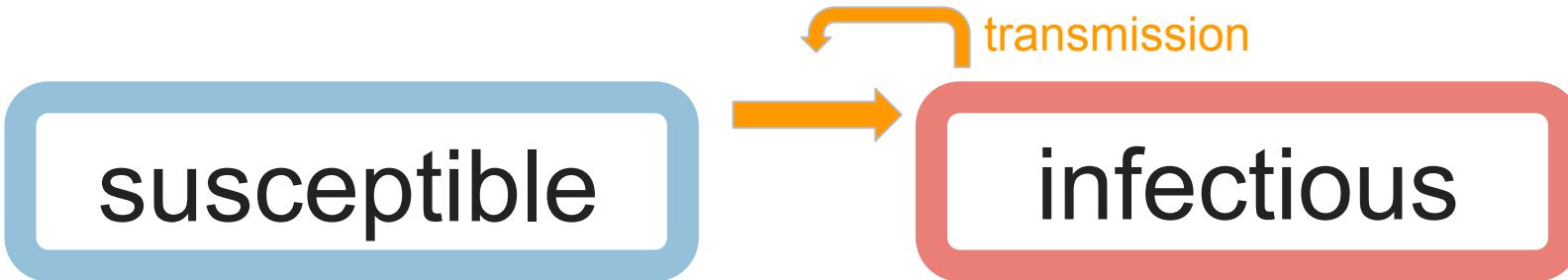
What is  $\gamma$  ?

**rate** =expected number of events per time unit



recovery **rate**  $\gamma$

# Flow diagrams as differential equations



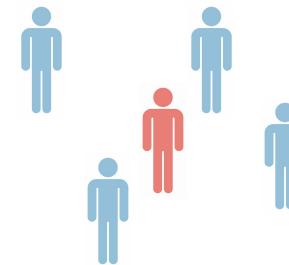
outflow **from** susceptible

$$S(t+1) - S(t) = -\beta \underbrace{\frac{I(t)}{S(t)+I(t)}}_{\text{probability of finding an infectious individual in the population}} S(t)$$

probability of finding an infectious individual  
in the population

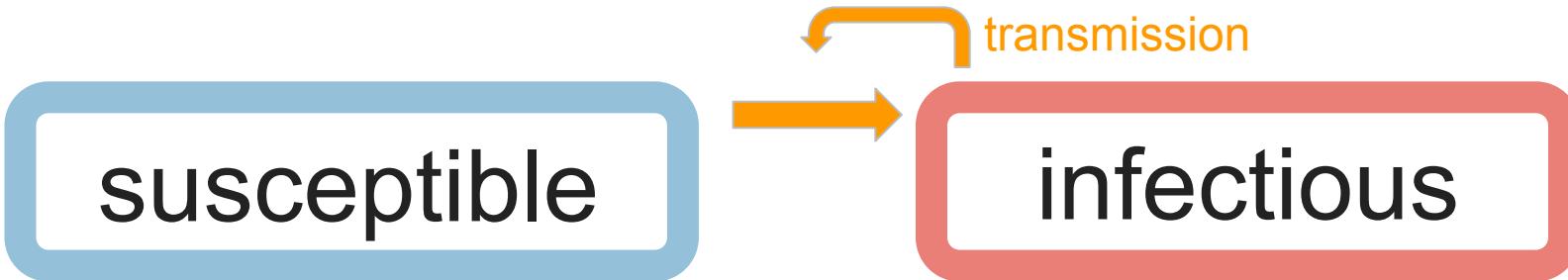
inflow **into** infectious

$$I(t+1) - I(t) = \beta \frac{I(t)}{S(t)+I(t)} S(t)$$



**assumption:** all individuals are well mixed

# Flow diagrams as differential equations



outflow **from** susceptible

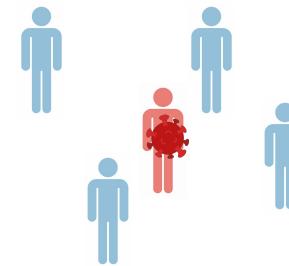
$$S(t+1) - S(t) = -\beta \frac{I(t)}{S(t)+I(t)} S(t)$$

transmission rate

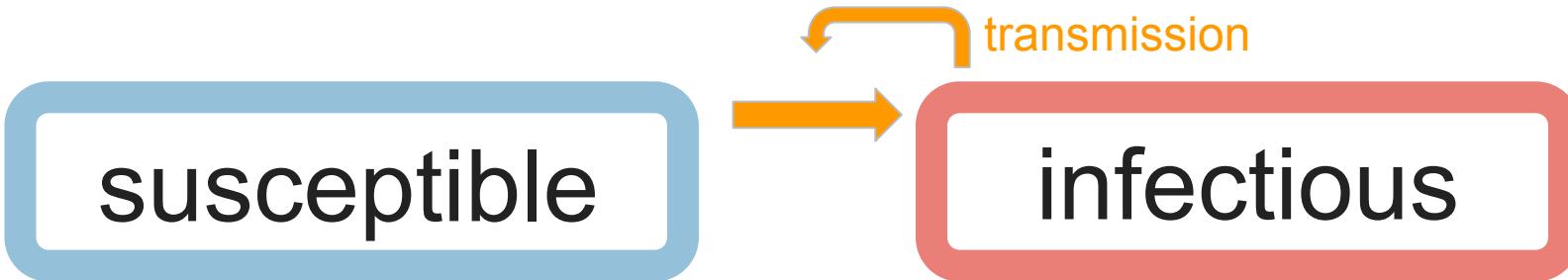
**assumption:** transmission is constant  
in time and across infectious  
individuals

inflow **into** infectious

$$I(t+1) - I(t) = \beta \frac{I(t)}{S(t)+I(t)} S(t)$$



# Flow diagrams as differential equations



outflow **from** susceptible

$$S(t+1) - S(t) = -\beta \frac{I(t)}{S(t)+I(t)} S(t)$$

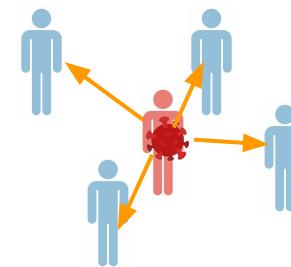


currently susceptible individuals

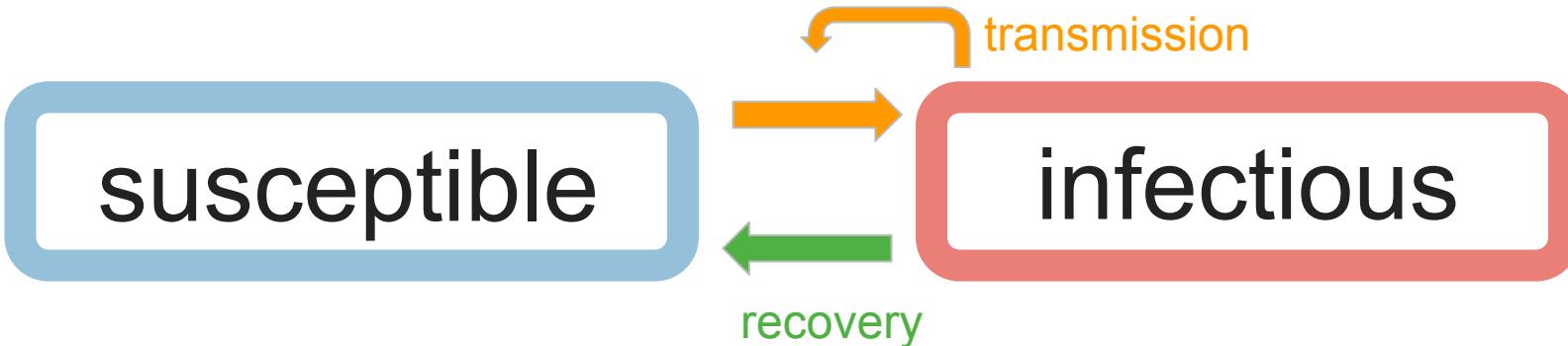
**assumption:** new infection happens  
**independently** and with **same probability**  
to each susceptible individual

inflow **into** infectious

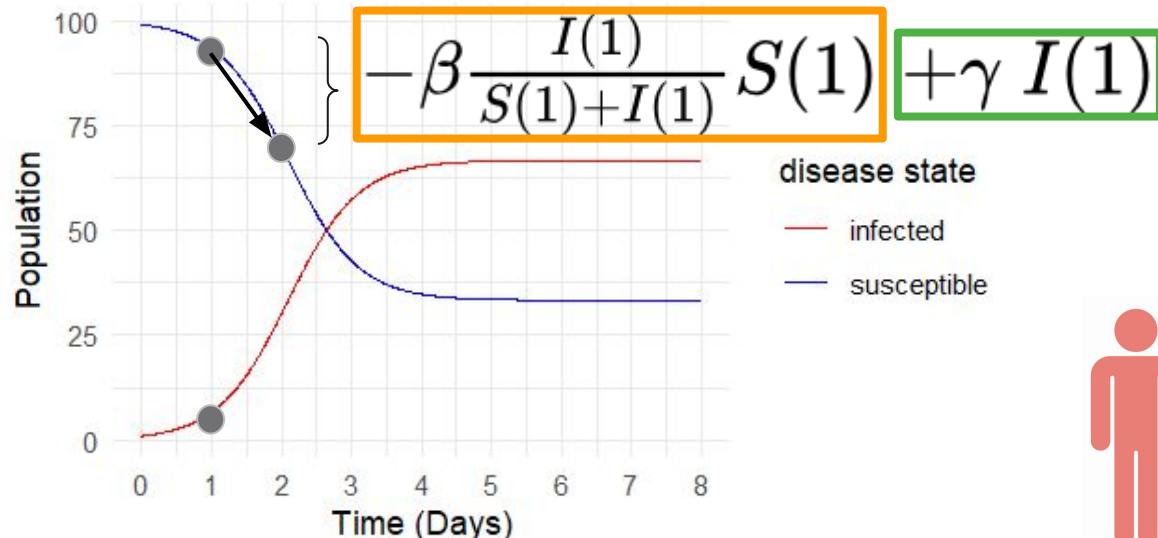
$$I(t+1) - I(t) = \beta \frac{I(t)}{S(t)+I(t)} S(t)$$



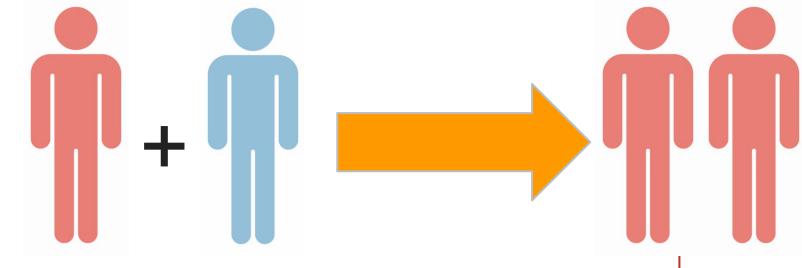
# Flow diagrams as differential equations



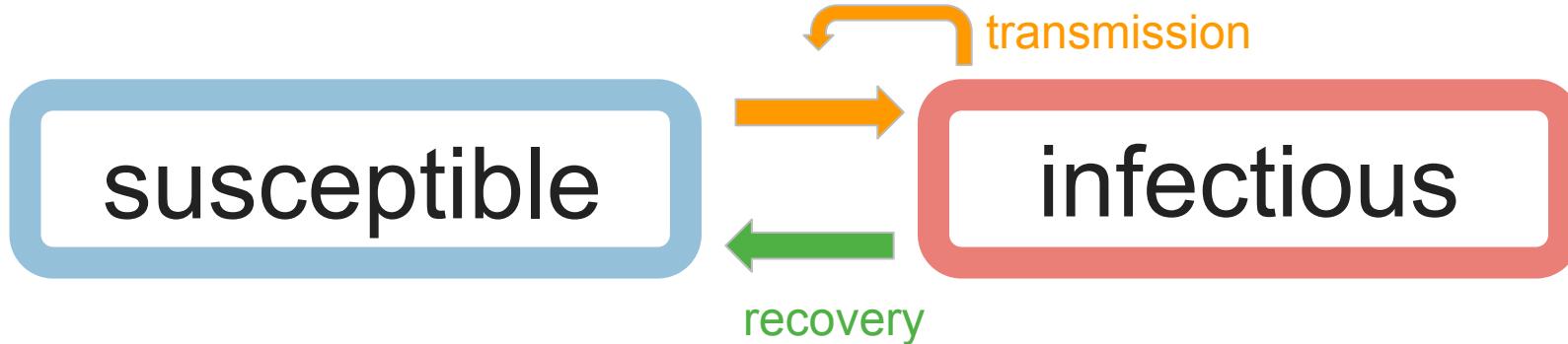
$$S(t + 1) - S(t) = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t)$$



*“coupled system”*



# Flow diagrams as differential equations



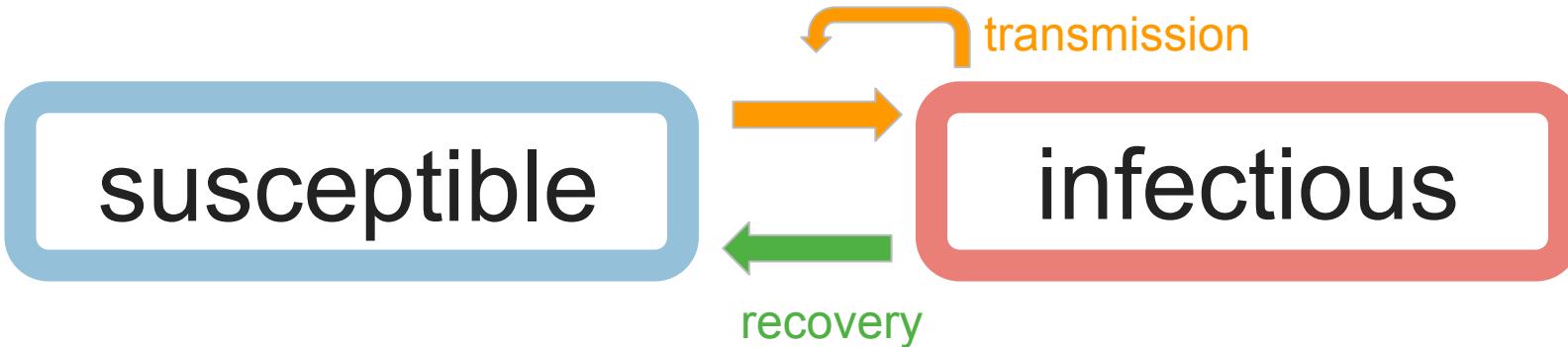
$$\left\{ \begin{array}{l} \frac{S(t+\Delta)-S(t)}{\Delta} = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \\ \frac{I(t+\Delta)-I(t)}{\Delta} = \beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \end{array} \right.$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \\ S(t) + I(t) &= N \end{aligned}$$

$$\Delta \rightarrow 0$$

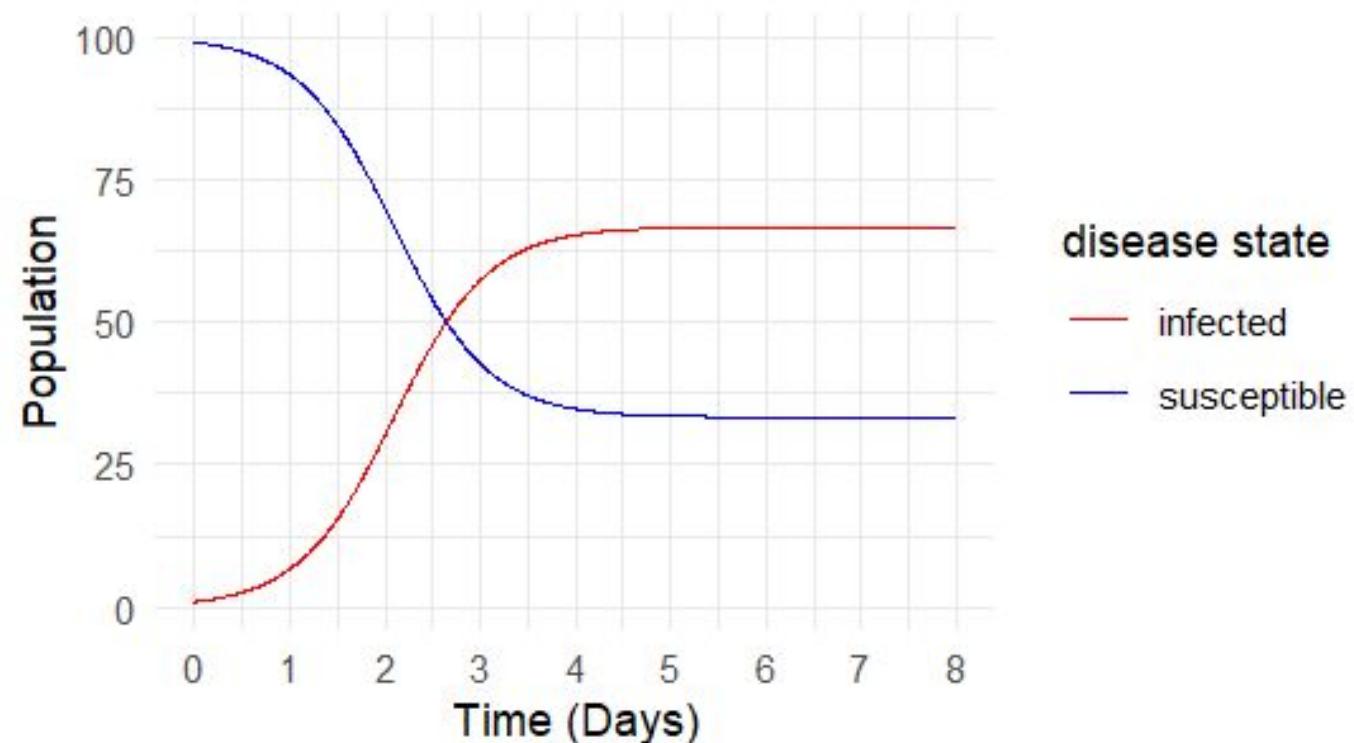
*SIS differential equation system*

# Flow diagrams as differential equations



## transmission kernel:

- **depletion** of susceptibles: as there are fewer susceptible individuals in the population, infections also stall
- **non-linearity**: new infections are generally **not proportional** to current infections



# Solving differential equations with computers

- many **nonlinear differential equations** require advanced mathematical methods to be solved explicitly (e.g. characteristic polynomial, monodromy matrix, variation of parameters, special functions)
- certain **qualitative** properties of SIS differential equations can be obtained by **linearizing** the system (e.g. basic reproduction number)
- we will use **computers** to calculate **numerical solution** curves that approximate exact solution to SIS differential equations

# Solving differential equations with computers

$$\begin{cases} \frac{S(t+\Delta)-S(t)}{\Delta} = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \\ \frac{I(t+\Delta)-I(t)}{\Delta} = \beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \end{cases}$$

Initial condition

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \end{aligned}$$

$$S(t) + I(t) = N; \quad t \geq 0$$

$$\Delta \rightarrow 0$$

# Solving differential equations with computers

$$\left\{ \begin{array}{l} \frac{S(t+\Delta)-S(t)}{\Delta} = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \\ \frac{I(t+\Delta)-I(t)}{\Delta} = \beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \end{array} \right. \quad \begin{array}{l} S(0) = S_0 \\ I(0) = I_0 \end{array}$$

$S(t) + I(t) = N; t \geq 0$

Constant population size

$\Delta \rightarrow 0$

# Solving differential equations with computers

$$\begin{cases} \frac{S(t+\Delta)-S(t)}{\Delta} = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \\ \frac{I(t+\Delta)-I(t)}{\Delta} = \beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \end{cases}$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \\ S(t) + I(t) &= N; \quad t \geq 0 \end{aligned}$$

Step size

$\Delta \rightarrow 0$

# Solving differential equations with computers

$$\left\{ \begin{array}{l} \frac{S(t+\Delta)-S(t)}{\Delta} = -\beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \\ \frac{I(t+\Delta)-I(t)}{\Delta} = \beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \end{array} \right. \quad \begin{array}{l} S(0) = S_0 \\ I(0) = I_0 \\ S(t) + I(t) = N; \quad t \geq 0 \end{array}$$

Rewrite the equation system:

recurrence

$$\left\{ \begin{array}{l} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{array} \right. \quad 0 < \Delta \ll 1$$

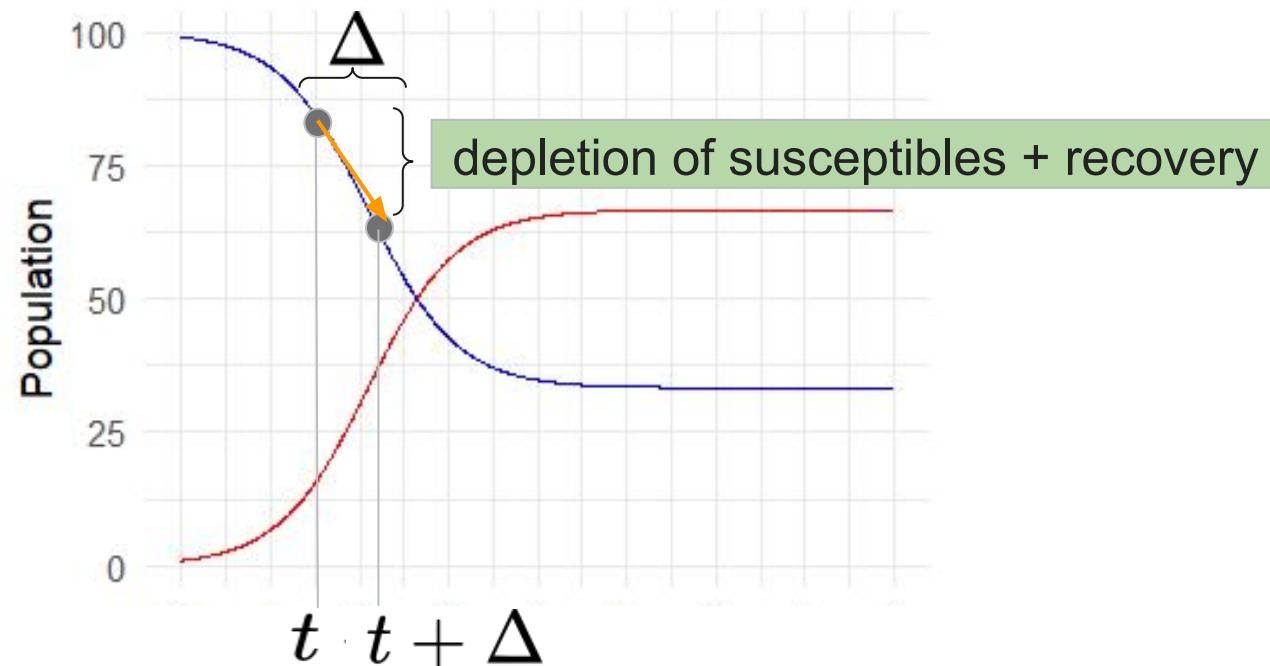
# Solving differential equations with computers

Recurrence:

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \end{aligned}$$

$$S(t) + I(t) = N; \quad t \geq 0$$



$$0 < \Delta \ll 1$$

# Solving differential equations with computers

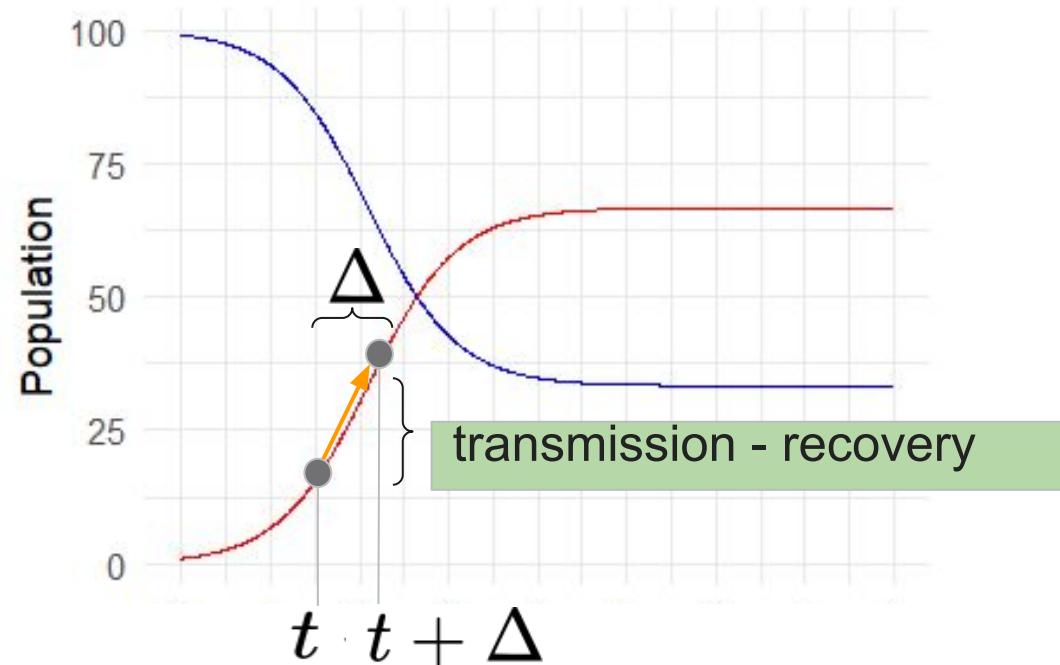
Recurrence:

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \end{aligned}$$

$$S(t) + I(t) = N; \quad t \geq 0$$

$$0 < \Delta \ll 1$$



# Solving differential equations with computers

Recurrence:

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \end{aligned}$$

$$S(t) + I(t) = N; \quad t \geq 0$$

$\cdots$	$t$	$t + \Delta$	$t + 2\Delta$	$\cdots$

$$0 < \Delta \ll 1$$

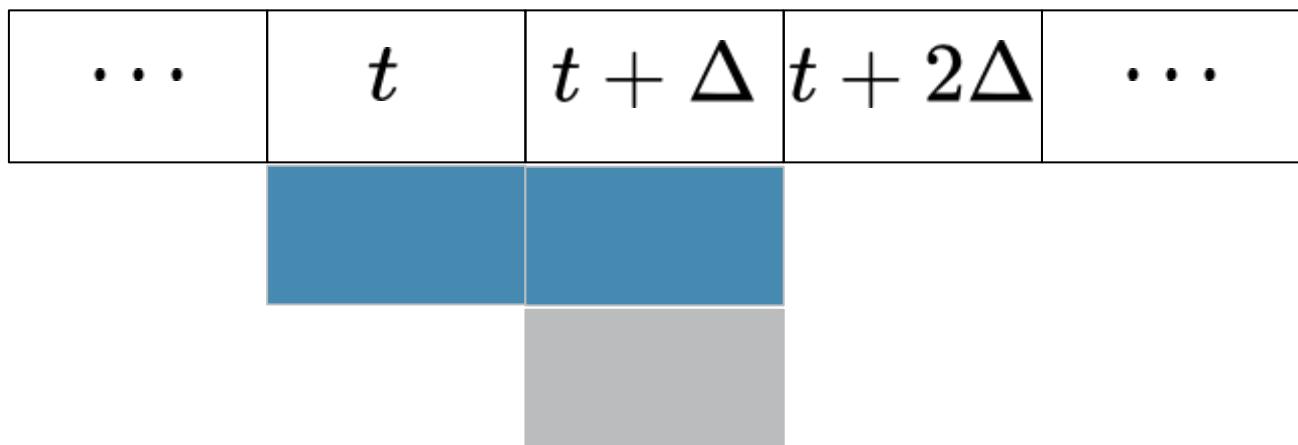
# Solving differential equations with computers

Recurrence:

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

$$\begin{aligned} S(0) &= S_0 \\ I(0) &= I_0 \end{aligned}$$

$$S(t) + I(t) = N; \quad t \geq 0$$



$$0 < \Delta \ll 1$$

# Solving differential equations with computers

Simulation setup

$$0 < \Delta \ll 1$$

$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t) + I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t) + I(t)} - \gamma I(t) \right) \end{cases}$$

# Solving differential equations with computers

Simulation setup

$$0 < \Delta \ll 1$$
$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

```
Delta=0.01
timeHorizon=100
timesteps<-seq(0,timeHorizon,Delta)

S=I=rep(0,length(timesteps))
N=500
I0=1
S[1]=N-I0
I[1]=I0
```

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

# Solving differential equations with computers

Simulation setup

$$0 < \Delta \ll 1$$
$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

```
Delta=0.01
timeHorizon=100
timesteps<-seq(0,timeHorizon,Delta)

S=I=rep(0,length(timesteps))
N=500
I0=1
S[1]=N-I0
I[1]=I0
```

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} + \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} - \gamma I(t) \right) \end{cases}$$

# Solving differential equations with computers

$$0 < \Delta \ll 1$$

$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

$$\begin{cases} S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t) + I(t)} S(t) - \gamma I(t) \right) \\ I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t) + I(t)} S(t) + \gamma I(t) \right) \end{cases}$$

infectivity parameter      recovery parameter

# Solving differential equations with computers

$$0 < \Delta \ll 1$$

$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

beta=0.3

gamma=0.1

$$S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \right)$$

$$I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \right)$$

# Solving differential equations with computers

$$0 < \Delta \ll 1$$

$$S(0) = S_0$$

$$I(0) = I_0$$

$$S(t) + I(t) = N; \quad t \geq 0$$

Update next time step

$$S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \right)$$

$$I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \right)$$

# Solving differential equations with computers

loop over time steps index set

```
for (i in c(1:(length(timesteps)-1))) {  
  
    S[i+1] = S[i] + Delta*(-beta*I[i]/N*S[i] + gamma*I[i])  
    I[i+1] = I[i] + Delta*(beta*I[i]/N*S[i] - gamma*I[i])  
  
}
```

$$S(t + \Delta) = S(t) + \Delta \left( -\beta \frac{I(t)}{S(t)+I(t)} S(t) - \gamma I(t) \right)$$
$$I(t + \Delta) = I(t) + \Delta \left( \beta \frac{I(t)}{S(t)+I(t)} S(t) + \gamma I(t) \right)$$

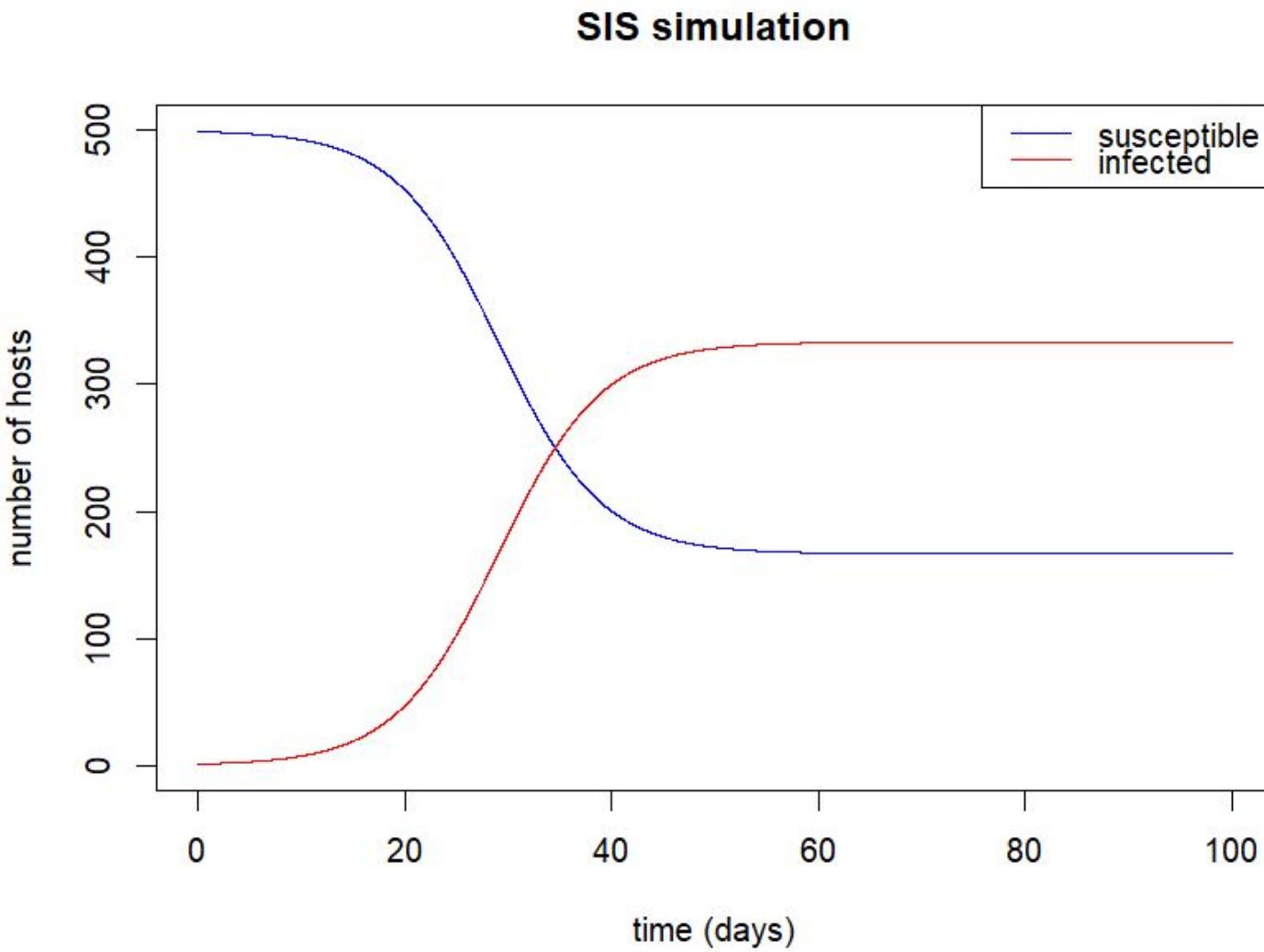
# Solving differential equations with computers

```
plot(timesteps,S,col="blue",type="l",
      xlab="time steps (days)",ylab="number of hosts",
      main="SIS simulation",ylim = c(0, N))

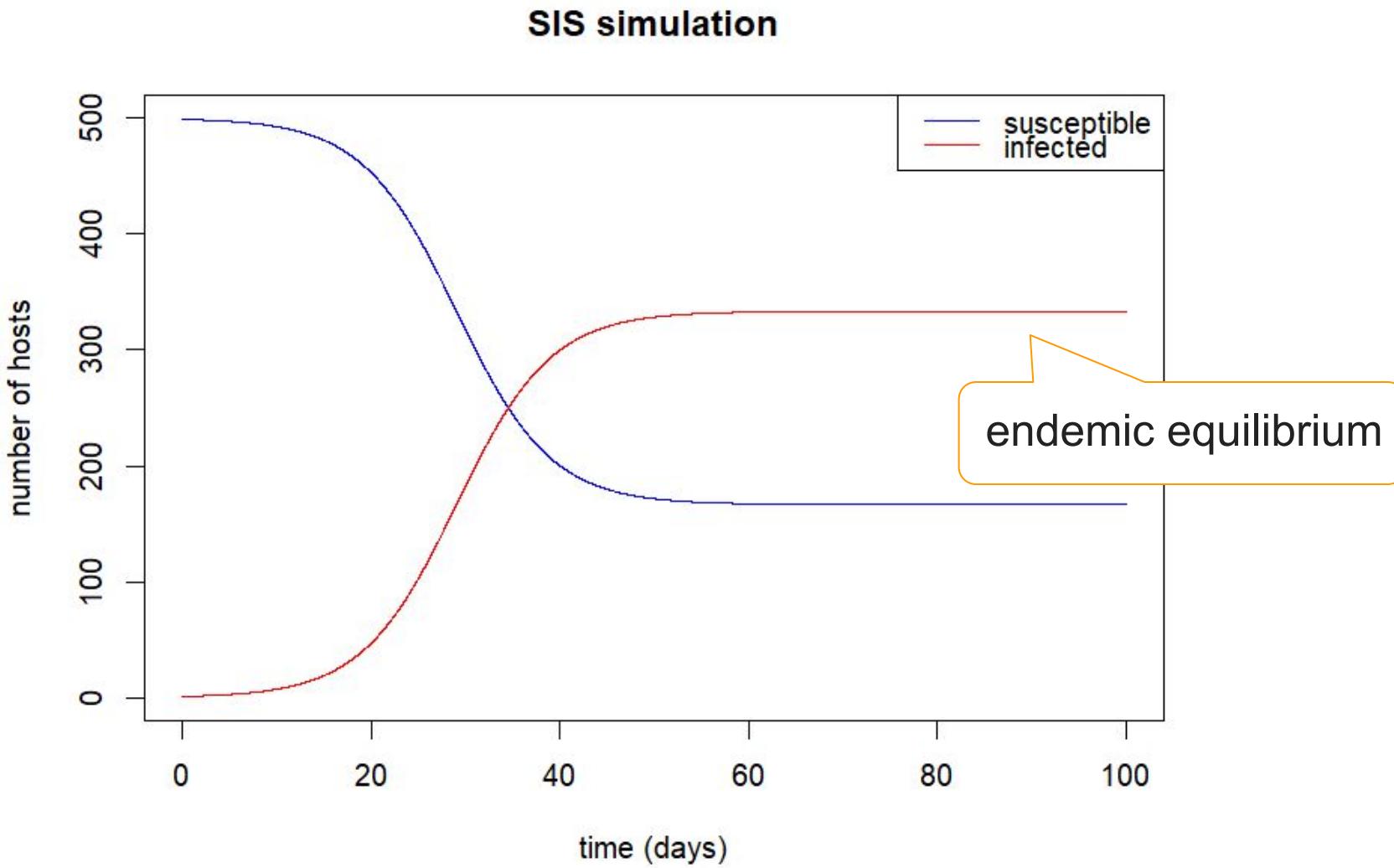
lines(timesteps,I,col="red")

legend( x="topright",
        legend=c("susceptible","infected"),
        col=c("blue","red"),lwd=1)
```

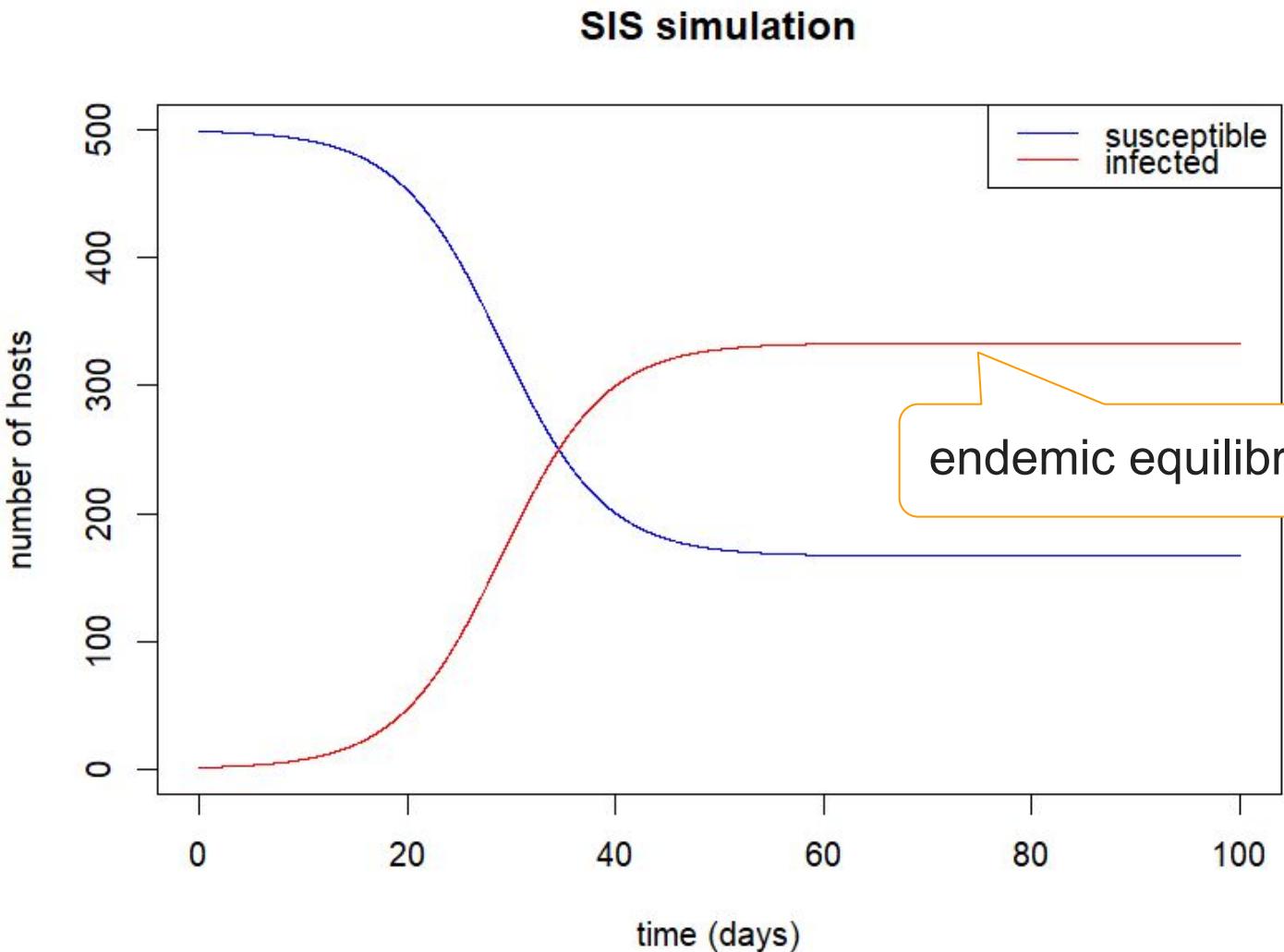
# Solving differential equations with computers



# Solving differential equations with computers



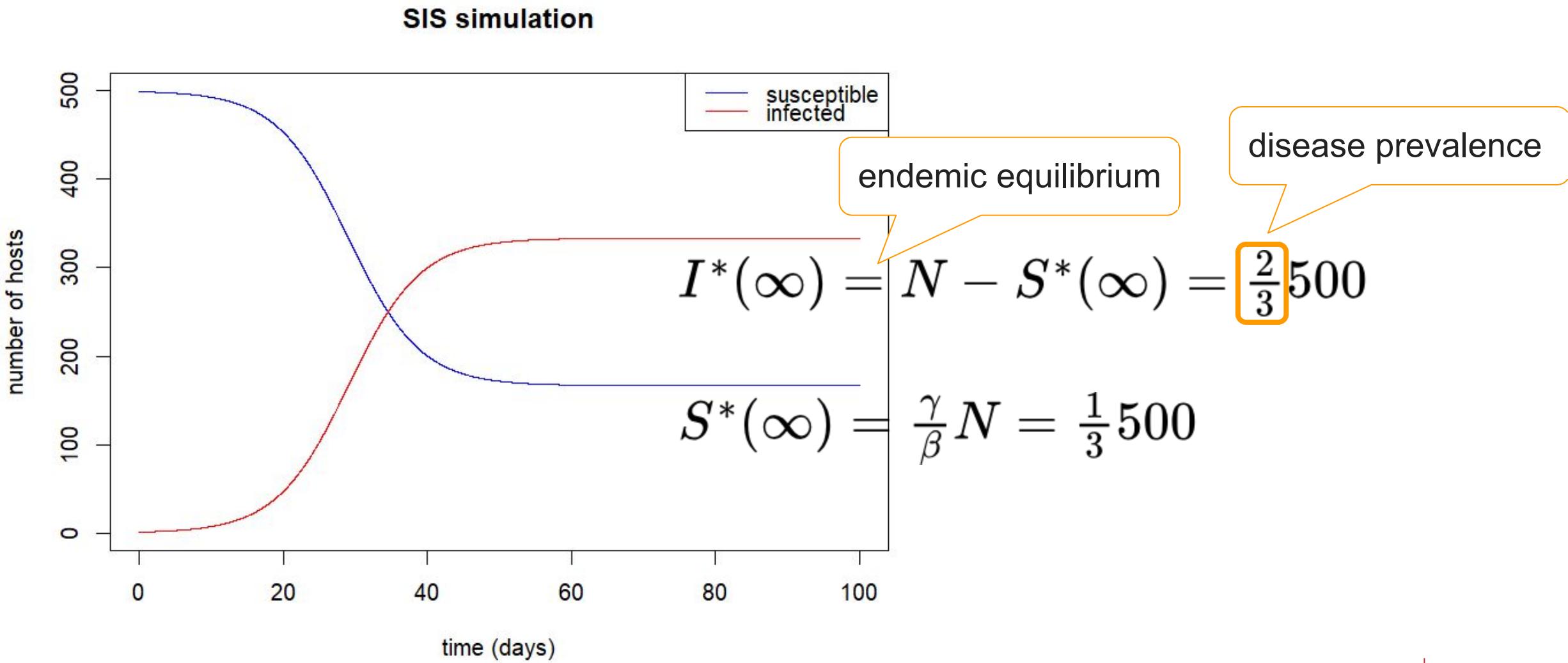
# Solving differential equations with computers



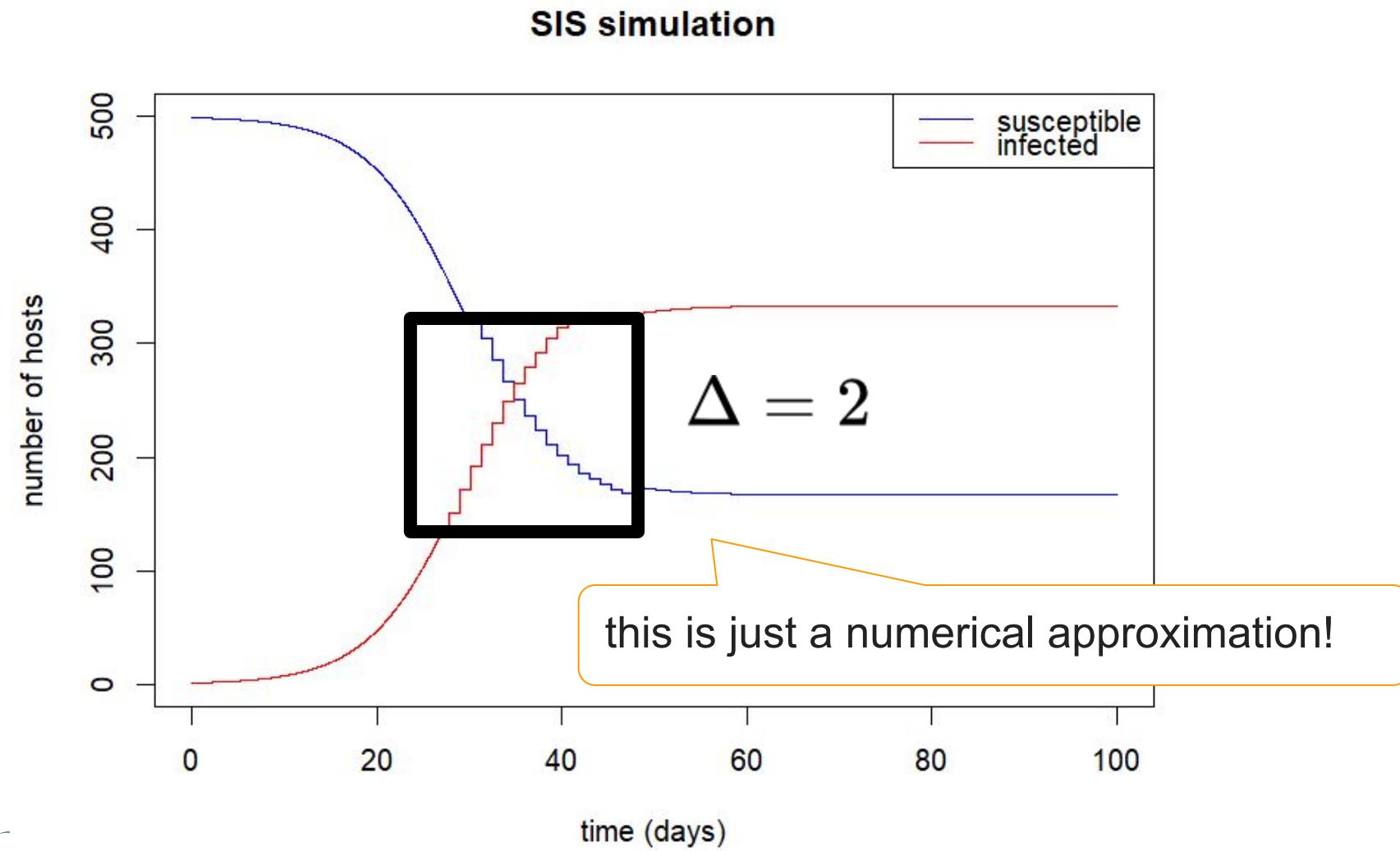
equilibrium = no change between time points !

$$\begin{aligned} 0 &= \frac{S^*(t+\Delta) - S^*(t)}{\Delta} \\ &= -\beta \frac{I^*(t)}{N} S^*(t) + \gamma I^*(t) \\ &= I^*(t) \left( -\beta \frac{S^*(t)}{N} + \gamma \right) \\ &\Leftrightarrow -\beta \frac{S^*(t)}{N} + \gamma = 0 \\ &\Leftrightarrow S^*(\infty) = \frac{\gamma}{\beta} N \end{aligned}$$

# Solving differential equations with computers



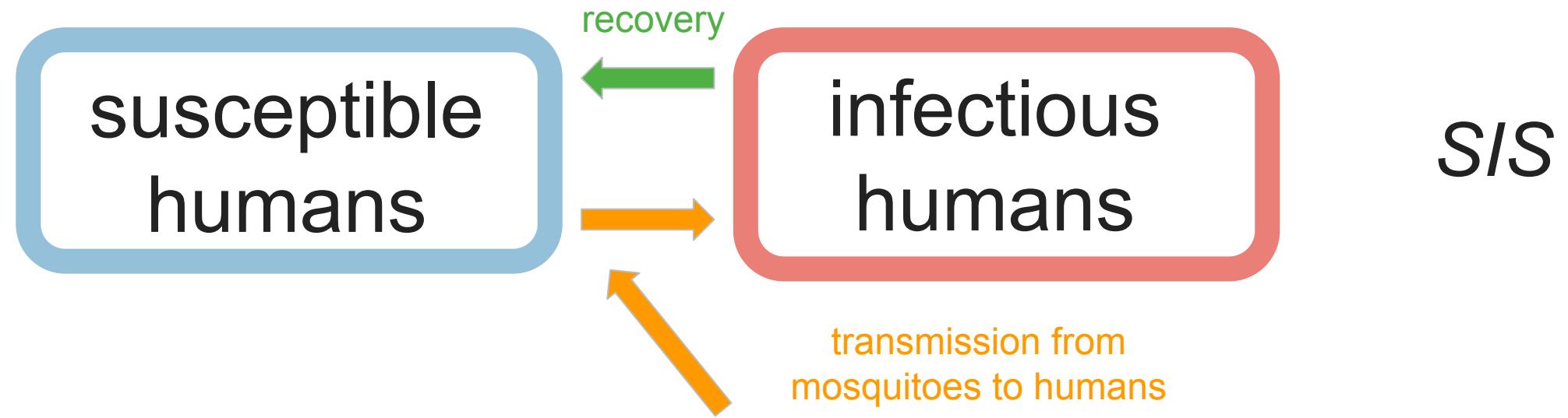
# Solving differential equations with computers



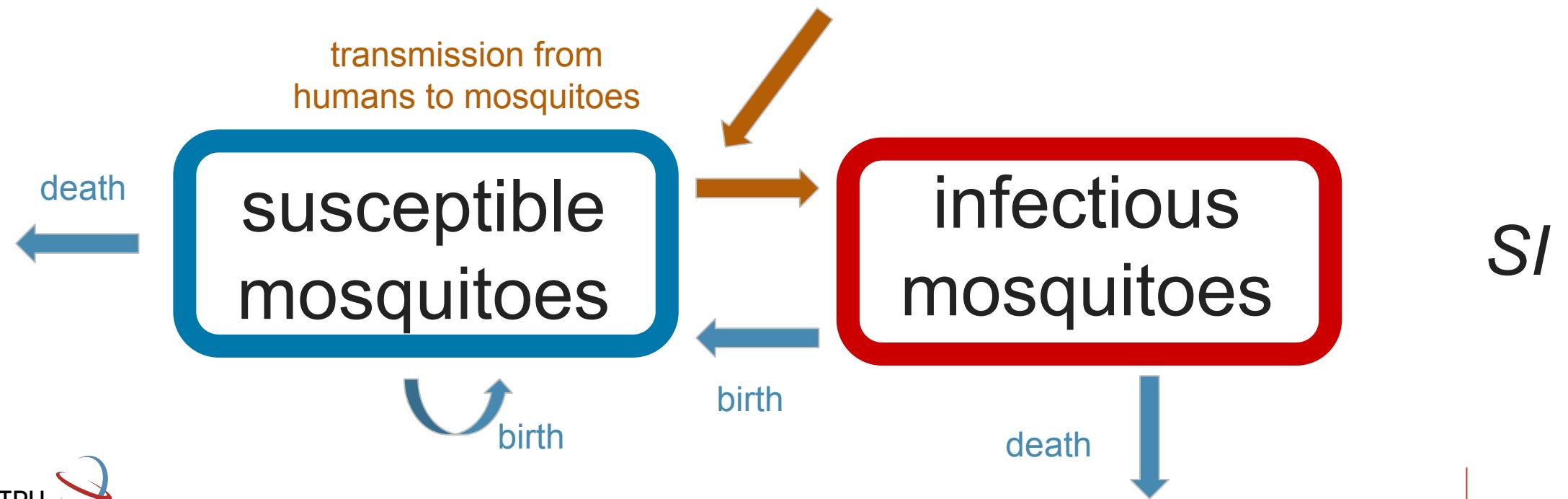


## 1 - Malaria model parameters

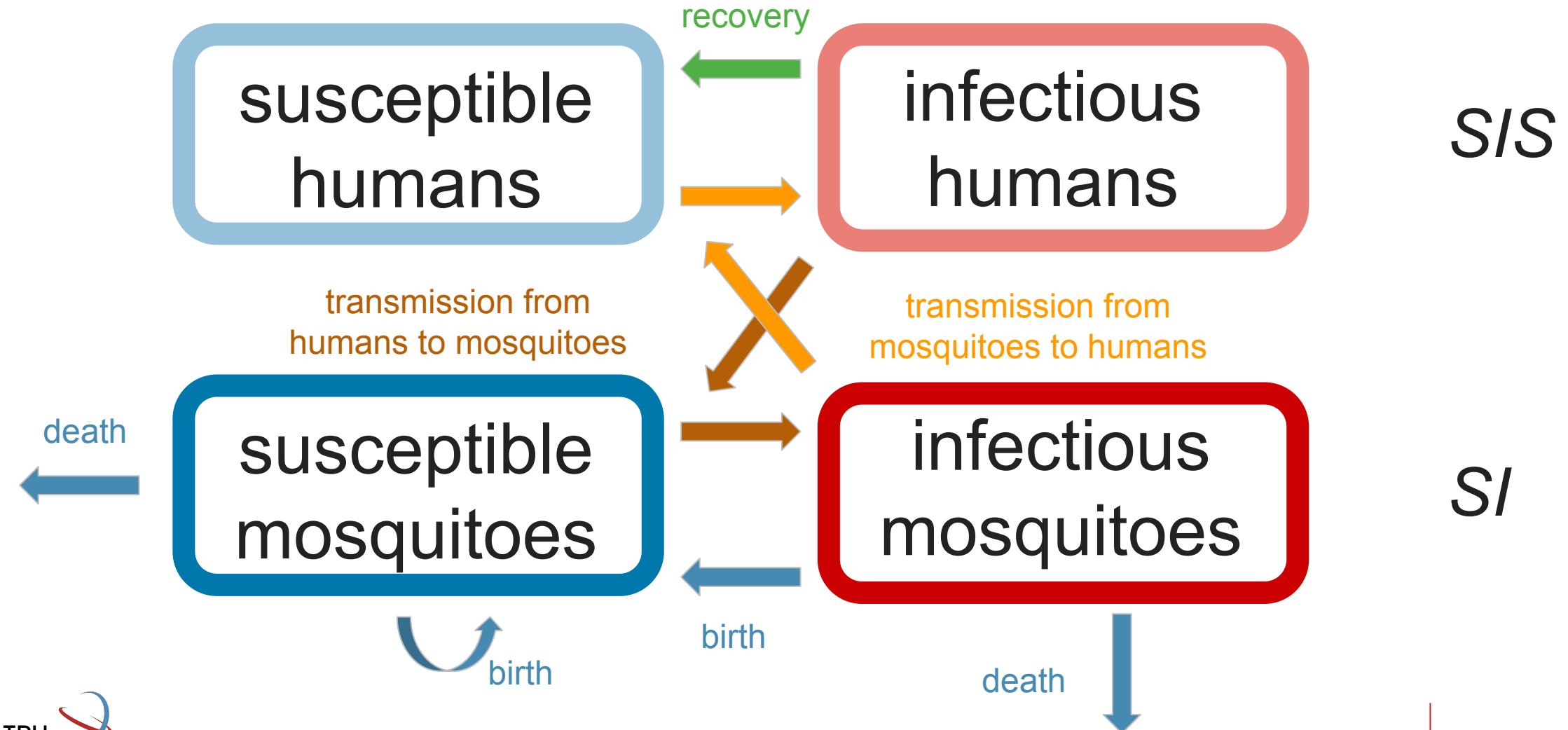
# Malaria transmission between host and vector



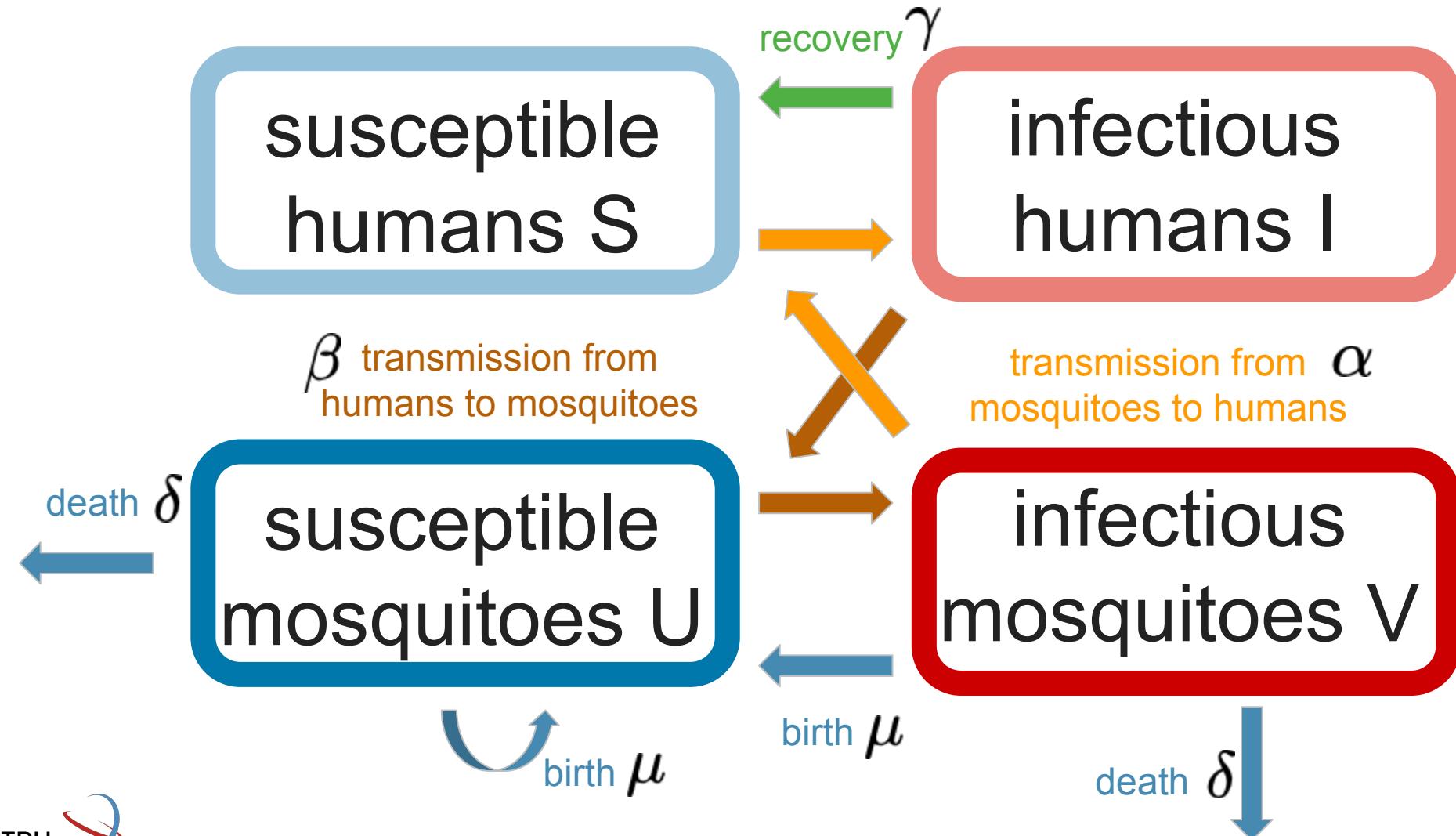
# Malaria transmission between host and vector



# Malaria transmission between host and vector



# Malaria transmission between host and vector



# Where can we get parameter values from?

 $\frac{1}{\mu}$ 

time span from ovipositing to emergence of adult female mosquito from pupa stage, given ovipositing

 $\frac{1}{\delta}$ 

life span of adult female mosquito

 $\frac{1}{\gamma}$ 

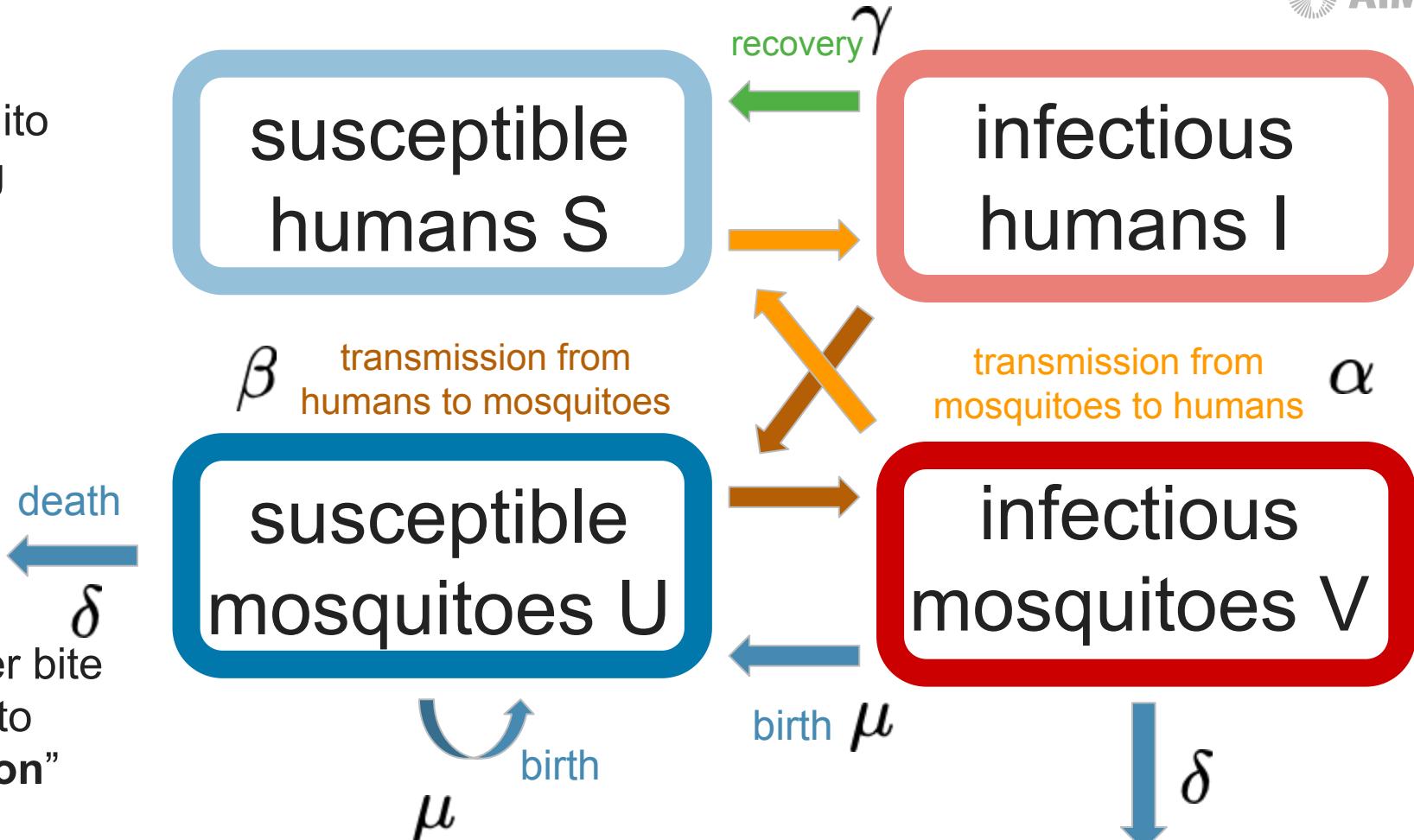
duration of infection in human

 $\alpha$ 

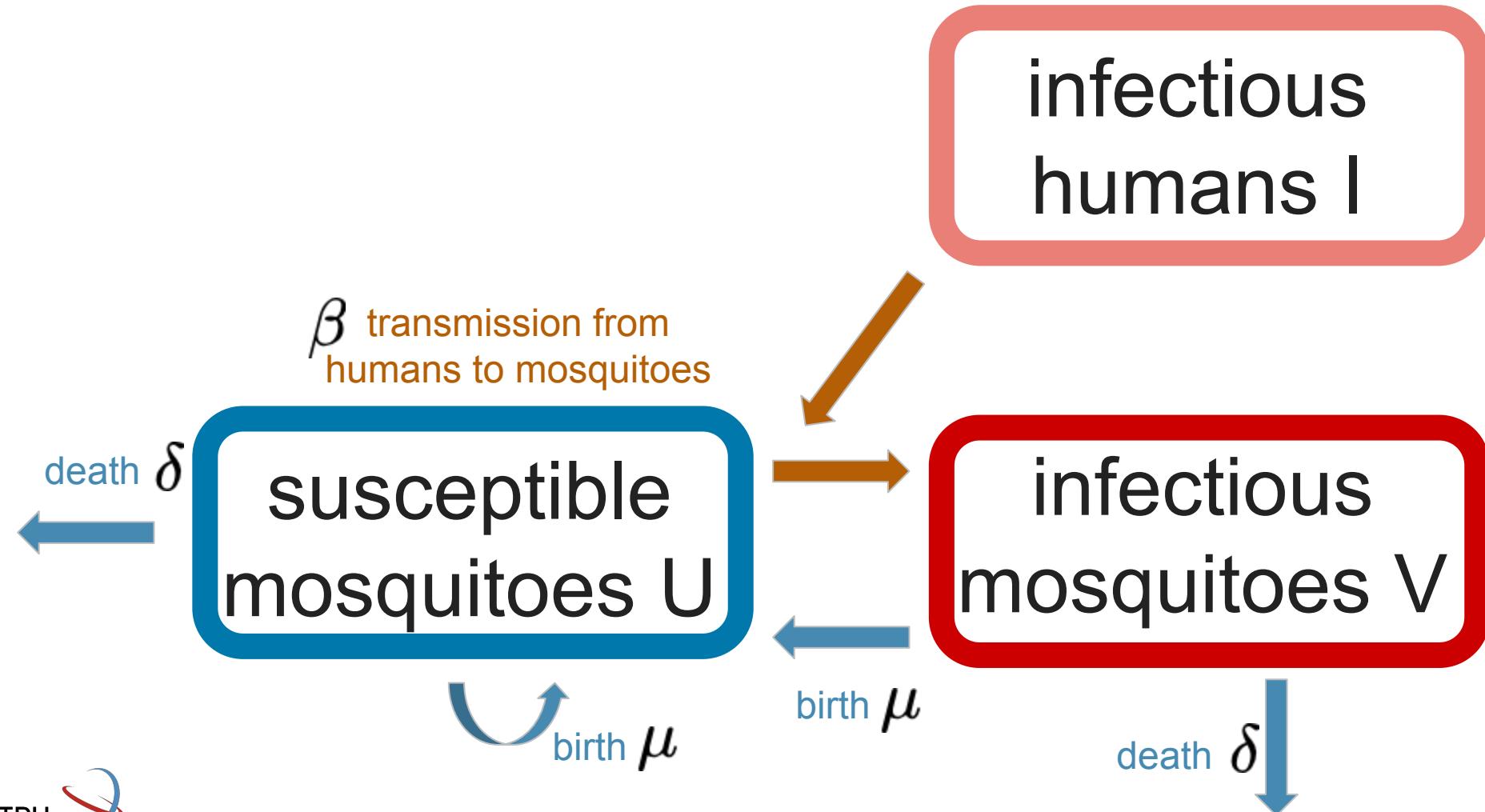
host seeking/biting X probability per bite for infectious mosquito to transmit to human: "**entomological inoculation**"

 $\beta$ 

host seeking/biting X acquisition rate from infectious human with gametocytemia to mosquito



# Malaria transmission between host and vector

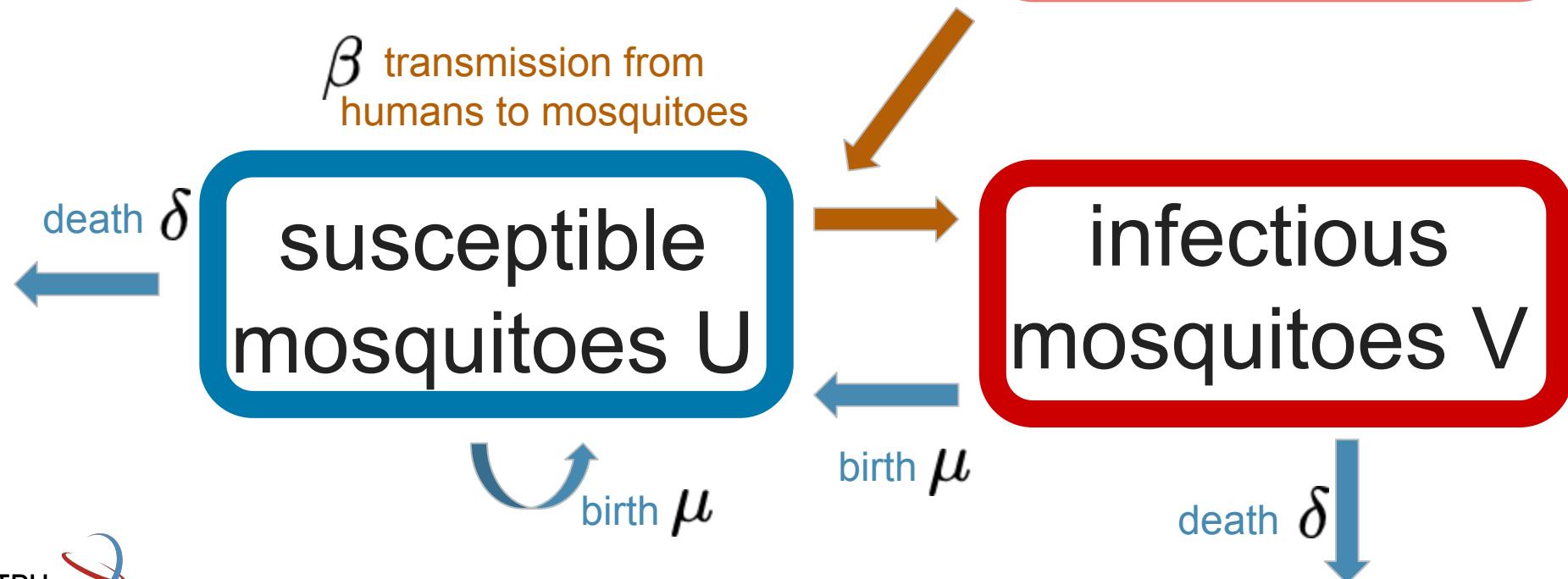


# Malaria transmission between host and vector

$$\frac{dU}{dt}(t) = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t)$$

$$\frac{dV}{dt}(t) = \beta \frac{I(t)}{H} U(t) - \delta V(t)$$

infectious  
humans I



# Malaria transmission between host and vector

$$\frac{dU}{dt}(t) = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t)$$

$$\frac{dV}{dt}(t) = \beta \frac{I(t)}{H} U(t) - \delta V(t)$$

$M = U(0) + V(0)$  constant mosquito population size at equilibrium  $\mu = \delta$

infectious humans I

$\beta$  transmission from humans to mosquitoes

susceptible mosquitoes U

infectious mosquitoes V



birth  $\mu$



death  $\delta$



# Malaria transmission between host and vector

$$\frac{dU}{dt}(t) = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t)$$

$$\frac{dV}{dt}(t) = \beta \frac{I(t)}{H} U(t) - \delta V(t)$$

$H = S(0) + I(0)$  constant human populations size

$\frac{I(t)}{H}$  infectious human density

infectious humans I

$\beta$  transmission from humans to mosquitoes

susceptible mosquitoes U

infectious mosquitoes V



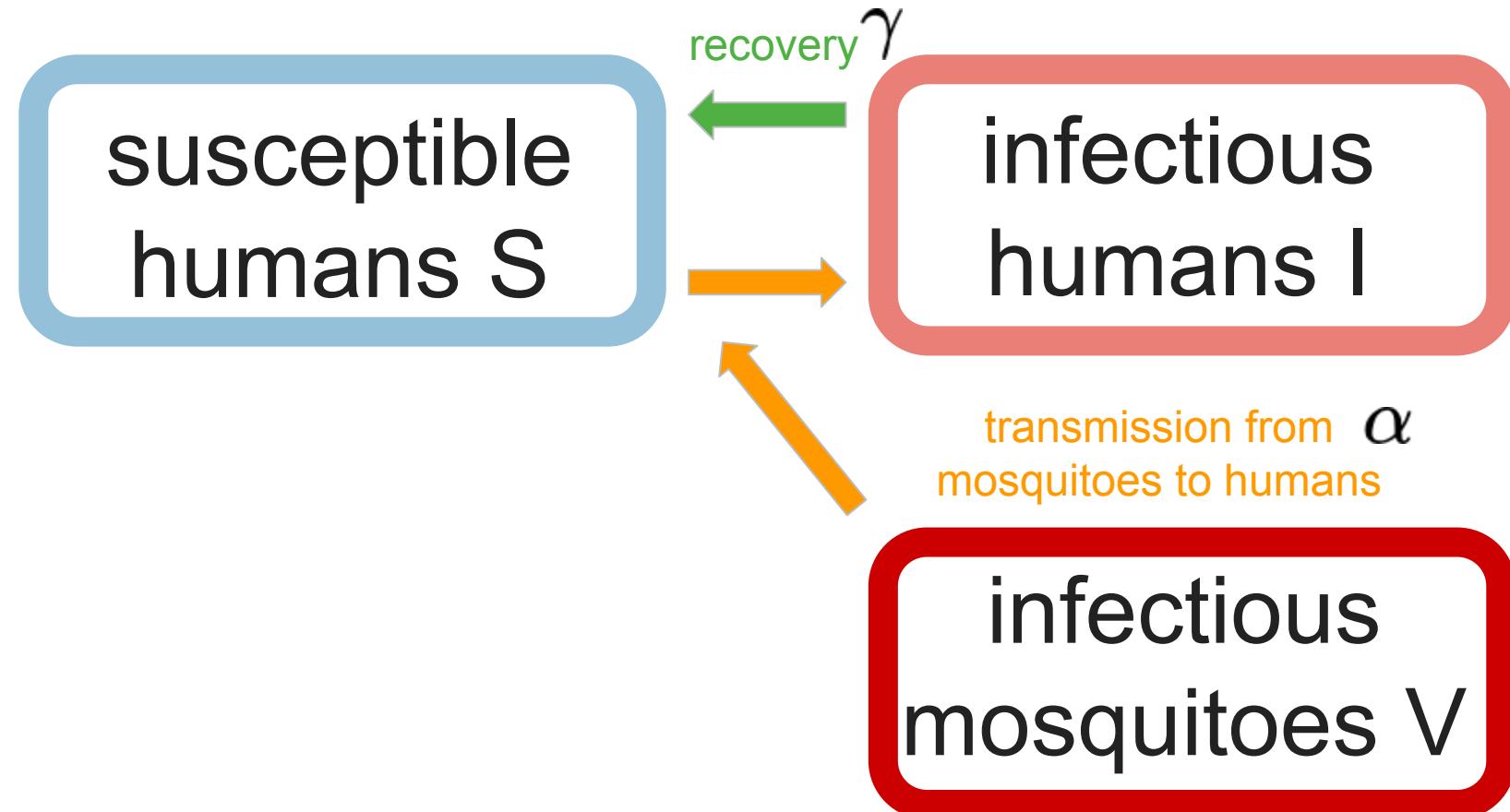
birth  $\mu$



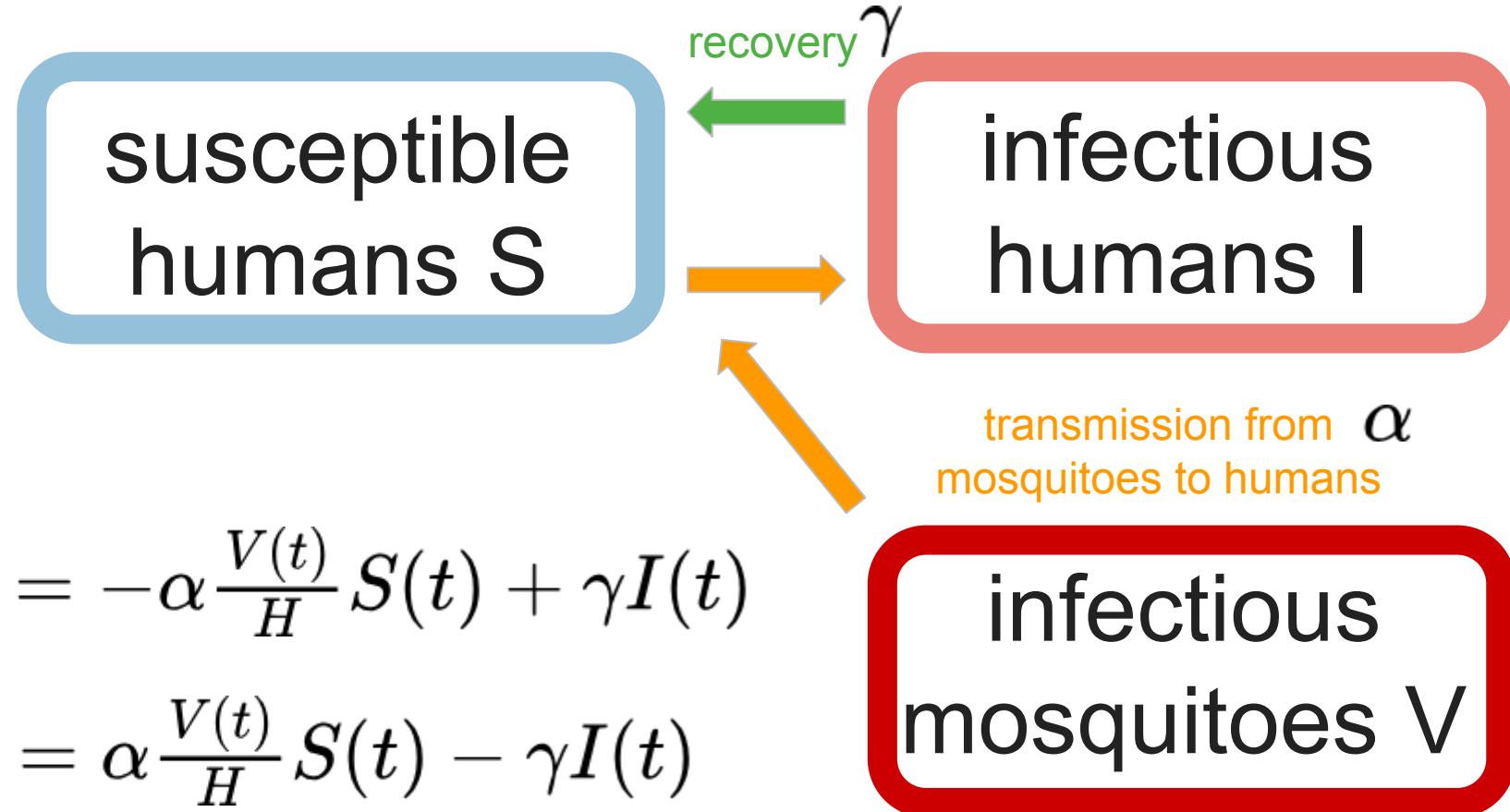
death  $\delta$



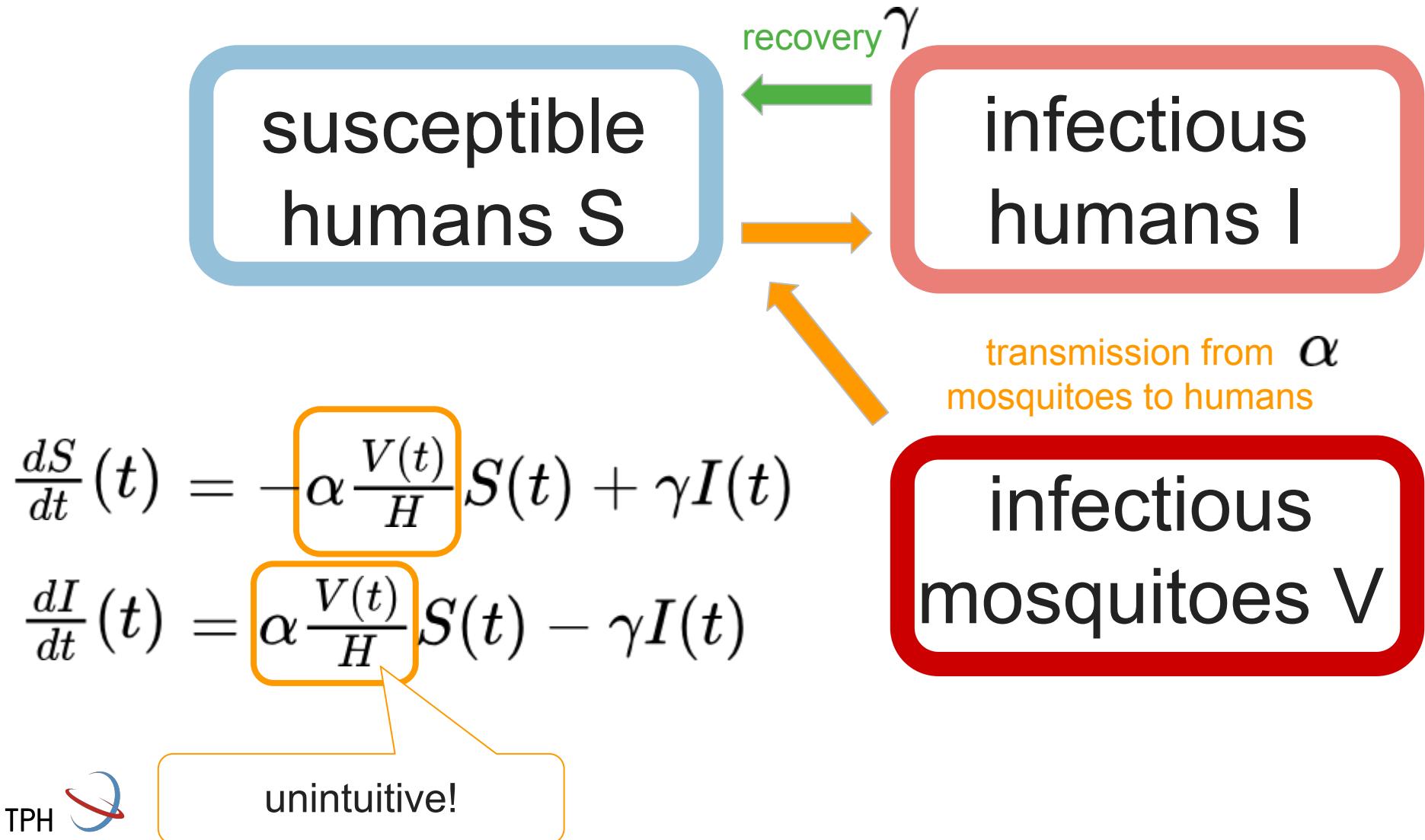
# Malaria transmission between host and vector



# Malaria transmission between host and vector



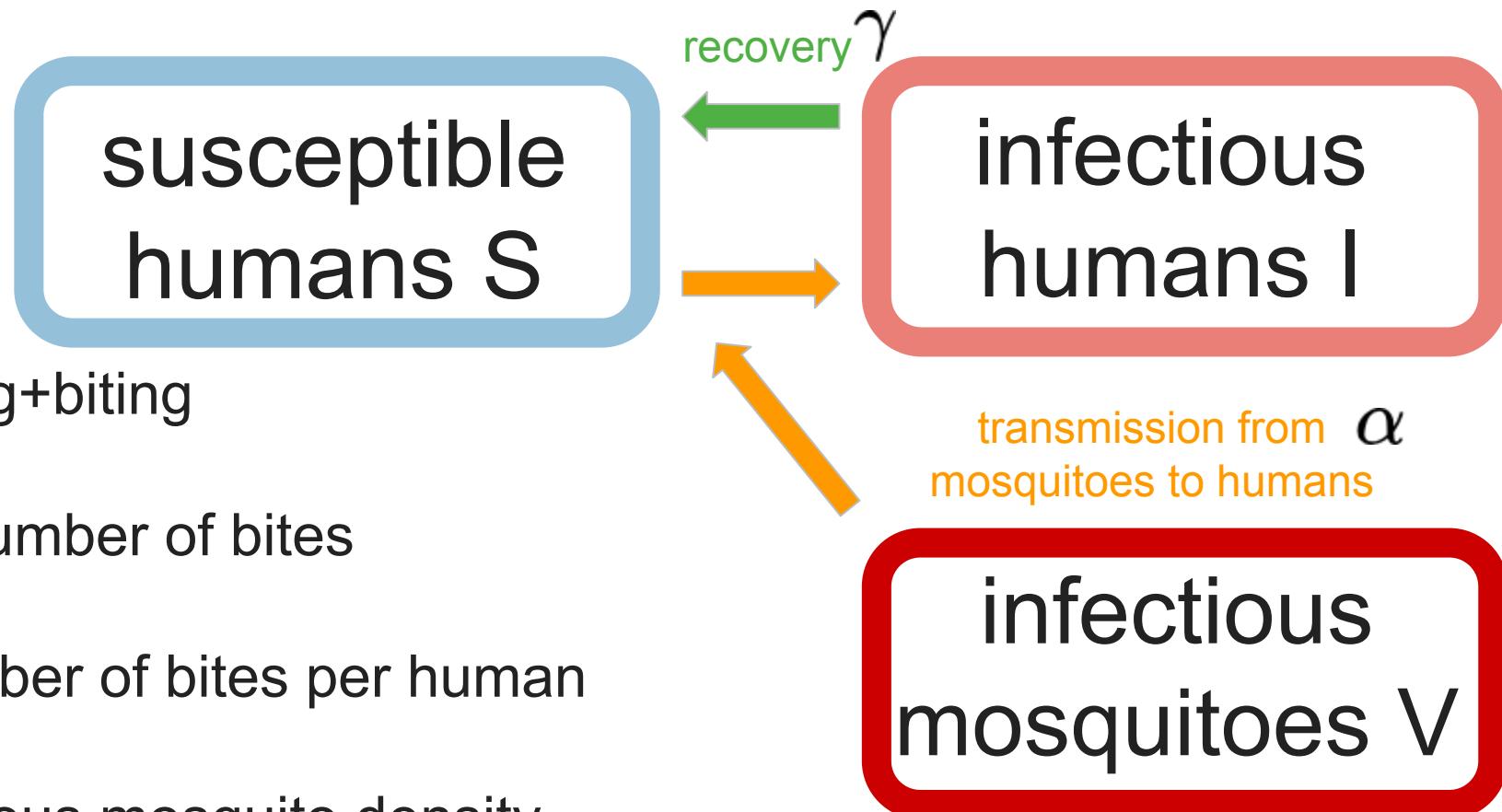
# Malaria transmission between host and vector



# Malaria transmission between host and vector

$$\alpha \frac{V(t)}{H} = ?$$

- $\alpha$  host seeking+biting
- $\alpha(U + V)$  number of bites
- $\frac{\alpha(U+V)}{H}$  number of bites per human
- $\frac{V}{U+V}$  infectious mosquito density
- $\frac{V}{U+V} \alpha \frac{U+V}{H} = \alpha \frac{V}{H}$



# Malaria transmission between host and vector

```
###SIS-SI host-vector model
Delta=0.01 #step size
timeHorizon=200 #maximum number of days/weeks/months to simulate
timesteps<-seq(0,timeHorizon,Delta) #time points to simulate

S=I=U=V=rep(0,length(timesteps))
H=1000
K=5 #VectorHumanRatio
M=H*K
I0=1
V0=8
S[1]=H-I0
I[1]=I0
U[1]=N*K-V0
V[1]=V0
```

setup & initial  
conditions

# Malaria transmission between host and vector

```
###SIS-SI host-vector model
Delta=0.01 #step size
timeHorizon=200 #maximum number of days/weeks/months to simulate
timesteps<-seq(0,timeHorizon,Delta) #time points to simulate
```

S= We have 5 times as many mosquitoes as human hosts!

```
H=100
K=5 #VectorHumanRatio
M=H*K
I0=1
V0=8
S[1]=H-I0
I[1]=I0
U[1]=N*K-V0
V[1]=V0
```

# Malaria transmission between host and vector

```
alpha=0.05
gamma=1/20
beta=0.08
delta=1/10
mu=delta
```

```
for (i in c(1:(length(timesteps)-1))) {
  S[i+1] = S[i] + Delta*(- alpha*V[i]*S[i]/H + gamma*I[i])
  I[i+1] = I[i] + Delta*( alpha*V[i]*S[i]/H - gamma*I[i])
  U[i+1] = U[i] + Delta*(- beta*U[i]*I[i]/H + mu*M - delta*U[i])
  V[i+1] = V[i] + Delta*( beta*U[i]*I[i]/H - delta*V[i])
}
```

# Malaria transmission between host and vector

```
alpha=0.05
gamma=1/20
beta=0.08
delta=1/10
mu=delta
```

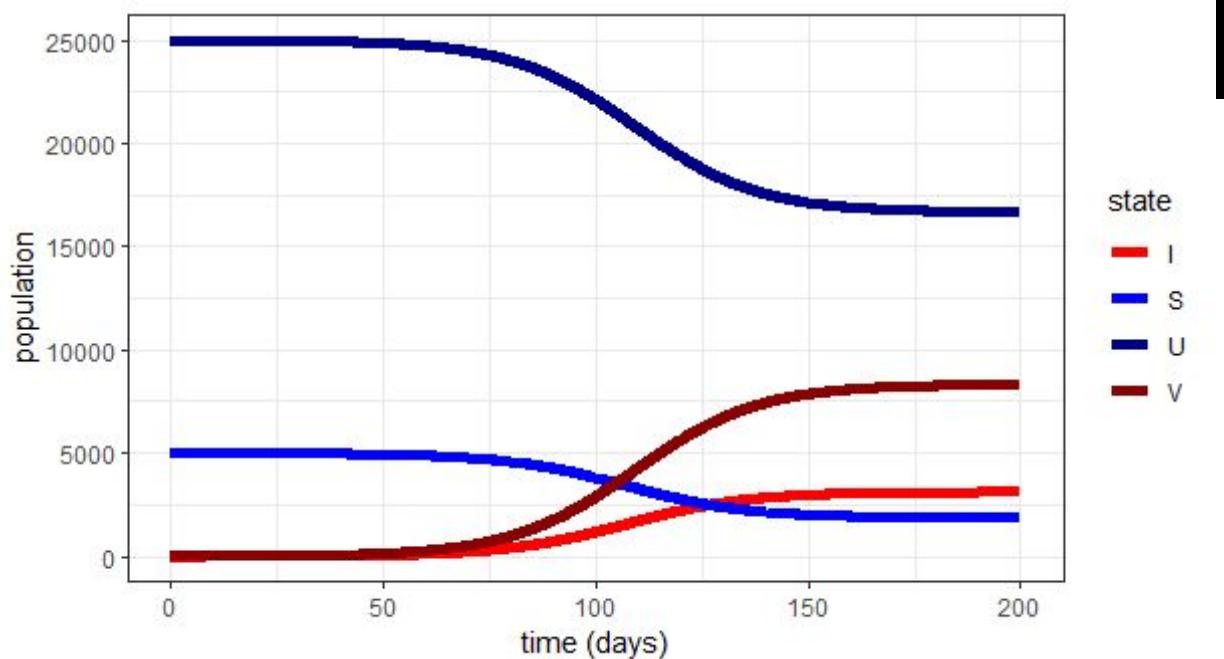
$$\frac{U(t+\Delta) - U(t)}{\Delta} = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t)$$

$$\frac{V(t+\Delta) - V(t)}{\Delta} = \beta \frac{I(t)}{H} U(t) - \delta V(t)$$

```
for (i in c(1:(length(timesteps)-1))) {
  S[i+1] = S[i] + Delta*(- alpha*V[i]*S[i]/H + gamma*I[i])
  I[i+1] = I[i] + Delta*( alpha*V[i]*S[i]/H - gamma*I[i])
  U[i+1] = U[i] + Delta*(- beta*U[i]*I[i]/H + mu*M - delta*U[i])
  V[i+1] = V[i] + Delta*( beta*U[i]*I[i]/H - delta*V[i])
}
```

# Malaria transmission between host and vector

```
cbind(timesteps, S, I, U, V) %>%
  as.data.frame %>%
  pivot_longer(cols = -timesteps) %>%
  ggplot() +
  geom_line(aes(x=timesteps, y=value, color=name), linewidth=2) +
  scale_color_manual(values=c("red", "blue", "darkblue", "darkred"), name="state") +
  scale_y_continuous(name="population") +
  scale_x_continuous(name="time (days)") +
  theme_bw()
```



# Malaria transmission between host and vector

**For more complex models, use the `deSolve` R package!**

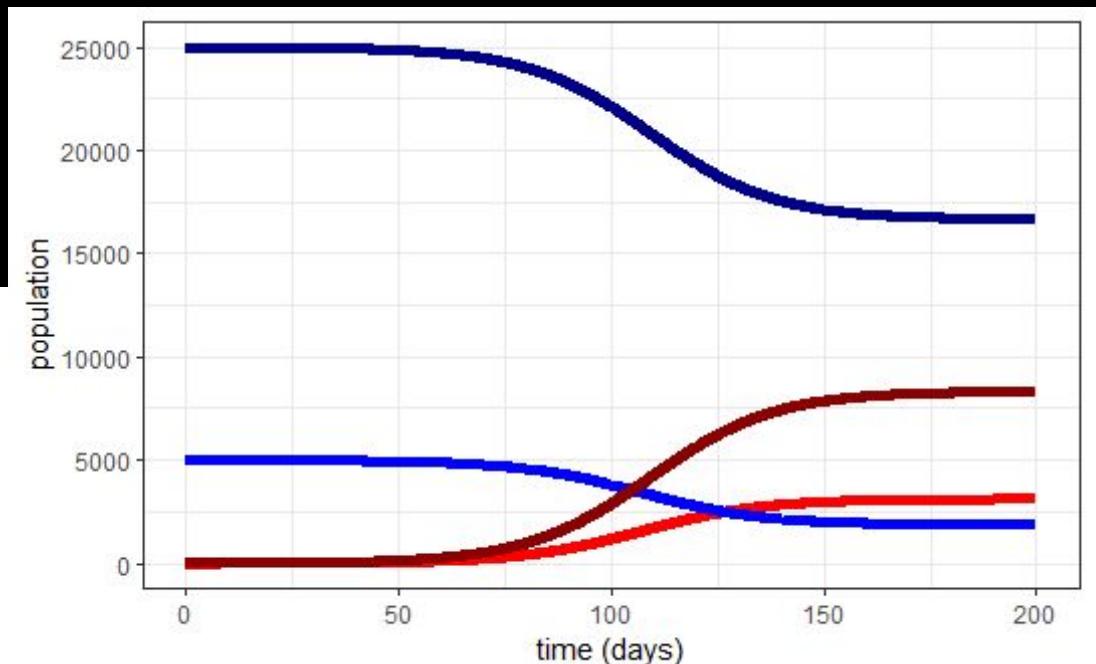
```
## Use R package deSolve for ODEs
library("deSolve")
RossMcDonald.model<-function(time,state,parms) {
  with(as.list(c(state,parms)), {
    alpha=parms[1]; gamma=parms[2]; beta=parms[3]; mu=parms[4]; delta=parms[5]
    S=state[1]; I=state[2]; U=state[3]; V=state[4];
    H=5000; K=5; M=H*K

    dS= - alpha*V*S/H + gamma*I
    dI= alpha*V*S/H - gamma*I
    dU= -beta*U*I/H + mu*M-delta*U
    dV= beta*U*I/H-delta*V

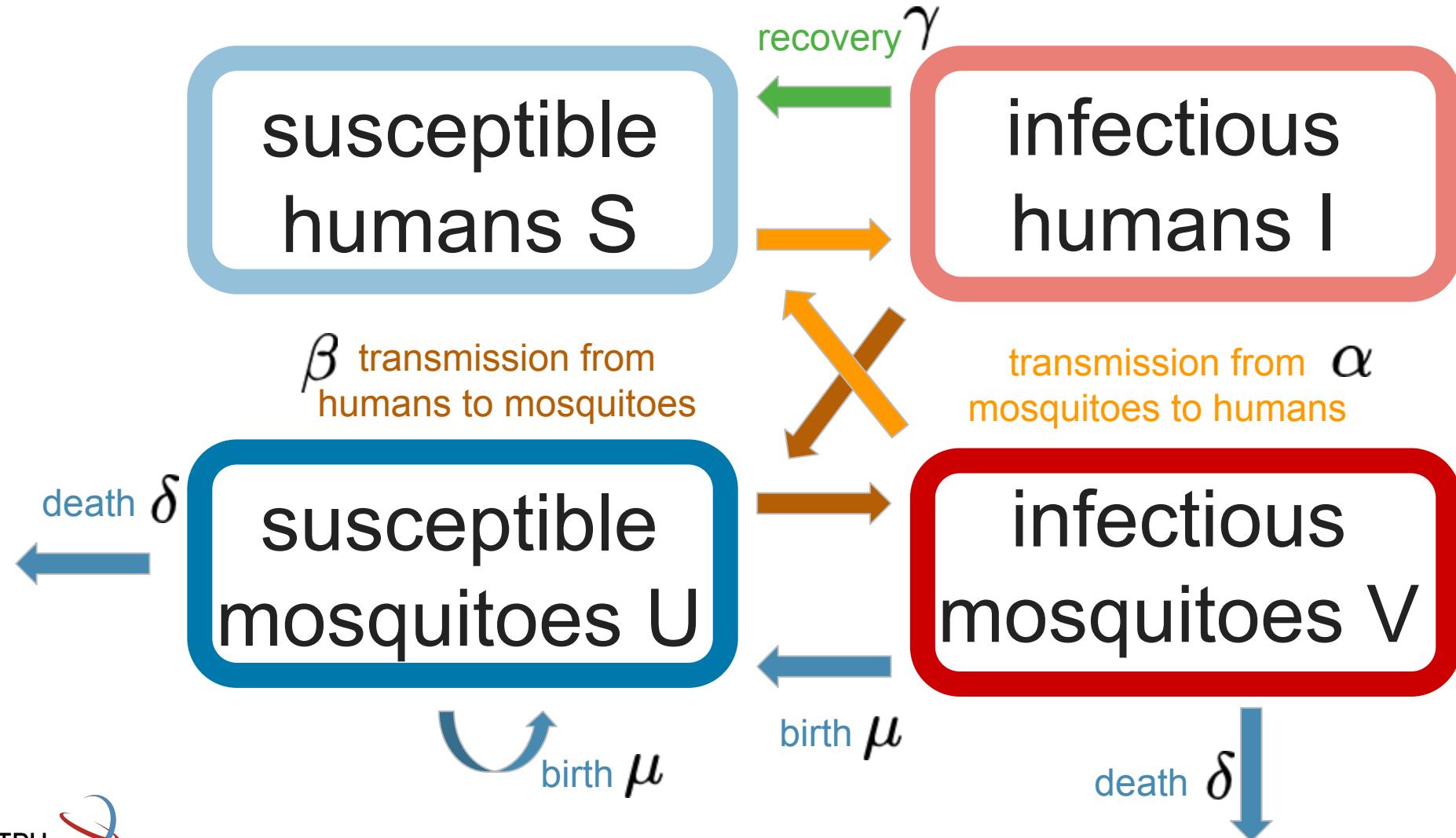
    return(list(c(dS,dI,dU,dV)))
  })
}
K<-5
x0 <- c(S=H-I0,I=I0,U=H*K-V0,V=V0)##initial condition
timesteps<-seq(0,200,1)##time unit in days
parms <- c(alpha=0.05, gamma=1/20, beta=0.08, mu=1/10,delta=1/10) #parameters
output<-ode(x0,timesteps,RossMcDonald.model,parms)
```

# Malaria transmission between host and vector

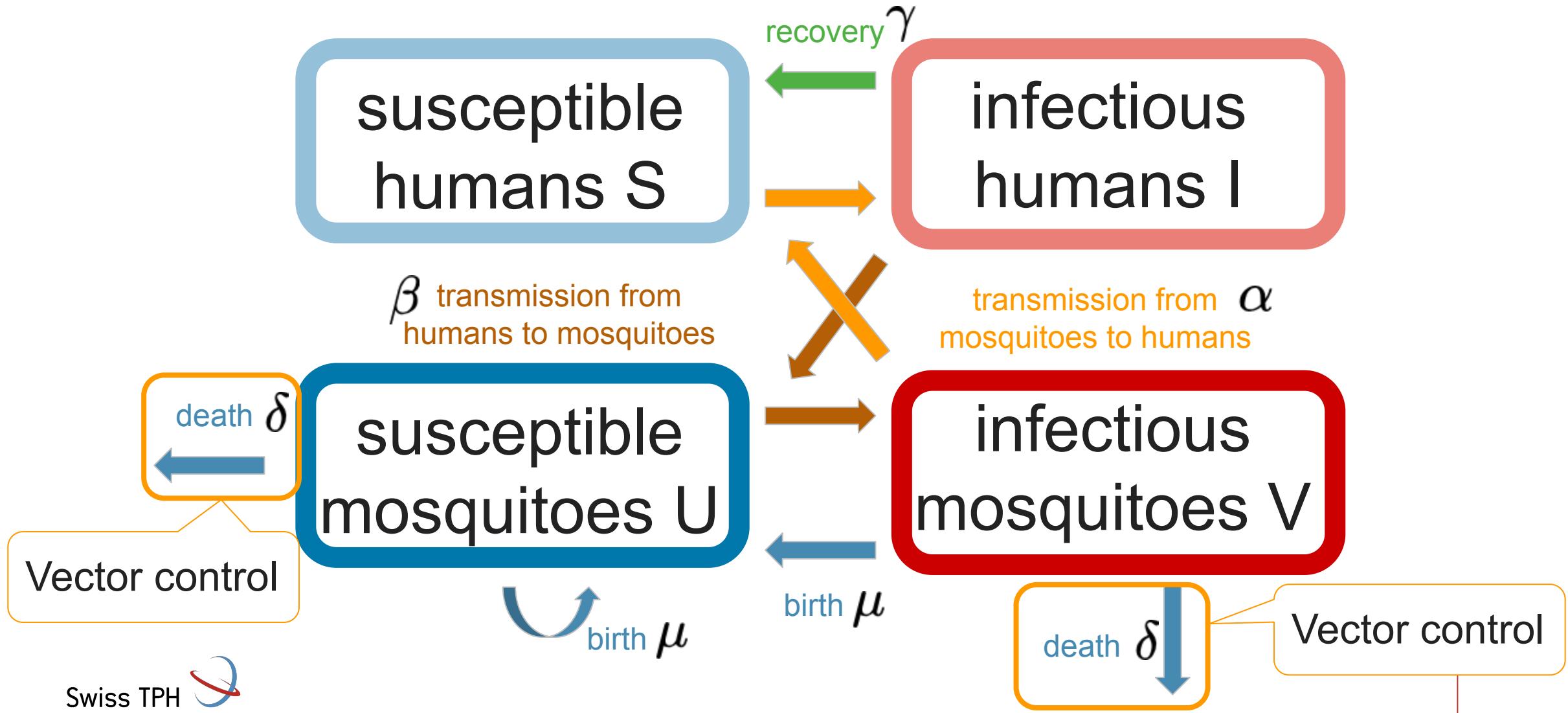
```
output %>%
  as.data.frame %>%
  pivot_longer(cols = -timesteps) %>%
  ggplot() +
  geom_line(aes(x=timesteps, y=value, color=name), linewidth=2) +
  scale_color_manual(values=c("red", "blue", "darkblue", "darkred"), name="disease state") +
  scale_y_continuous(name="population") +
  theme_bw()
```



# Malaria transmission model with vector control



# Malaria transmission model with vector control



# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

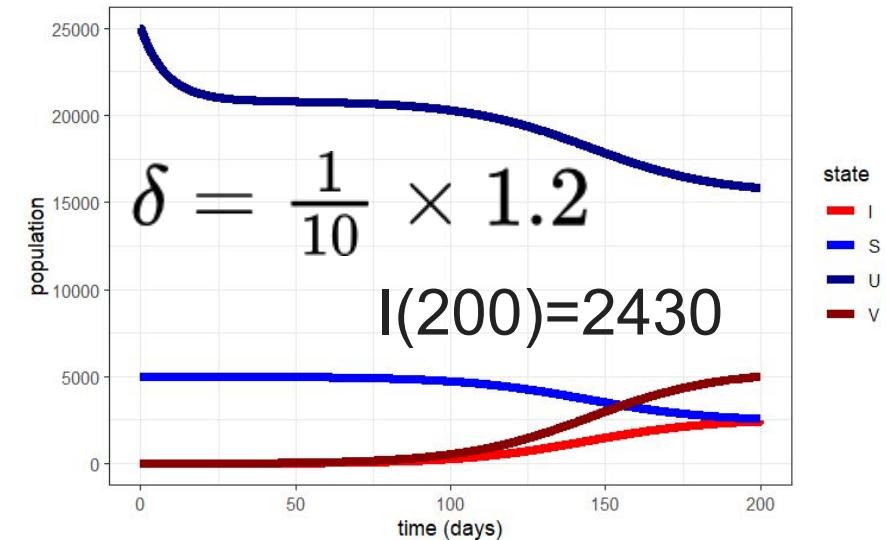
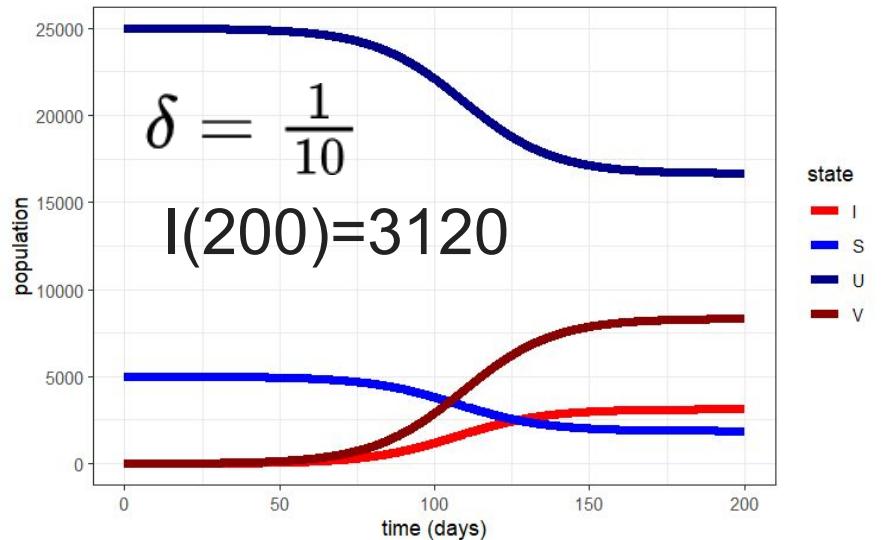
- simulation 1 as in the slides
- simulation 2 starts from simulation 1, but adds vector control as an 20% increase in the default mosquito mortality rate  $\delta = \frac{1}{10}$ , all other parameters are kept the same
- **compare** the endemic equilibrium of infected hosts between simulation 1 & 2

# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulation 2 starts from simulation 1, but adds vector control as an 20% increase in the default mosquito mortality rate  $\delta = \frac{1}{10}$ , all other parameters are kept the same
- **compare** the endemic equilibrium of infected hosts between simulation 1 & 2



# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

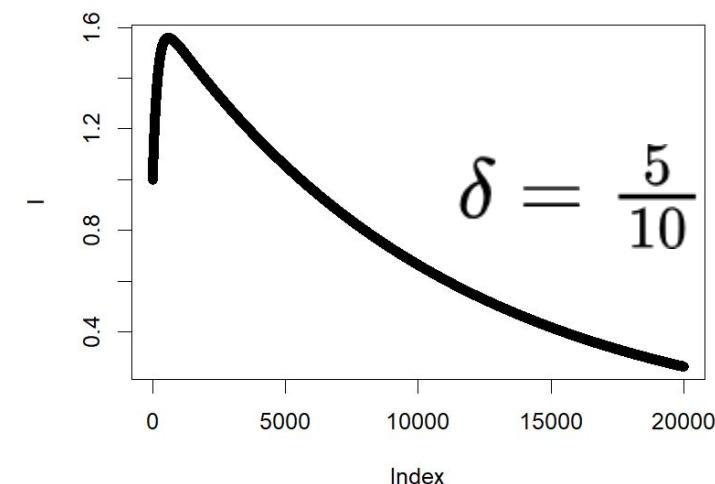
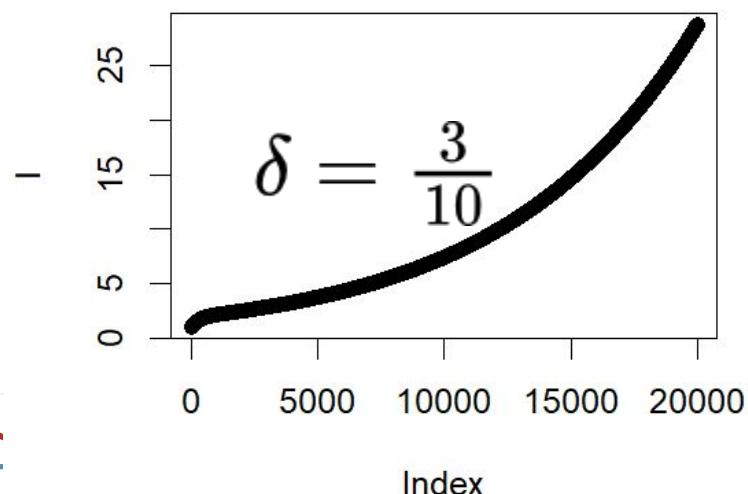
- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$     $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!

# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$   $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!

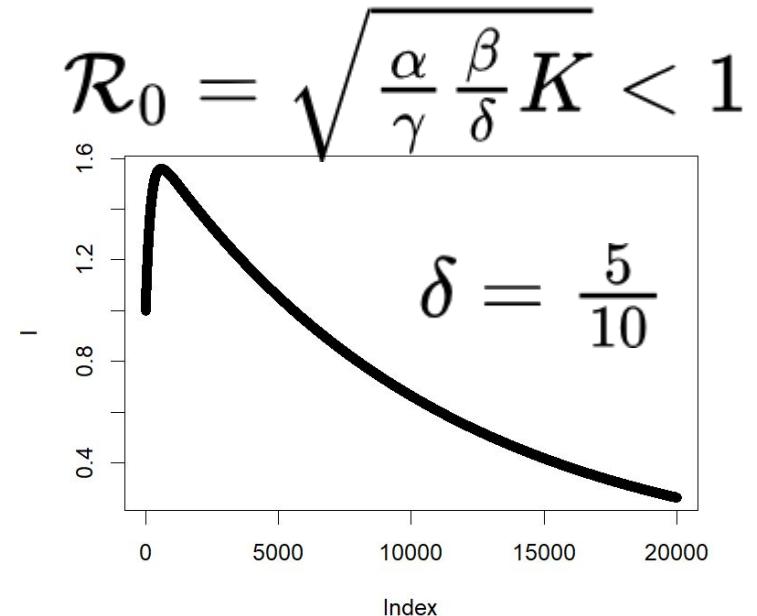
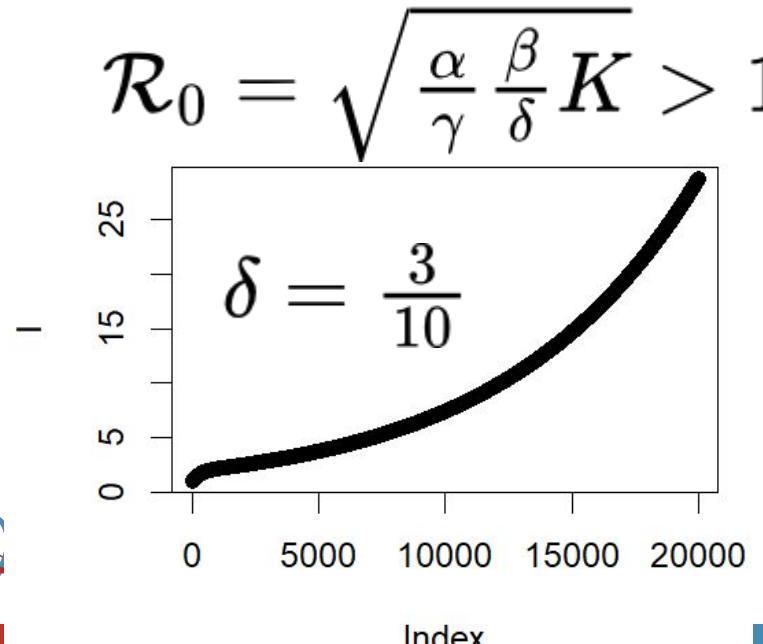


# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$     $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!

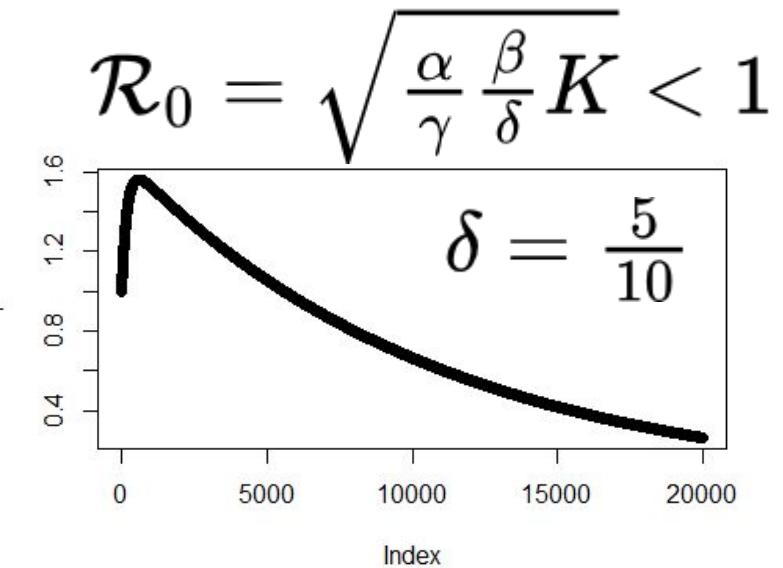
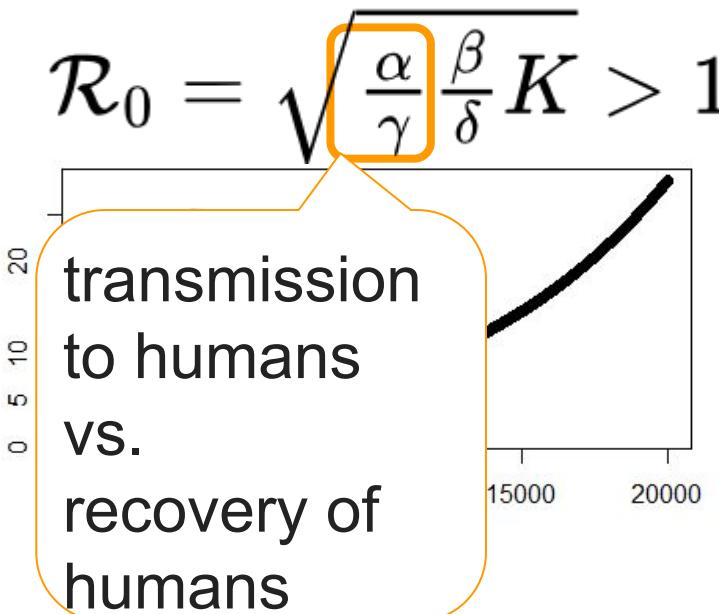


# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$   $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!



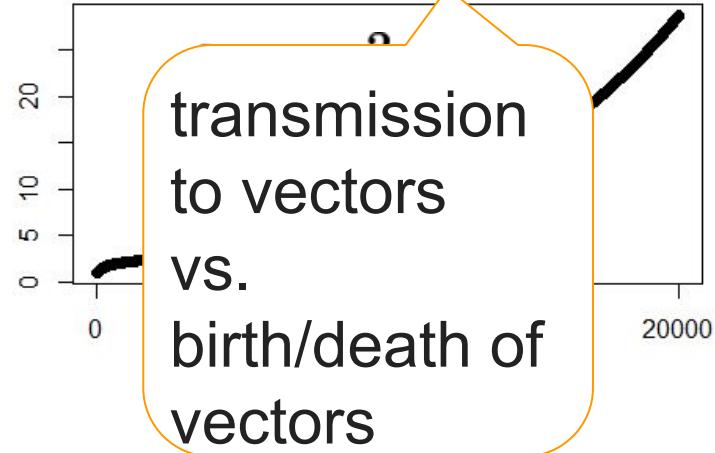
# Practical

## SIS-SI host-vector model with vector control

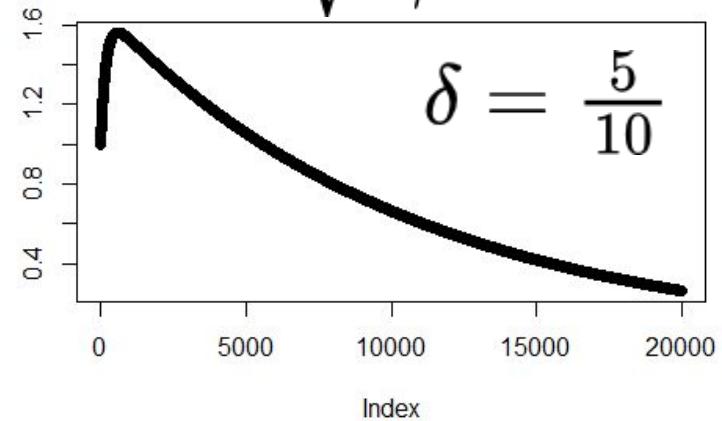
We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$   $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!

$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} > 1$$



$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} < 1$$

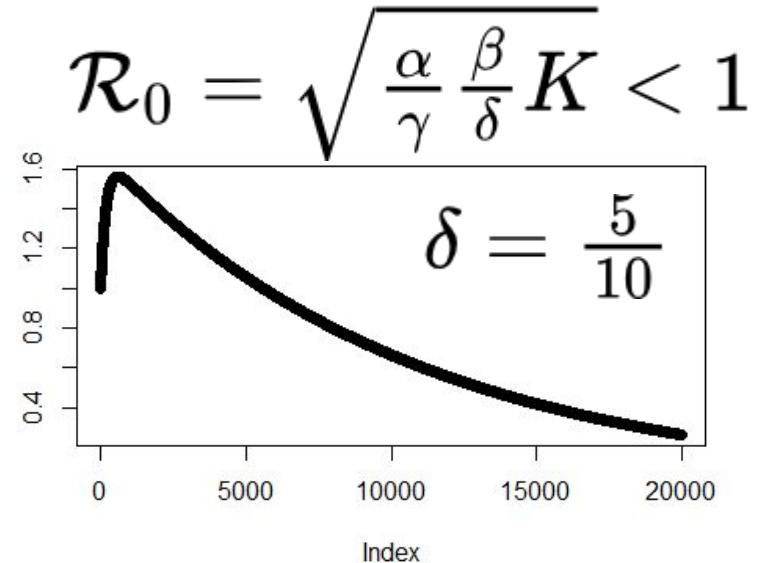
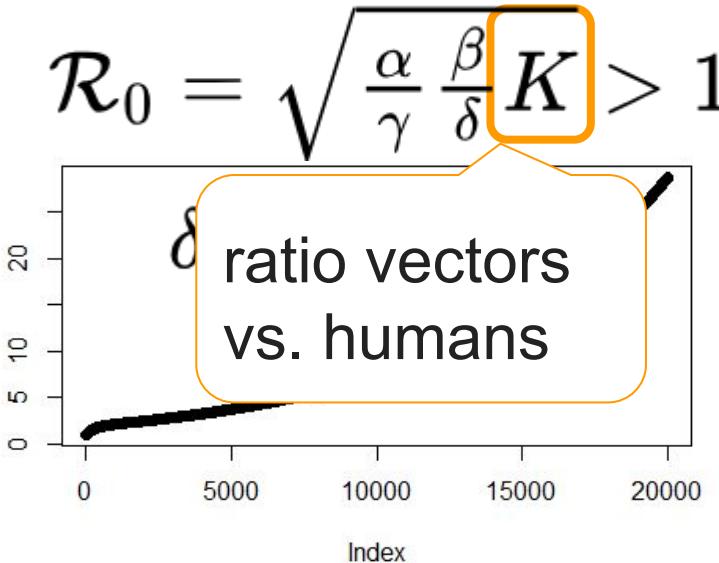


# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- simulation 1 as in the slides
- simulations 2&3 start from simulation 1, but add vector control as 3 times and 5 times the default mosquito mortality rate  $\delta = \frac{1}{10}$   $\mu = \delta$
- **compare** the solution curves for infections in humans! discuss!



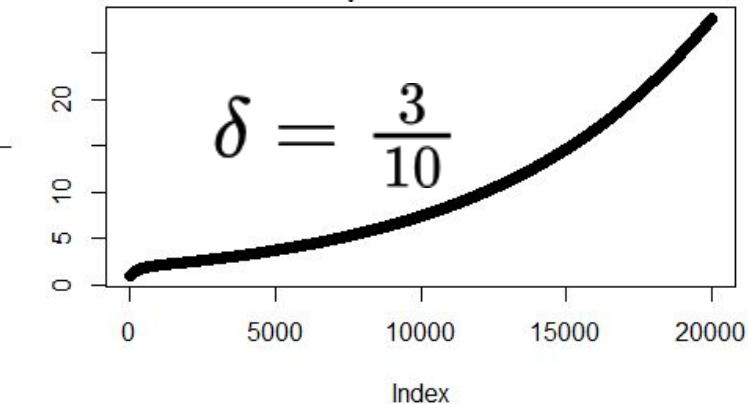
# Practical

## SIS-SI host-vector model with vector control

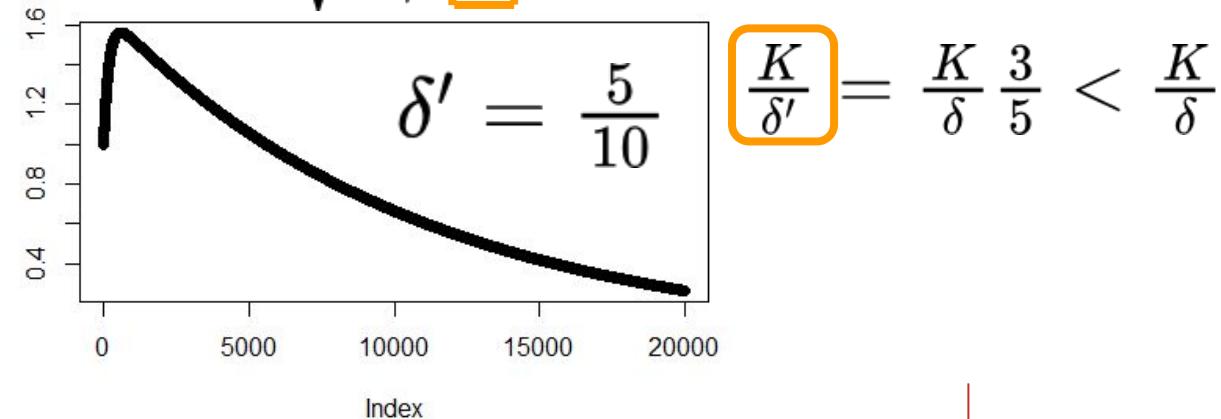
### Mosquito Theorem:

vector control is a sufficient condition for malaria elimination in humans

$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} > 1$$



$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta'} K} < 1$$



# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- Sweep through 20, 40, ..., 300% increase in mosquito mortality  $\delta = \frac{1}{10}$ .
- Visualize the relationship between mosquito mortality and prevalence at the endemic equilibrium!
- **Hint:** write a function in R with  $\delta$  as input parameter and **prevalence** at the endemic equilibrium as output

```
prevalence_EE<-function(delta) {  
  mu<-delta  
  delta<-delta  
  #... all the simulation code goes here!  
  prev_EE=tail(I,1)/H  
  output=data.frame(delta=delta,EE=prev_EE)  
  return(output)  
}
```

# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- Sweep through 20, 40, ..., 300% increase in mosquito mortality  $\delta = \frac{1}{10}$ .
- Visualize the relationship between mosquito mortality and prevalence at the endemic equilibrium!
- **Hint:** write a function in R with  $\delta$  as input parameter and **prevalence** at the endemic equilibrium as output

```
delta_increase<-seq(1,5,0.2)

lapply(1/10*delta_increase,prevalence_EE) %>%
  bind_rows () ->df

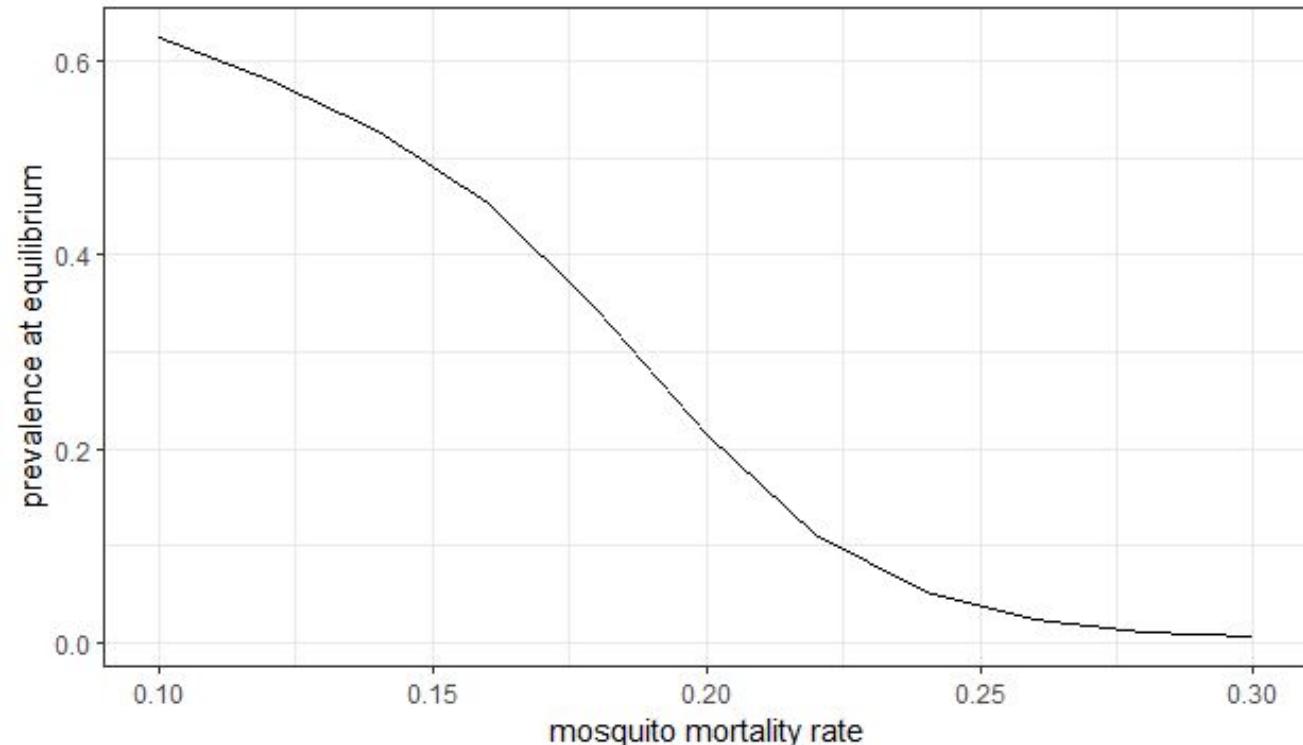
df%>%
  ggplot() + geom_line(aes(x=delta,y=EE)) +
  scale_x_continuous(name="mosquito mortality rate") +
  scale_y_continuous(name="prevalence at equilibrium") +
  theme_bw()
```

# Practical

## SIS-SI host-vector model with vector control

We want to run and **compare** two simulations in R:

- Sweep through 20, 40, ..., 300% increase in mosquito mortality  $\delta = \frac{1}{10}$ .
- Visualize the relationship between mosquito mortality and prevalence at the endemic equilibrium!



# Ramifications of our vector-host model

**What mechanisms should be included and how should those be implemented numerically?**

mechanism	model ramification	numerical considerations
<i>seasonality</i>	transmission to humans is not constant	time-dependent coefficients from periodic function
<i>immunity</i>	number of past episodes influences rate for productive infection	exposed compartments for human hosts
<i>climate</i>	vector birth/death parameters are not constant	monotone functions from calendar time to climate variable to vector parameters
<i>health system</i>	compartments for humans in various stages of treatment cascade	multi-stage Markov model on top of disease transmission model

# Ramifications of our vector-host model

**What mechanisms should be included and how should those be implemented numerically?**

**Israel's course!**

mechanism	model ramification	numerical considerations
<i>seasonality</i>	transmission to humans is not constant	time-dependent coefficients from periodic function
<i>immunity</i>	number of past episodes influences rate for productive infection	exposed compartments for human hosts
<i>climate</i>	vector birth/death parameters are not constant	monotone functions from calendar time to climate variable to vector parameters
<i>health system</i>	compartments for humans treatment cascade	multi-stage Markov model on top of disease transmission model

## Key takeaway points:

---

- **host-vector** models take into account vector population dynamics and biting
- **ratio between vector and host population** can switch between two distinct quantitative infection dynamics (extinction vs endemicity)
- vector control modeled as increased vector mortality allows to **link entomological** parameters to **infection** outcomes in humans

# Where can we get parameter values from?

$\frac{1}{\mu}$

time span from ovipositing to emergence of adult female mosquito from pupa stage, given ovipositing

$\frac{1}{\delta}$

life span of adult female mosquito

$\frac{1}{\gamma}$

duration of infection in human

$\alpha$

host seeking/biting X probability per bite for infectious mosquito to transmit to human: "**entomological inoculation**"

$\beta$

host seeking/biting X acquisition rate from infectious human with gametocytemia to mosquito

death  
 $\delta$

susceptible humans S

susceptible mosquitoes U

recovery  
 $\gamma$

transmission from humans to mosquitoes  
 $\beta$

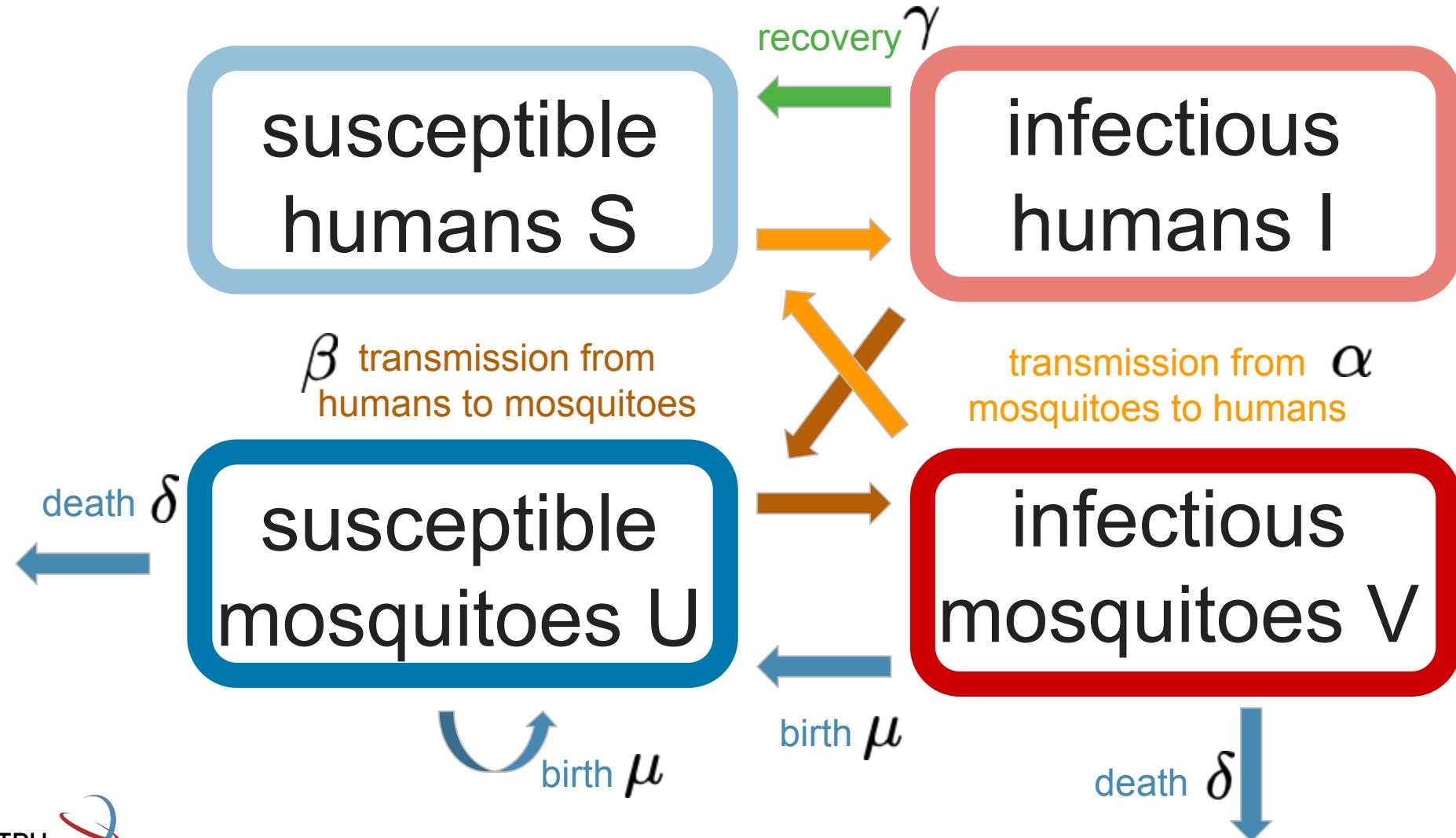
transmission from mosquitoes to humans  
 $\alpha$

birth  
 $\mu$

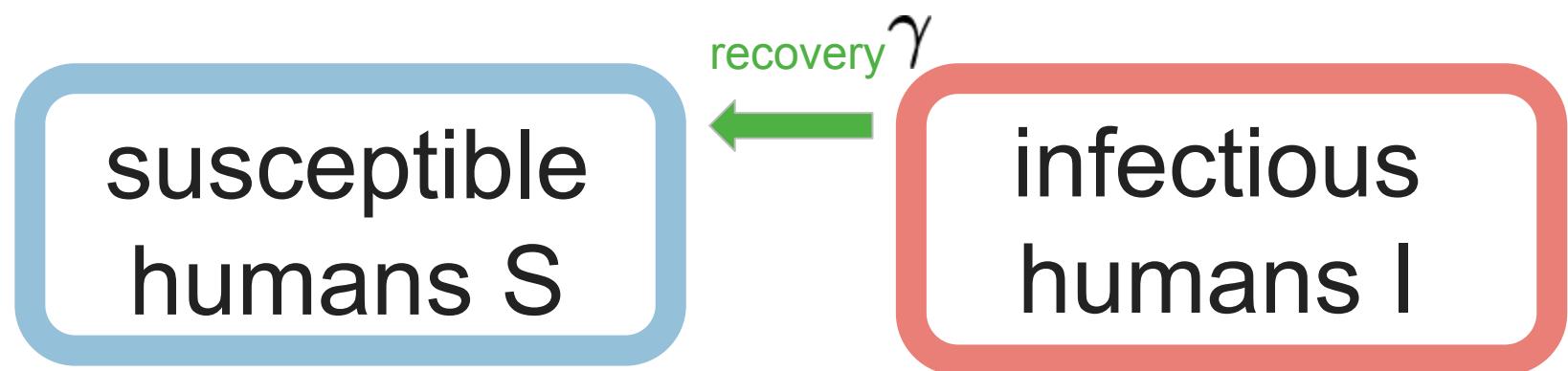
birth  
 $\mu$

death  
 $\delta$

# Malaria transmission model with vector control



# From infection duration to recovery rates

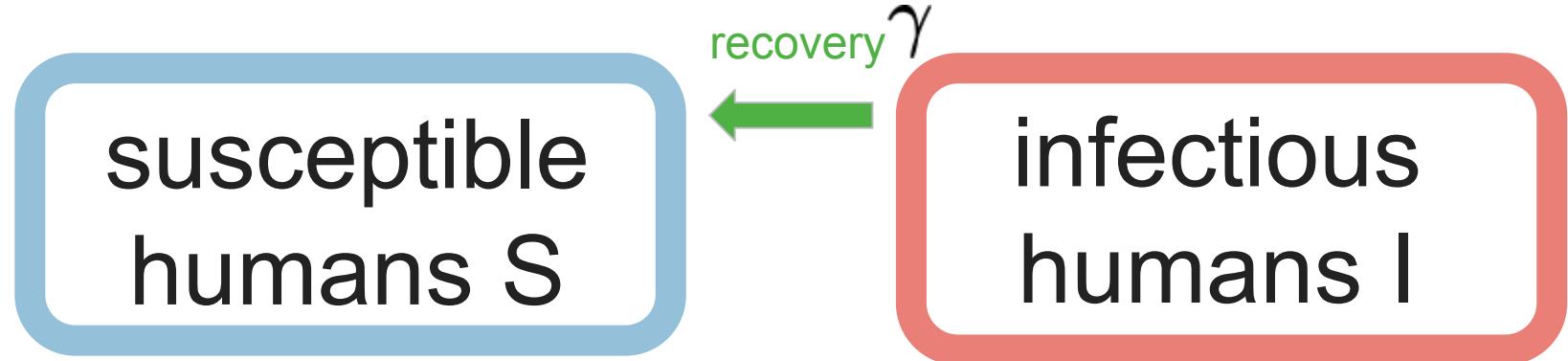


$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t)$$

$$\Delta \ll 1$$

**rate**  $\gamma$  = expected number of **recovery** events between time  $t$  and  $t + \Delta$

# From infection duration to recovery rates



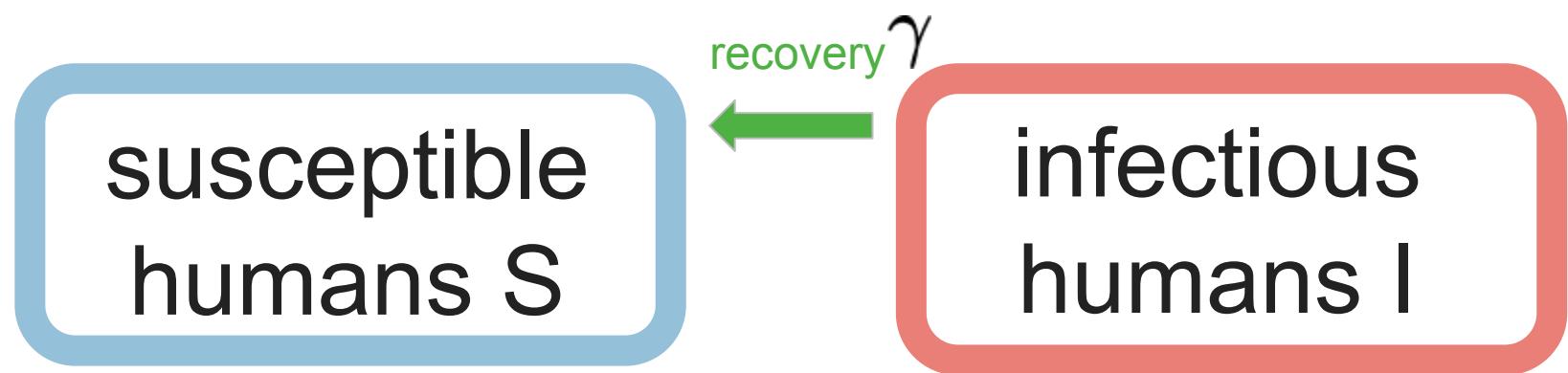
$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

$\Delta \ll 1$

**rate  $\gamma$**  = expected number of **recovery** events between time  $t$  and  $t + \Delta$

**recovery** events in individuals are independent from each other

# From infection duration to recovery rates



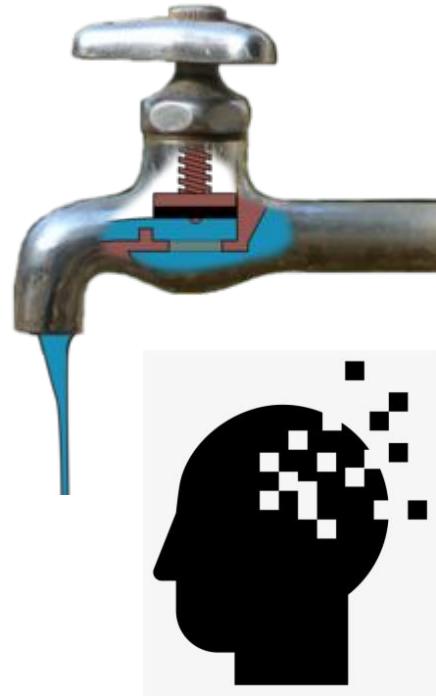
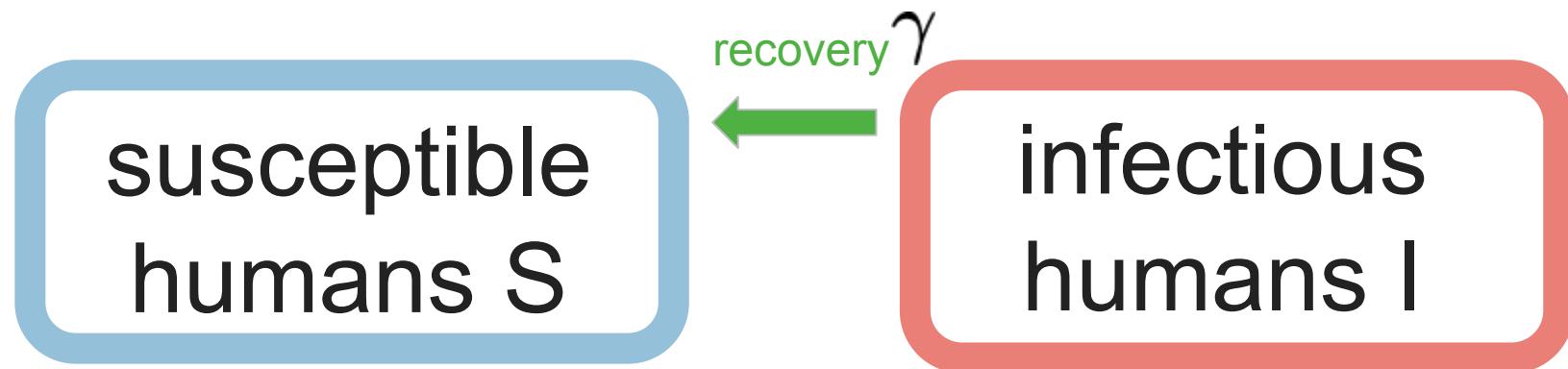
$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

$\Delta \ll 1$

**rate  $\gamma$**  = expected number of **recovery** events between time  $t$  and  $t + \Delta$

**R** random variable: **time to recovery** for infected hosts

# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

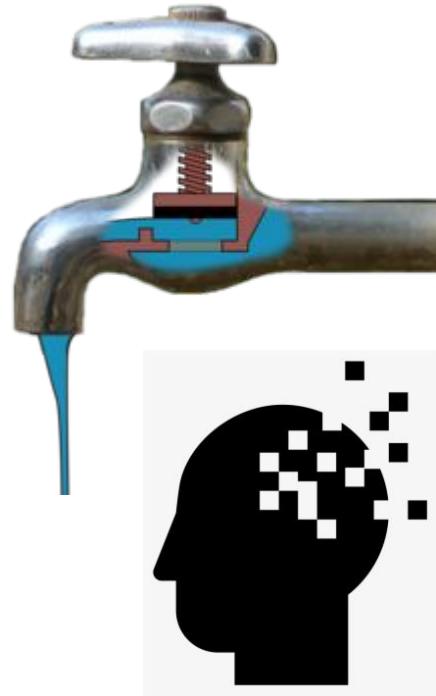
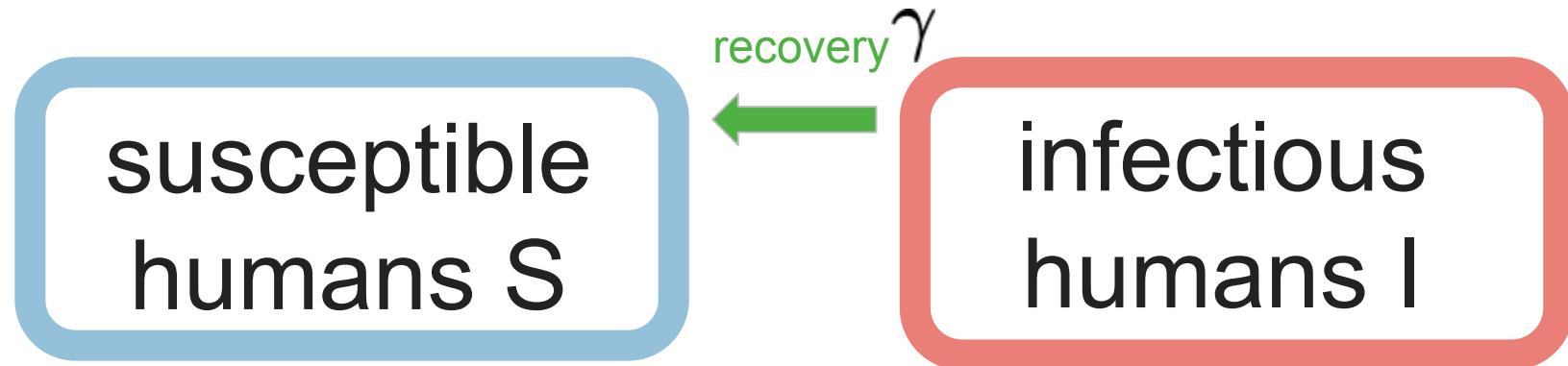
$$\Delta \ll 1$$

$R$  random variable: time to recovery for infected hosts

**Assumption:** time to recovery  $R$  is **memory-less**

$$\mathbb{P}(R > t + \Delta | R > t) = \mathbb{P}(R > \Delta)$$

# From infection duration to recovery rates



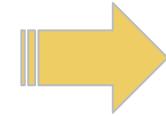
$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \dots + \gamma)}_{I(t) \text{ times}}$$

$\Delta \ll 1$

$R$  random variable: time to recovery for infected hosts

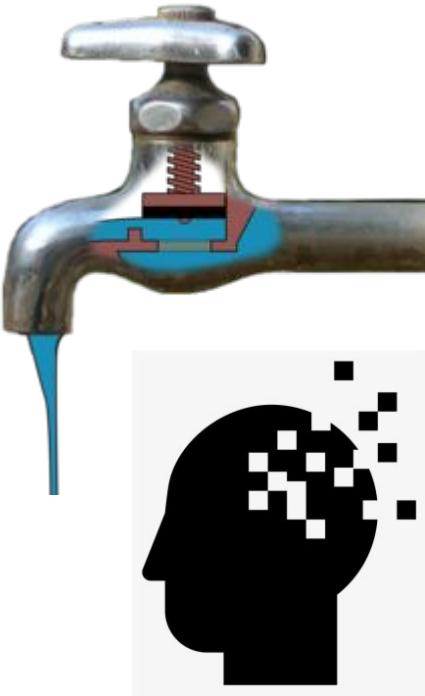
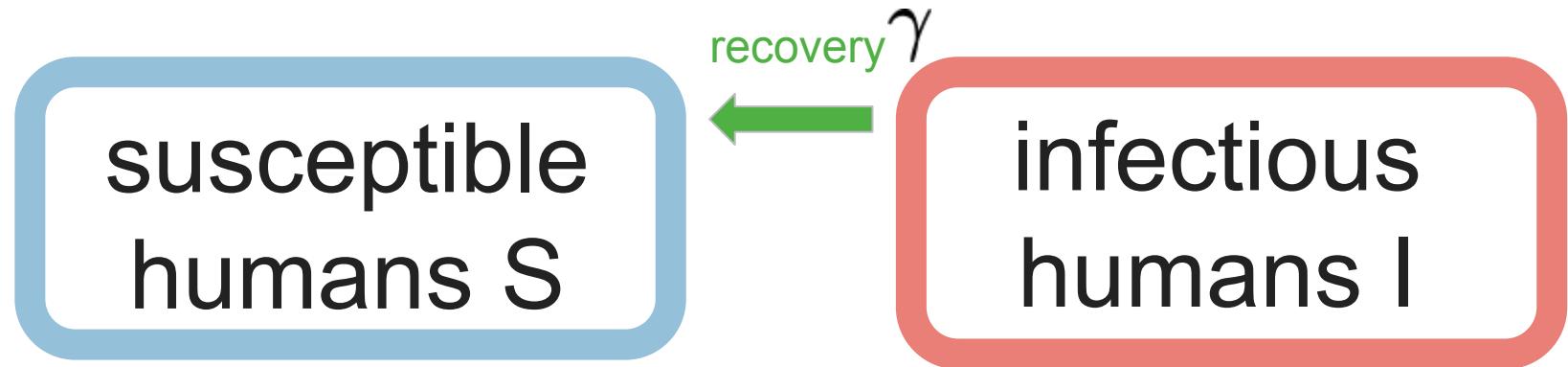
**Assumption:** time to recovery  $R$  is **memory-less**

$$\mathbb{P}(R > t + \Delta | R > t) = \mathbb{P}(R > \Delta)$$



$R$  must follow exponential distribution!

# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

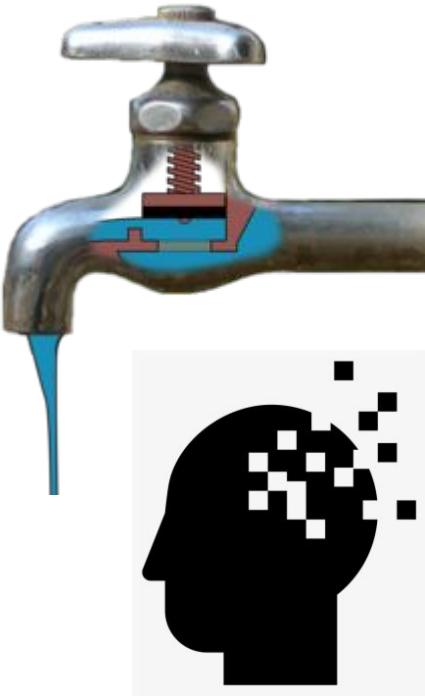
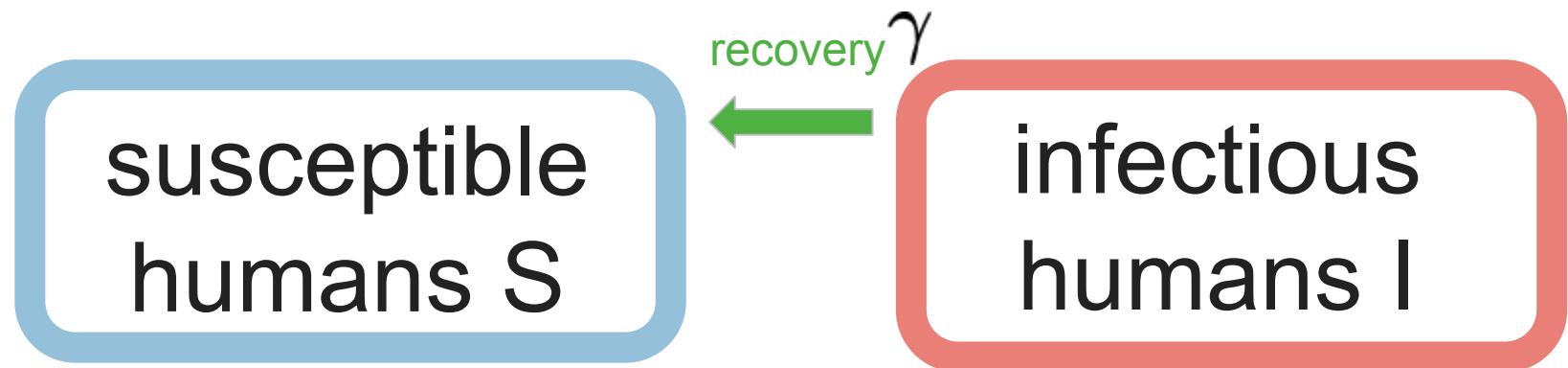
$$\Delta \ll 1$$

$R$  random variable: time to recovery for infected hosts

$R$  has exponential distribution

$$\mathbb{P}(R > \Delta) = e^{-\gamma\Delta} \quad \text{for } \gamma = -\ln \mathbb{P}(R > 1)$$

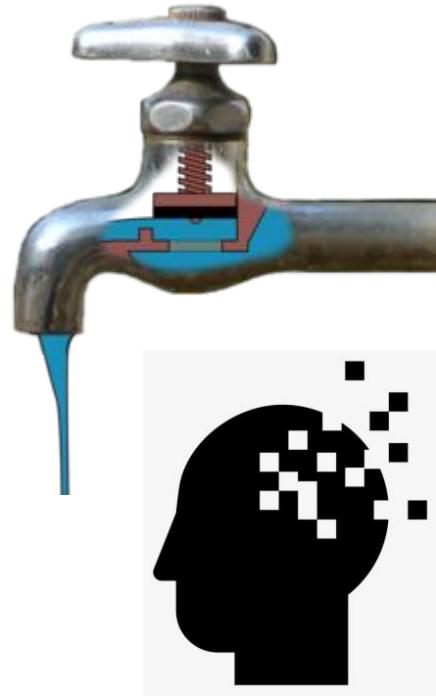
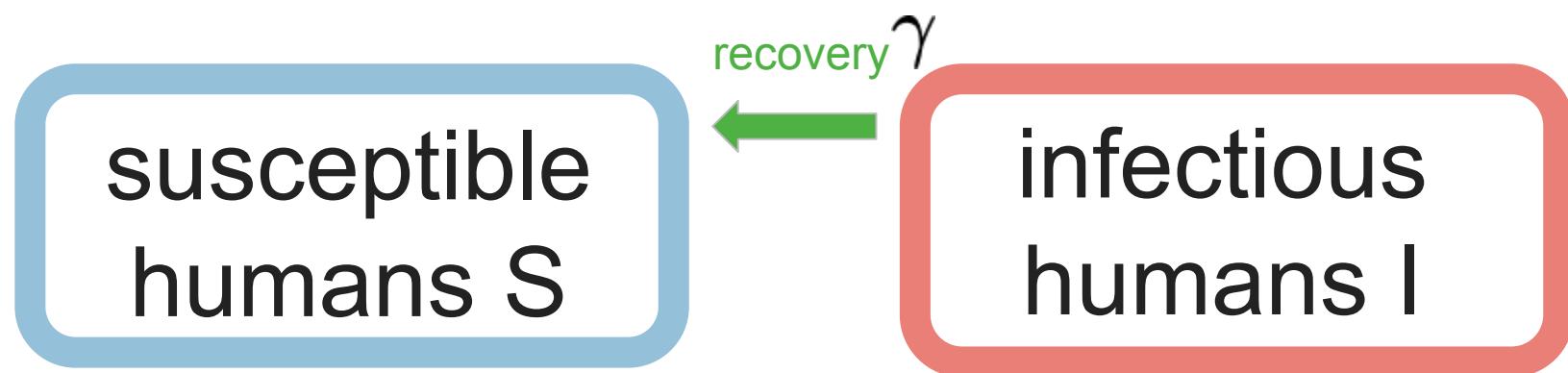
# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

**rate  $\gamma$**  = expected number of **recovery** events between time  $t$  and  $t + \Delta$

# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

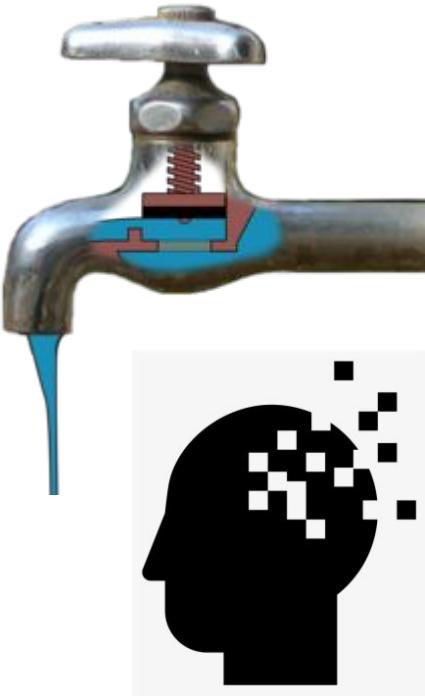
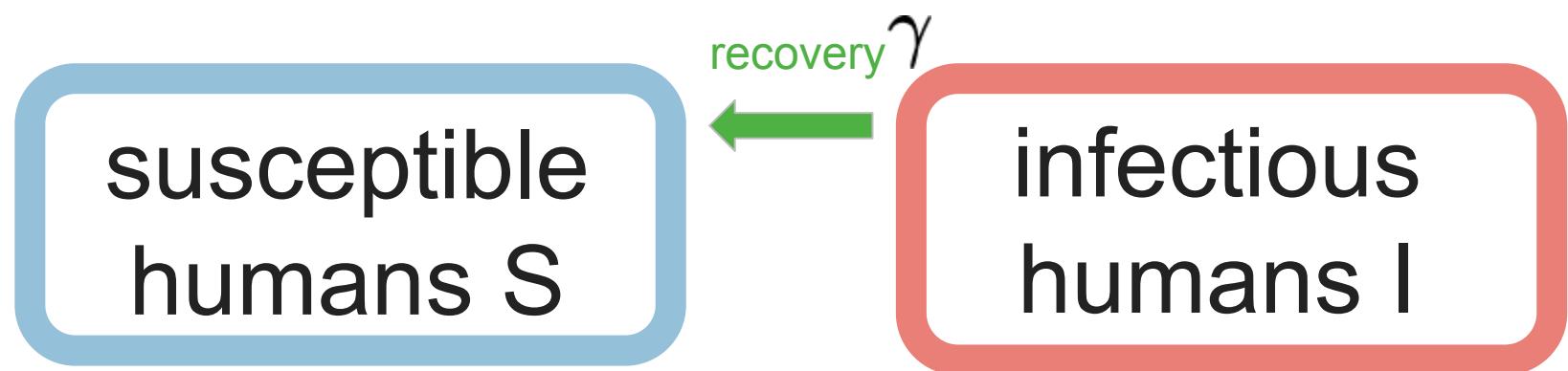
**rate**  $\gamma$  = expected number of **recovery** events between time  $t$  and  $t + \Delta$

$$\frac{1}{\Delta} \mathbb{P}(t < R < t + \Delta | R > t) = \frac{1}{\Delta} \frac{\int_t^{t+\Delta} \gamma e^{-\gamma x} dx}{e^{-\gamma t}} \approx \gamma \quad \Delta \ll 1$$

frequency of instantaneous recovery events given exponential distribution

"hazard function" or "intensity of recovery"

# From infection duration to recovery rates

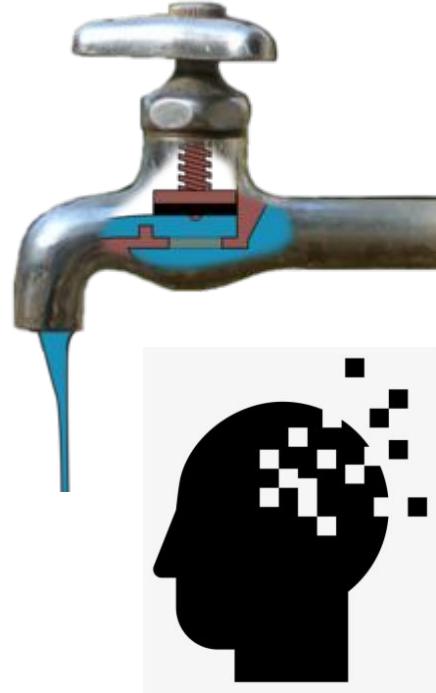
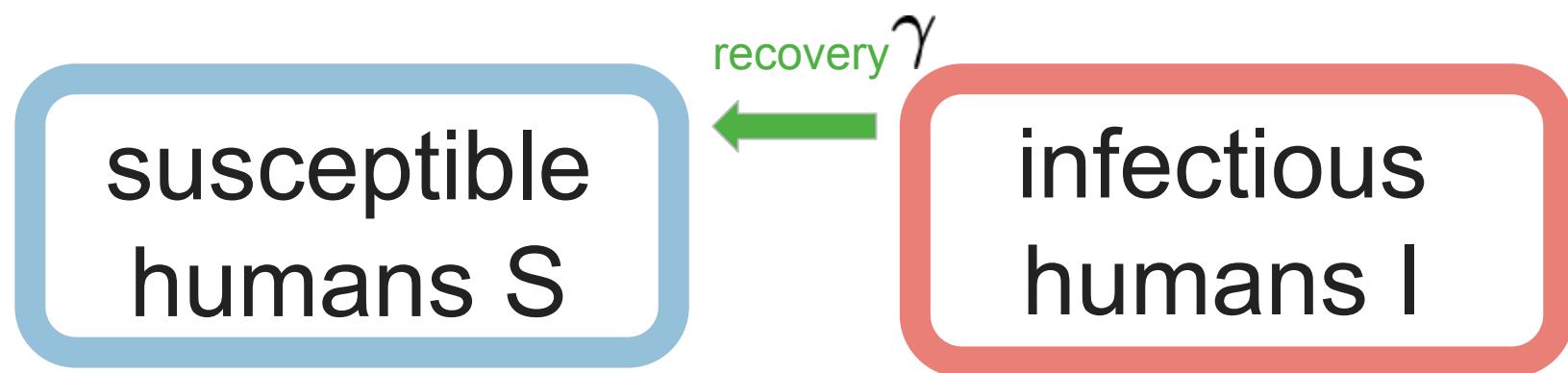


$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

$\gamma$  = intensity of recovery

Which data can be used to get a numerical value for  $\gamma$  ?

# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

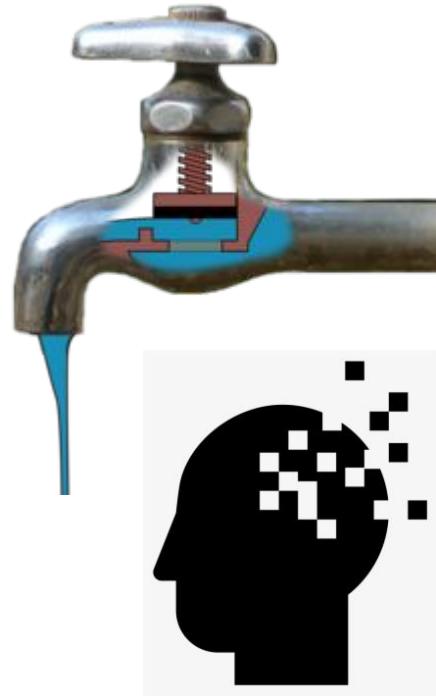
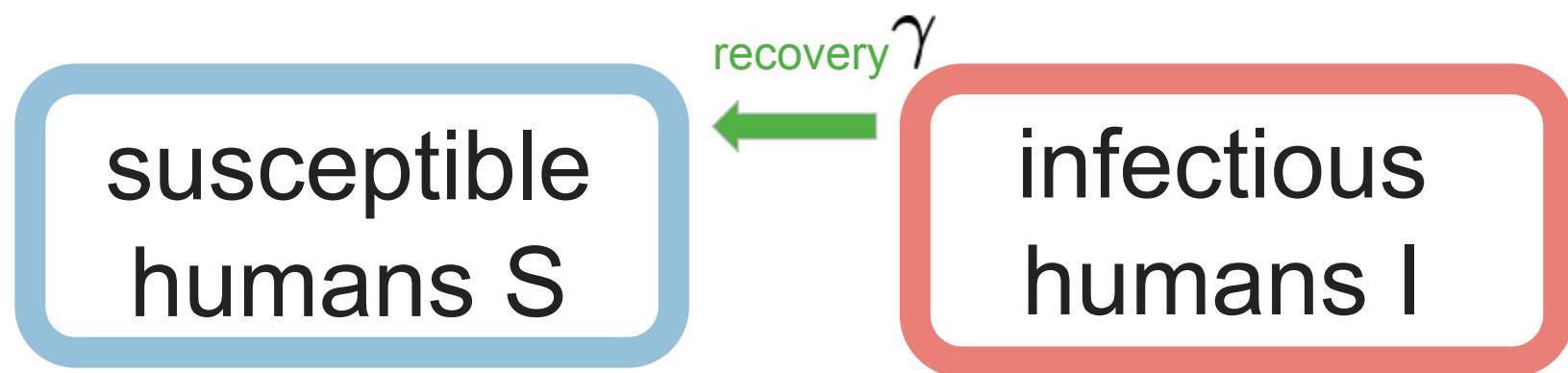
$\gamma$  = intensity of recovery

Which data can be used to get a numerical value for  $\gamma$  ?

time to recovery = infection duration

$R$  has exponential distribution:  $\mathbb{E}(R) = \frac{1}{\gamma}$

# From infection duration to recovery rates



$$\frac{I(t+\Delta) - I(t)}{\Delta} = -\gamma I(t) = -\underbrace{(\gamma + \cdots + \gamma)}_{I(t) \text{ times}}$$

$$\gamma = \text{intensity of recovery} = \frac{1}{\text{average infection duration}}$$

Which data can be used to get a numerical value for  $\gamma$  ?

time to recovery = infection duration

$R$  has exponential distribution:  $\mathbb{E}(R) = \frac{1}{\gamma}$

# Practical: Numerical values for parameters

## Infection duration

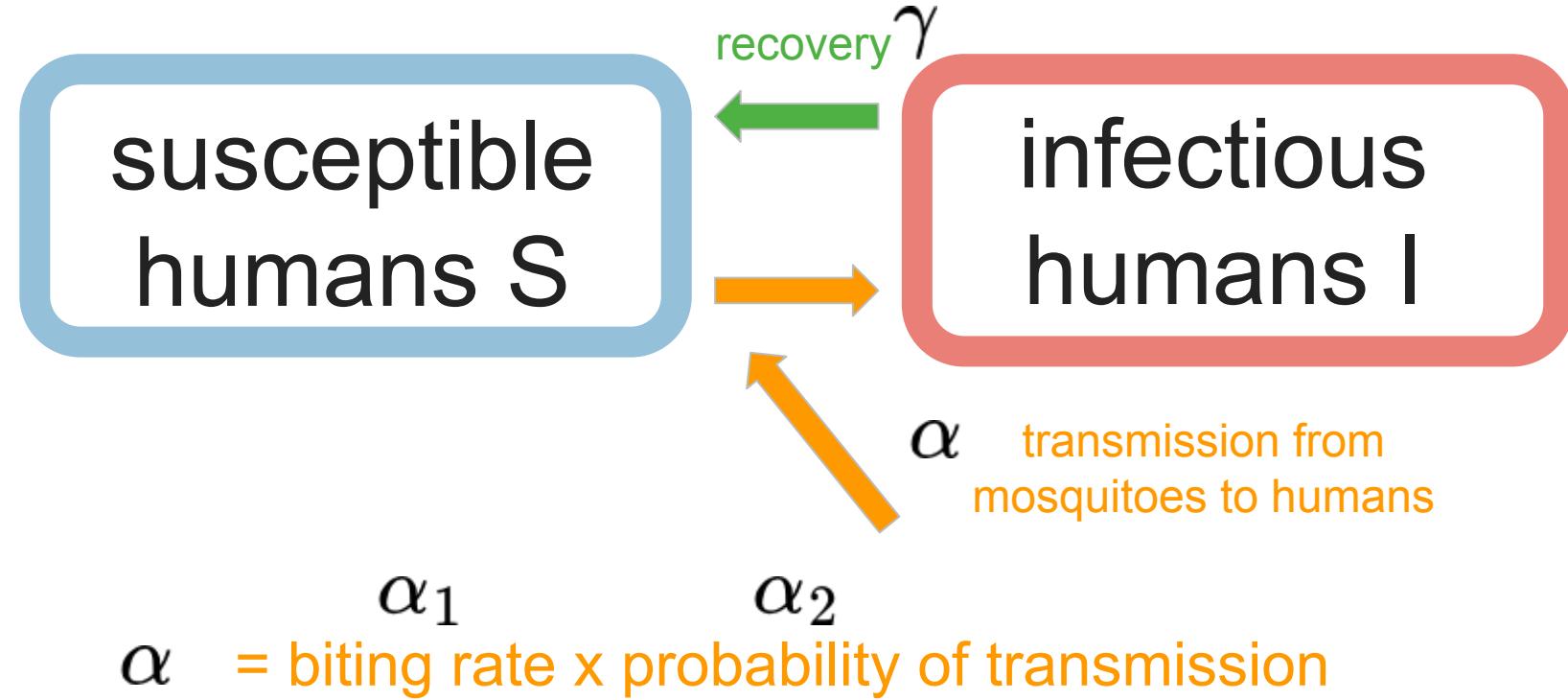
- Download the paper:

Chitnis et al.: *Determining Important Parameters in the Spread of Malaria Through the Sensitivity Analysis of a Mathematical Model*,  
Bull Math Bio 2008

- Discuss difference between our simplified Ross-McDonald model and equations 1a) - 1g) in the paper
- Search in the appendix A for possible values of **infection duration** in human hosts



# Malaria transmission model with vector control



host seeking  
number of bites mosquitoes can make  
number of bites human can receive

infectiousness of pathogen strain  
parasite sporozoite load

# Practical: Numerical values for parameters

## Biting rates and infection probability

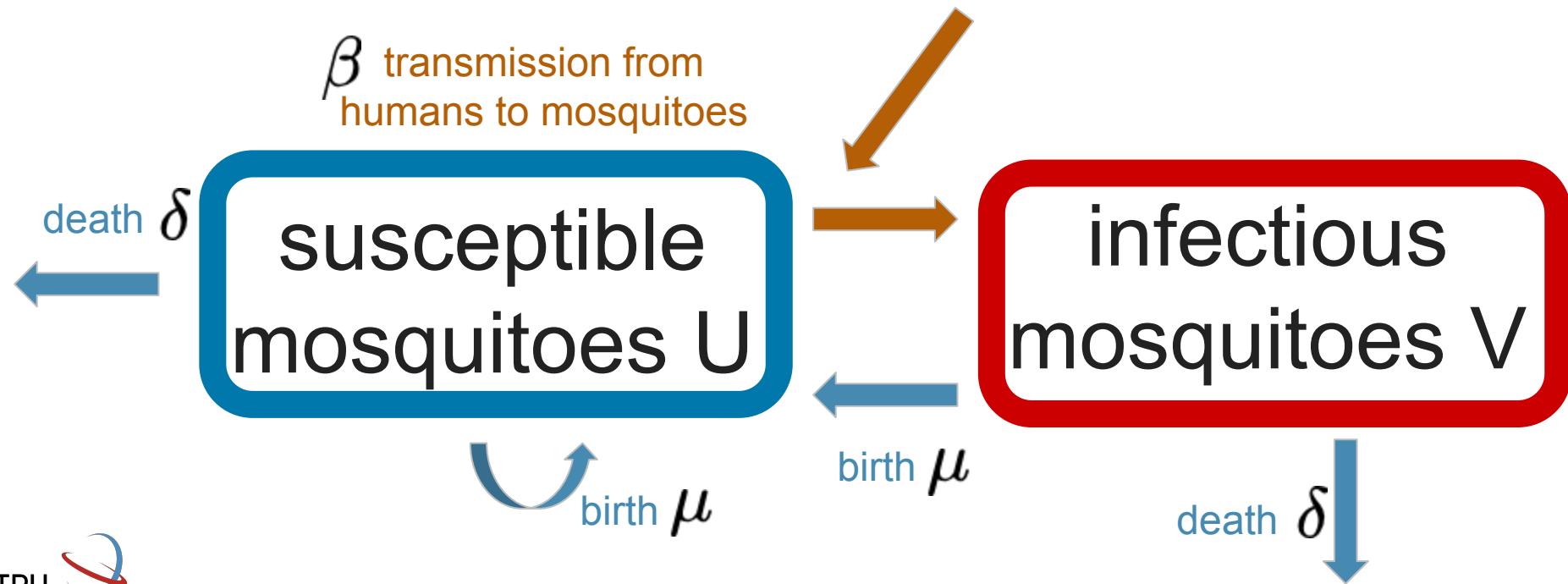
- Download the paper:

Chitnis et al.: *Determining Important Parameters in the Spread of Malaria Through the Sensitivity Analysis of a Mathematical Model*,  
Bull Math Bio 2008

- Search in the table 2&3 (baseline high) possible values of **biting rates** and **probability** of infection



# Malaria transmission model with vector control



# Practical: Numerical values for parameters

## Mosquito life cycle and infections to humans

- Download the paper:

Chitnis et al.: *Determining Important Parameters in the Spread of Malaria Through the Sensitivity Analysis of a Mathematical Model*,  
Bull Math Bio 2008

- Search in the table 2&3 (baseline high) possible values of **mosquito birth and death rates** and **probability** of infection from humans to mosquitoes



# Practical 6d: Simulate with numerical values!

parameter	description	value	unit
gamma	reciprocal of untreated infection duration		1/day
alpha_1	biting rate within gonotrophic cycle		1/day
alpha_2	probability of transmission to humans		1
delta_1	density-independent mosquito mortality rate		1/day
delta_2	density-dependent mosquito mortality rate		1/mosquito 1/day
mu	per capita mosquito birth rate		1/day
beta	probability of transmission to mosquitoes		1

# Practical 6d: Simulate with numerical values!

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	0.022	1
delta_1	density-independent mosquito mortality rate	0.033	1/day
delta_2	density-dependent mosquito mortality rate	0.00002	1/mosquito 1/day
mu	per capita mosquito birth rate	0.13	1/day
beta	probability of transmission to mosquitoes	0.48	1

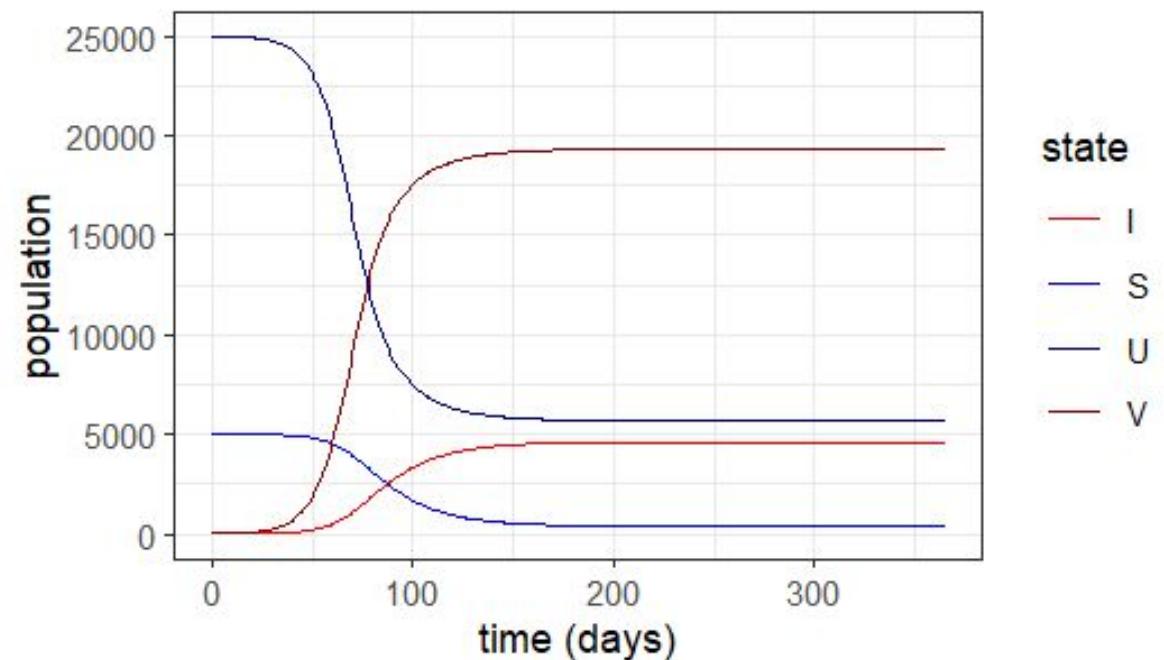
# Practical 6d: Simulate with numerical values!

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	0.022	1
delta_1	density-independent mosquito mortality rate	0.033	1/day
delta_2	density-dependent mosquito mortality rate	0.00002	1/mosquito 1/day
mu	per capita mosquito birth rate	0.13	1/day
beta	probability of transmission to mosquitoes	0.48	1

```
H=5000; I0=1; V0=8; VectorHumanRatio=5; finalT=1*365
x0 <- c(S=H-I0, I=I0, U=H*VectorHumanRatio-V0, V=V0)##initial condition
timesteps<-seq(0,finalT,1)##time unit in days
alpha1<- 0.5; alpha2<-0.022; alpha=alpha1*alpha2; delta1=0.033; delta2=2*10^-05
delta=delta1+H*VectorHumanRatio*delta2
parms <- c(alpha=alpha, gamma=1/285,beta=0.48, mu=0.13,delta=0.13)
results<-ode(x0,timesteps,RossMcDonald.model,parms) %>%
  as.data.frame
```

# Practical 6d: Simulate with numerical values!

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	0.022	1
delta_1	density-independent mosquito mortality rate	0.033	1/day
delta_2	density-dependent mosquito mortality rate	0.00002	1/mosquito 1/day
mu	per capita mosquito birth rate	0.1	
beta	probability of transmission to mosquitoes	0.4	



## Key takeaway points:

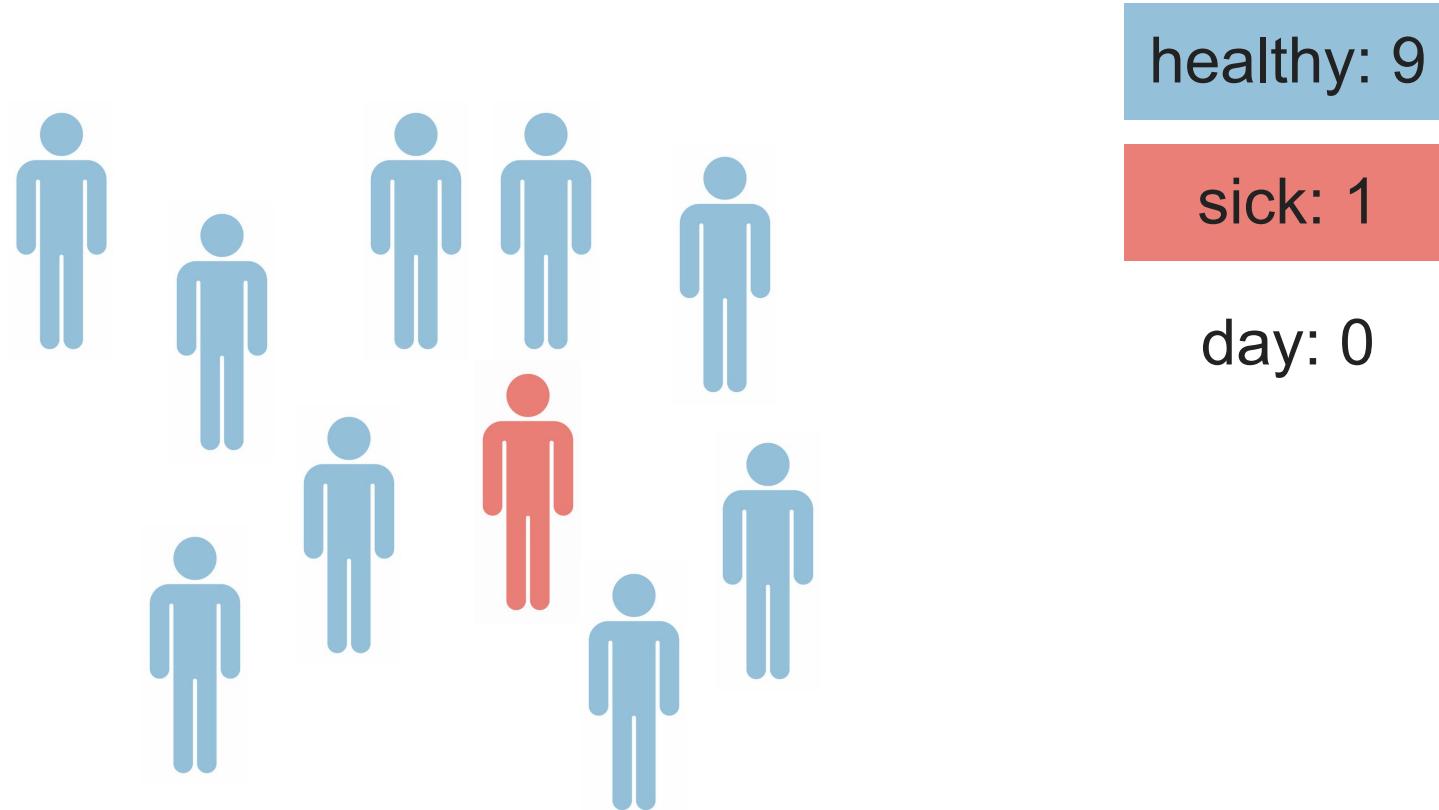
---

- ordinary differential equations are **memory-less**
- **time to event** in compartment (e.g. recovery) is interpreted as **exponentially distributed** random variable
- **rate** of such event is reciprocal of expected value, such that numerical values for rates can be derived from **average duration** data (e.g. infection, life,...)

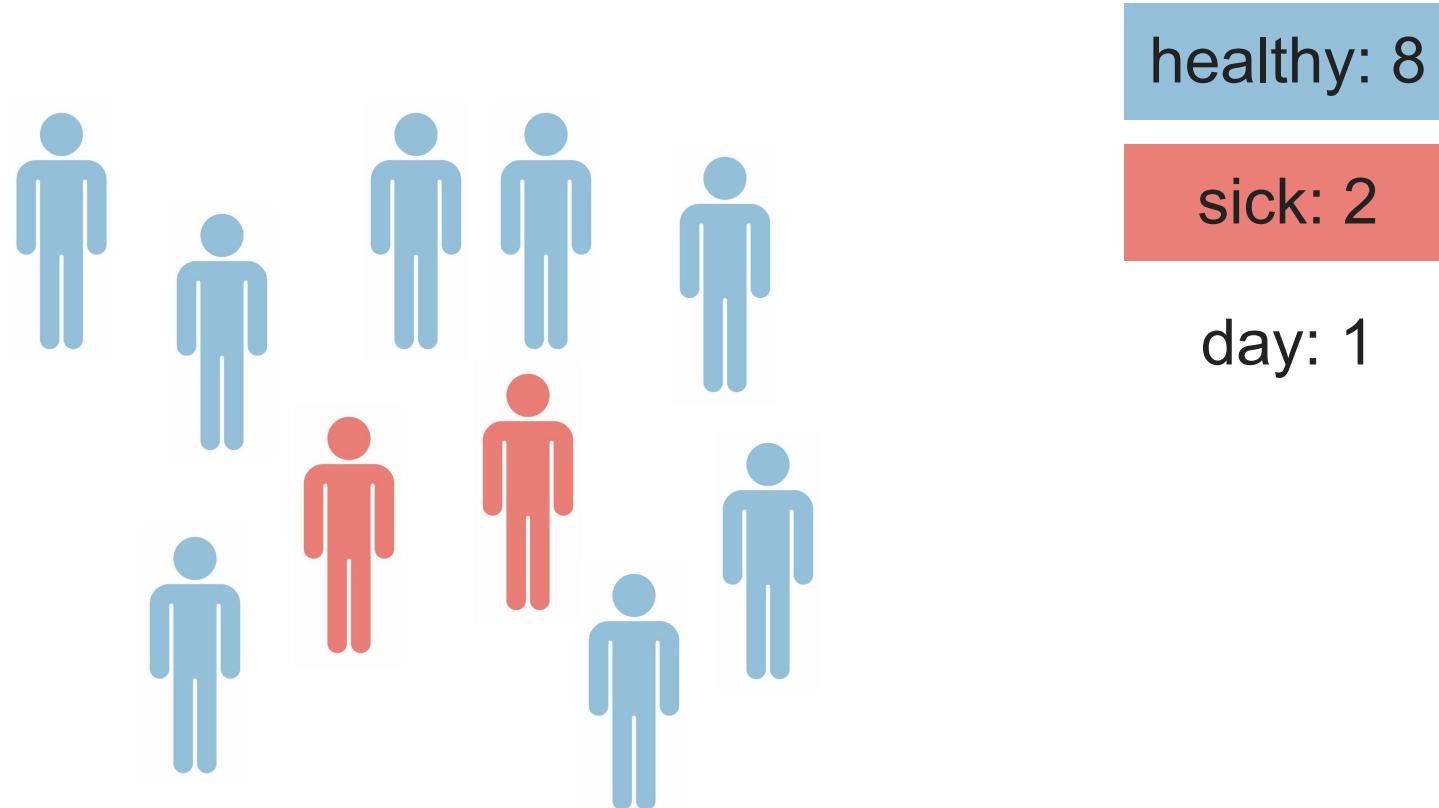


2 - R0

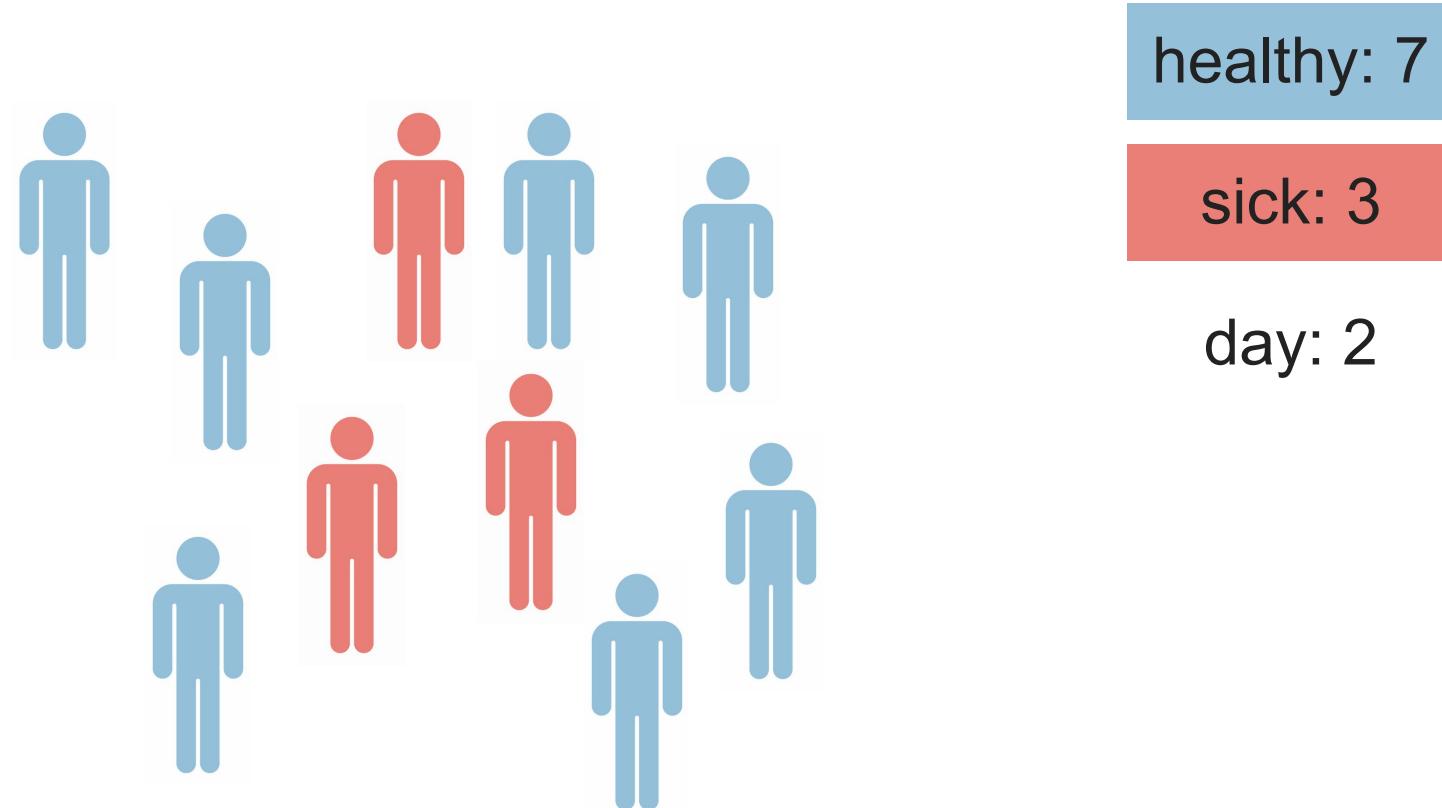
# Snapshots of disease prevalence



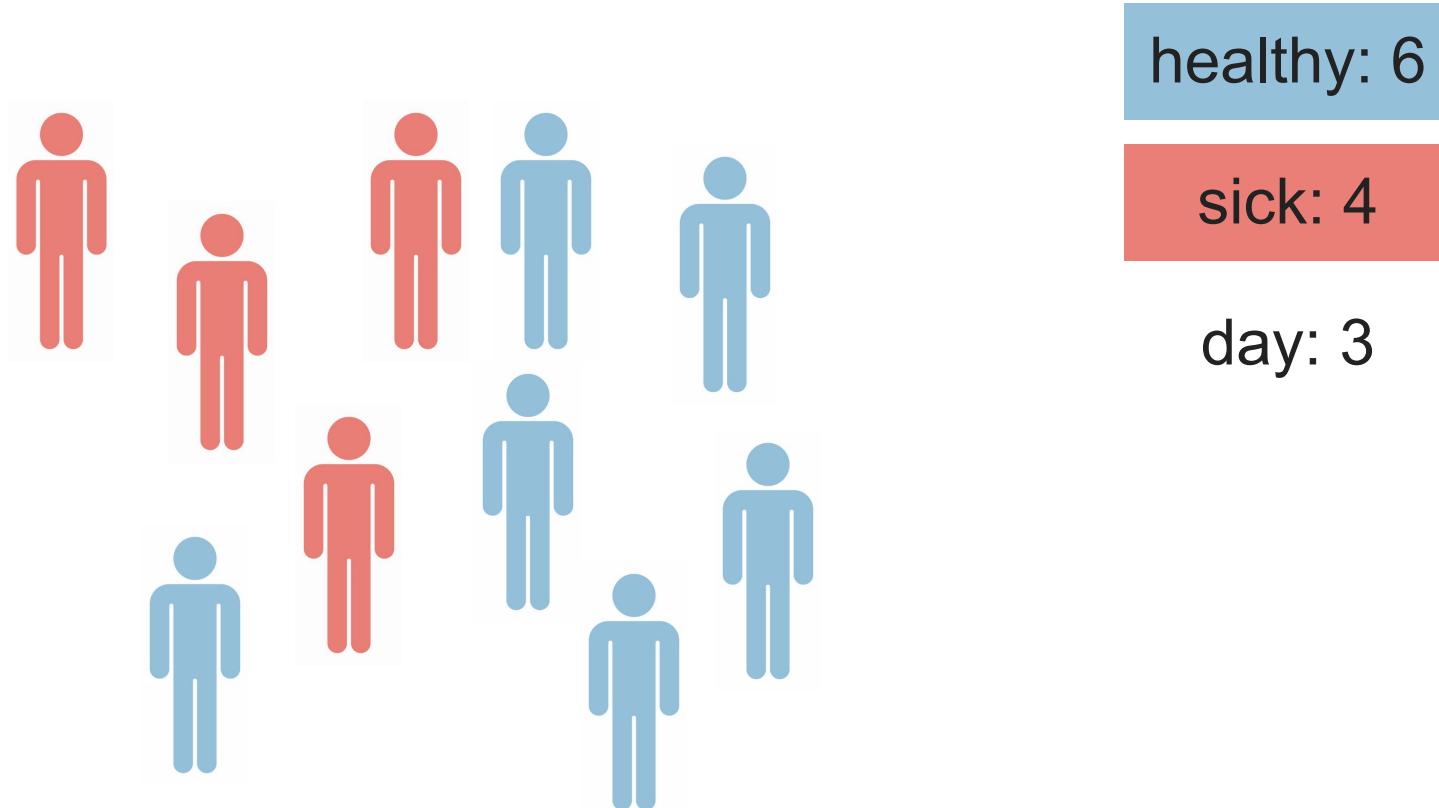
# Snapshots of disease prevalence



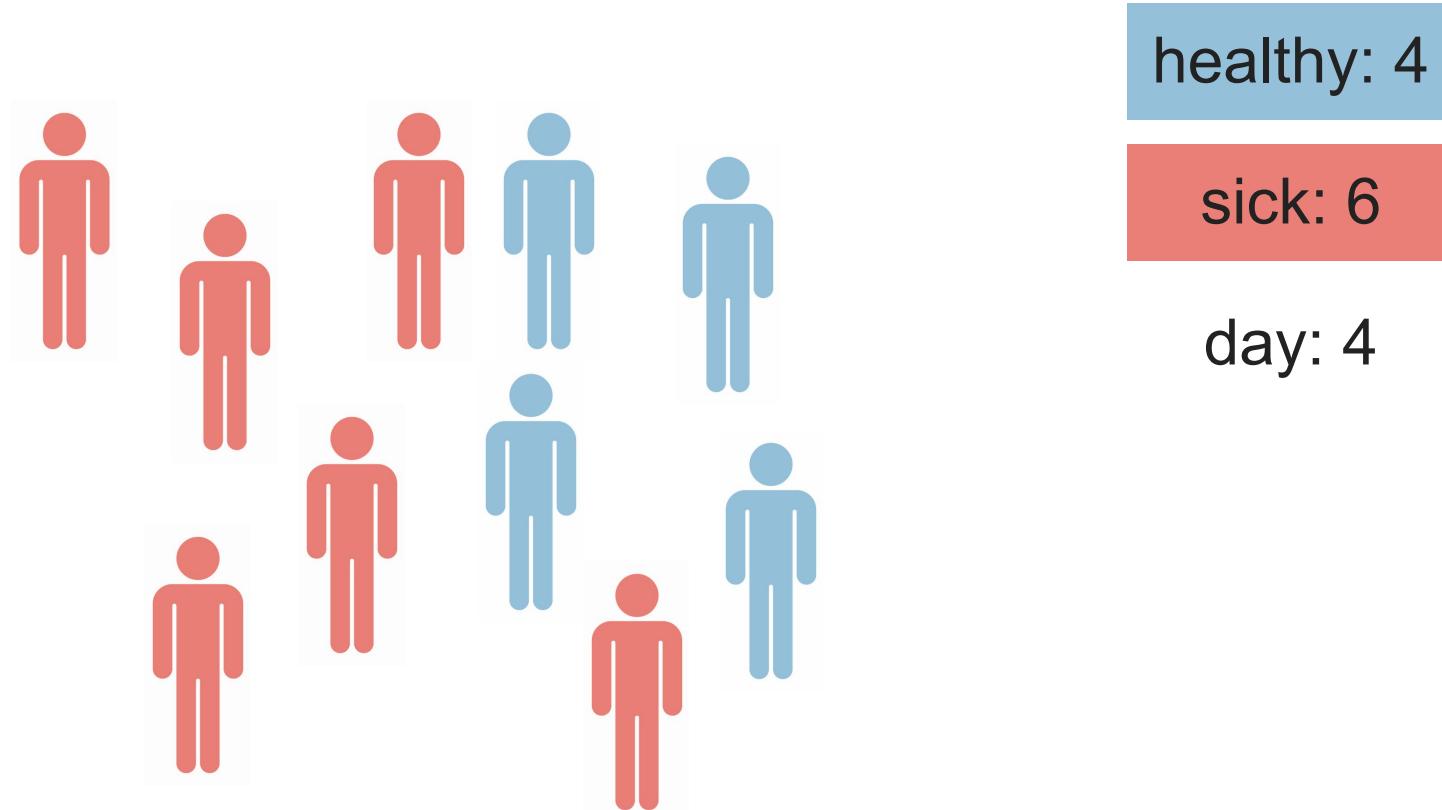
# Snapshots of disease prevalence



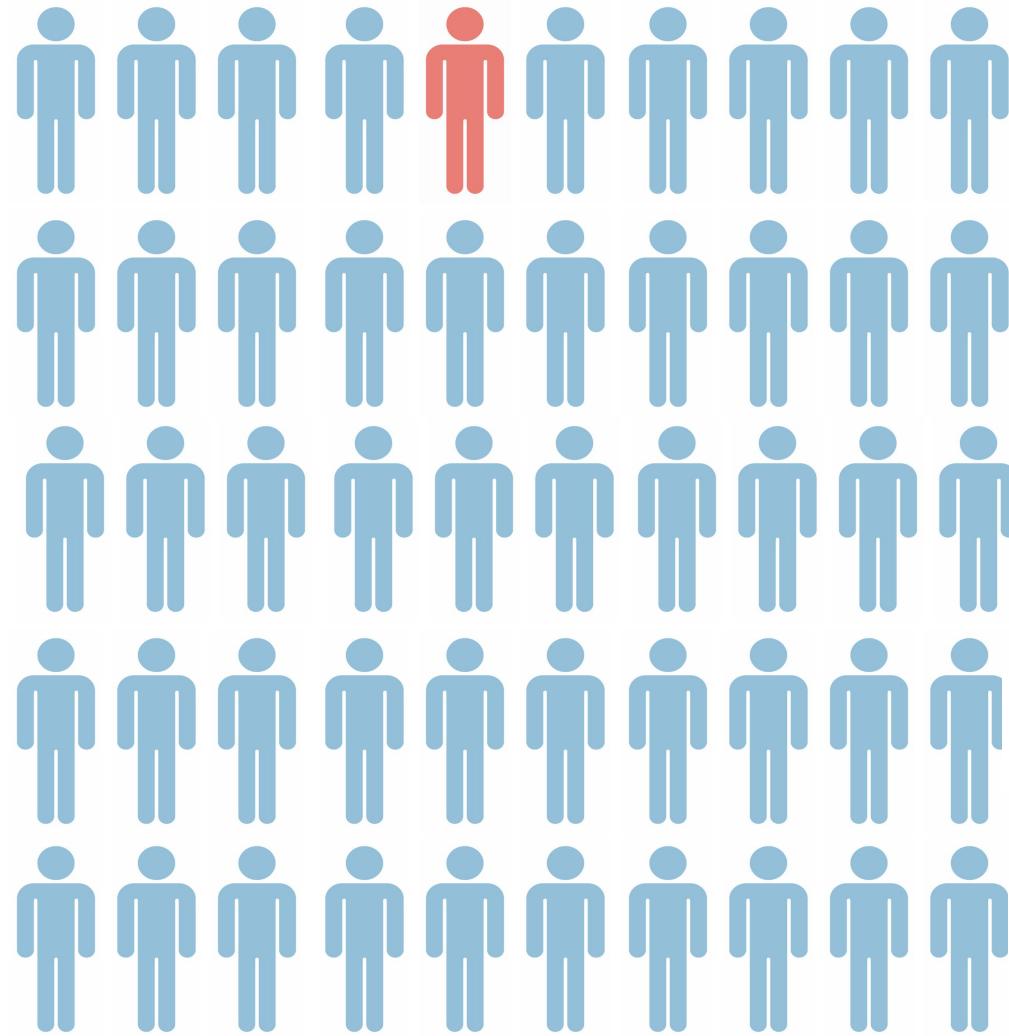
# Snapshots of disease prevalence



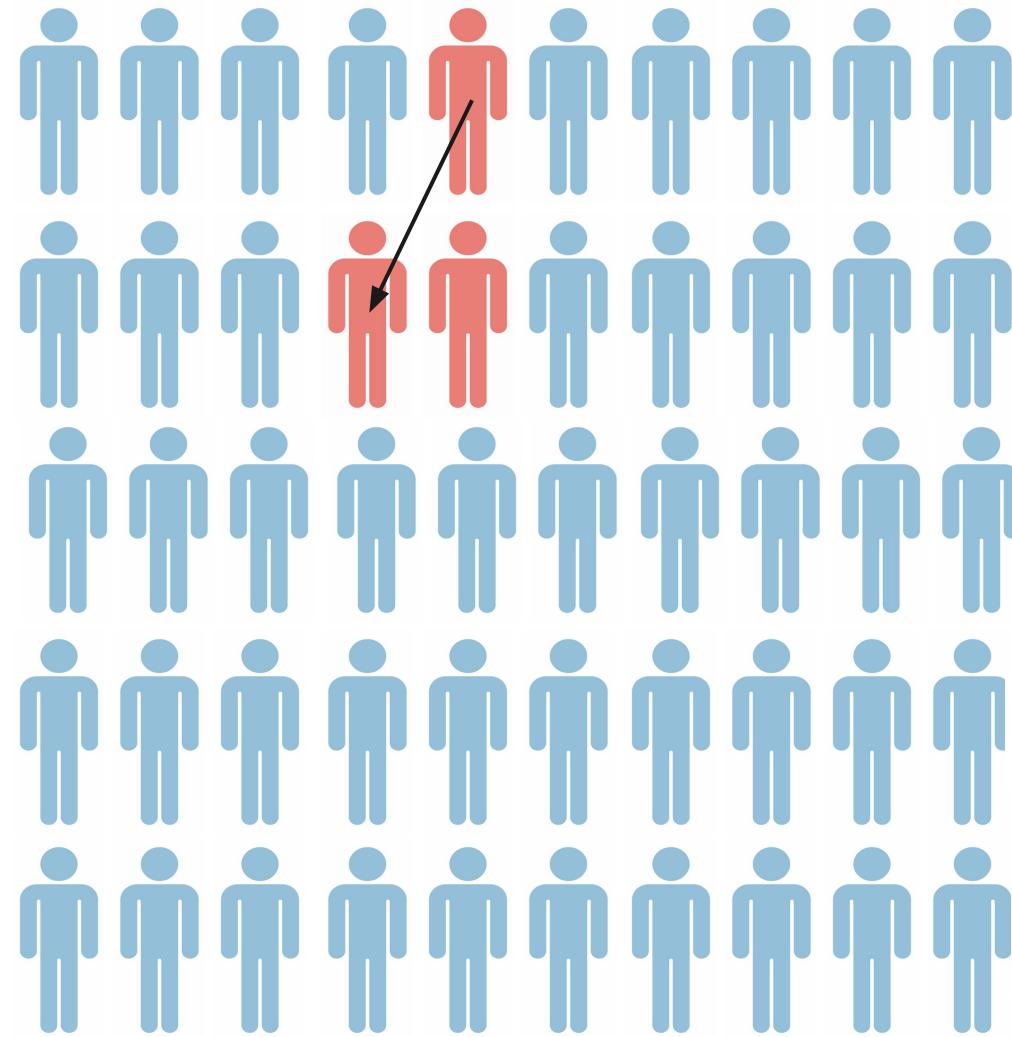
# Snapshots of disease prevalence



# Generations



# Generations



day: 0

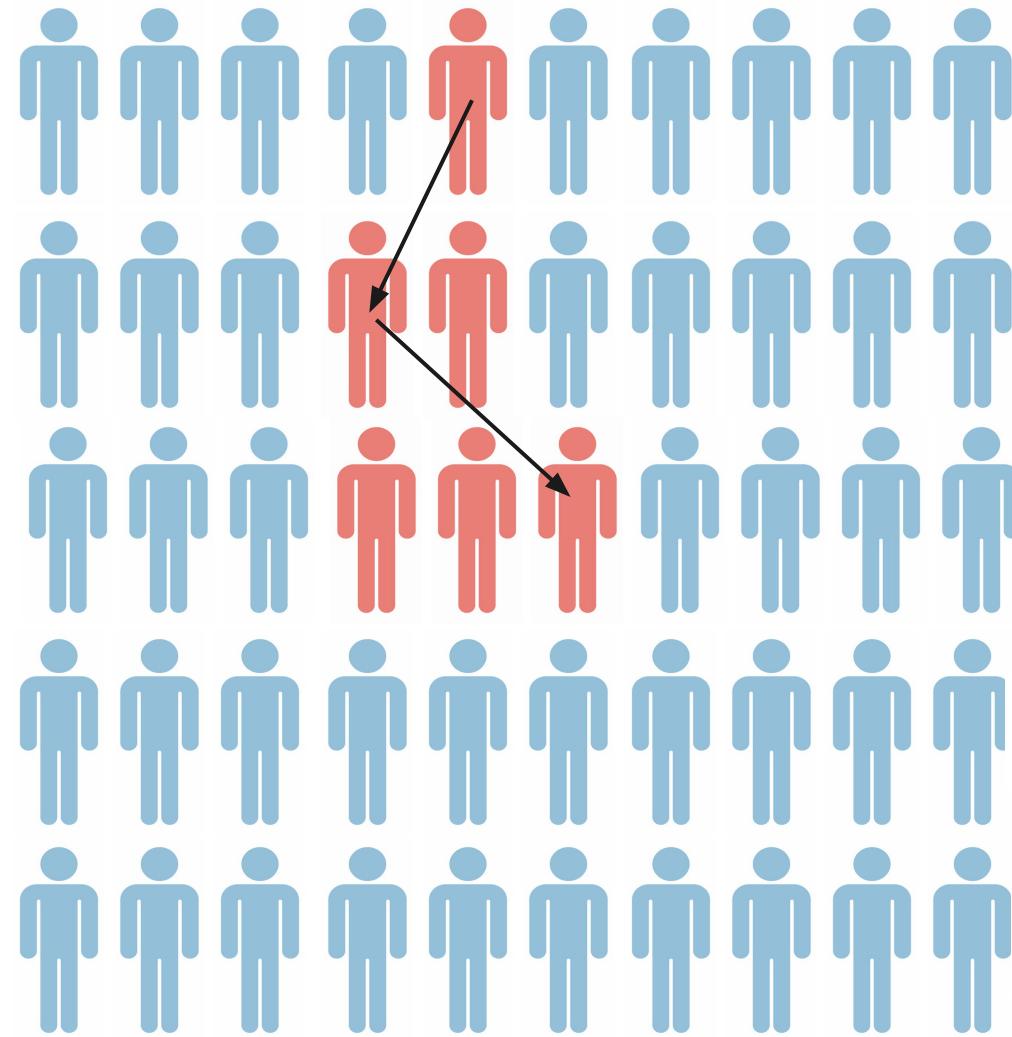
day: 1      new infection: +1

day: 2

day: 3

day: 4

# Generations



day: 0

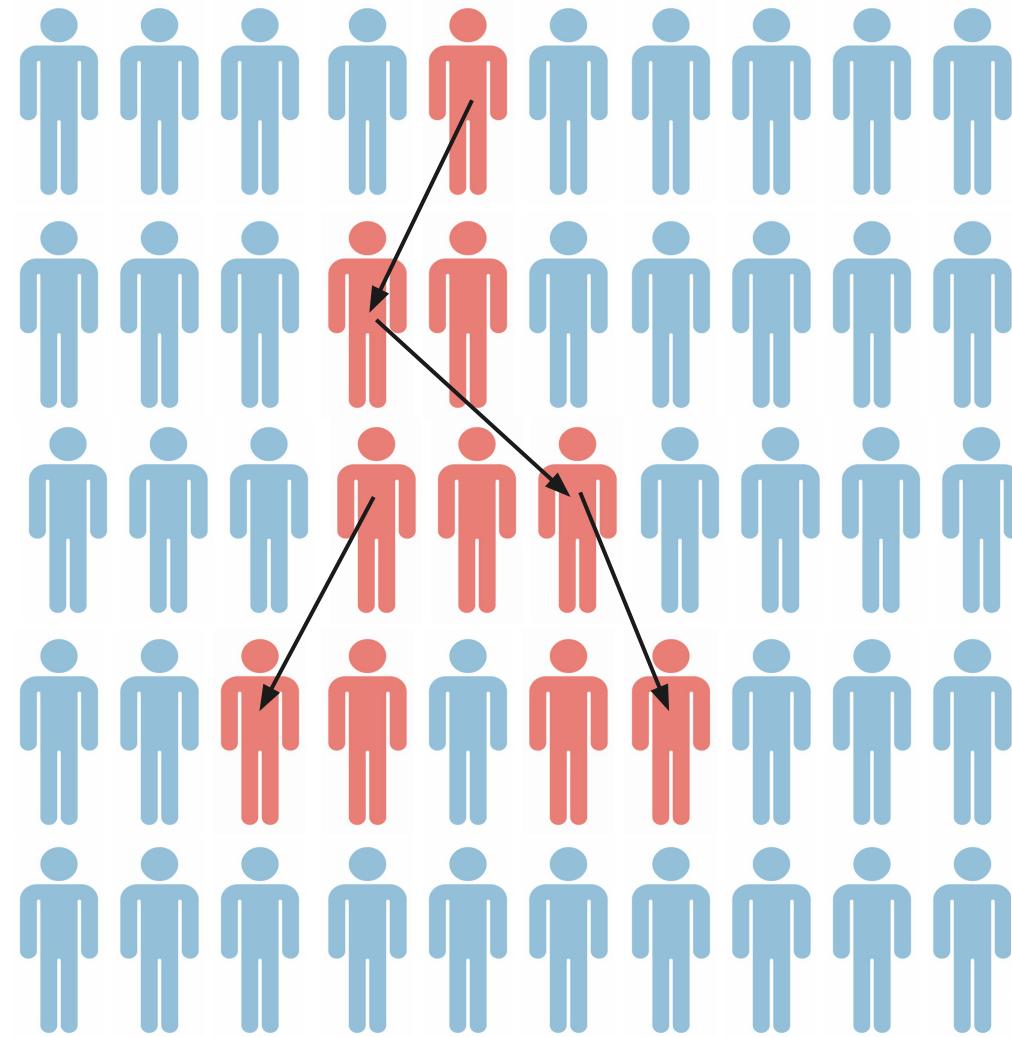
day: 1      new infection: +1

day: 2      new infection: +1

day: 3

day: 4

# Generations



day: 0

day: 1

new infection: +1

day: 2

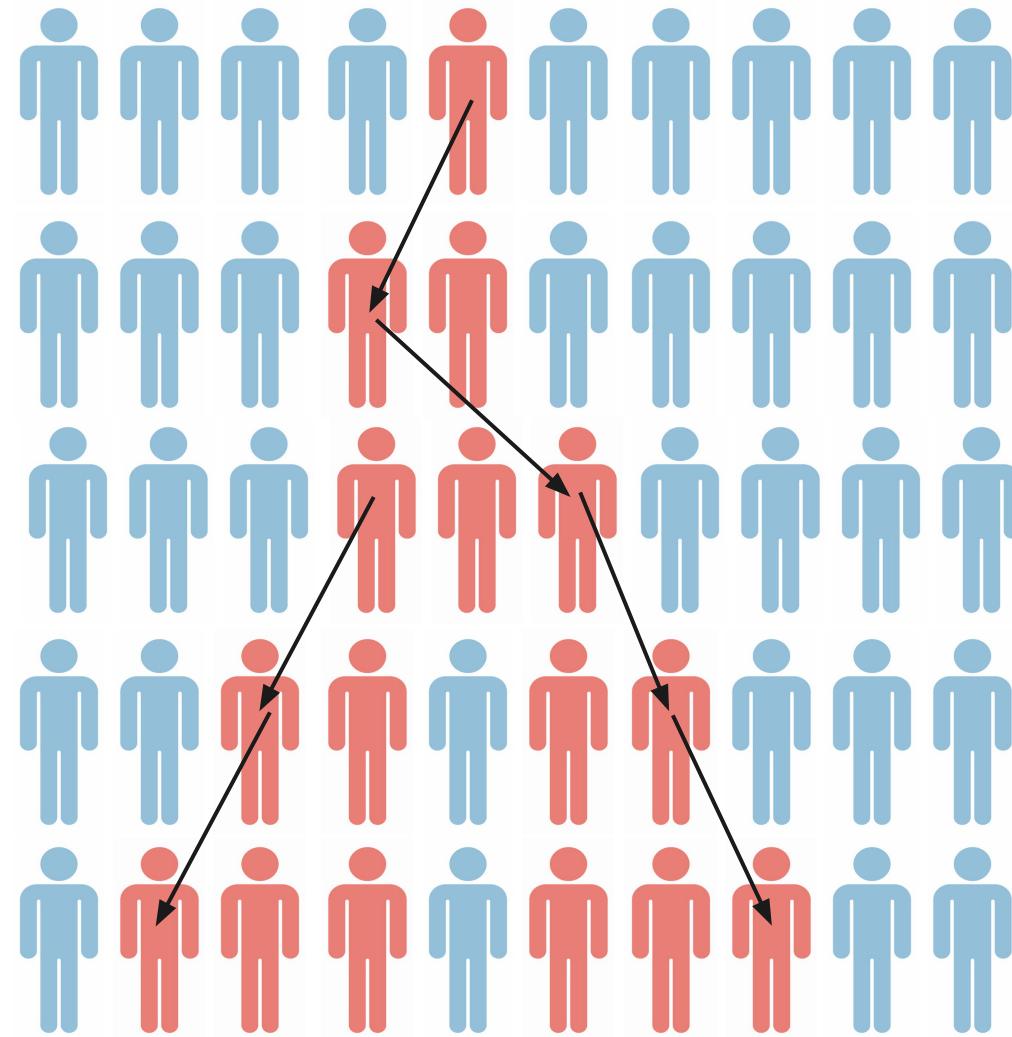
new infection: +1

day: 3

new infection: +2 recovery: -1

day: 4

# Generations



day: 0

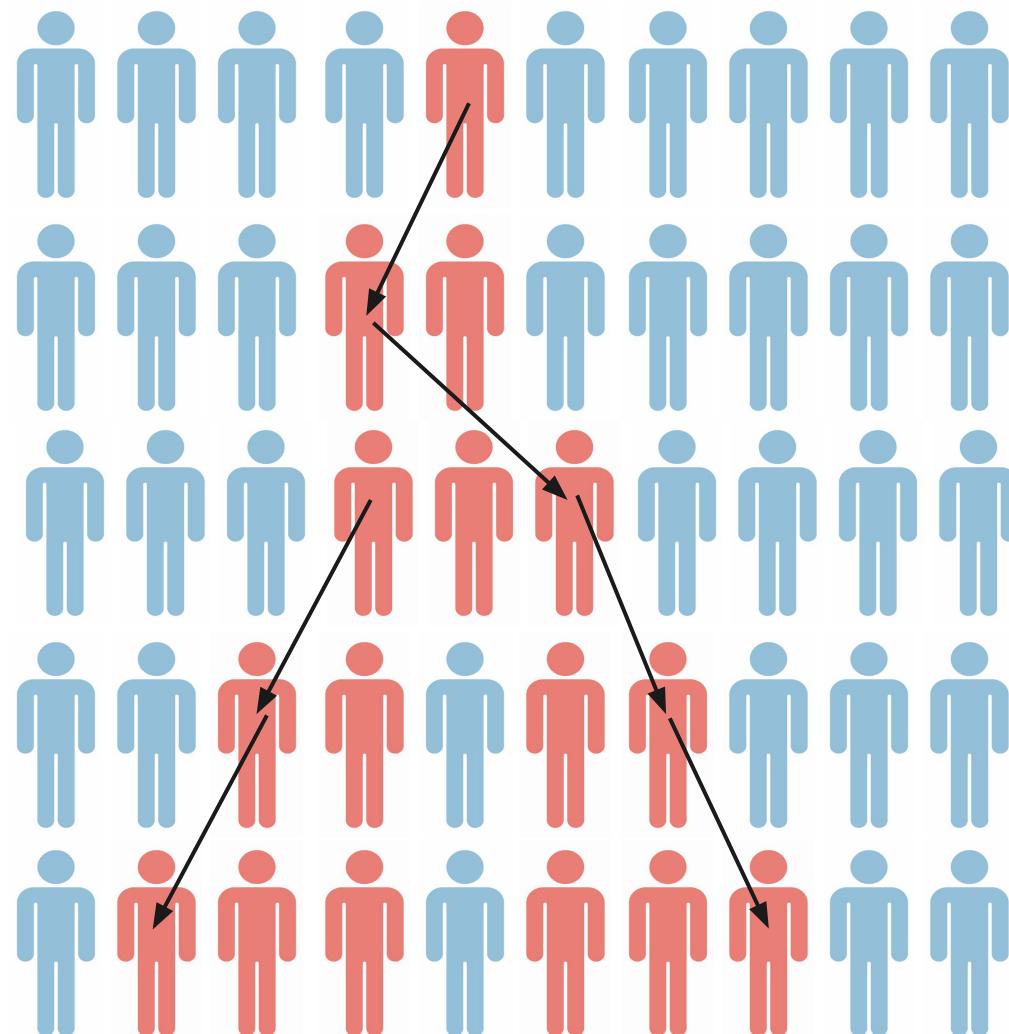
day: 1      new infection: +1

day: 2      new infection: +1

day: 3      new infection: +2      recovery: -1

day: 4      new infection: +2

# Generations



day: 0

day: 1

day: 2

day: 3

day: 4

new infection: +1

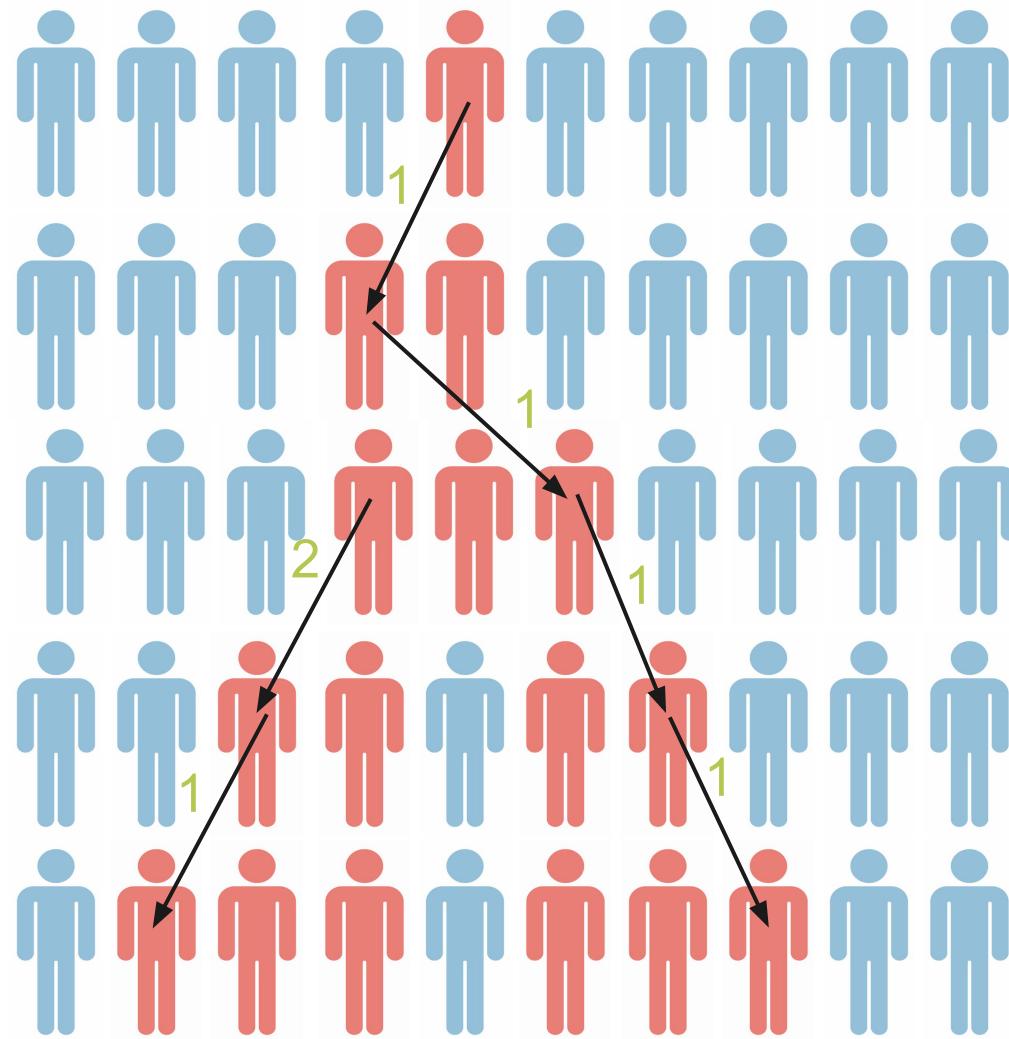
new infection: +1

new infection: +2 recovery: -1

new infection: +2

growth rate  $r$ :  
per capita change in number of  
new cases per unit of time  
 $r = (1+1+2-1+2)/4/10 = 0.125$

# Generations



day: 0

day: 1

day: 2

day: 3

day: 4

new infection: +1

new infection: +1

new infection: +2      recovery: -1

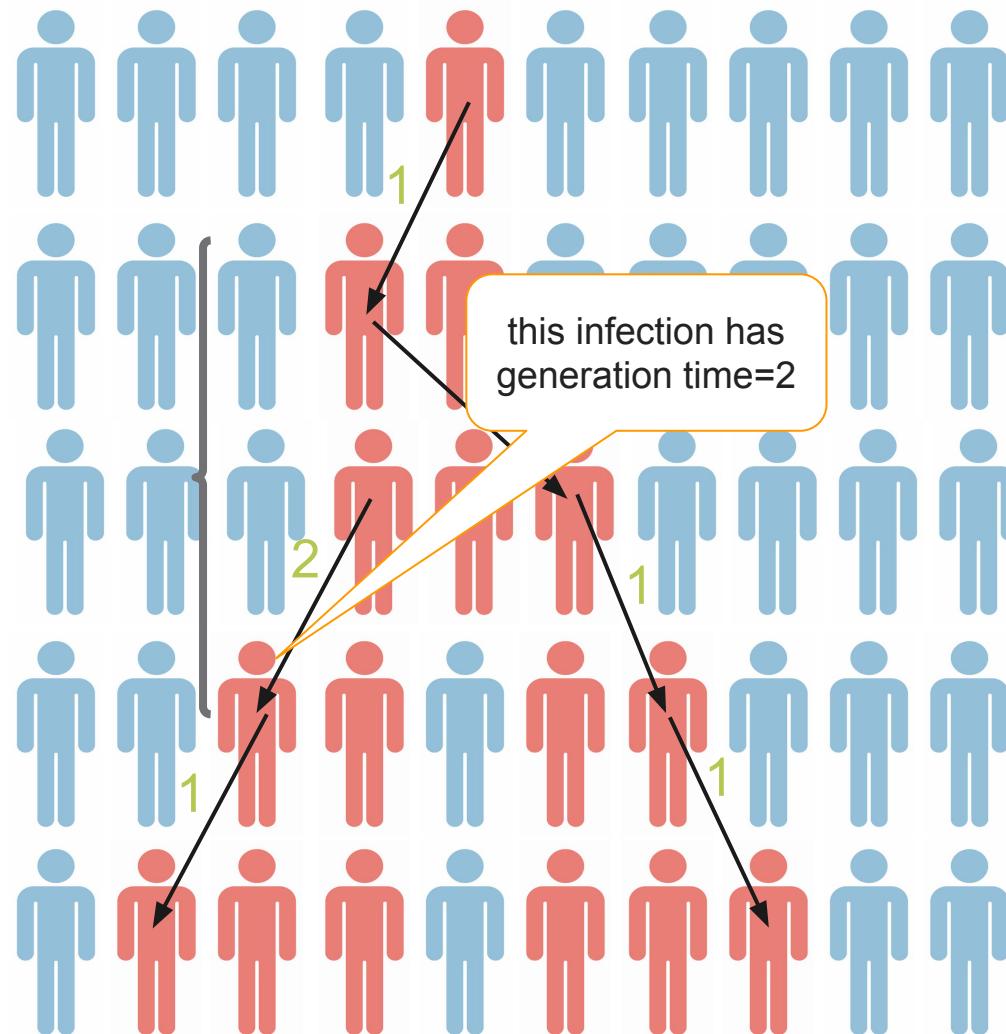
new infection: +2

generation time T:

mean duration between time of infection of a secondary infectee and the time of infection of its primary infector

$$T = (1+1+2+1+1)/6 = 1.16$$

# Generations



day: 0

day: 1

new infection: +1

day: 2

new infection: +1

day: 3

new infection: +2      recovery: -1

day: 4

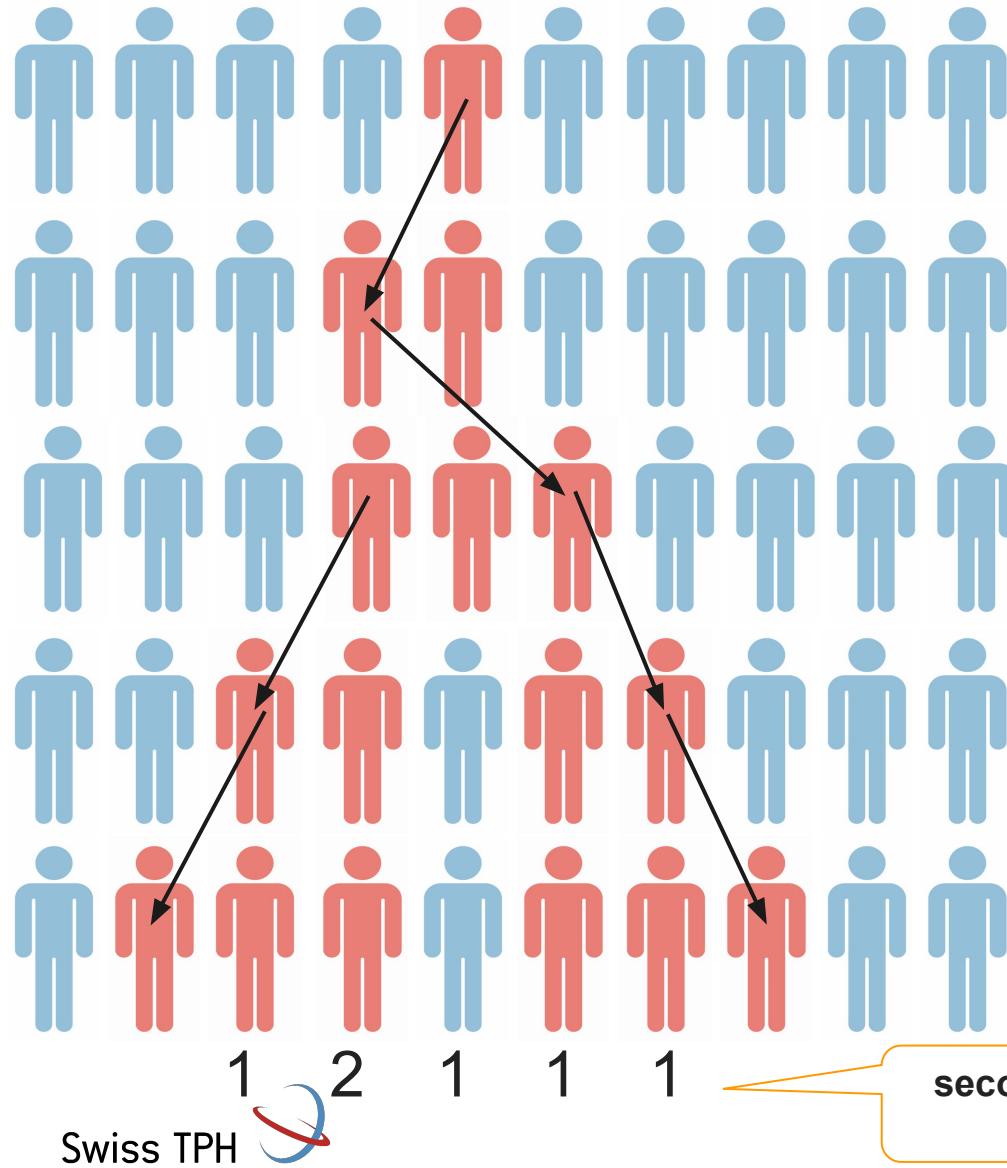
new infection: +2

generation time T:

mean duration between time of infection of a secondary infectee and the time of infection of its primary infector

$$T = (1+1+2+1+1)/6 = 1.16$$

# Basic reproduction number



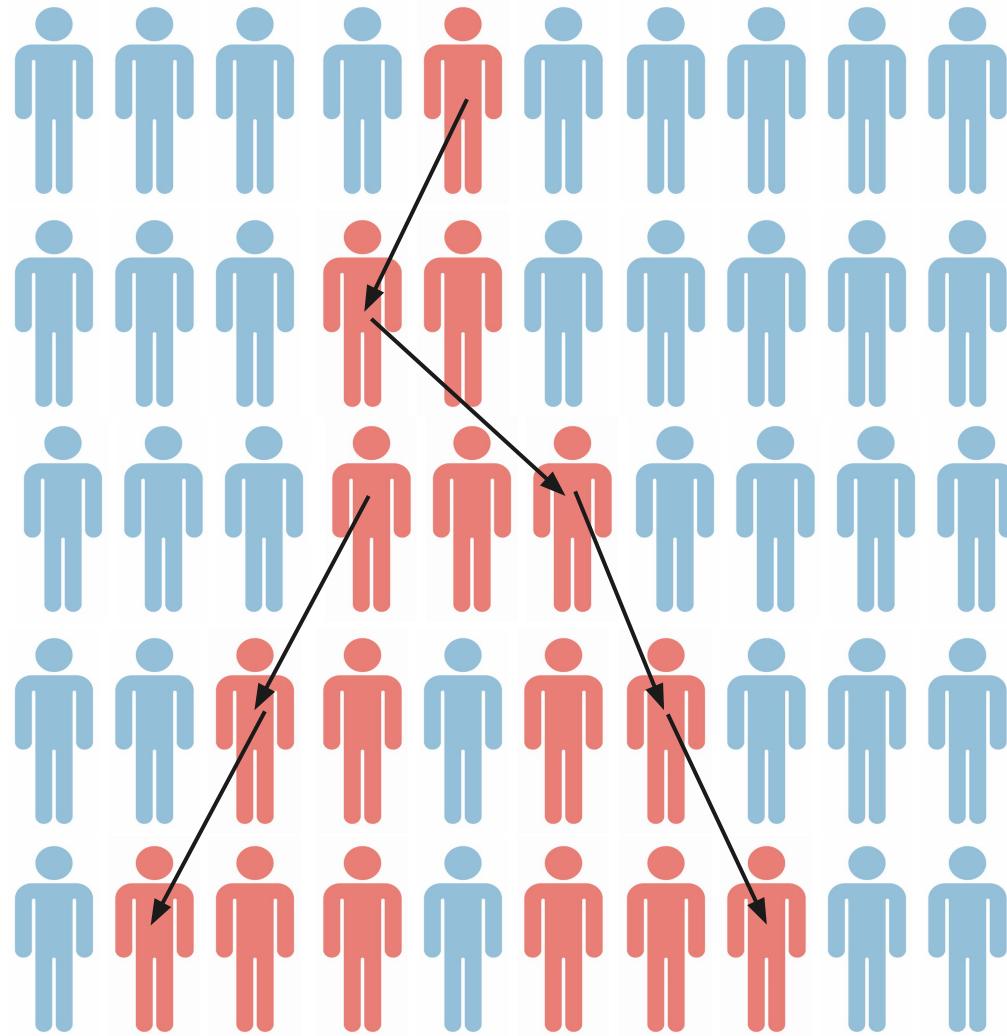
**basic reproduction number  $\mathcal{R}_0$**  :

“expected number of secondary infections that arise from a typical primary case during its entire period of infectiousness in a completely susceptible population”

$$\mathcal{R}_0 > 1$$

initial outbreak will result in full-scale epidemic!

# Basic reproduction number



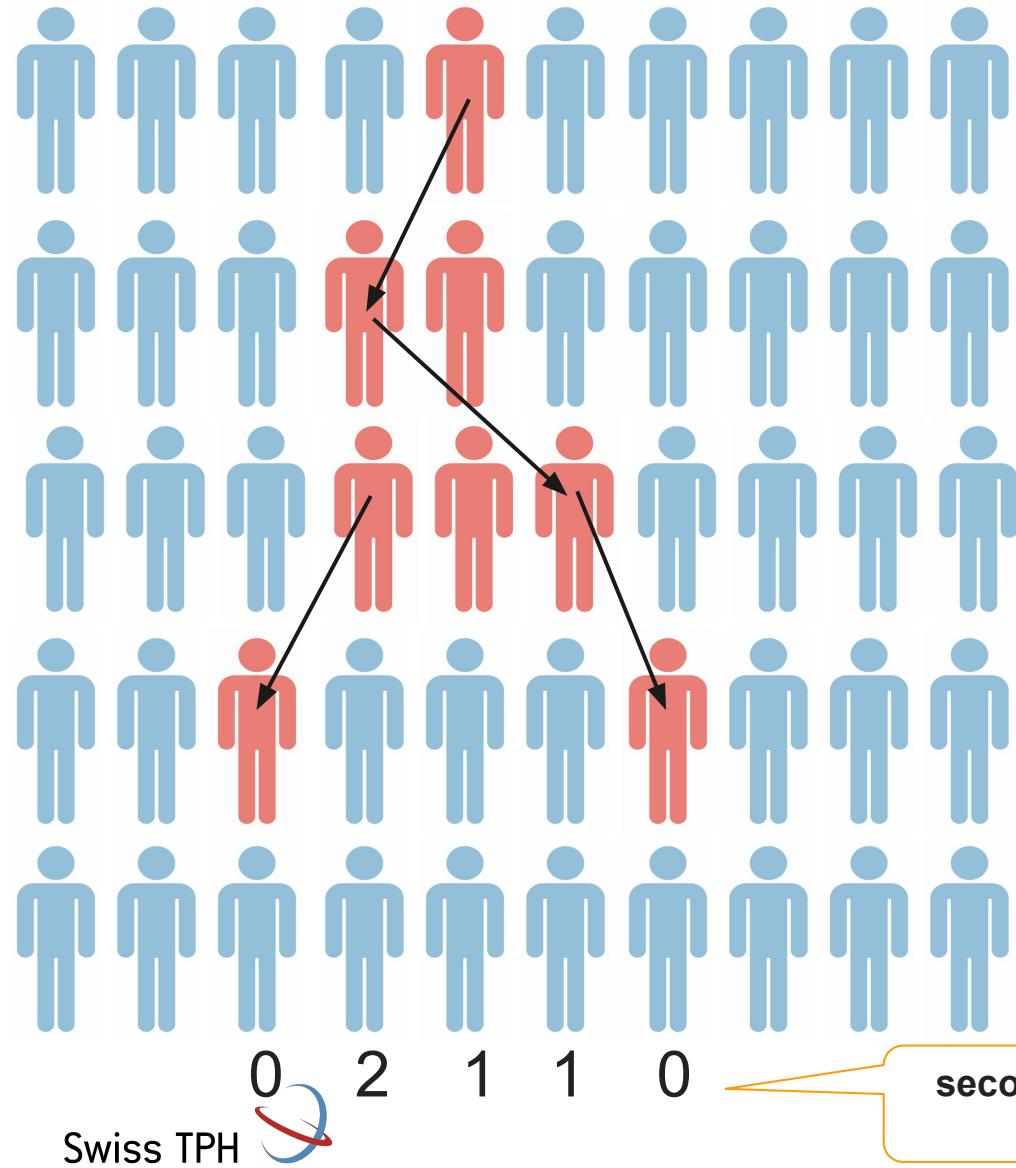
**basic reproduction number  $\mathcal{R}_0$**  :

“expected number of secondary infections that arise from a typical primary case during its entire period of infectiousness in a completely susceptible population”

$$\mathcal{R}_0 = 1 + rT$$

$$\mathcal{R}_0 = 1 + 0.125 \times 1.16 = 1.1458$$

# Basic reproduction number



**basic reproduction number  $\mathcal{R}_0$**  :

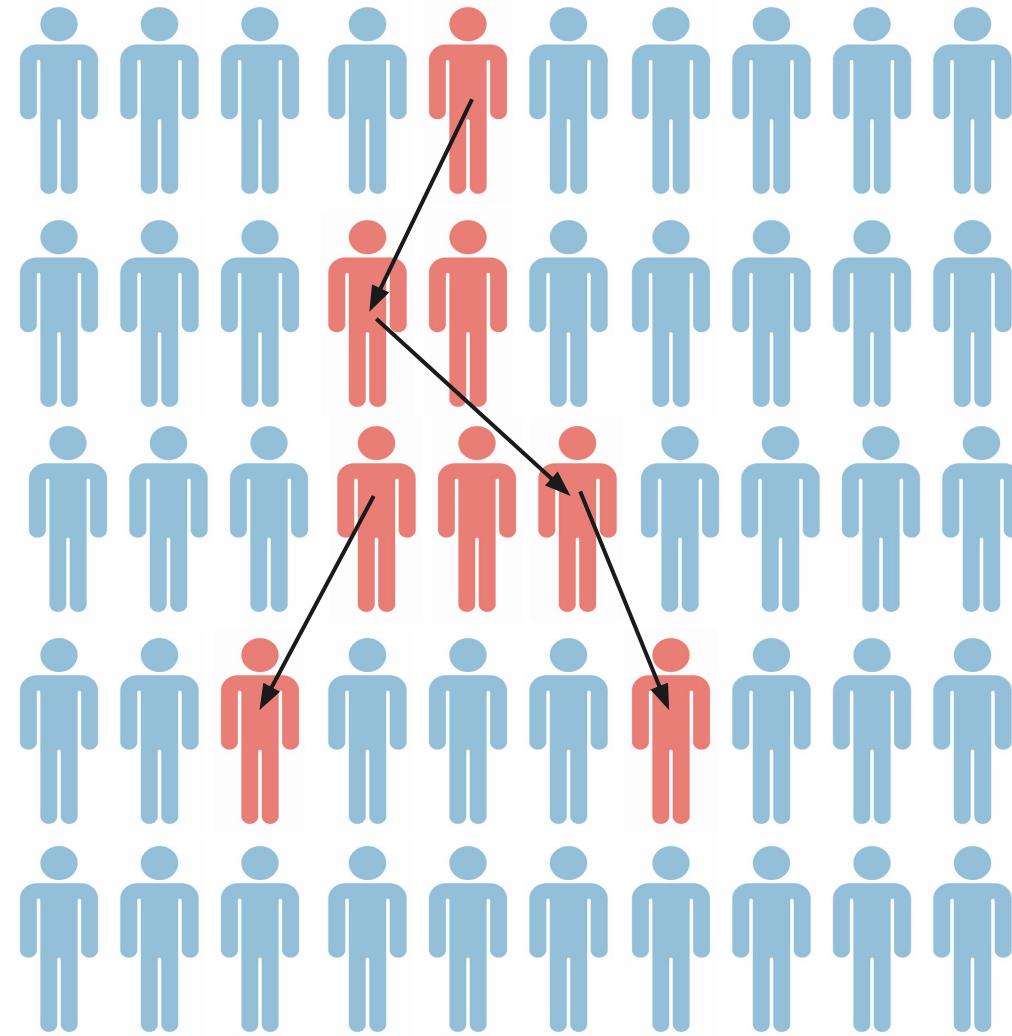
“expected number of secondary infections that arise from a typical primary case during its entire period of infectiousness in a completely susceptible population”

$$\mathcal{R}_0 < 1$$

initial outbreak quickly dies out!



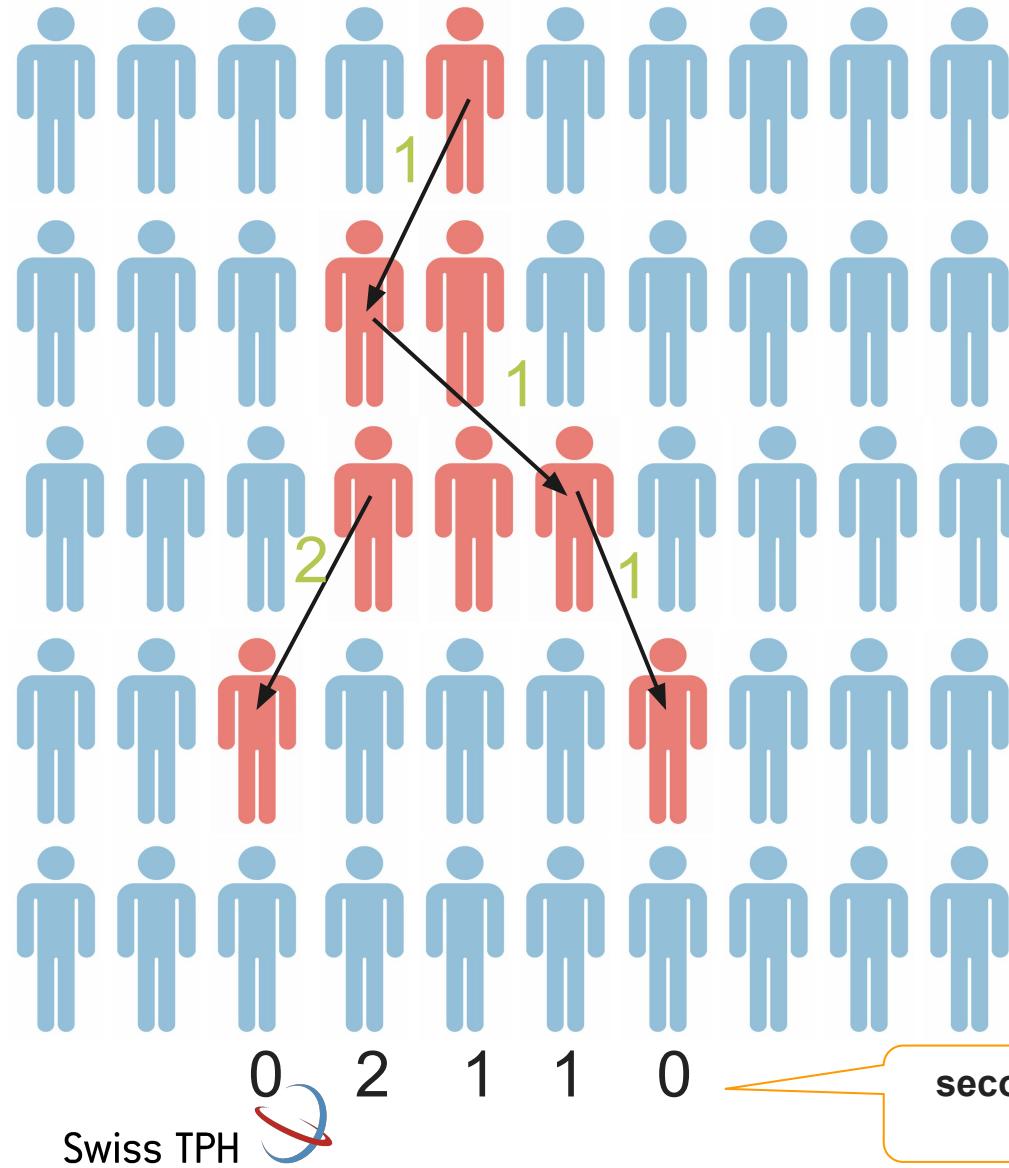
# Basic reproduction number



secondary infections caused by each host,  
average:  $4/5 < 1$



# Basic reproduction number



Calculate growth rate  $r$  and generation time  $T$  based on this observation!

$$r = (1+1+3+2-2)/4/10 = -1/40$$

$$T = (1+1+2+1)/4 = 5/4$$

$$\mathcal{R}_0 = 1 + rT = 0.96875$$

# Lotka-Euler equation and basic reproduction number

$n(a)$  rate of new infection at generation time  $a$

$$b(t) = \int_{a \geq 0} b(t-a)n(a)da \quad \text{"renewal equation for new infections"}$$

$$b(t) = b(t-a)e^{ra} \quad \text{exponential growth with rate } r$$

$$b(t) = \int_{a \geq 0} b(t)e^{-ra}n(a)da$$

$$1 = \int_{a \geq 0} e^{-ra}n(a)da \quad \text{"Lotka-Euler equation"}$$

$$\mathcal{R}_0 = \int_{a \geq 0} n(a)da \quad \text{total number of secondary infections from a host with generation times } a$$

generation time = infection duration probability distribution:  $g(a) = \frac{n(a)}{\mathcal{R}_0}$

from Lotka-Euler:  $\frac{1}{\mathcal{R}_0} = \int_{a \geq 0} e^{-ra}g(a)da$

# Lotka-Euler equation and basic reproduction number

from Lotka-Euler:  $\frac{1}{\mathcal{R}_0} = \int_{a \geq 0} e^{-ra} g(a) da$

$M_f(z) = \int_{a \geq 0} e^{za} f(a) da$  “moment-generating function of probability distribution  $f$ ”

$$\mathcal{R}_0 = \frac{1}{M_g(-r)}$$

Now back to **ODE models**: we have seen that from the **memory-less property of time to recovery** it follows that infection duration must be exponentially distributed with intensity:  $\gamma = 1/T$

$$g(a) = \gamma e^{-\gamma a}$$

$$M_g(z) = \frac{\gamma}{\gamma - z} \quad \mathcal{R}_0 = \frac{\gamma + r}{\gamma} = 1 + \frac{r}{\gamma} = 1 + rT$$

# Lotka-Euler equation and basic reproduction number

from Lotka-Euler:  $\frac{1}{\mathcal{R}_0} = \int_{a \geq 0} e^{-ra} g(a) da$

$M_f(z) = \int_{a \geq 0} e^{za} f(a) da$  “moment-generating function of probability distribution f”

$$\mathcal{R}_0 = \frac{1}{M_g(-r)}$$

Now back to **ODE models**: we have seen that from the **memory-less property of time to recovery** it follows that infection duration must be exponentially distributed with intensity:  $\gamma = 1/T$

$$g(a) = \gamma e^{-\gamma a}$$

$$M_g(z) = \frac{\gamma}{\gamma - z}$$

Poisson counting process defined as counting process with exponentially distributed time to next event (=recovery); Poisson process has mean b  
gamma+r is rate of new infections

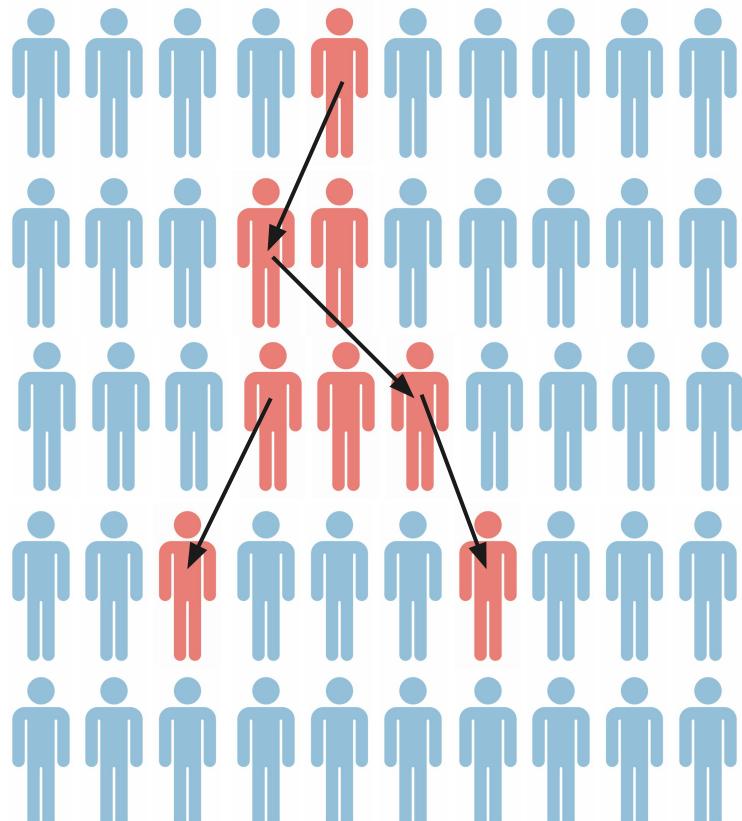
$$\mathcal{R}_0 = \frac{\gamma+r}{\gamma} = \frac{\beta}{\gamma}$$

# Linearization of differential equations

Let's be a bit more precise

**basic reproduction number:**

“expected number of secondary infections that arise from a typical primary case during its entire period of infectiousness in a completely susceptible population”



This population is not completely susceptible any more!  
Depletion of susceptible!

$$\left. \begin{array}{l} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \end{array} \right\}$$

**linearize at the disease-free equilibrium!**

# Linearization of differential equations

$$\left. \begin{array}{l} \frac{dS}{dt} = -\beta S \frac{I}{N} \\ \frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I \end{array} \right\} \text{linearize at the disease free-equilibrium!}$$

$$X = \begin{bmatrix} S \\ I \end{bmatrix} \quad f(X) = \begin{bmatrix} -\beta X_1 \frac{X_2}{N} \\ \beta X_1 \frac{X_2}{N} - \gamma X_2 \end{bmatrix} \quad \left. \frac{dX}{dt} = f(X) \right\} \text{non-linear differential equation}$$

$$(Df)(X) = (\partial_{X_j} f(X)^i)_{i,j=1,2} \quad \text{Jacobian of vector field } f$$

$$\left. \frac{dX}{dt} = (Df)(X) \right\} \text{linear differential equation}$$



# Linearization of differential equations

Find the disease-free equilibrium (DFE) of the SIR system!

Linearize the SIR system at the DFE!

$$X = \begin{bmatrix} S \\ I \end{bmatrix} \quad f(X) = \begin{bmatrix} -\beta X_1 \frac{X_2}{N} \\ \beta X_1 \frac{X_2}{N} - \gamma X_2 \end{bmatrix} \quad (Df)(X) = (\partial_{X_j} f(X)^i)_{i,j=1,2}$$



# Linearization of differential equations

Find the disease-free equilibrium (DFE) of the SIR system!

Linearize the SIR system at the DFE!

$$X = \begin{bmatrix} S \\ I \end{bmatrix} \quad f(X) = \begin{bmatrix} -\beta X_1 \frac{X_2}{N} \\ \beta X_1 \frac{X_2}{N} - \gamma X_2 \end{bmatrix} \quad (Df)(X) = (\partial_{X_j} f(X)^i)_{i,j=1,2}$$

$$\frac{dX}{dt} = f(X) = 0 \Rightarrow X^* = \begin{bmatrix} S^* \\ 0 \end{bmatrix} \quad S^* = N \text{ w.l.o.g.}$$

$$(Df)(X^*) = \begin{bmatrix} -\beta \times 0 & -\beta \frac{S^*}{N} \\ \beta \times 0 - 0 & \beta \frac{S^*}{N} - \gamma \end{bmatrix}$$



# Linearization of differential equations

Find the disease-free equilibrium (DFE) of the SIR system!

Linearize the SIR system at the DFE!

$$X = \begin{bmatrix} S \\ I \end{bmatrix} \quad f(X) = \begin{bmatrix} -\beta X_1 \frac{X_2}{N} \\ \beta X_1 \frac{X_2}{N} - \gamma X_2 \end{bmatrix}$$

$$\frac{dX}{dt} = f(X) = 0 \Rightarrow X^* = \begin{bmatrix} S^* \\ 0 \end{bmatrix} \quad S^* = N \text{ w.l.o.g.}$$

$$(Df)(X^*) = \begin{bmatrix} -\beta \times 0 & -\beta \frac{S^*}{N} \\ \beta \times 0 - 0 & \boxed{\beta \frac{S^*}{N} - \gamma} \end{bmatrix}$$

This is the only part that contributes to changes in infections!  
If it is positive, then epidemic increases!!

$$\beta > \gamma \Leftrightarrow \frac{\beta}{\gamma} > 1$$

# Next generation matrix theory

- **linearized system** (depletion of susceptibles are not taken into consideration at the beginning of an epidemic)
- **disease-free equilibrium** (we are only interested whether a single infected host is able to invade the population of susceptibles)
- **only infectious** compartments are considered for **epidemic growth** condition
- **decomposition** into inflow to/ outflow from infectious compartments

# Next generation matrix theory

0. Calculate disease-free equilibrium  $\frac{dX}{dt} = f(X) = 0$
1. Determine infectious compartments
2. Decompose inflow and outflow for infectious compartments  $f = \mathbb{F} - \mathbb{V}$
3. Linearize inflow and outflow at disease-free equilibrium
$$(Df)(X^*) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix}$$
$$F = \partial_j \mathbb{F}^i(X^*) \quad V = \partial_j \mathbb{V}^i(X^*)$$
4. Next-generation matrix  $FV^{-1}$
5. Spectral radius  $\rho(A) = \max\{|\lambda_i| : \lambda \in \text{eig}(A)\}$ 

$F$  non-negative  
 $F(i,j)$  rate at which infected individuals in compartment  $j$  produce new infections in compartment  $i$
6.  $\mathcal{R}_0 = \rho(FV^{-1})$

# Next generation matrix theory

0. Calculate disease-free equilibrium  $\frac{dX}{dt} = f(X) = 0$
1. Determine infectious compartments
2. Decompose inflow and outflow for infectious compartments  $f = \mathbb{F} - \mathbb{V}$
3. Linearize inflow and outflow at disease-free equilibrium
$$(Df)(X^*) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix} \quad F = \partial_j \mathbb{F}^i(X^*) \quad V = \partial_j \mathbb{V}^i(X^*)$$
4. Next-generation matrix  $FV^{-1}$
5. Spectral radius  $\rho(A) = \max\{|\lambda_i|; A\lambda_i = \lambda_i a_i\}$ 

V non-singular  
 $V^{-1}(i,j)$  average time  
individual spends in  
compartment j during its  
lifetime
6.  $\mathcal{R}_0 = \rho(FV^{-1})$

# Next generation matrix theory

0. Calculate disease-free equilibrium  $\frac{dX}{dt} = f(X) = 0$
1. Determine infectious compartments
2. Decompose inflow and outflow for infectious compartments  $f = \mathbb{F} - \mathbb{V}$
3. Linearize inflow and outflow at disease-free equilibrium
$$(Df)(X^*) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix} \quad F = \partial_i \mathbb{F}^i(X^*) \quad V = \partial_j \mathbb{V}^i(X^*)$$

FV<sup>-1</sup> FV<sup>-1</sup> FV<sup>-1</sup>(i,j) expected number of new infections in compartment i produced by the infected individual originally introduced into compartment j
4. Next-generation matrix
5. Spectral radius  $\rho(A) = \max\{|\lambda_i|; A\lambda_i = \lambda_i a_i\}$
6.  $\mathcal{R}_0 = \rho(FV^{-1})$

# Next generation matrix theory

0. Calculate disease-free equilibrium  $\frac{dX}{dt} = f(X) = 0$
1. Determine infectious compartments
2. Decompose inflow and outflow for infectious compartments  $f = \mathbb{F} - \mathbb{V}$
3. Linearize inflow and outflow at disease-free equilibrium
$$(Df)(X^*) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix} \quad F = \partial_j \mathbb{F}^i(X^*) \quad V = \partial_j \mathbb{V}^i(X^*)$$
4. Next-generation matrix  $FV^{-1}$
5. Spectral radius  $\rho(A) = \max\{|\lambda_i|; A\lambda_i = \lambda_i a_i\}$
6.  $\mathcal{R}_0 = \rho(FV^{-1})$

# Next generation matrix theory

Where does the spectral radius come from?

$M := FV^{-1}$  next generation matrix

$M(X^*) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  bounded operator with  $\|M\| = \max_{i,j} M_{ij}$

After  $k$  generations, infected population in compartment  $i$  is:  $\sum_j M_{ij}^k$

Per-generation growth factor of infections is:  $\|M^k\|^{\frac{1}{k}}$

Long-term growth factor of infections:  $\lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}} = \rho(M)$

$\rho(M)$  is calculated by maximum of the modulus of eigenvalues of  $M$

$\rho(M) < \|M\|$  spectral radius is **lower bound** for any matrix norm

# Next generation matrix theory

Where does the spectral radius come from?

$M := FV^{-1}$  next generation matrix

$M(X^*) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  bounded operator with  $\|M\| = \max_{i,j} M_{ij}$

After  $k$  generations, infected population in compartment  $i$  is:  $\sum_j M_{ij}^k$

Per-generation growth factor of infections is:  $\|M^k\|^{\frac{1}{k}}$

think of binary tree size:  $2^k$   
growth factor: 2

Long-term growth factor of infections:  $\lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}} = \rho(M)$

$\rho(M)$  is calculated by maximum of the modulus of eigenvalues of  $M$

$\rho(M) < \|M\|$  spectral radius is **lower bound** for any matrix norm

# Next generation matrix theory

Where does the spectral radius come from?

$M := FV^{-1}$  next generation matrix

$M(X^*) : \mathbb{R}^m \rightarrow \mathbb{R}^m$  bounded operator with  $\|M\| = \max_{i,j} M_{ij}$

After  $k$  generations, infected population in compartment  $i$  is:  $\sum_j M_{ij}^k$

Per-generation growth factor of infections is:  $\|M^k\|^{\frac{1}{k}}$

Long-term growth factor of infections:  $\lim_{k \rightarrow \infty} \|M^k\|^{\frac{1}{k}} = \boxed{\rho(M)}$

$\rho(M)$  is calculated by maximum of the modulus of eigenvalues of  $M$

$\rho(M) < \|M\|$  spectral radius is **lower bound** for any matrix norm

spectral radius

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{dS}{dt}(t) = -\alpha \frac{V(t)}{H} S(t) + \gamma I(t)$$

$$\frac{dI}{dt}(t) = \alpha \frac{V(t)}{H} S(t) - \gamma I(t)$$

$$\frac{dU}{dt}(t) = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t)$$

$$\frac{dV}{dt}(t) = \beta \frac{I(t)}{H} U(t) - \delta V(t)$$

1. Define infectious compartments:  $X = \begin{bmatrix} I \\ V \end{bmatrix}$
2. Decompose inflow and outflow for infectious compartments:

$$\mathbb{F} = \begin{pmatrix} \alpha S \frac{V}{H} \\ \beta U \frac{I}{H} \end{pmatrix} \quad \mathbb{V} = \begin{pmatrix} \gamma I \\ \delta V \end{pmatrix}$$

# R<sub>0</sub> for vector-host model with next-generation matrix

1. Define infectious compartments:  $X = \begin{bmatrix} I \\ V \end{bmatrix}$

2. Decompose inflow and outflow for infectious compartments:

$$\mathbb{F} = \begin{pmatrix} \alpha S \frac{V}{H} \\ \beta U \frac{I}{H} \end{pmatrix} \quad \mathbb{V} = \begin{pmatrix} \gamma I \\ \delta V \end{pmatrix}$$

3. Linearize inflow and outflow at disease-free equilibrium, calculate Jacobian matrix:

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$\left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\mathcal{R}_0 = \rho \left( \boxed{\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1}} \right)$$

Diekmann-Heesterbeek-Metz 1990:  
Next Generation Matrix

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$\left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\mathcal{R}_0 = \rho \left( \boxed{\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1}} \right)$$

Diekmann-Heesterbeek-Metz 1990:  
**Next Generation Matrix**

Let's do the  $\mathcal{R}_0$  maths!

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$\left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix} = \boxed{\begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta M}{\gamma H} & 0 \end{bmatrix}}$$

Next Generation Matrix

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$\left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix} = \boxed{\begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix}}$$

Next Generation Matrix

Characteristic polynomial for eigenvalues

$$\det \left( \begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix} - \lambda \mathbb{I} \right) = 0$$

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$(\frac{\partial \mathbb{V}}{\partial X})^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix} = \boxed{\begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix}}$$

Next Generation Matrix

Eigenvalues

Characteristic polynomial for eigenvalues

$$\lambda_{1/2} = \pm \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

$$\det \left( \begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix} - \lambda \mathbb{I} \right) = 0$$

# R<sub>0</sub> for vector-host model with next-generation matrix

$$\frac{\partial \mathbb{F}}{\partial X} = \begin{bmatrix} 0 & \alpha S^* \frac{1}{H} \\ \beta U^* \frac{1}{H} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \quad \frac{\partial \mathbb{V}}{\partial X} = \begin{bmatrix} \gamma & 0 \\ 0 & \delta \end{bmatrix}$$

4. Calculate next-generation matrix, spectral radius of NGM:

$$(\frac{\partial \mathbb{V}}{\partial X})^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix}$$

$$\frac{\partial \mathbb{F}}{\partial X} \left( \frac{\partial \mathbb{V}}{\partial X} \right)^{-1} = \begin{bmatrix} 0 & \alpha \\ \beta \frac{M}{H} & 0 \end{bmatrix} \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \delta^{-1} \end{bmatrix} = \boxed{\begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix}}$$

Next Generation Matrix

Eigenvalues

Characteristic polynomial for eigenvalues

$$\det \left( \begin{bmatrix} 0 & \frac{\alpha}{\delta} \\ \frac{\beta}{\gamma} \frac{M}{H} & 0 \end{bmatrix} - \lambda \mathbb{I} \right) = 0$$

$$\lambda_{1/2} = \pm \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

Spectral radius

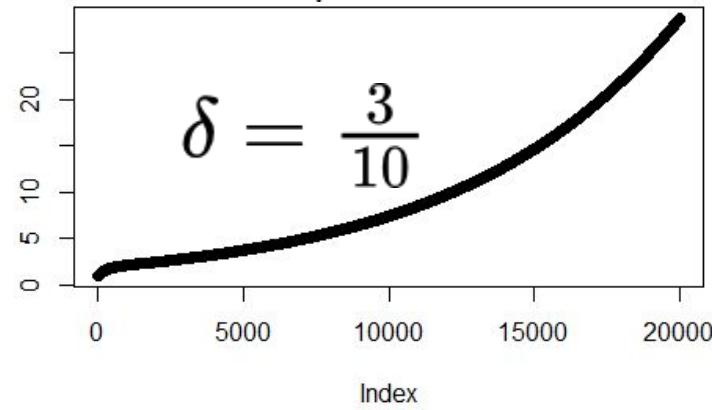
$$\mathcal{R}_0 = \max |\lambda_i| = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

# $R_0$ for vector-host model with next-generation matrix

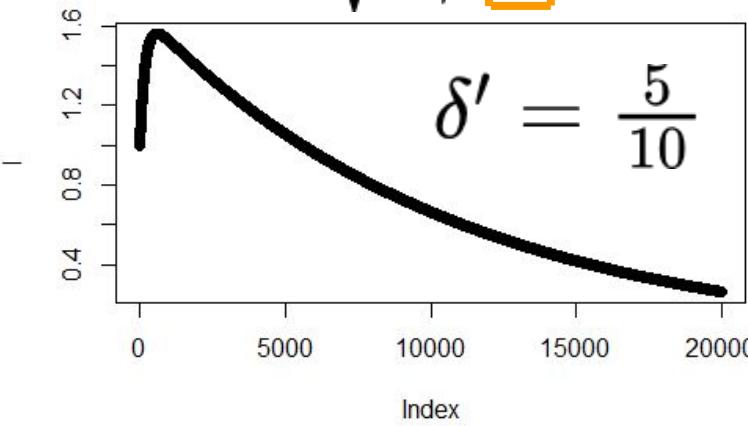
## Mosquito Theorem:

vector control is a sufficient condition for malaria elimination in humans

$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} > 1$$



$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta'} K} < 1$$



$$\frac{K}{\delta'} = \frac{K}{\delta} \frac{3}{5} < \frac{K}{\delta}$$

ratio of vectors to hosts

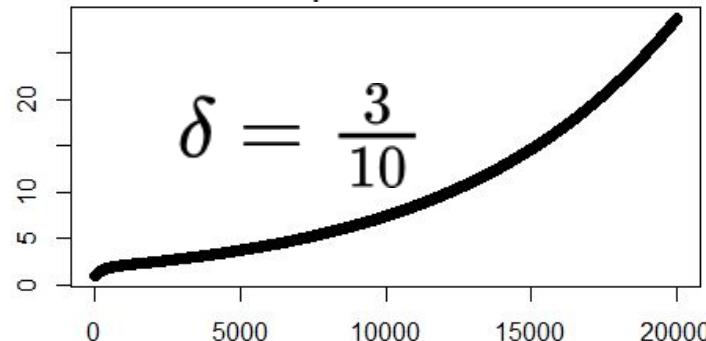
$$R_0 = \max |\lambda_i| = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

# $R_0$ for vector-host model with next-generation matrix

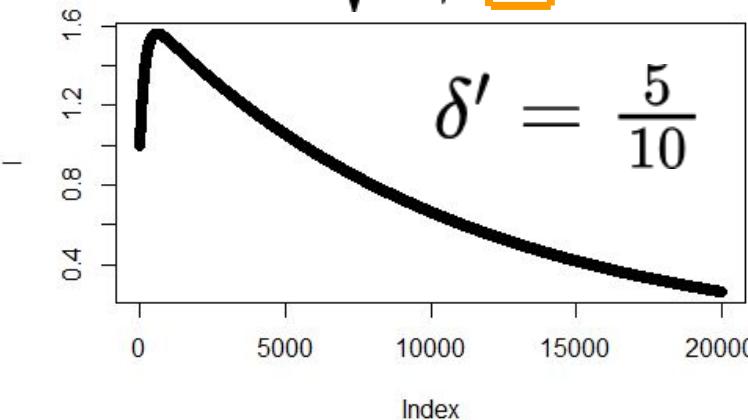
**Mosquito Theorem:**

vector control is a sufficient condition for malaria elimination in humans

$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} > 1$$



$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta'} K} < 1$$



$$\frac{K}{\delta'} = \frac{K}{\delta} \frac{3}{5} < \frac{K}{\delta}$$

Clara Champagne's course:

$$R_0 = \frac{m a^2 b c e^{-gv}}{rg}$$

$$R_0 = \max |\lambda_i| = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

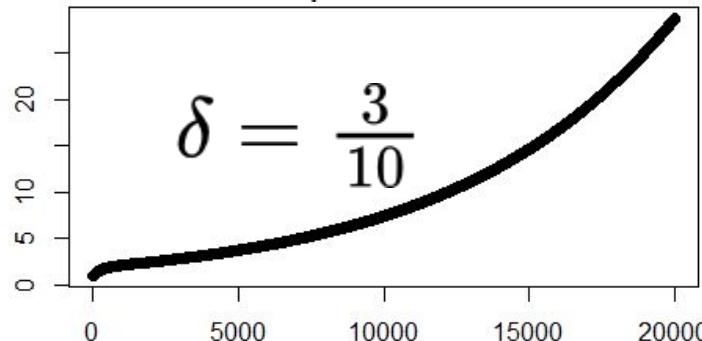
ratio of vectors to hosts

# $R_0$ for vector-host model with next-generation matrix

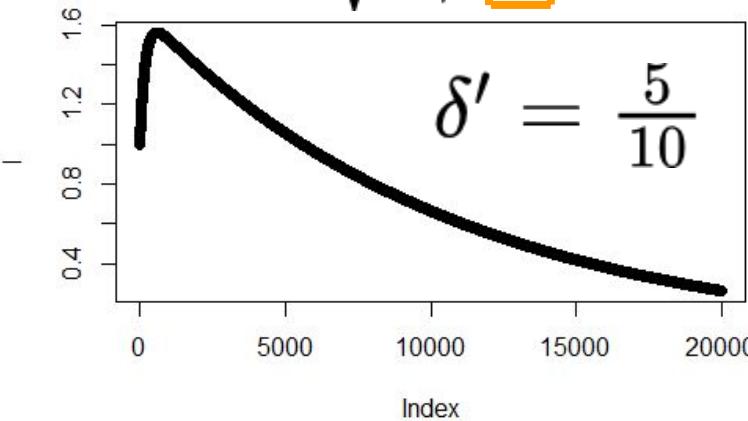
**Mosquito Theorem:**

vector control is a sufficient condition for malaria elimination in humans

$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K} > 1$$



$$R_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta'} K} < 1$$



$$\frac{K}{\delta'} = \frac{K}{\delta} \frac{3}{5} < \frac{K}{\delta}$$

Clara Champagne's course:

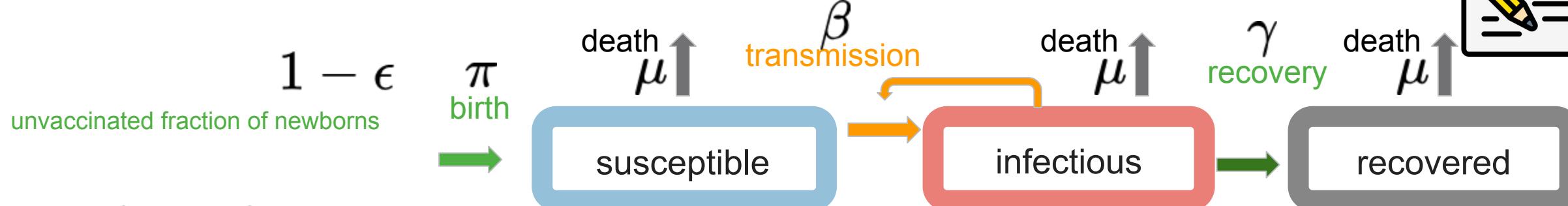
$$R_0 = \frac{m a^2 b c e^{-gv}}{r g}$$

recovery/death

transmission

$$R_0 = \max |\lambda_i| = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} \frac{M}{H}}$$

# R<sub>0</sub> for SIR model with newborn vaccination

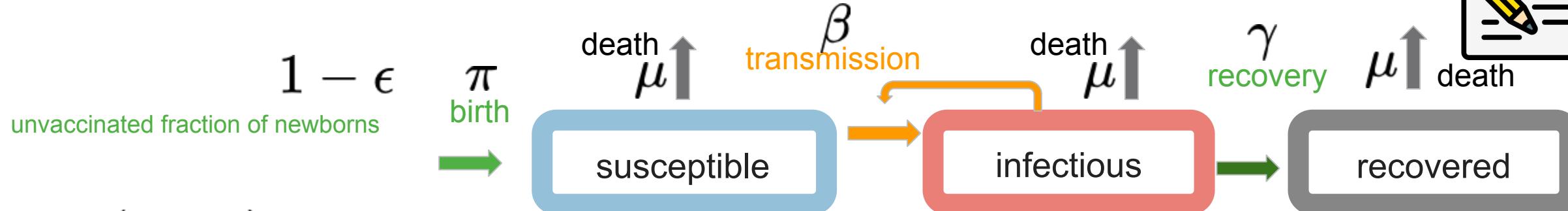


$$\frac{dS}{dt} = \pi(1 - \epsilon) - \beta SI - \mu S$$

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

$$\frac{dR}{dt} = \gamma I - \mu R$$

# R<sub>0</sub> for SIR model with newborn vaccination



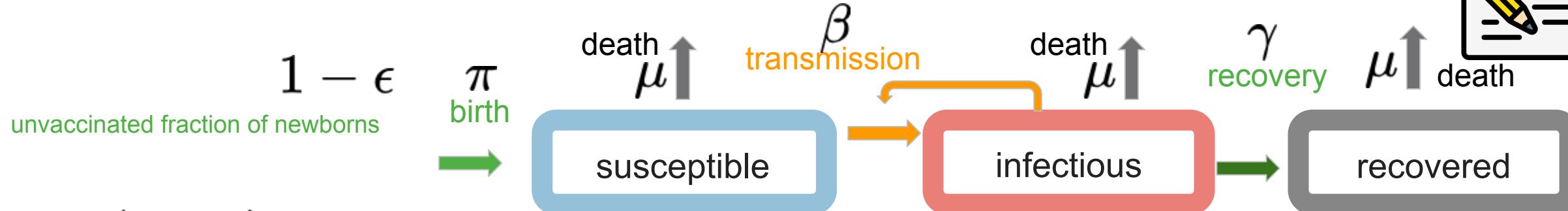
$$\frac{dS}{dt} = \pi(1 - \epsilon) - \beta SI - \mu S$$

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

$$\frac{dR}{dt} = \gamma I - \mu R$$

1. DFE:  $S^* = (1 - \epsilon) \frac{\pi}{\mu} S_0, I = R = 0$
2. Infectious compartment & decomposition:  $\mathbb{F} = \beta SI$   
 $\mathbb{V} = (\gamma + \mu)I$
3. Linearization at DFE  $F = (1 - \epsilon) \frac{\pi}{\mu} \beta$   $V = \gamma + \mu$
4. Next-generation matrix  $FV^{-1} = (1 - \epsilon) \frac{\pi}{\mu} \frac{\beta}{\gamma + \mu}$
5. Spectral radius  $\mathcal{R}_0^\epsilon = (1 - \epsilon) \frac{\pi}{\mu} \frac{\beta}{\gamma + \mu} = (1 - \epsilon) \mathcal{R}_0$

# R<sub>0</sub> for SIR model with newborn vaccination



$$\frac{dS}{dt} = \pi(1 - \epsilon) - \beta SI - \mu S$$

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

$$\frac{dR}{dt} = \gamma I - \mu R$$

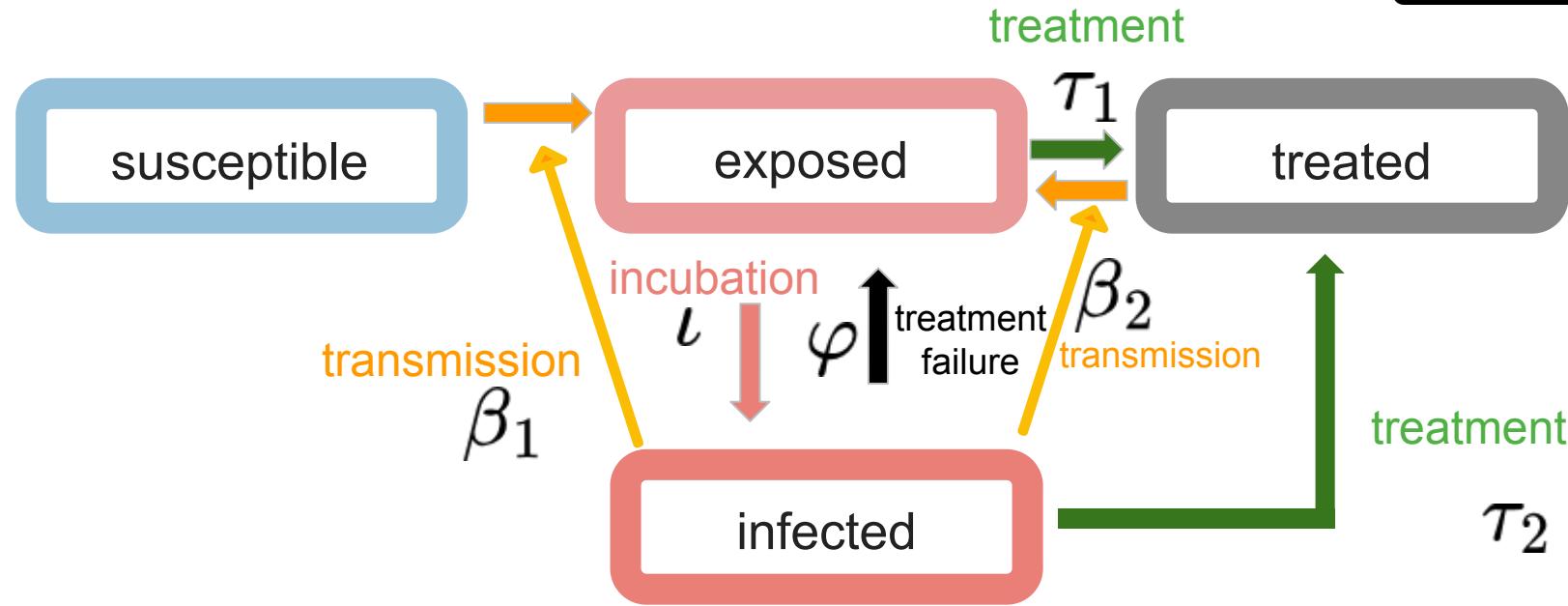
Increasing the vaccination coverage of newborns will control disease spread

$$\mathcal{R}_0^\epsilon = (1 - \epsilon) \frac{\pi}{\mu} \frac{\beta}{\gamma + \mu} = (1 - \epsilon) \mathcal{R}_0$$

Required vaccination coverage to control disease spread

$$\mathcal{R}_0^\epsilon < 1 \Leftrightarrow \epsilon > 1 - \frac{1}{\mathcal{R}_0}$$

# R<sub>0</sub> for SEI with treatment failure





# R<sub>0</sub> for SEI with treatment failure

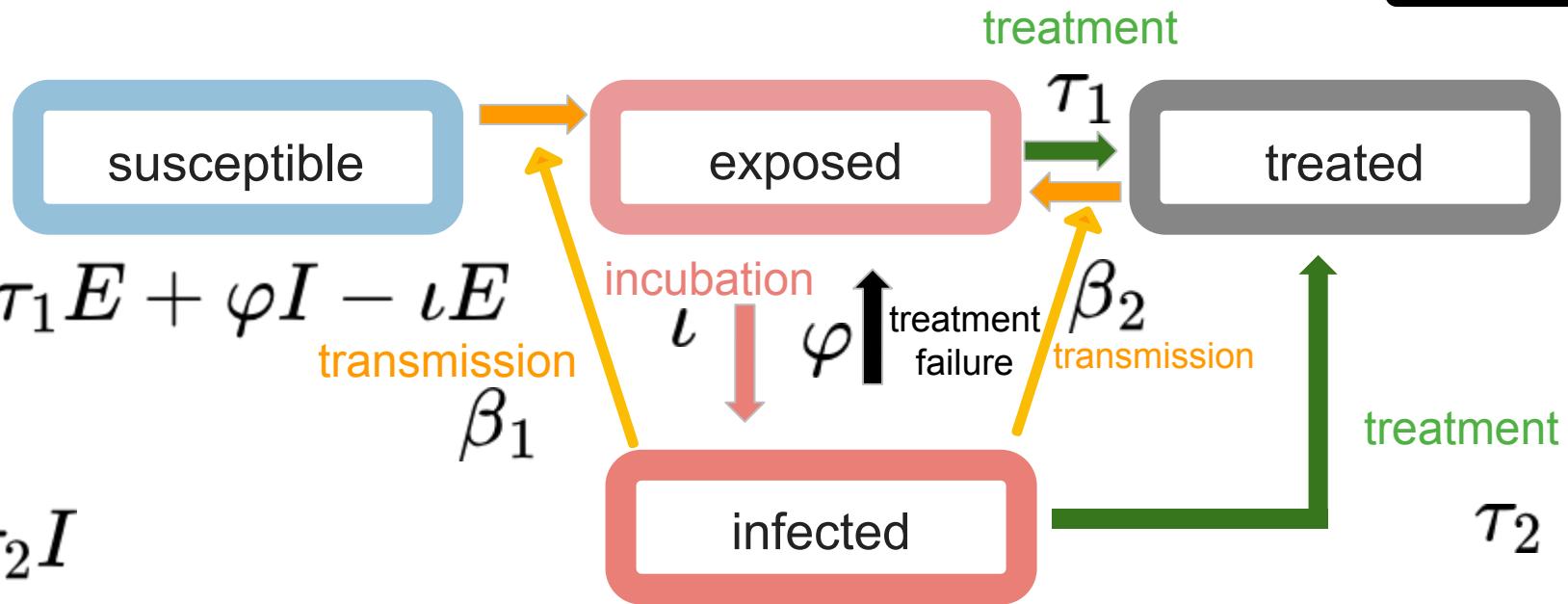
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$





# R<sub>0</sub> for SEI with treatment failure

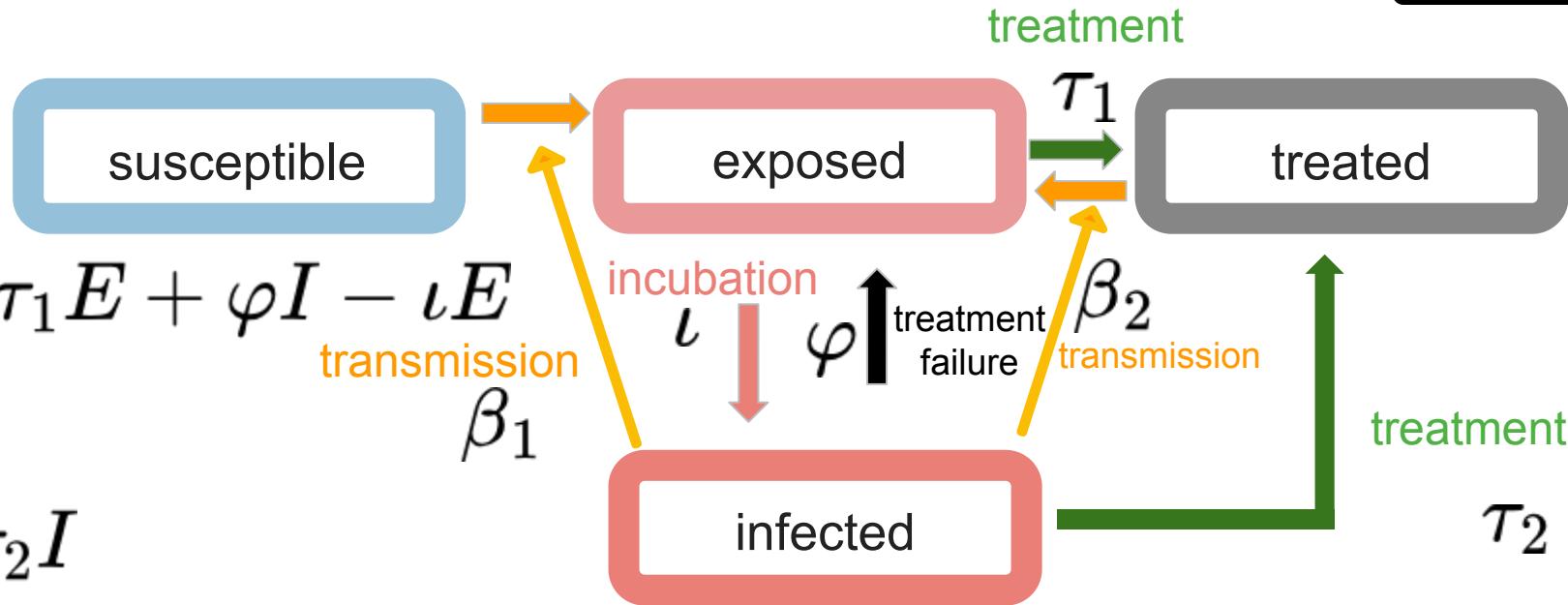
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



Calculate **basic reproduction number** following the steps:

1. Disease-free equilibrium
2. Infectious compartments & decomposition
3. Linearization
4. Next-generation matrix
5. Spectral radius



# R<sub>0</sub> for SEI with treatment failure

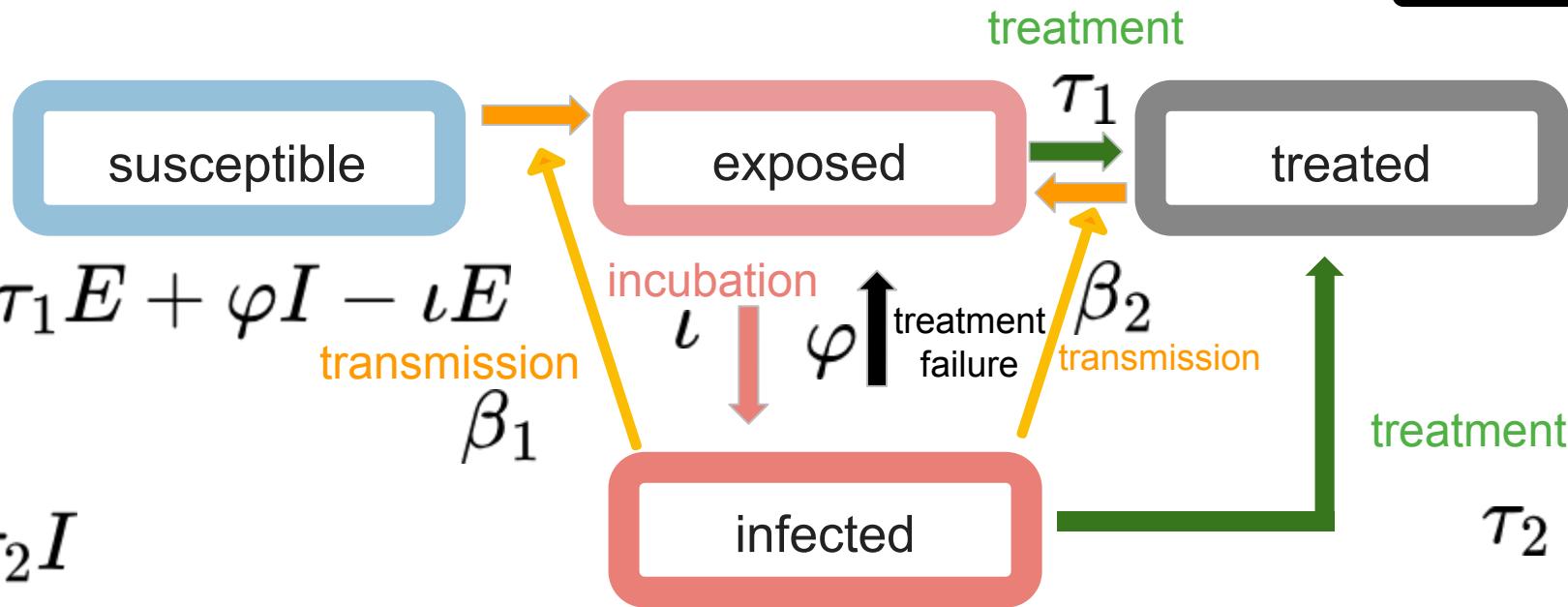
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



## 1. Disease-free equilibrium

$$S^* = S(0), E = I = T = 0$$

# R<sub>0</sub> for SEI with treatment failure



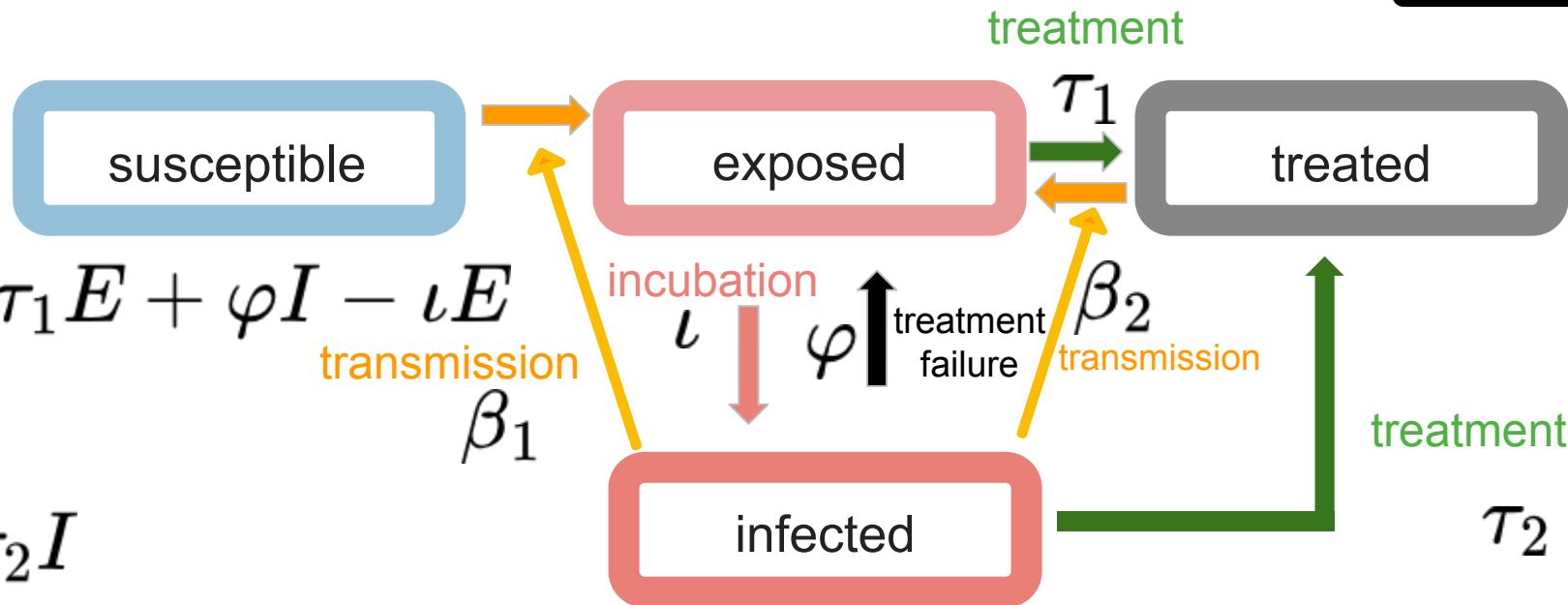
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



## 2. Infectious compartments & decomposition

E, I are infectious compartments  
incubation and treatment failure **are not** considered to be new infections

$$\mathbb{F} = \begin{bmatrix} \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} \\ 0 \end{bmatrix} \quad \mathbb{V} = \begin{bmatrix} \tau_1 E - \varphi I + \iota E \\ -\iota E + \varphi I + \tau_2 I \end{bmatrix}$$

$$\frac{dX}{dt} = f(X) = \mathbb{F}(X) - \mathbb{V}(X)$$



# R<sub>0</sub> for SEI with treatment failure

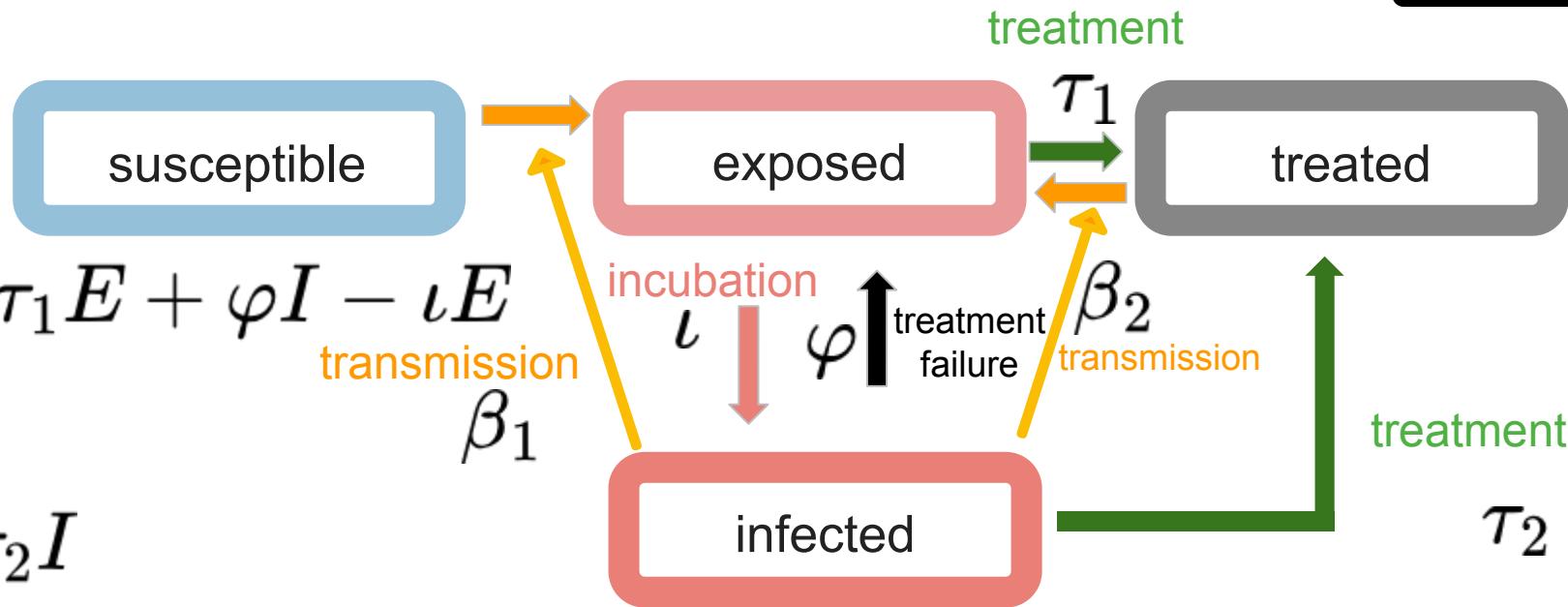
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



### 3. Linearization at DFE

$$F = \begin{bmatrix} 0 & \beta_1 \\ 0 & 0 \end{bmatrix} \quad V = \begin{bmatrix} \tau_1 + \iota & -\varphi \\ -\iota & \varphi + \tau_2 \end{bmatrix}$$



# R<sub>0</sub> for SEI with treatment failure

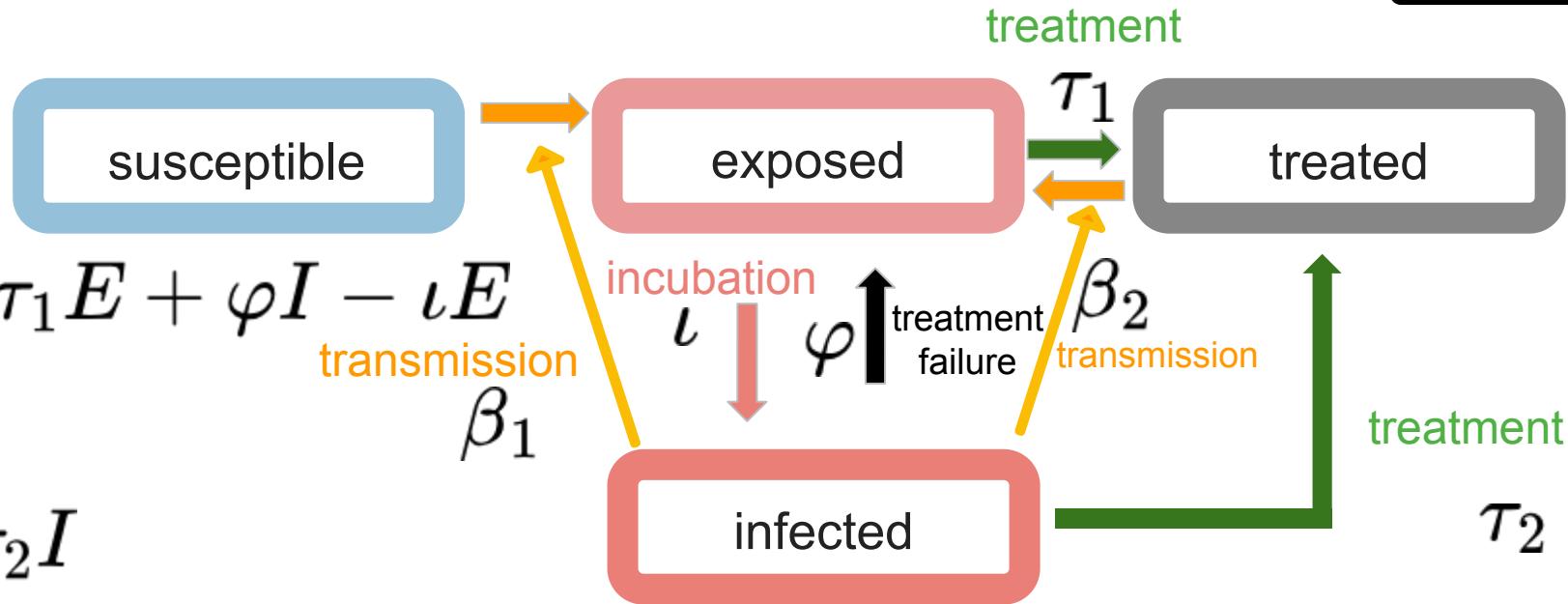
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



## 4. Next-generation matrix

$$F = \begin{bmatrix} 0 & \beta_1 \\ 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \tau_1 + \iota & -\varphi \\ -\iota & \varphi + \tau_2 \end{bmatrix} \quad V^{-1} = \frac{1}{(\tau+\iota)(\varphi+\tau_2)-\iota\varphi} \begin{bmatrix} \varphi + \tau_2 & \varphi \\ \iota & \tau_1 + \iota \end{bmatrix}$$

$$M = FV^{-1} = \frac{1}{(\tau+\iota)(\varphi+\tau_2)-\iota\varphi} \begin{bmatrix} \iota\beta_1 & (\tau_1 + \iota)\beta_1 \\ 0 & 0 \end{bmatrix}$$



# R<sub>0</sub> for SEI with treatment failure

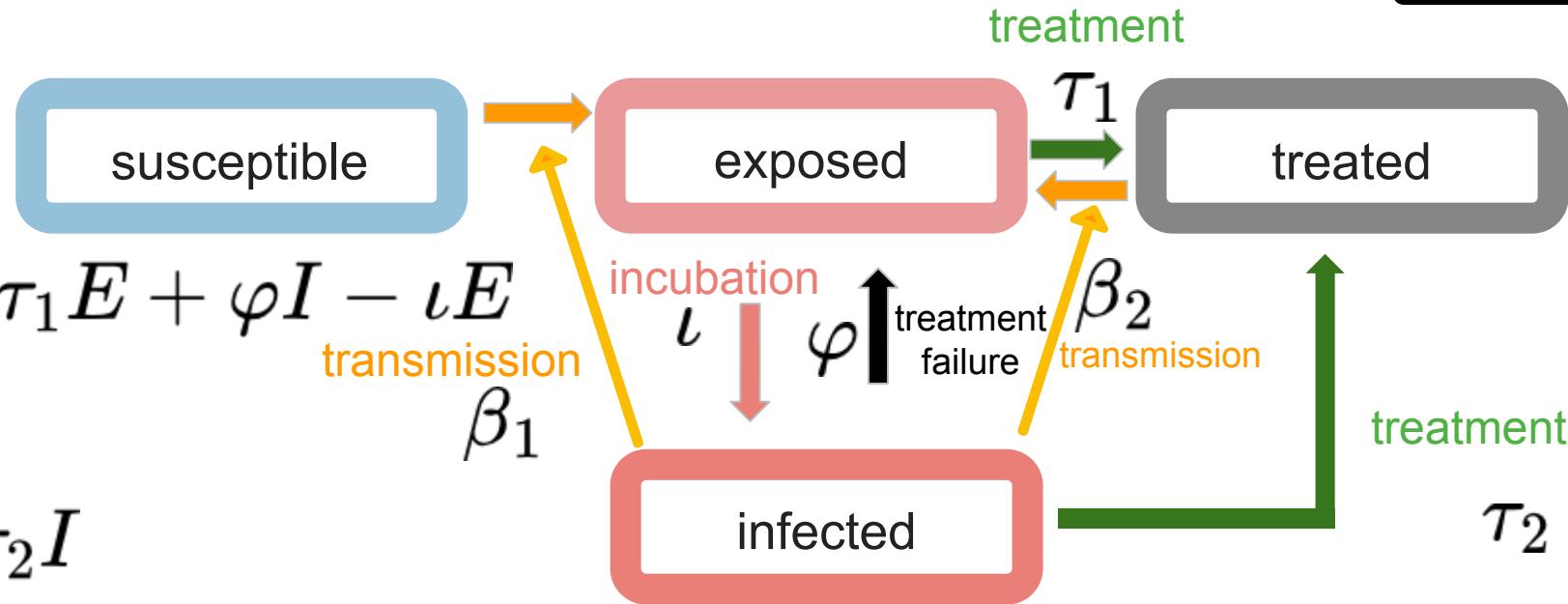
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



## 5. Basic reproduction number

$$\mathcal{R}_0 = \frac{\iota \beta_1}{(\tau + \iota)(\varphi + \tau_2) - \iota \varphi}$$



# R<sub>0</sub> for SEI with treatment failure

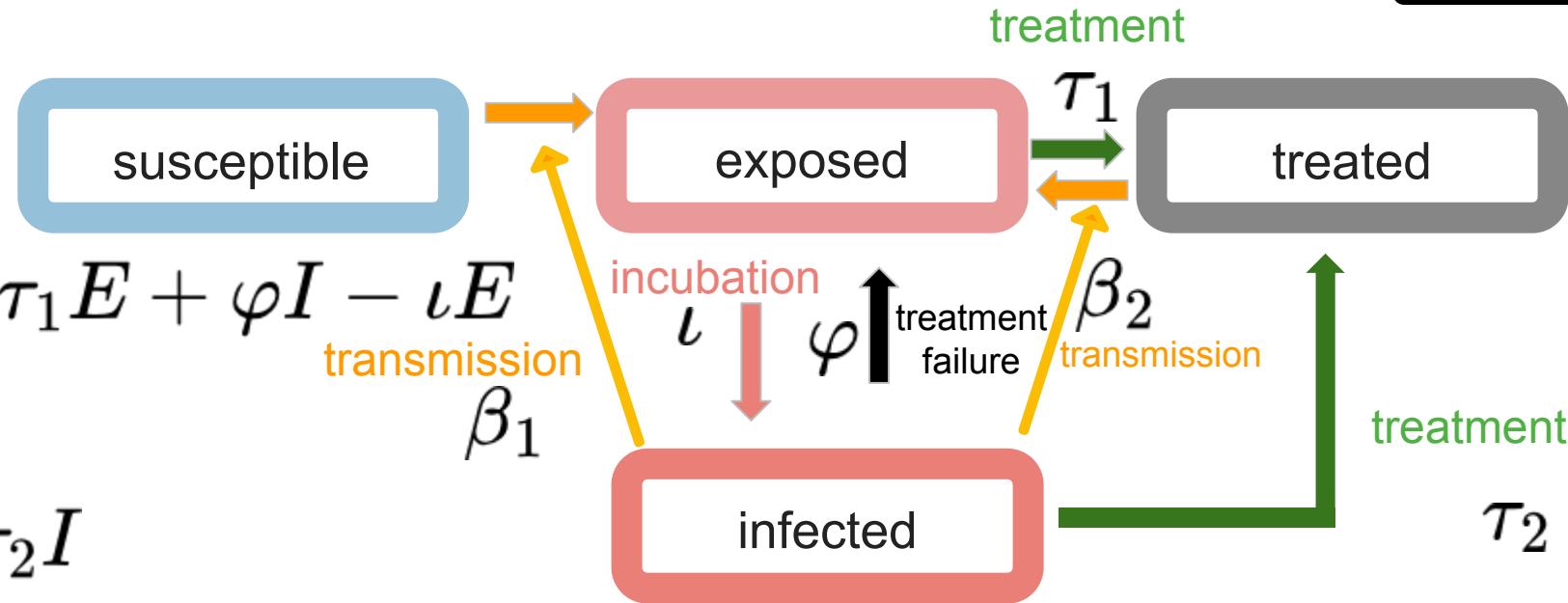
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



E, I are infectious compartments  
**both** incubation and treatment failure **are** considered to be new infections

$$\mathbb{F} = \begin{bmatrix} \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} + \varphi I \\ \iota E \end{bmatrix} \quad \mathbb{V} = \begin{bmatrix} \tau_1 E + \iota E \\ \tau_2 I + \varphi I \end{bmatrix}$$

Decomposition is not unique!



# R<sub>0</sub> for SEI with treatment failure

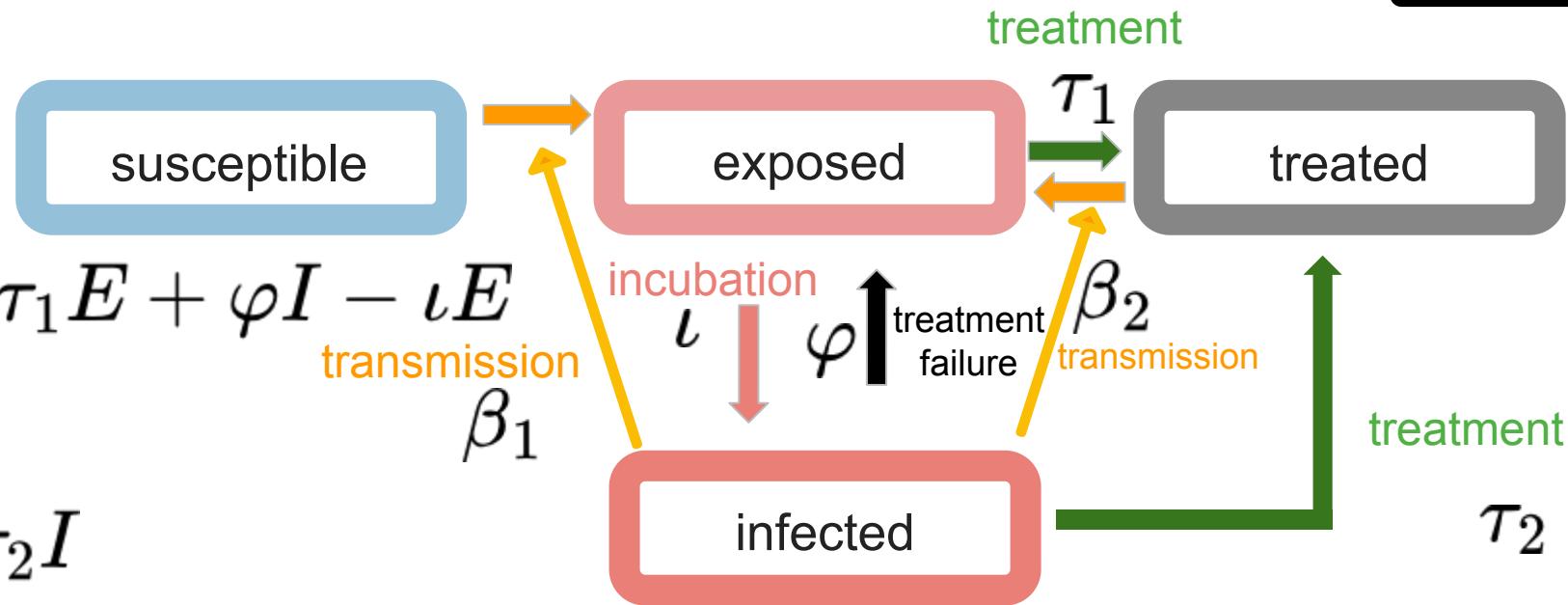
## Tuberculosis

$$\frac{dS}{dt} = -\beta_1 S \frac{I}{N}$$

$$\frac{dE}{dt} = \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} - \tau_1 E + \varphi I - \iota E$$

$$\frac{dI}{dt} = \iota E - \varphi I - \tau_2 I$$

$$\frac{dT}{dt} = -\beta_2 T \frac{I}{N} + \tau_1 E + \tau_2 I$$



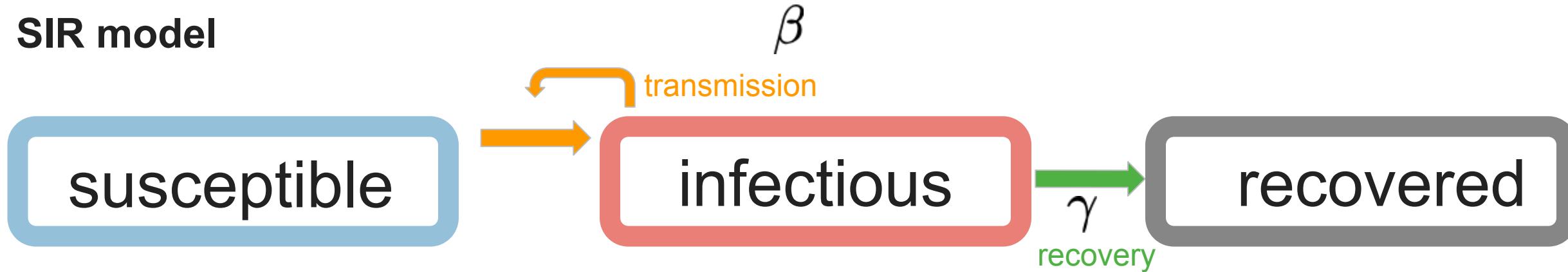
E, I are infectious compartments  
**both** incubation and treatment failure **are** considered to be new infections

$$\mathbb{F} = \begin{bmatrix} \beta_1 S \frac{I}{N} + \beta_2 T \frac{I}{N} + \varphi I \\ \iota E \end{bmatrix} \quad \mathbb{V} = \begin{bmatrix} \tau_1 E + \iota E \\ \tau_2 I + \varphi I \end{bmatrix}$$

**Assignment:** Calculate basic reproduction number in this case and compare!

# R<sub>0</sub> and the final epidemic size

SIR model



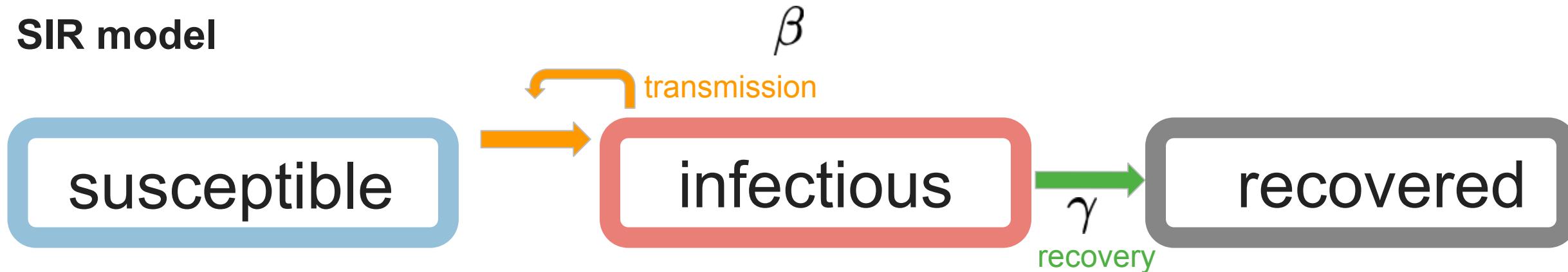
$$\left. \begin{array}{l} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \\ S + I + R = 1 \end{array} \right\} \Rightarrow$$

**final epidemic size**  
How many hosts had been infected (and recovered) during the course of the epidemic?

$$I(\infty) = 0 \quad R(\infty) = ?$$

# R<sub>0</sub> and the final epidemic size

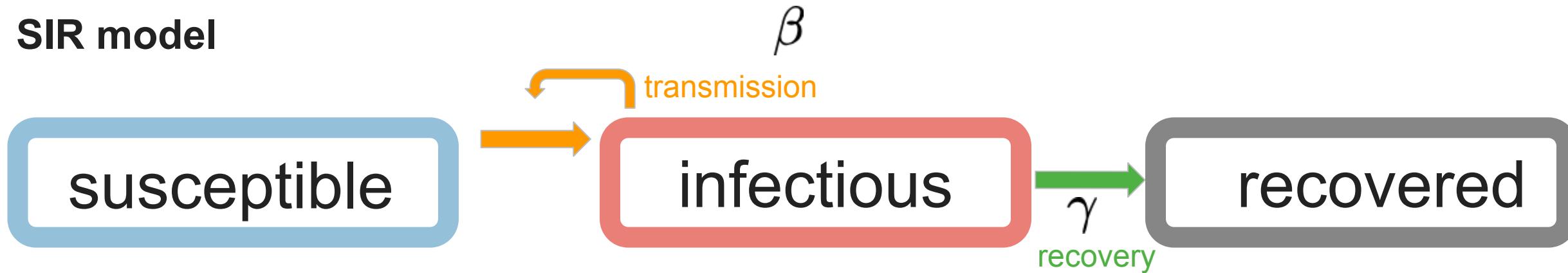
SIR model



$$\left. \begin{array}{l} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \\ S + I + R = 1 \end{array} \right\} \Rightarrow \begin{array}{l} \frac{dI}{dS} = \frac{\beta SI - \gamma I}{-\beta SI} = -1 + \frac{\gamma}{\beta} \frac{1}{S} \\ dI = -dS + \frac{\gamma}{\beta} \frac{dS}{S} = -dS + \frac{\gamma}{\beta} d \log S \end{array}$$

# R<sub>0</sub> and the final epidemic size

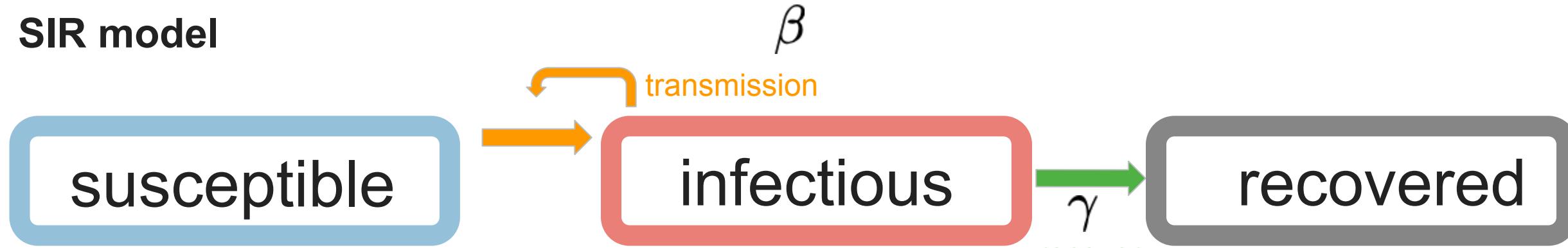
SIR model



$$\left. \begin{array}{l} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \\ S + I + R = 1 \end{array} \right\} \Rightarrow \begin{array}{l} \frac{dI}{dS} = \frac{\beta SI - \gamma I}{-\beta SI} = -1 + \frac{\gamma}{\beta} \frac{1}{S} \\ dI = -dS + \frac{\gamma}{\beta} \frac{dS}{S} = -dS + \frac{\gamma}{\beta} d \log S \\ I(t) = -S(t) + \frac{\gamma}{\beta} \log(S(t)) + C \end{array}$$

# R<sub>0</sub> and the final epidemic size

SIR model



$$I(t) = -S(t) + \frac{\gamma}{\beta} \log(S(t)) + C \quad t \geq 0$$

$$I(0) \approx 0$$

$$S(0) = 1$$

$$I(\infty) = 0$$

$$-S(\infty) + \frac{\gamma}{\beta} \log(S(\infty))) = -S(0) + \frac{\gamma}{\beta} \log(S(0))$$

$$\log S(\infty) = \frac{\beta}{\gamma}(S(\infty) - 1)$$

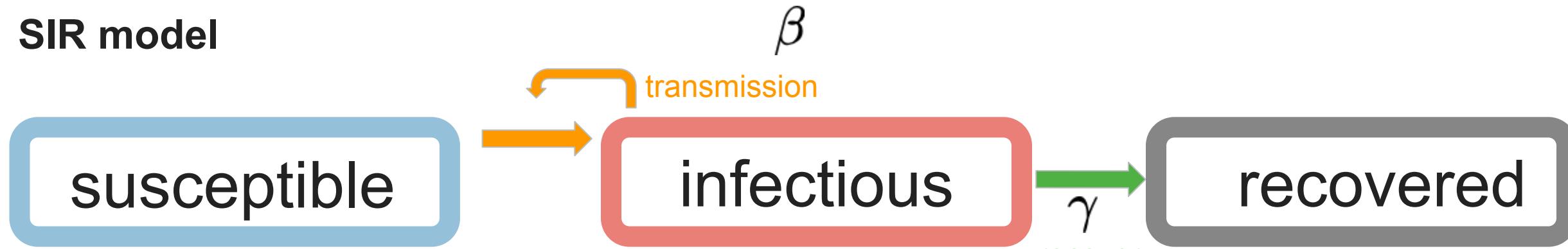
$$S(\infty) = e^{-\frac{\beta}{\gamma}(1-S(\infty))}$$

final epidemic size

$$R(\infty) = 1 - e^{-\frac{\beta}{\gamma} R(\infty)}$$

# R<sub>0</sub> and the final epidemic size

SIR model



$$I(t) = -S(t) + \frac{\gamma}{\beta} \log(S(t)) + C \quad t \geq 0$$

$$I(0) \approx 0$$

$$S(0) = 1$$

$$I(\infty) = 0$$

$$-S(\infty) + \frac{\gamma}{\beta} \log(S(\infty))) = -S(0) + \frac{\gamma}{\beta} \log(S(0))$$

$$\log S(\infty) = \frac{\beta}{\gamma}(S(\infty) - 1)$$

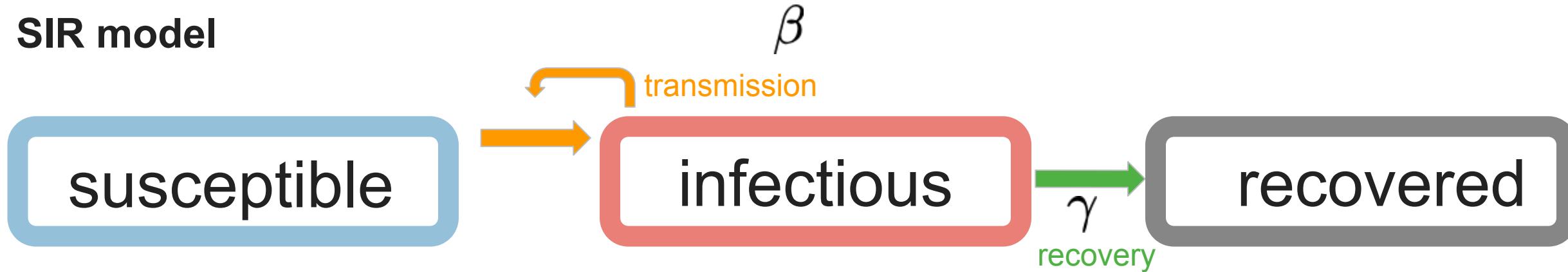
$$S(\infty) = e^{-\frac{\beta}{\gamma}(1-S(\infty))}$$

final epidemic size  
depends on R<sub>0</sub>!

$$R(\infty) = 1 - e^{-\mathcal{R}_0 R(\infty)}$$

# R<sub>0</sub> and the final epidemic size

## SIR model



**Assignment:** Calculate final epidemic size as a function of basic reproduction number using !

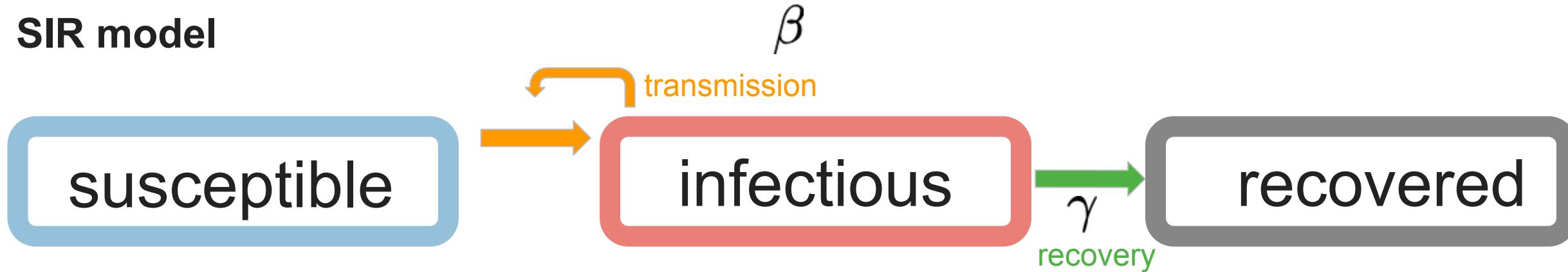
```
f<-function(x,y) {x-1+exp(-y*x) }
FES<-function(R0) uniroot(function(x) f(x,R0),c(0.001,1)) [[1]]
FES(3)
```

final epidemic size, R<sub>0</sub>>1

$$R(\infty) = 1 - e^{-\mathcal{R}_0 R(\infty)}$$

# R<sub>0</sub> and the final epidemic size

## SIR model



Assume a case-fatality ratio of 5%!



With the Rwandan population in 2020, how many deaths would have occurred in total during the uncontrolled epidemic?

```
sapply(c(1.2, 2, 3), FES) * 14 * 10^6 * 0.05
```

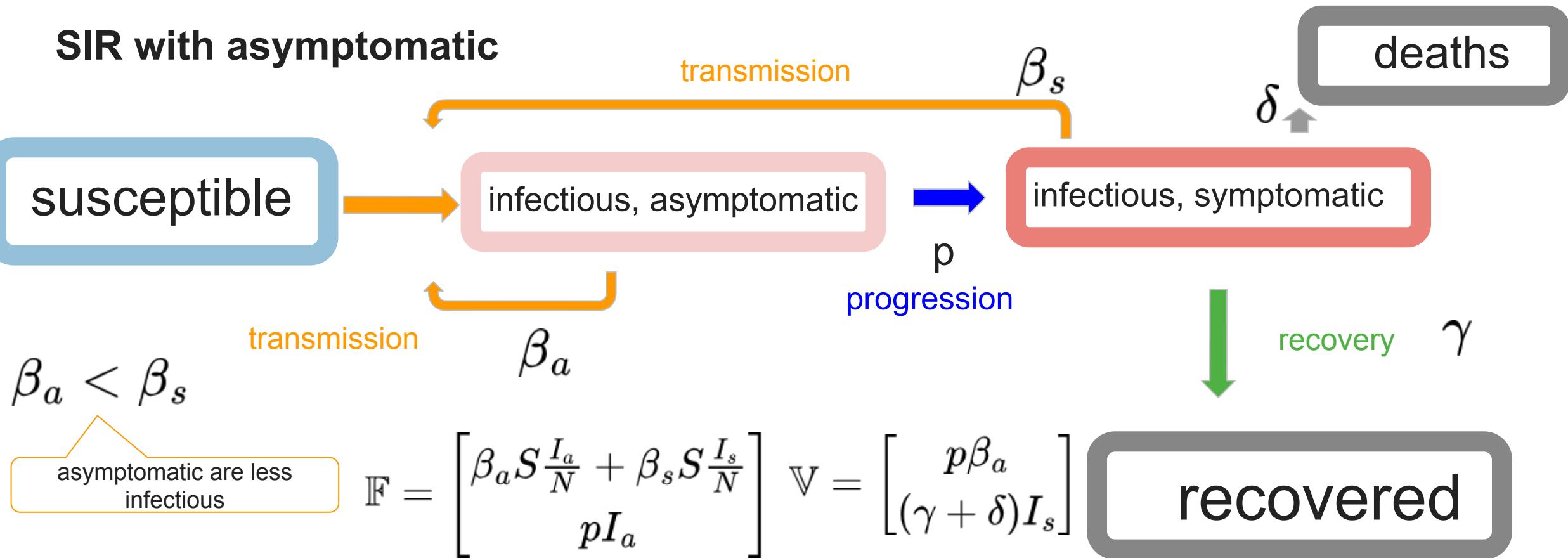
## Key takeaway points:

---

- $R_0$  can be derived as an approximation from **growth rate** and **generation time** data
- $R_0$  is a **threshold** summarizing model parameters providing information whether a small outbreak will become an epidemic
- **next-generation matrix** method allows to calculate  $R_0$  for more complex models
- $R_0$  is an important **performance indicator** in public health (e.g. vector control, vaccination, final epidemic size)

# More examples for next-generation matrix calculations

## SIR with asymptomatic



$$F = \begin{bmatrix} \beta_a & \beta_s \\ p & 0 \end{bmatrix} \quad V = \begin{bmatrix} p & 0 \\ 0 & \gamma + \delta \end{bmatrix} \quad FV^{-1} = \begin{bmatrix} \frac{\beta_a}{p} & \frac{\beta_s}{\gamma + \delta} \\ 1 & 0 \end{bmatrix}$$

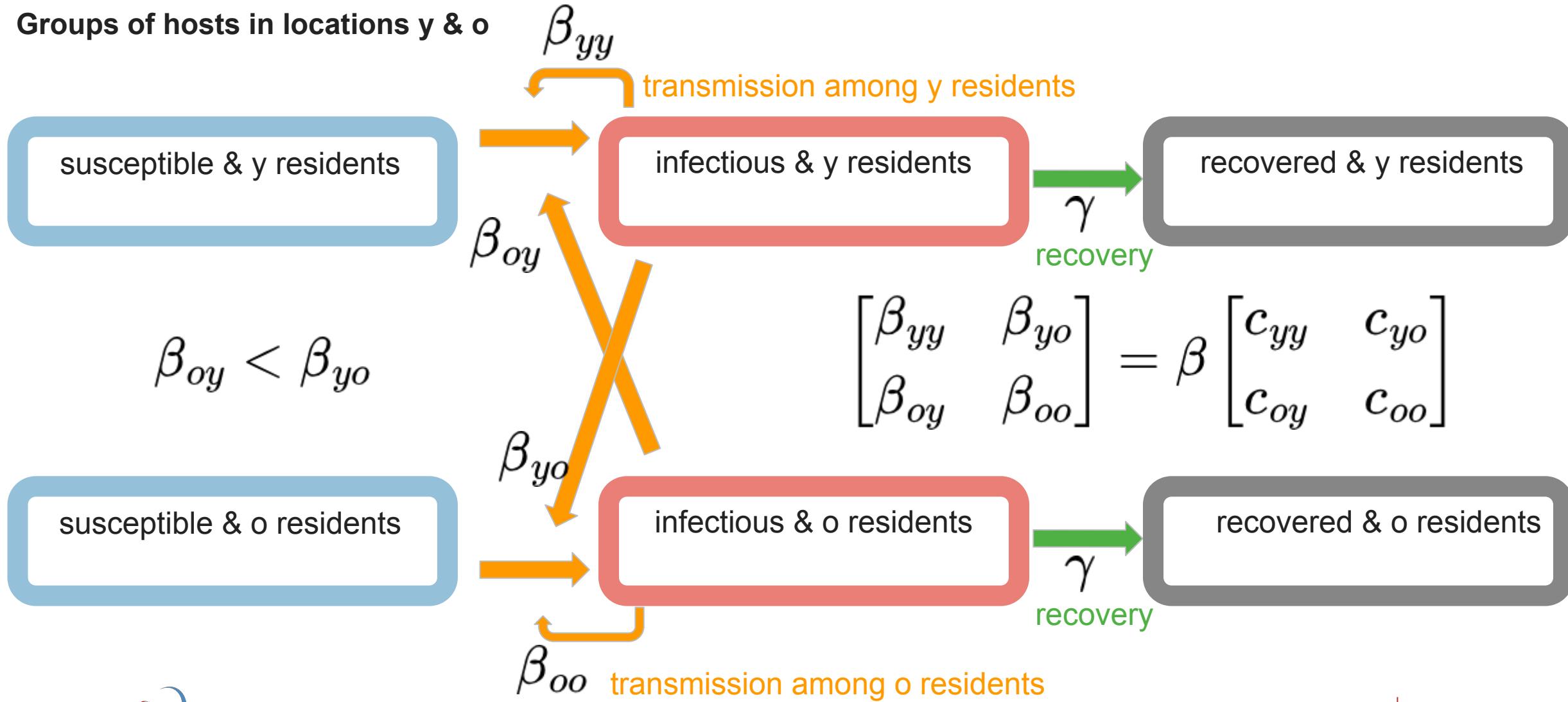
$$(\frac{\beta_a}{p} - \lambda)(-\lambda) - \frac{\beta_s}{\gamma + \delta} = 0$$

$$\lambda^2 - \lambda \frac{\beta_a}{p} - \frac{\beta_s}{\gamma + \delta} = 0$$

$$\mathcal{R}_0 = \frac{\beta_a}{2p} + \sqrt{\frac{\beta_a^2}{4p^2} + \frac{\beta_s}{\gamma + \delta}}$$

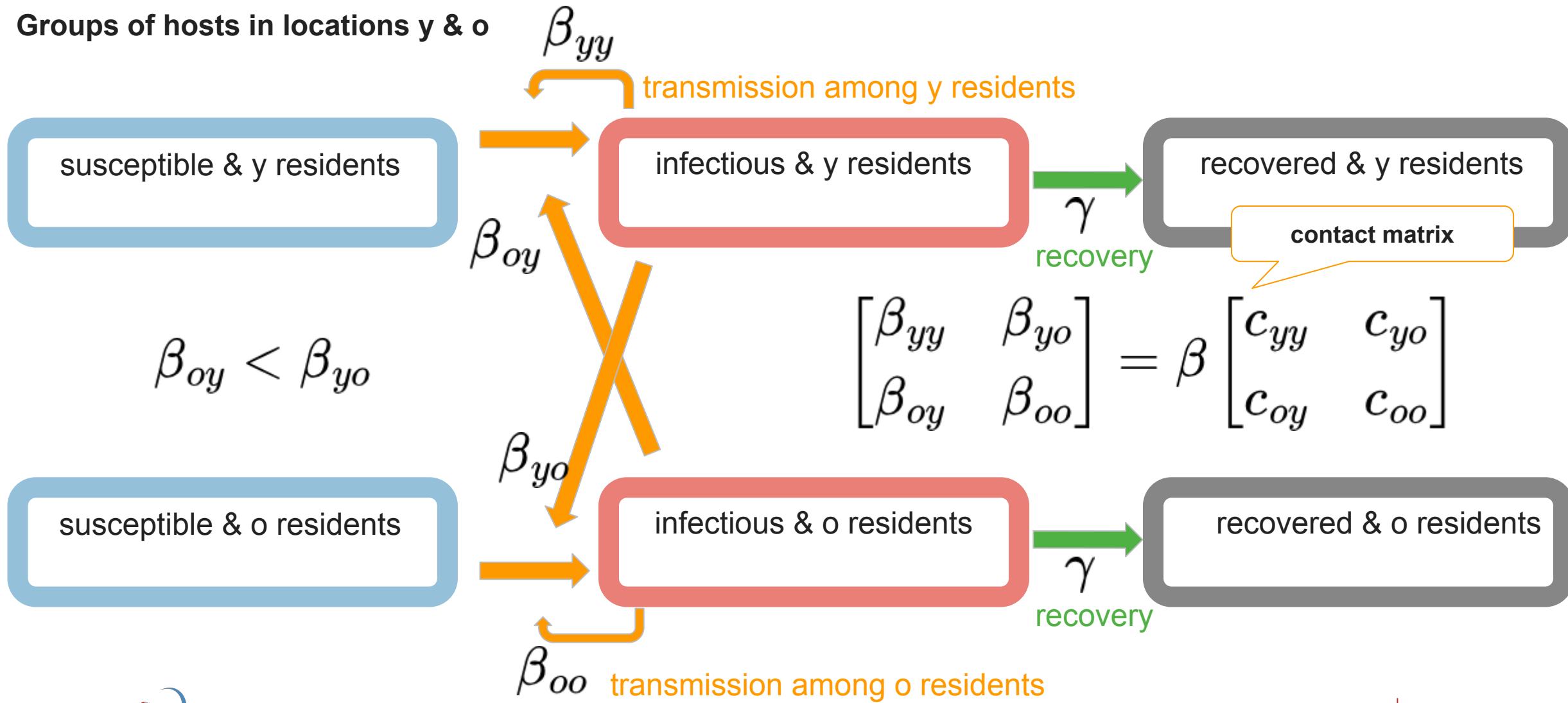
# More examples for next-generation matrix calculations

Groups of hosts in locations y & o



# More examples for next-generation matrix calculations

Groups of hosts in locations y & o



# More examples for next-generation matrix calculations

## Age groups

$$\mathbb{F} = \begin{bmatrix} \beta \frac{S^y}{N} \sum_i c_{iy} I^i \\ \beta \frac{S^o}{N} \sum_i c_{io} I^i \end{bmatrix}$$

$$\mathbb{V} = \begin{bmatrix} \gamma I^y \\ \gamma I^o \end{bmatrix}$$

$$F = \beta \left( \frac{(S^i)^*}{N} c_{ji} \right)_{i,j}$$

$$V = \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix}$$

$$V^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \gamma^{-1} \end{bmatrix}$$

$$\mathcal{R}_0 = \frac{1}{2} \frac{\beta}{\gamma} \rho(C^T) = \frac{1}{2} \frac{\beta}{\gamma} \rho(C)$$

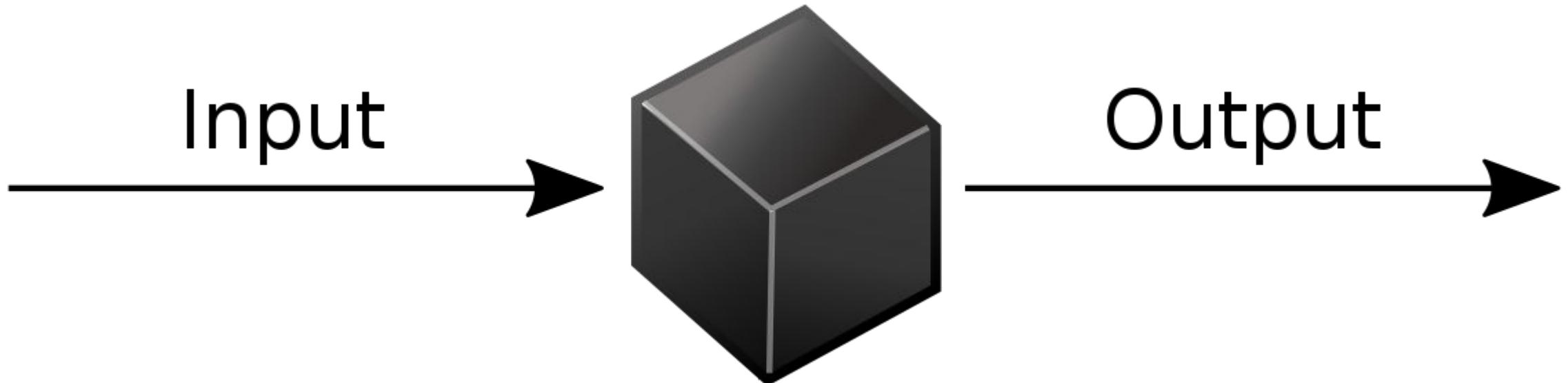
assuming that  
 $(S^y)^* = (S^o)^* = N/2$

spectral radius of  
contact matrix



### 3 - Sensitivity analysis

# Sensitivity analysis



# Sensitivity analysis

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

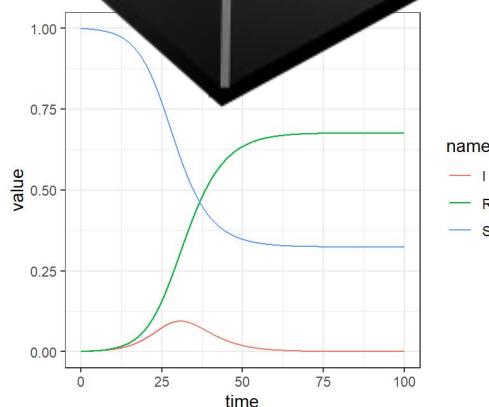
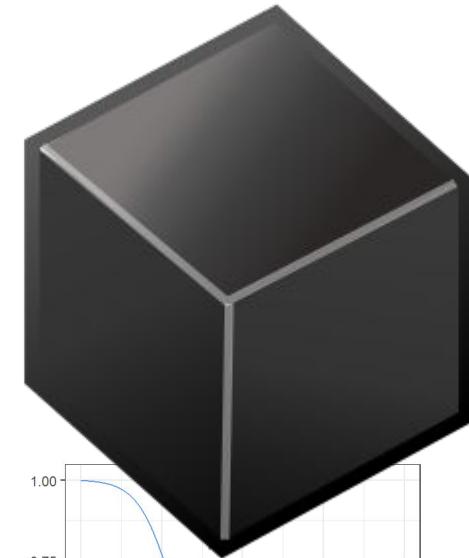
$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

Input



Output



# Sensitivity analysis

$$\beta = 0.5$$

Input

$$\gamma = 0.3$$

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

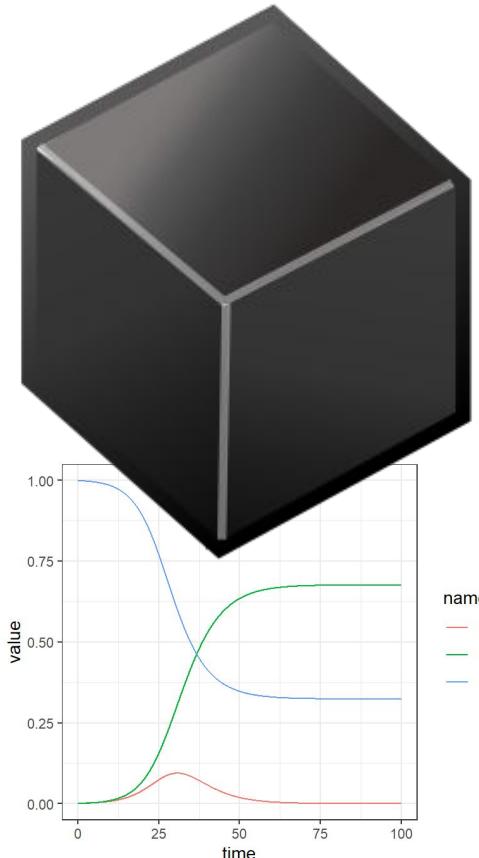
$$\mathcal{R}_0 = \frac{5}{3}$$

$$R(\infty) = 0.68$$

Output

$$\max_{t \geq 0}(I(t)) = 0.09 = \hat{I}$$

$$\operatorname{argmin}_{t \geq 0}\{I(t) = \hat{I}\} = 30.78 \text{ days}$$



# Sensitivity analysis

$$\beta = 0.9$$

Input

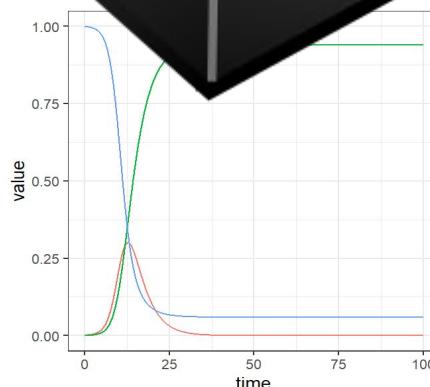
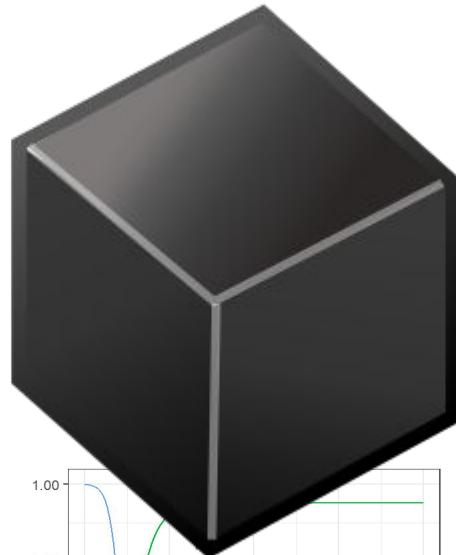
$$\gamma = 0.3$$

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

$$\mathcal{R}_0 = 3$$

$$R(\infty) = 0.94$$

Output



$$\max_{t \geq 0} (I(t)) = 0.3 = \hat{I}$$

$$\operatorname{argmin}_{t \geq 0} \{I(t) = \hat{I}\} = 12.78 \text{ days}$$

# Sensitivity analysis

Important questions after we have built our model:

- how do small **variations** in input values **influence** model output? [gradient]
- is there a general (e.g. linear) **trend** between input and output? [regression]
- if we consider **input values as random**, how do output values depend on them? [Sobol]
- how are the **input parameters interacting** between each other towards the output? [Sobol]

# Sensitivity analysis

**Why** should we care about sensitivity?

- remove redundant parts of model
- fix insensitive parameters of model
- calibration should focus on sensitive parameters first

# Local sensitivity analysis

## Gradients

**Local sensitivity index** of output Q w.r.t. parameter p:

$$\chi_p^Q := \frac{\partial Q}{\partial p} \frac{p}{Q} \approx \frac{\frac{\partial Q}{\partial p}}{\frac{Q}{p}}$$

% change in output

% change in input

$$\chi_\gamma^{\mathcal{R}_0} := \frac{\partial \mathcal{R}_0}{\partial \gamma} \frac{\gamma}{\mathcal{R}_0} = -\beta \gamma^{-2} \frac{\gamma}{\beta} = -1$$

for 1 % increase in recovery rate, we have 1% decrease in R0

$$\chi_\beta^{\mathcal{R}_0} := \frac{\partial \mathcal{R}_0}{\partial \beta} \frac{\beta}{\mathcal{R}_0} = \gamma^{-1} \frac{\beta}{\beta} = \gamma^{-1} \gamma = 1$$

for 1 % increase in transmission rate, we have 1% increase in R0

# Local sensitivity analysis

## Gradients

**Local sensitivity index** of output Q w.r.t. parameter p:

$$\chi_p^Q := \frac{\partial Q}{\partial p} \frac{p}{Q} \approx \frac{\frac{\partial Q}{\partial p}}{\frac{Q}{p}}$$

% change in output

% change in input

$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K}$$

$$\chi_{\beta}^{\mathcal{R}_0} := \frac{\partial \mathcal{R}_0}{\partial \beta} \frac{\beta}{\mathcal{R}_0} =$$

# Local sensitivity analysis

## Gradients

**Local sensitivity index** of output Q w.r.t. parameter p:

$$\chi_p^Q := \frac{\partial Q}{\partial p} \frac{p}{Q} \approx \frac{\frac{\partial Q}{\partial p}}{\frac{Q}{p}}$$

% change in output

% change in input

$$\mathcal{R}_0 = \sqrt{\frac{\alpha}{\gamma} \frac{\beta}{\delta} K}$$

$$\chi_{\beta}^{\mathcal{R}_0} := \frac{\partial \mathcal{R}_0}{\partial \beta} \frac{\beta}{\mathcal{R}_0} = \frac{1}{2}$$

# Local sensitivity analysis

## Gradients

**Local sensitivity index** of output Q w.r.t. parameter p:

$$\chi_p^Q := \frac{\partial Q}{\partial p} \frac{p}{Q} \approx \frac{\frac{\partial Q}{\partial p}}{\frac{Q}{p}}$$

% change in output

% change in input

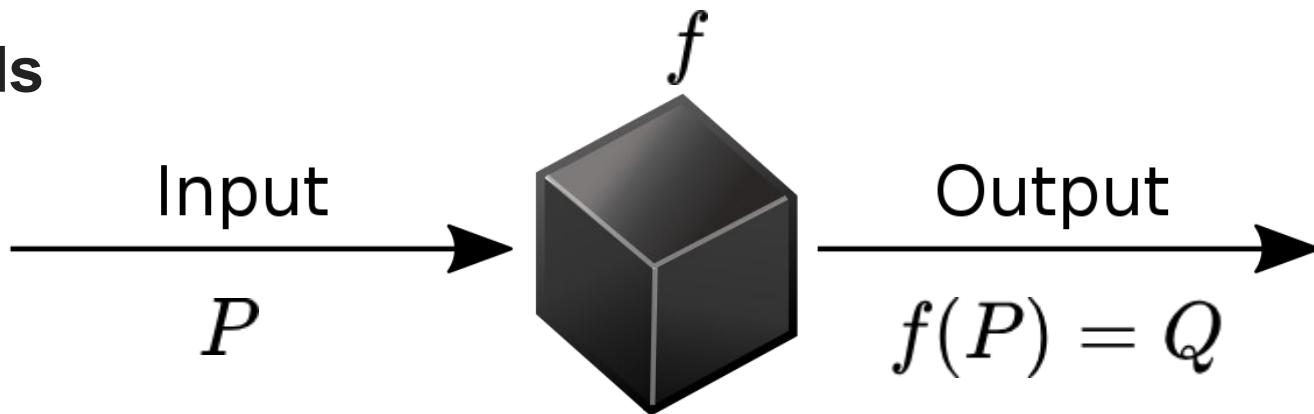
$$R(\infty) = 1 - e^{-\mathcal{R}_0 R(\infty)}$$

$\chi_\beta^{R(\infty)} =$  Use automated differentiation for the root function of  $x \mapsto 1 - e^{-\mathcal{R}_0 x} - x$

# Global sensitivity analysis

## Variance-based methods

- model as black box



- decompose variance in output based on variance in input

$$f_0(P) = \mathbb{E}(f(P))$$

average effect on output

$$f_i(P_i) = \mathbb{E}(f(P)|P_i) - \mathbb{E}(f(P))$$

main effect on output by varying P\_i

$$f_{ij}(P_i, P_j) = \mathbb{E}(f(P)|P_i, P_j) - f_0 - f_i - f_j$$

main effect on output  
by varying P\_i and P\_j

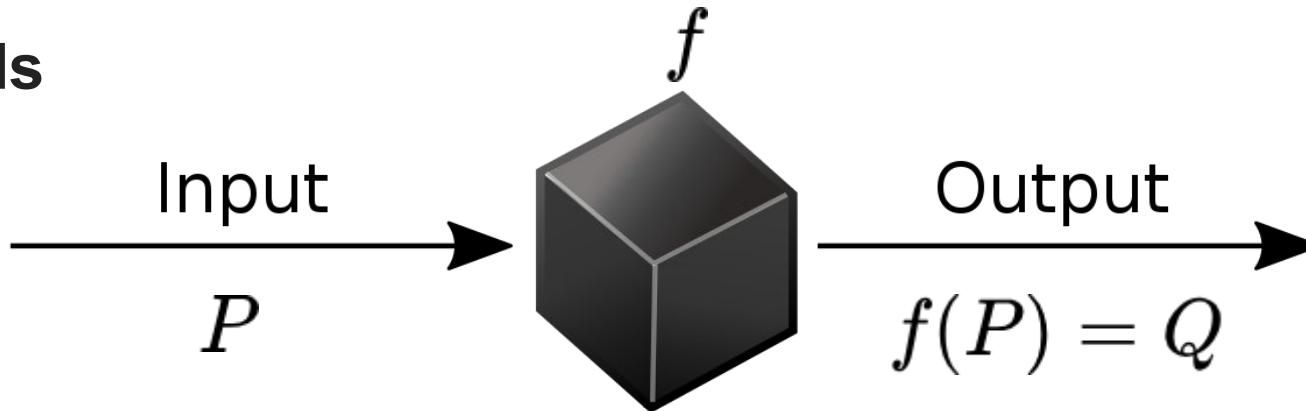
$$f(P) = f_0 + \sum_i f_i(P_i) + \sum_{i,j} f_{ij}(P_i, P_j) + \dots$$

orthogonal  
decomposition

# Global sensitivity analysis

## Variance-based methods

- model as black box



- decompose **variance** in output based on variance in input

$$f(P) = f_0 + \sum_i f_i(P_i) + \sum_{i,j} f_{ij}(P_i, P_j) + \dots$$

$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

$$\text{Var}(f(P)) = \sum_i V_i + \sum_{i < j} V_i V_j + \dots$$

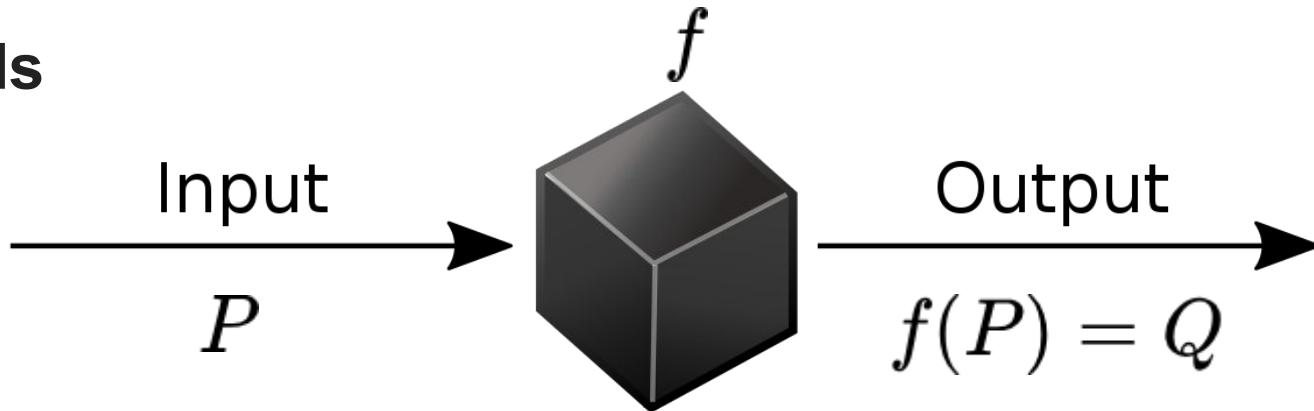
$$V_i = \text{Var}(f_i(P_i)) = \text{Var}(\mathbb{E}(f(P)|P_i))$$

$$V_{ij} = \text{Var}(f_{ij}(P_i, P_j)) = \text{Var}(\mathbb{E}(f(P)|P_i, P_j)) - V_i - V_j$$

# Global sensitivity analysis

## Variance-based methods

- model as black box



- decompose **variance** in output based on variance in input

$$f(P) = f_0 + \sum_i f_i(P_i) + \sum_{i,j} f_{ij}(P_i, P_j) + \dots$$

$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

$$\text{Var}(f(P)) = \sum_i V_i + \sum_{i < j} V_i V_j + \dots$$

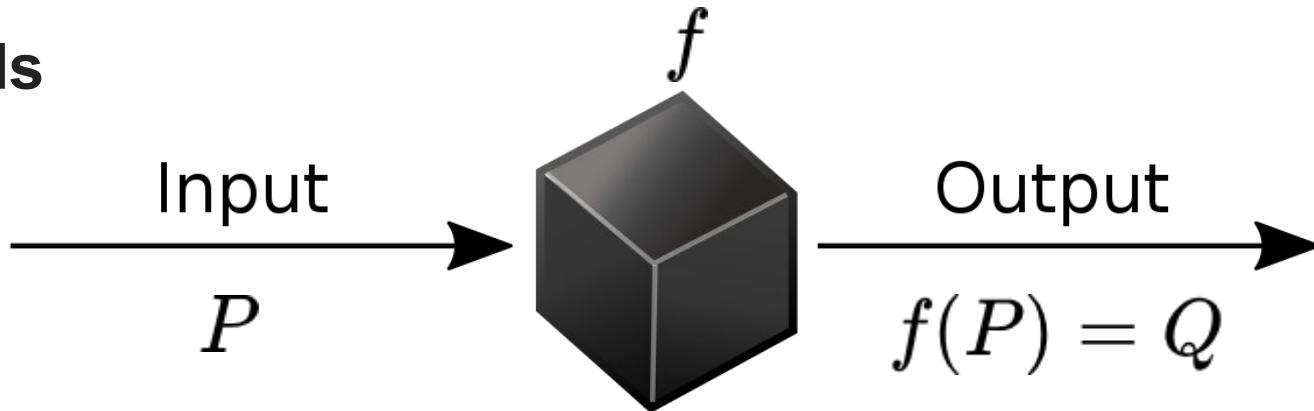
interaction between  $P_i$  and  $P_j$   $\text{Var}(\mathbb{E}(f(P)|P_i))$

$$V_{ij} = \text{Var}(f_{ij}(P_i, P_j)) = \text{Var}(\mathbb{E}(f(P)|P_i, P_j)) - V_i - V_j$$

# Global sensitivity analysis

## Variance-based methods

- model as black box



- decompose **variance** in output based on variance in input

$$V_i = \text{Var}(f_i(P_i)) = \text{Var}(\mathbb{E}(f(P)|P_i))$$

$$V_{ij} = \text{Var}(f_{ij}(P_i, P_j)) = \text{Var}(\mathbb{E}(f(P)|P_i, P_j)) - V_i - V_j$$

$$S_i = \frac{V_i}{\text{Var}(f(P))}$$

**first-order Sobol index**

relative contribution of from  $P_i$  to the total variance of the output

# Global sensitivity analysis

## Variance-based methods

$$V_i = \text{Var}(f_i(P_i)) = \text{Var}(\mathbb{E}(f(P)|P_i))$$

$$S_i = \frac{V_i}{\text{Var}(f(P))} \quad \text{first-order Sobol index}$$

- need to sample from  $P_i$
- two independent sample sets of size N from parameter distribution A, B
- conditional probability:
- estimator for  $V_i$ :

$$\hat{V}_i = \frac{1}{N} \sum_j f(B)_j (\underbrace{f(A_B^i)_j - f(A)_j}_{\text{conditional probability}})$$

conditional probability

$A_B^i$ : i column of B into A

# Global sensitivity analysis

## Variance-based methods

$$V_i = \text{Var}(f_i(P_i)) = \text{Var}(\mathbb{E}(f(P)|P_i))$$

$$S_i = \frac{V_i}{\text{Var}(f(P))}$$

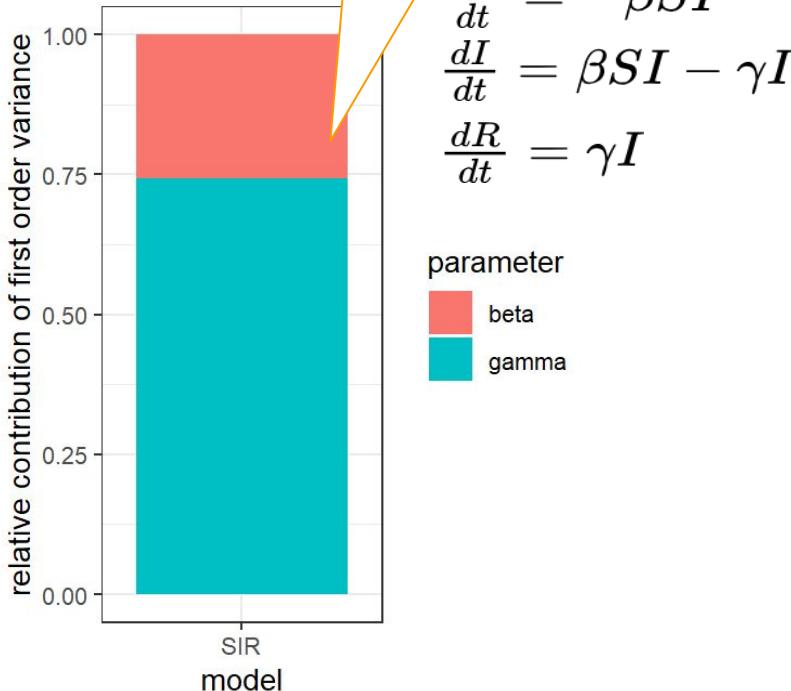
**first-order Sobol index**

Need to fix lower/upper bounds for parameter sampling here: 0.1-0.3

$$\hat{V}_i = \frac{1}{N} \sum_j f(B)_j (f(A_B^i)_j - f(A)_j)$$

Use the `sobolSalt` function in the R package `sensitivity` with  $f(P)$ = infection peak for the SIR model!

Download the AIDM2.R script from the webpage!



# Global sensitivity analysis

## Variance-based methods

$$V_i = \text{Var}(f_i(P_i)) = \text{Var}(\mathbb{E}(f(P)|P_i))$$

$$S_i = \frac{V_i}{\text{Var}(f(P))}$$

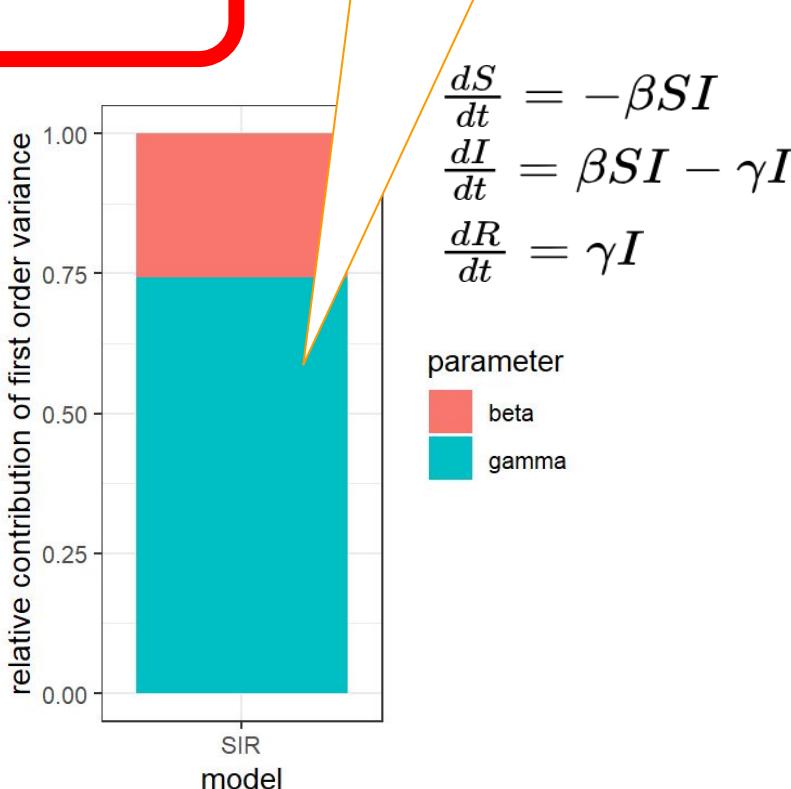
**first-order Sobol index**

Contribution to peak of infection is dominated by recovery rate

$$\hat{V}_i = \frac{1}{N} \sum_j f(B)_j (f(A_B^i)_j - f(A)_j)$$

Use the `sobolSalt` function in the R package `sensitivity` with  $f(P)$ = infection peak for the SIR model!

Download the AIDM2.R script from the webpage!



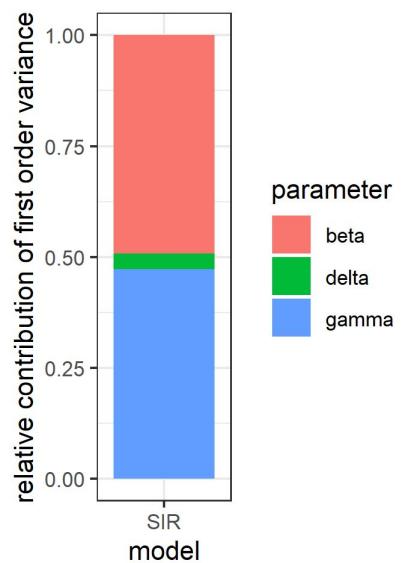
# Global sensitivity analysis

## Variance-based methods

$$S_i = \frac{V_i}{\text{Var}(f(P))} \quad \text{first-order Sobol index}$$

Use the `sobolSalt` function in the R package `sensitivity` with  $f(P)$ = infection peak for the model!

Adapt the **AIDM\_02.R** script for an SIR model with virulence (death from disease)!



Virulence = mortality due to the infectious disease

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

# Global sensitivity analysis

## Variance-based methods

Use the `sobolSalt` function in the R package `sensitivity` with  $f(P)$ = endemic equilibrium for the malaria model!

Adapt the script in AIDM\_02.R!

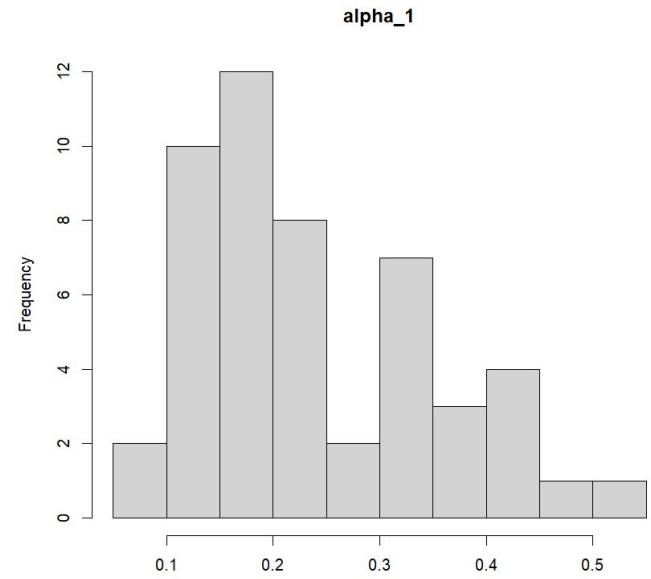
# Sensitivity vs. Uncertainty

## Forward-propagation of uncertainty

- Parameters usually carry **uncertainty**, i.e. the numeric value for simulation is drawn randomly from probability distribution
- Assume that the biting rate ***alpha\_1*** follows a Gamma probability distribution with shape=5, scale=0.05 and simulate 500 trajectories:

```
alpha1_500<-rgamma (n=500, shape=5, scale=0.05)
```

How **uncertain is the prevalence** at the endemic equilibrium given parameter uncertainty?



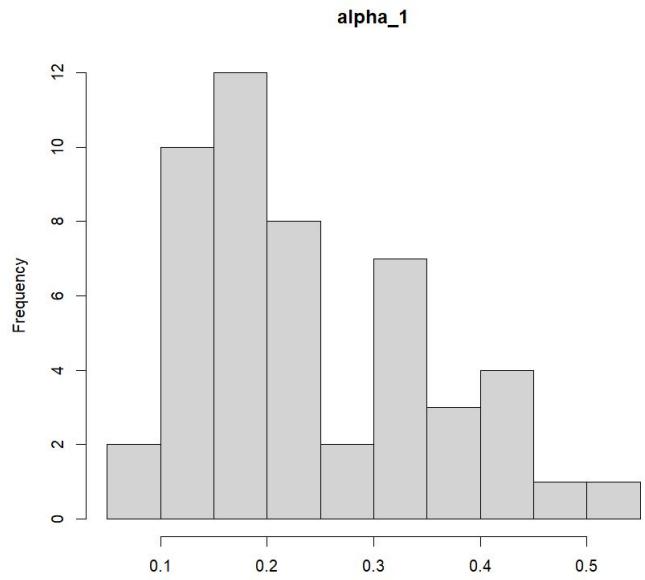
# Forward-propagation of uncertainty

How **uncertain is the prevalence** at the endemic equilibrium of our malaria model given parameter uncertainty?

- Write a function

```
runRM<-function(alpha1) {  
  ...  
  results$prevalence<-results$I/H  
  return(results) }
```

- Use `lapply(alpha_500, runRM)` in R
- See the AIDM2.R script from the webpage!

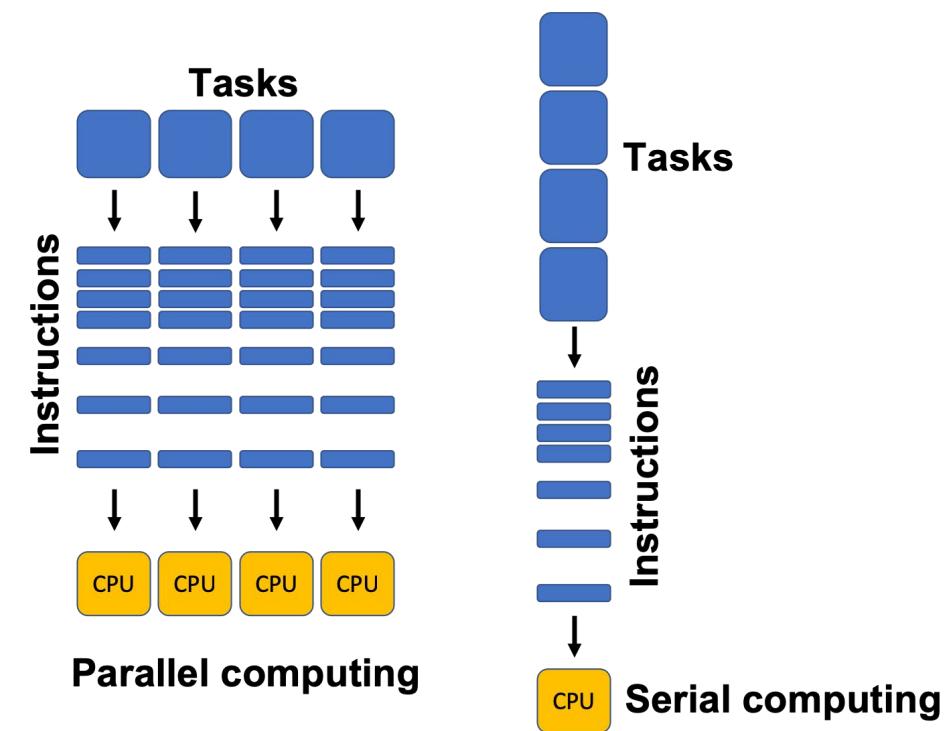


# Forward-propagation of uncertainty

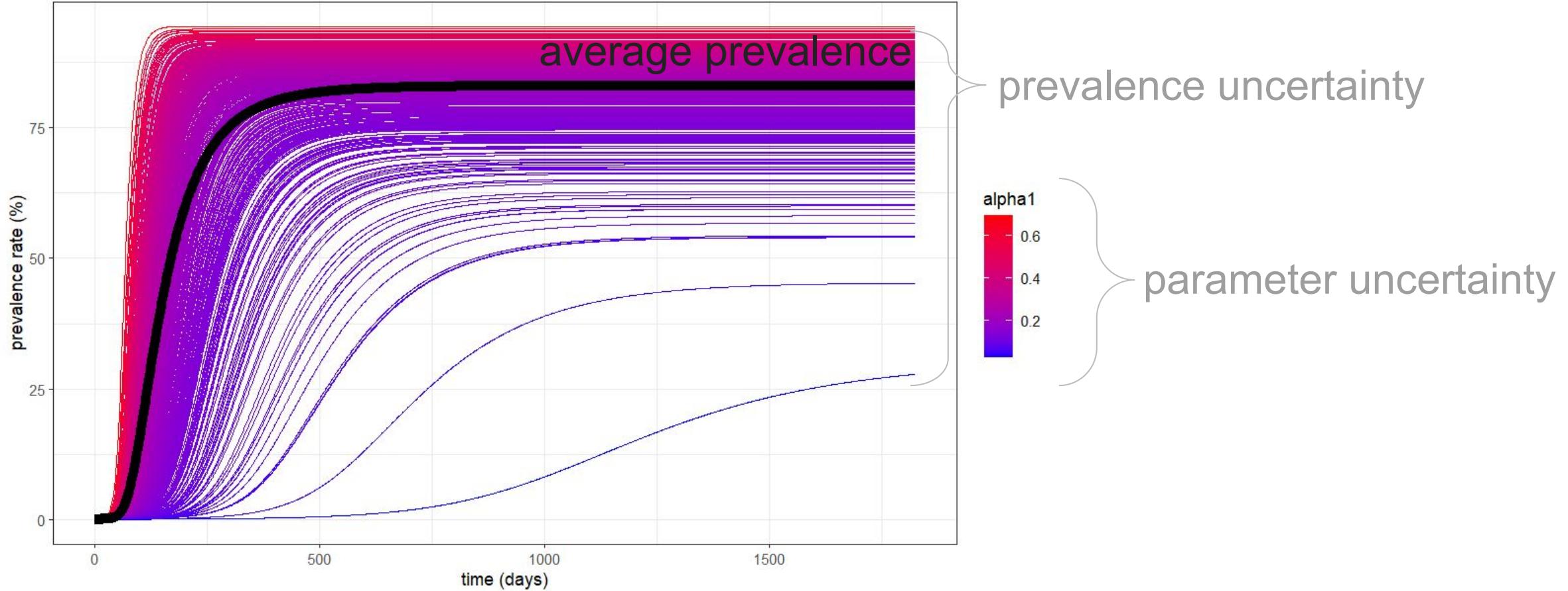
- Is `lapply(alpha1_500, runRM)` in R too slow?
- Parallelize!

```
library(foreach)
library(doParallel)

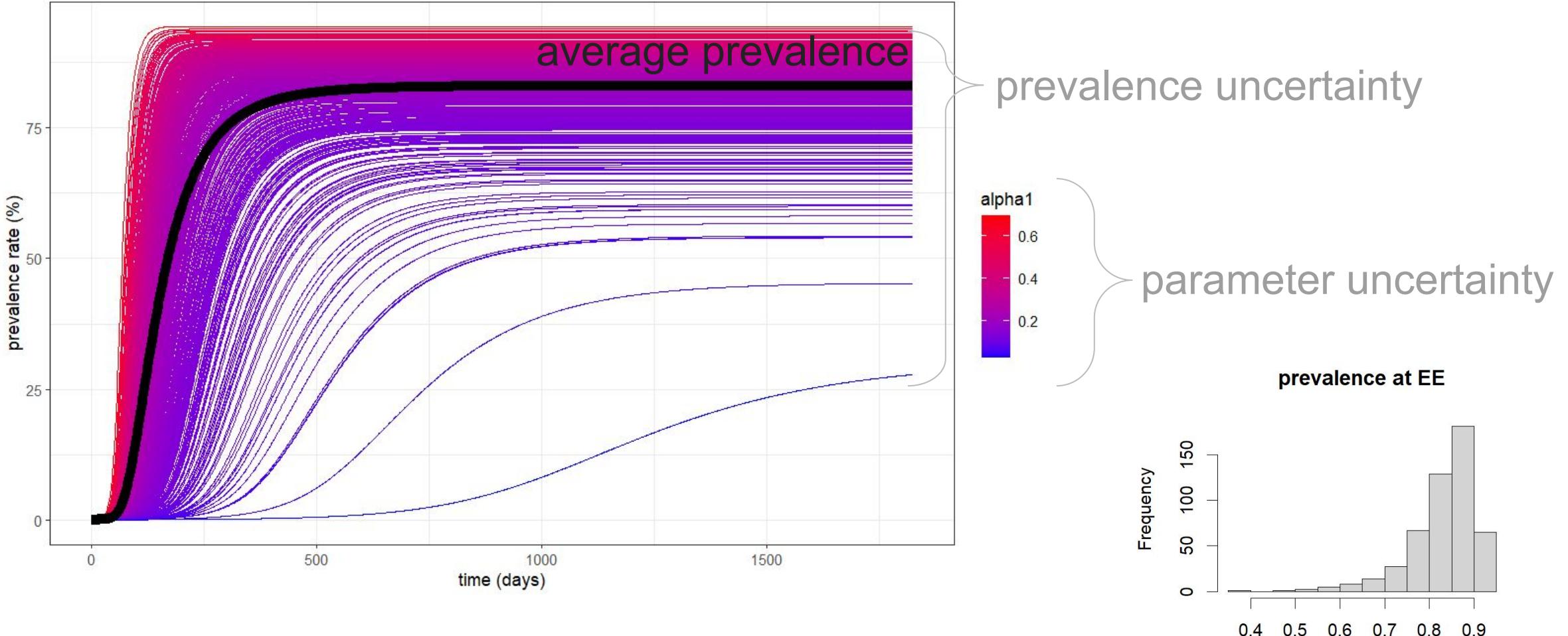
n_cores <- detectCores()
cluster <- makeCluster(n_cores - 1)
registerDoParallel(cluster)
n_iterations <- length(alpha1)
results <- list()
results <- foreach(i = 1:n_iterations) %dopar% {
  results[i] <- runRM(alpha1[i])
}
stopCluster(cl = cluster)
for(i in 1:n_iterations) {
  results[[i]]$alpha1<-alpha1[i]
  results[[i]]$iter<-i
}
results<-bind_rows(results)
```



# Forward-propagation of uncertainty



# Forward-propagation of uncertainty



## Key takeaway points:

---

- sensitivity analysis is an **important step in the design** of a model, it allows to understand the input-to-output relationship and interactions
- **gradient-based** sensitivity analysis considers small variations in parameter space
- **global sensitivity** analysis considers the relationship between variance of inputs and outputs
- **forward propagation of uncertainty** is important to illustrate the effect of parameter uncertainty on performance indicators of disease models



## 4 - Curve fitting

# Influenza outbreak in a boarding school

date	in bed	convalescent	total
1978-01-22	3	0	763
1978-01-23	8	0	763
1978-01-24	26	0	763
1978-01-25	76	0	763
1978-01-26	225	9	763
1978-01-27	298	17	763
1978-01-28	258	105	763
1978-01-29	233	162	763

## Main assumptions

- time series of **symptomatic cases** and **convalescent hosts**
- closed population
- well-mixed
- immunity upon recovery

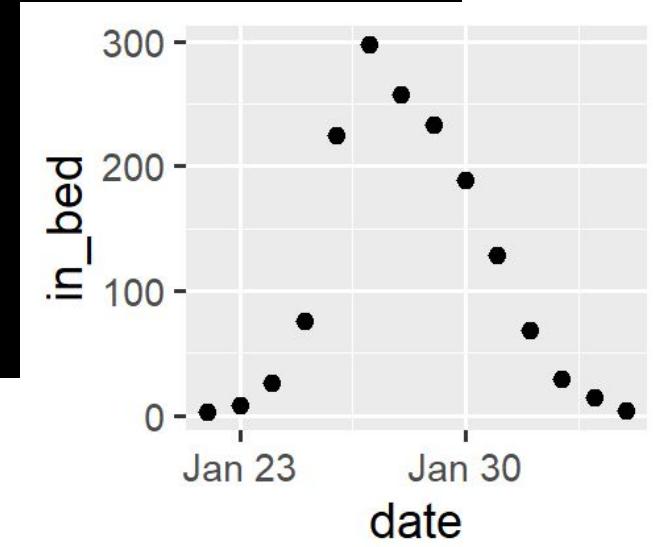
# Visualize the data

- install the R package **outbreaks**
- visualize with **ggplot2** the data frame **influenza\_england\_1978** as points

# Visualize the data

- install the R package **outbreaks**
- visualize with **ggplot2** the data frame **influenza\_england\_1978** as points

```
library(outbreaks)
##add time column to data from outbreaks package
influenza_england_1978_school<-influenza_england_1978_school%>%
  mutate(time=row_number(),
        total=763)
##plot the data
ggplot(influenza_england_1978_school) +
  geom_point(aes(x=date, y=in_bed))
```





# Visualize the data

- solve the ODE system with **deSolve** in R with  $\beta = 1.1$  and  $\gamma = 0.5$
- plot the solution curve on top of the data points

$$\frac{dS}{dt} = - \beta I \frac{S}{N}$$

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

$$\frac{dR}{dt} = \gamma I$$

# Visualize the data

- solve the ODE system with **deSolve** in R with  $\beta = 1.1$  and  $\gamma = 0.5$
- plot the solution curve on top of the data points

```
SIR.model<-function(time,state,parms) {
  with(as.list(c(state,parms)), {
    beta=parms[1]; gamma=parms[2];
    S=state[1]; I=state[2];R=state[3];N=S+I+R
    dS= - beta*I*S/N
    dI= beta*I*S/N - gamma*I
    dR= gamma*I
    return(list(c(dS,dI,dR)))
  })
}

time.points<-seq(0,nrow(influenza_england_1978_school),0.01)
initial.condition<-c(S=762,I=1,R=0)
parameters<-c(beta=1.1,gamma=0.5)

as.data.frame(ode(initial.condition,time.points,SIR.model,parameters))>solution_SIR

##lets plot the model output on top of the data
ggplot(merge(solution_SIR,influenza_england_1978_school))+  

  geom_line(aes(x=time,y=I),color="darkgreen")+
  geom_point(aes(x=time,y=in_bed))
```

$$\frac{dS}{dt} = - \beta I \frac{S}{N}$$

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

$$\frac{dR}{dt} = \gamma I$$

# Visualize the data

- solve the ODE system with **deSolve** in R with  $\beta = 1.1$  and  $\gamma = 0.5$
- plot the solution curve on top of the data points

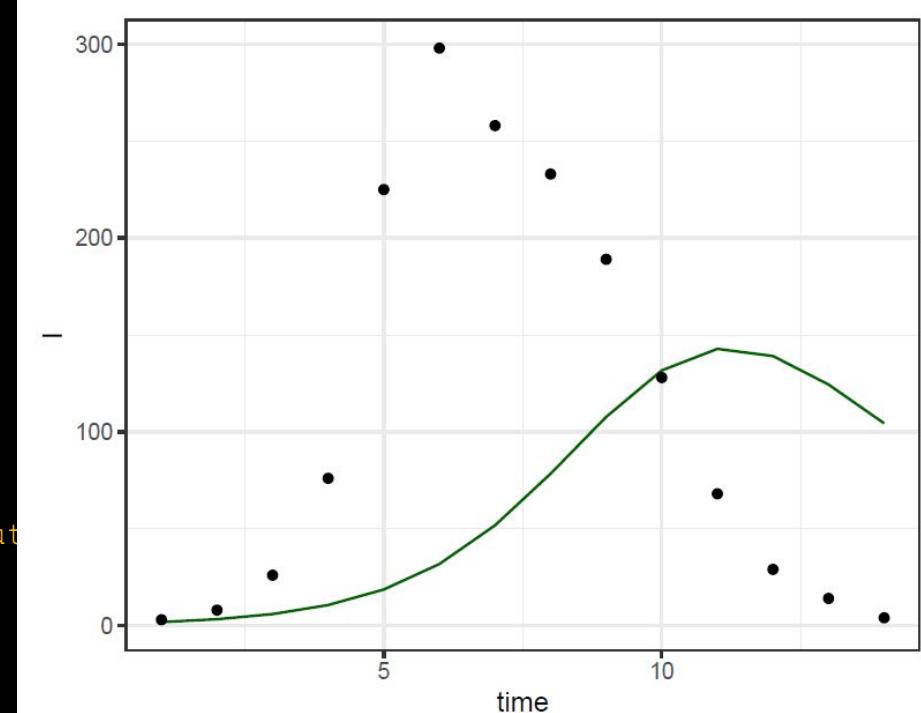
```
SIR.model<-function(time,state,parms) {
  with(as.list(c(state,parms)), {
    beta=parms[1]; gamma=parms[2];
    S=state[1]; I=state[2];R=state[3];N=S+I+R
    dS= - beta*I*S/N
    dI= beta*I*S/N - gamma*I
    dR= gamma*I
    return(list(c(dS,dI,dR)))
  })
}

time.points<-seq(0,nrow(influenza_england_1978_school),0.01)
initial.condition<-c(S=762,I=1,R=0)
parameters<-c(beta=1.1,gamma=0.5)

as.data.frame(ode(initial.condition,time.points,SIR.model,parameters))>solution

##lets plot the model output on top of the data
ggplot(merge(solution_SIR,influenza_england_1978_school))+  

  geom_line(aes(x=time,y=I),color="darkgreen")+
  geom_point(aes(x=time,y=in_bed))
```



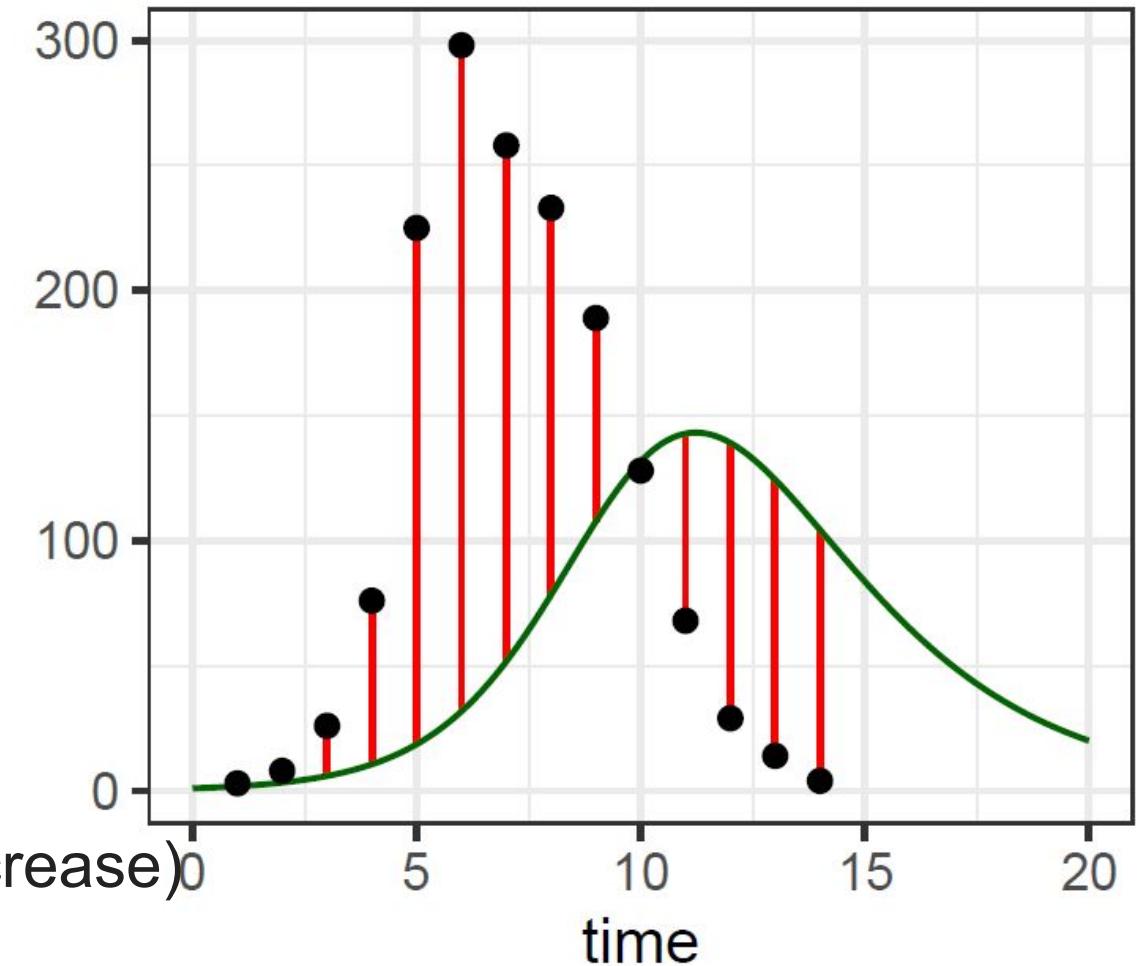
# Fitting the curve

Define error function between model output and data

$$\text{error}(\beta, \gamma) = \sum_{t=1}^{14} |I(t, \beta, \gamma) - d(t)|^2$$

$$\left\{ \begin{array}{l} \text{error}(\beta, \gamma) = \sum_{t=1}^{14} w_t |I(t, \beta, \gamma) - d(t)|^2 \\ \sum_{t=1}^{14} w_t = 1 \end{array} \right.$$

**weighted** error: certain time points in the data are more important (e.g. exponential increase)



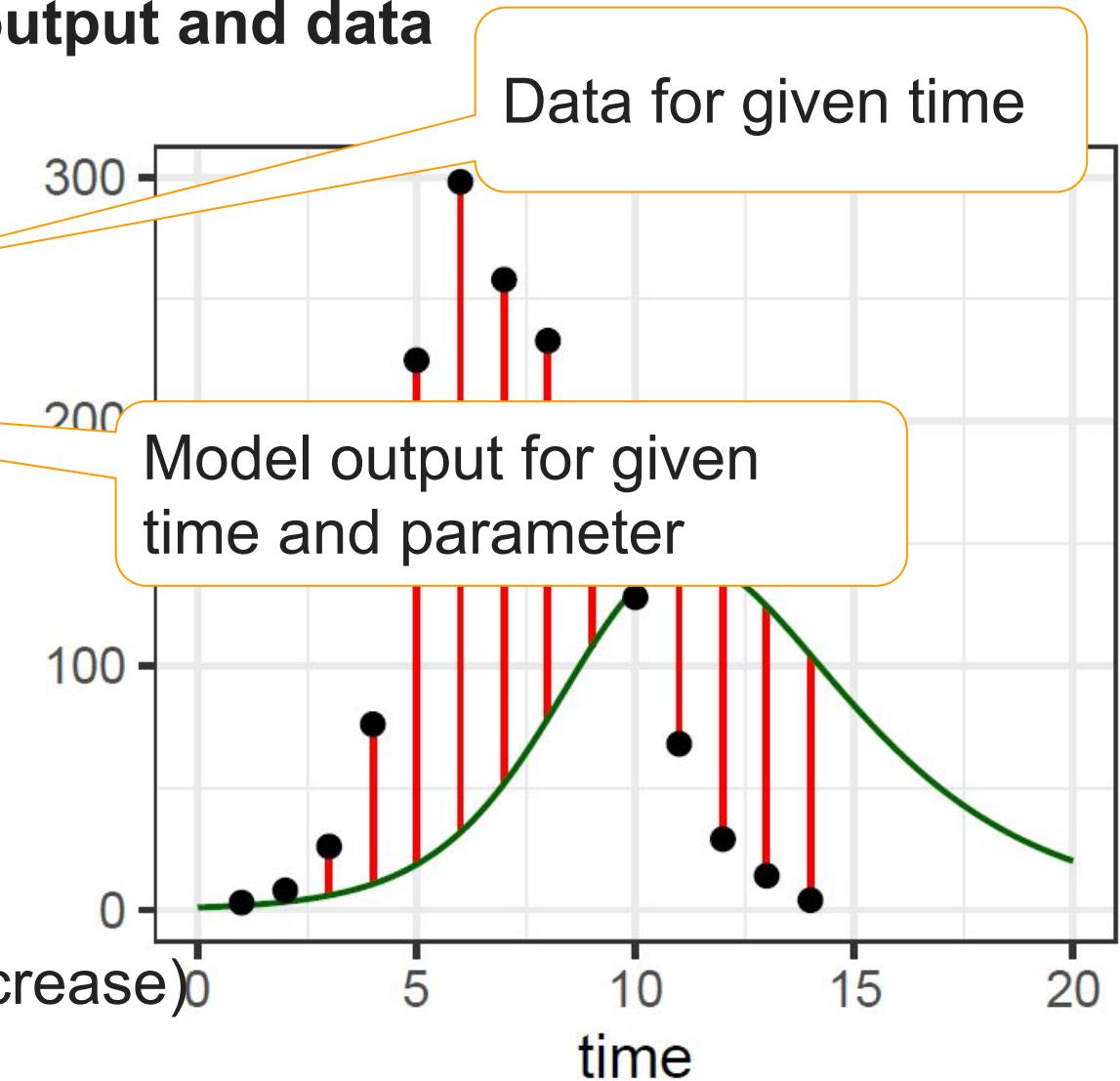
# Fitting the curve

Define error function between model output and data

$$\text{error}(\beta, \gamma) = \sum_{t=1}^{14} |I(t, \beta, \gamma) - d(t)|^2$$

$$\left\{ \begin{array}{l} \text{error}(\beta, \gamma) = \sum_{t=1}^{14} w_t |I(t, \beta, \gamma) - d(t)|^2 \\ \sum_{t=1}^{14} w_t = 1 \end{array} \right.$$

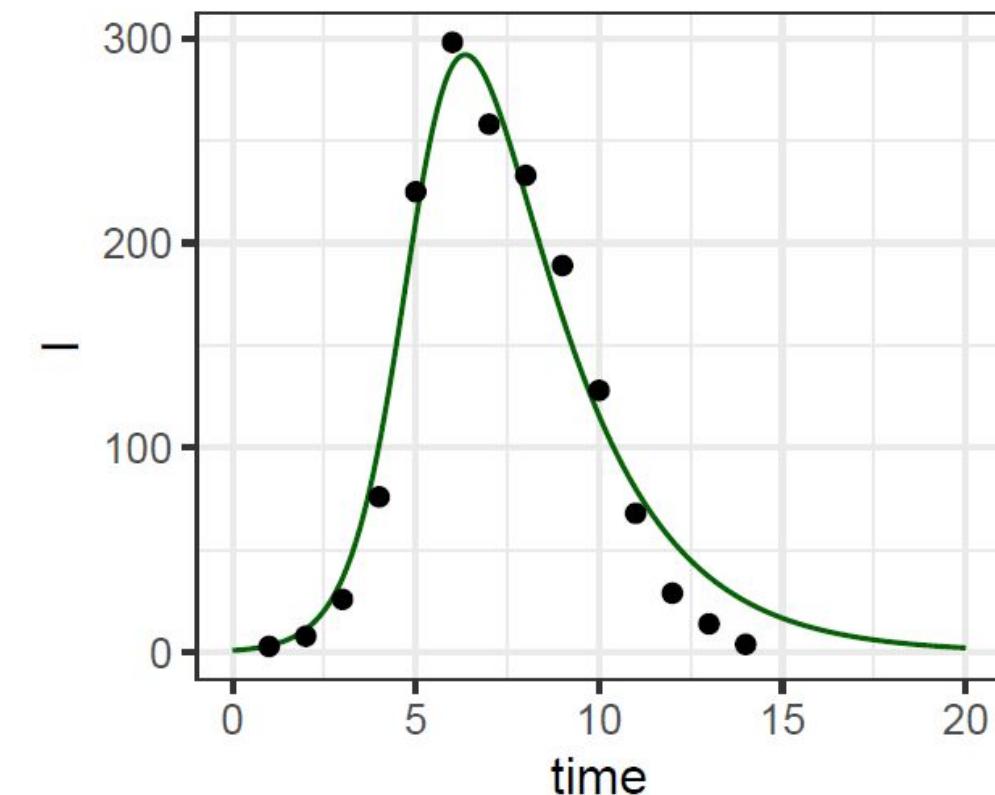
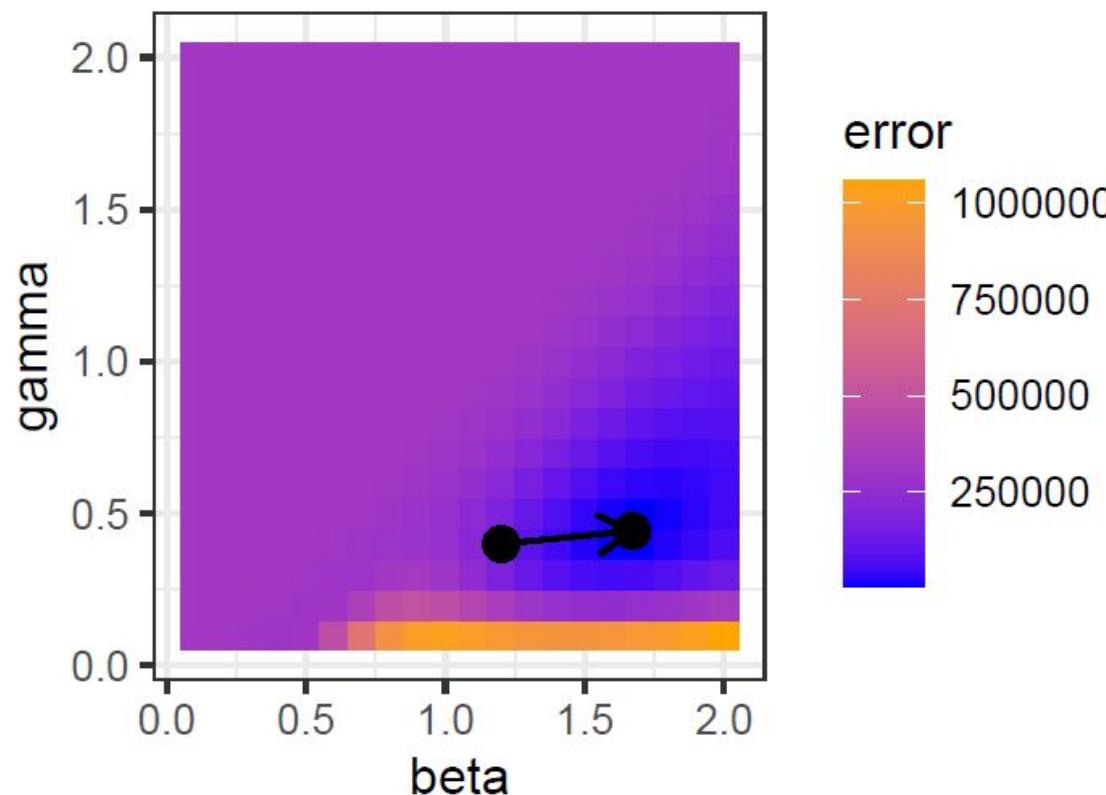
**weighted** error: certain time points in the data are more important (e.g. exponential increase)



# Fitting the curve

## Minimizing the error

- error function depends on model parameters
- “best” parameters are those for which error is minimal



# Fitting the curve

## Optimizing with Gauss-Newton

- residual  $r_i := d(i) - I(i, \beta)$
- least square is loss function  $S : \beta \mapsto \sum_{i=1}^{14} r_i^2$
- Gauss-Newton gradient equation:

$$\frac{\partial S}{\partial \beta} = 2 \sum_i r_i \frac{\partial r_i}{\partial \beta} = -2 \sum_i r_i \frac{\partial I(i, \beta)}{\partial \beta} = 0$$

- Taylor expansion at  $\beta^k$

$$I(i, \beta) = I(i, \beta^k) + \frac{\partial I(i, \beta^k)}{\partial \beta} (\beta - \beta^k)$$

- plug this into gradient equation, solve for  $\beta$  to obtain  $\beta^{k+1}$

# Fitting the curve

## Optimizing with Gauss-Newton

- residual  $r_i := d(i) - I(i, \beta)$
- least square is loss function  $S : \beta \mapsto \sum_{i=1}^{14} r_i^2$
- Gauss-Newton gradient equation:

$$\frac{\partial S}{\partial \beta} = 2 \sum_i r_i \frac{\partial r_i}{\partial \beta} = -2 \sum_i r_i \frac{\partial I(i, \beta)}{\partial \beta} = 0$$

- Taylor expansion at  $\beta^k$

$$I(i, \beta) = I(i, \beta^k) + \frac{\partial I(i, \beta^k)}{\partial \beta} (\beta - \beta^k)$$

Gauss-Newton will give you only a local minimum!

- plug this into gradient equation, solve for  $\beta$  to obtain  $\beta^{k+1}$

# Fitting the curve

## Optimizing with Gauss-Newton

$$\sum_i r_i(\beta) \frac{\partial I^i}{\partial \beta}(\beta) = 0$$

local minimum condition

$k = 0$  w.l.o.g. pick a start value  $\beta_0$

$$I(i, \beta) = I(i, \beta^k) + \frac{\partial I(i, \beta^k)}{\partial \beta} (\beta - \beta^k)$$

$$\sum_i \left\{ d_i - \left( I^i(\beta_0) + \frac{\partial I^i}{\partial \beta}(\beta_0)(\beta - \beta_0) \right) \right\} \frac{\partial I^i}{\partial \beta}(\beta) = 0 \quad J := \frac{\partial I}{\partial \beta}$$

$$J^T J \beta = J^T J \beta_0 - J^T r_i(\beta_0)$$

$$\beta = \beta_0 - (J^T J)^{-1} J^T r_i(\beta_0)$$

$$\beta_1 := \beta$$

iterate until gradient is  
very close to zero

# Fitting the curve

## Optimizing with Gauss-Newton in R

- use the function `optim` in the R package `stats`
- build a residual sum of squares function by running he ODE model and comparing its output to the data

```
##residual sum of squares
rss<-function(parameters) {
  time.points<-seq(0,20,0.1);
initial.condition<-c(S=762,I=1,R=0)
  ode(initial.condition,time.points,SIR.model,parameters) %>%
    as.data.frame->solution_SIR

solution_SIR<-merge(solution_SIR,influenza_england_1978_school)
  RSS<-sum((solution_SIR$I-solution_SIR$in_bed)^2)
  return(RSS)
}
```

# Fitting the curve

## Optimizing with Gauss-Newton in R

- assume  $\beta$  and  $\gamma$  are bounded between 0 and 2
- pick your favourite start value for  $\beta$  and  $\gamma$
- use the **optim** function with the “L-BFGS-B” method and provide **par**, **lower** and **upper** arguments

boundquasi-Newton method using also Hessianstarting valuebound

```
#optimize residual sum of squares over parameter space
start<-c(1.2, 0.4) #initial value
##run newton method for optimization
opt.param <- optim(par=start,rss,method = "L-BFGS-B",lower = c(0,0),upper = c(2,2),hessian = TRUE,control = list(parscale = c(10^-4,10^-4),factr=1))$par
```

# Fitting the curve

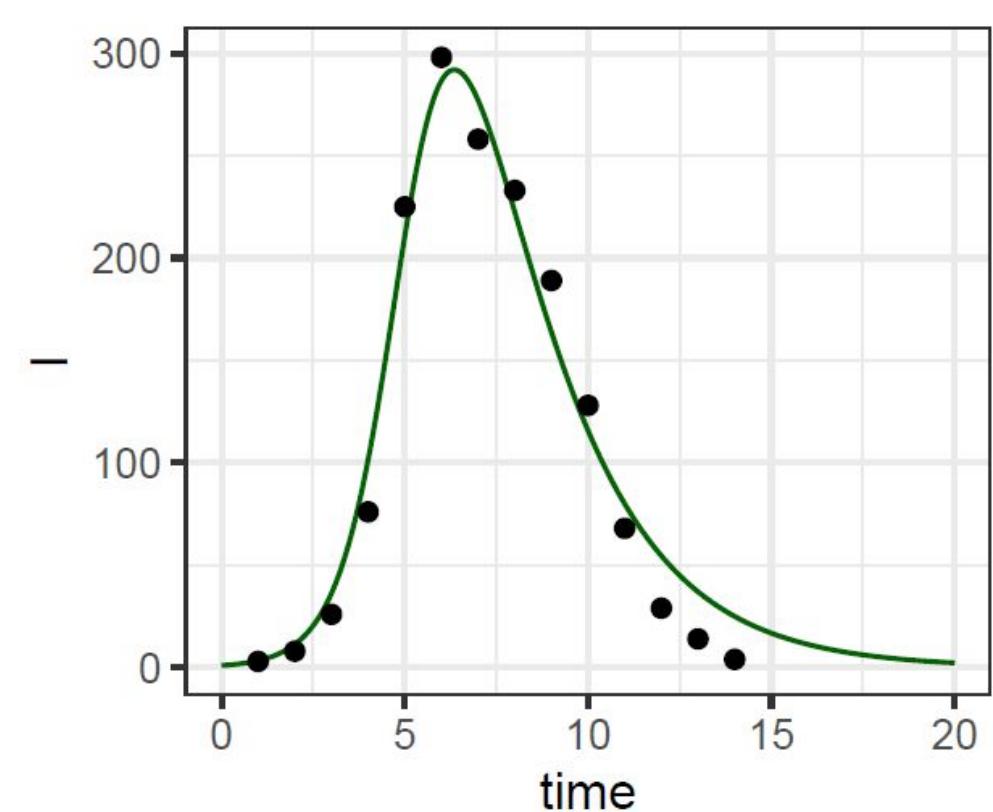
```
##let's simulate the optimal solution
ode(initial.condition,time.points,SIR.model,opt.param) %>%
  as.data.frame->solution_SIR_optim

##lets plot the model output on top of the data
ggplot(merge(solution_SIR_optim,influenza_england_1978_school))+
  geom_line(aes(x=time,y=I),color="darkgreen")+
  geom_point(aes(x=time,y=in_bed))
```

# Fitting the curve

```
##let's simulate the optimal solution
ode(initial.condition,time.points,SIR.model,opt.param) %>%
  as.data.frame->solution_SIR_optim

##lets plot the model output on top of
ggplot(merge(solution_SIR_optim,influenza),
  geom_line(aes(x=time,y=I),color="darkgreen"),
  geom_point(aes(x=time,y=in_bed)) )
```



# Is fitting the curve sufficient for the model to be informative?

Possible issues:

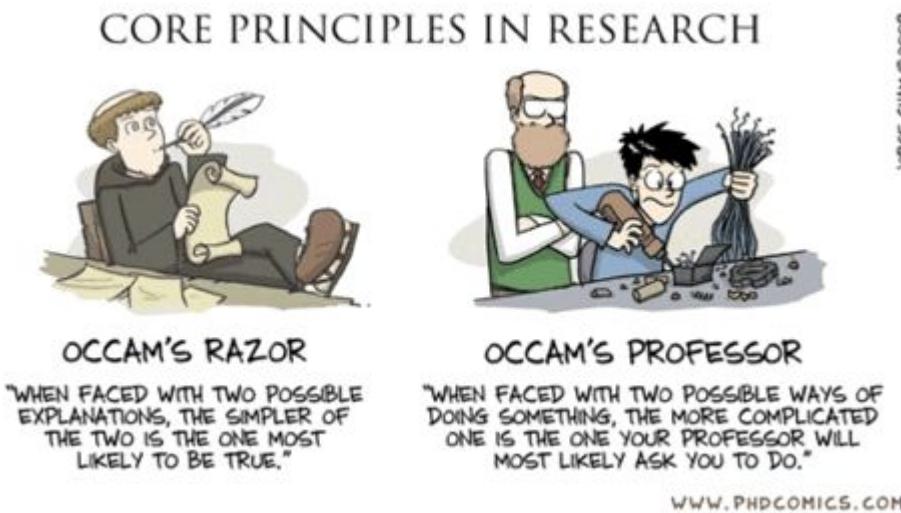
- influenza incubation period, spatial structure of school: **model misspecification**
- **prior knowledge** about clearance rate: narrow down bounds
- data recording has **measurement errors**: need statistical estimator
- **uncertainty** of parameters : credible intervals
- optimal solution not unique : **identifiability**

# What is a good model?

## Parsimony: explain data with model with as few as possible parameters

- Occam's razor:

*"It is futile to do with more things that which can be done with fewer."*



- which model should we prefer?
- model selection with Akaike Information Criterion (AIC):

$$AIC = \text{number of parameters} - \log \text{maximum likelihood}$$

- model with smaller AIC is the better model

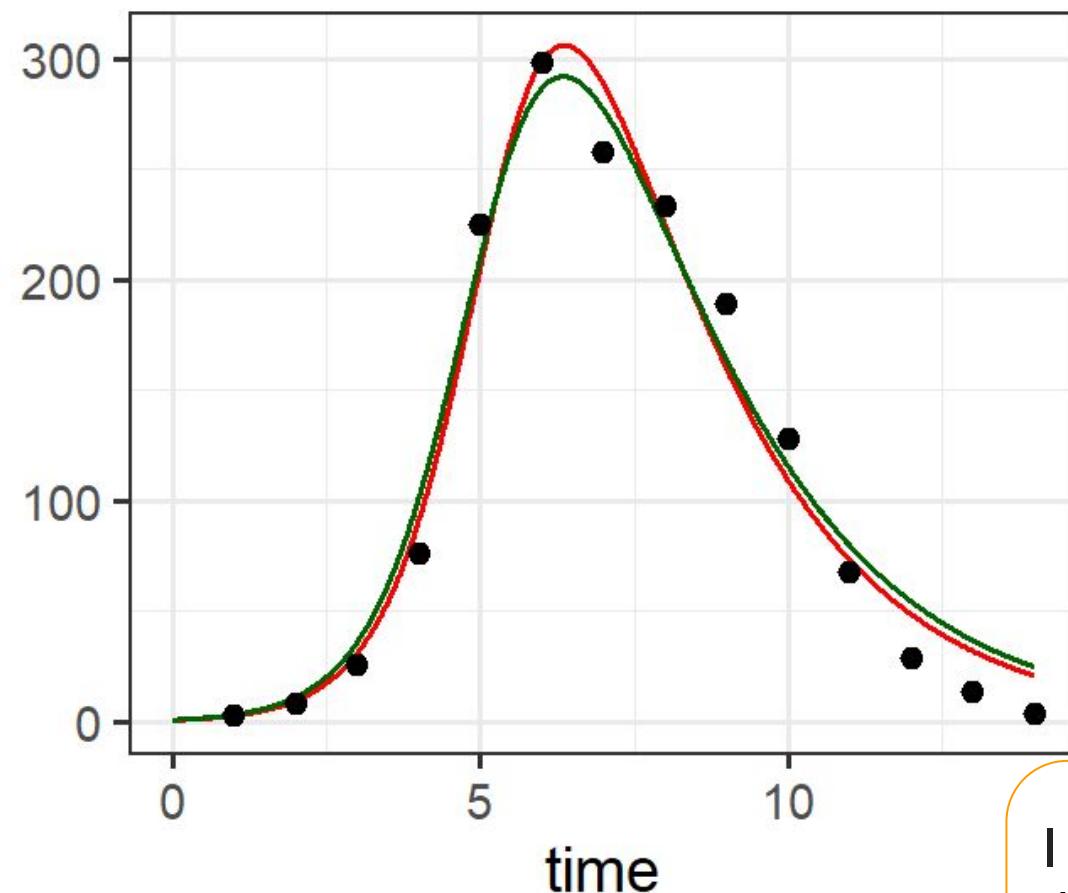
# What is a good model?

- add an exposed but not yet infectious compartment E to the SIR model
- write  $\alpha$  for the inverse incubation period (the rate from E to I )
- find minimal residual sum of squares solution for parameters  $\alpha, \beta, \gamma$  w.r.t. influenza data
- would you recommend Occam to rather use the SEIR model?



William of Occam, XIVth century

# What is a good model?



model

- SEIR
- SIR

AIC for SIR

$$2 - \log(rss(opt.param)) = -6.3$$

$$3 - \log(rss1(opt.param1)) = -5.3$$

AIC for SEIR

I would **choose the SIR model**,  
the additional parameter for SEIR  
does not justify the marginal  
decrease in error!



William of Occam, XIVth century

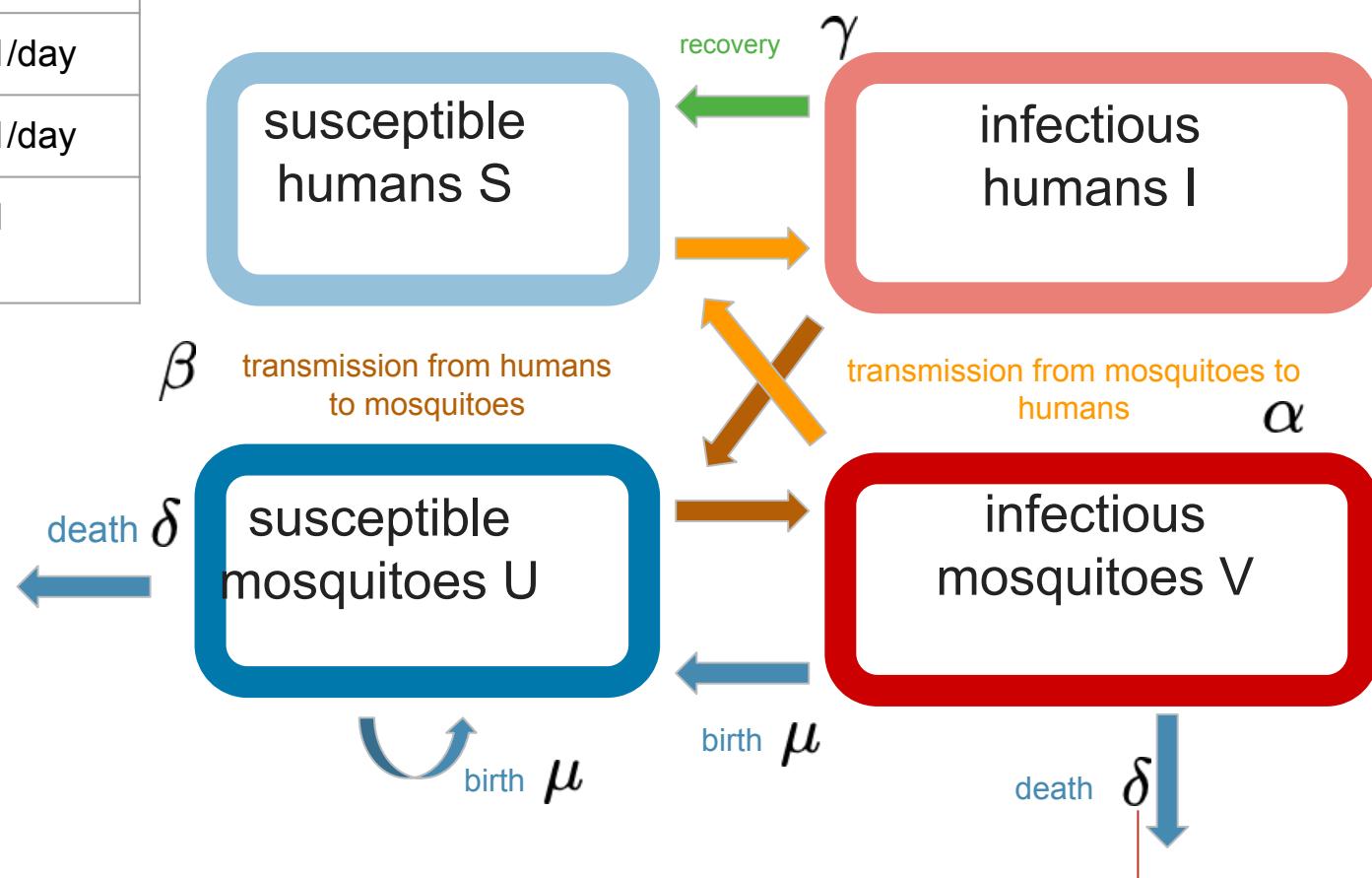
# Fitting the curve for the Ross-McDonald model

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	?	1
delta	mosquito mortality rate	0.13	1/day
mu	per capita mosquito birth rate	0.13	1/day
beta	probability of transmission to mosquitoes	?	1

prevalence at endemic equilibrium=0.35

alpha\_2=?

beta=?



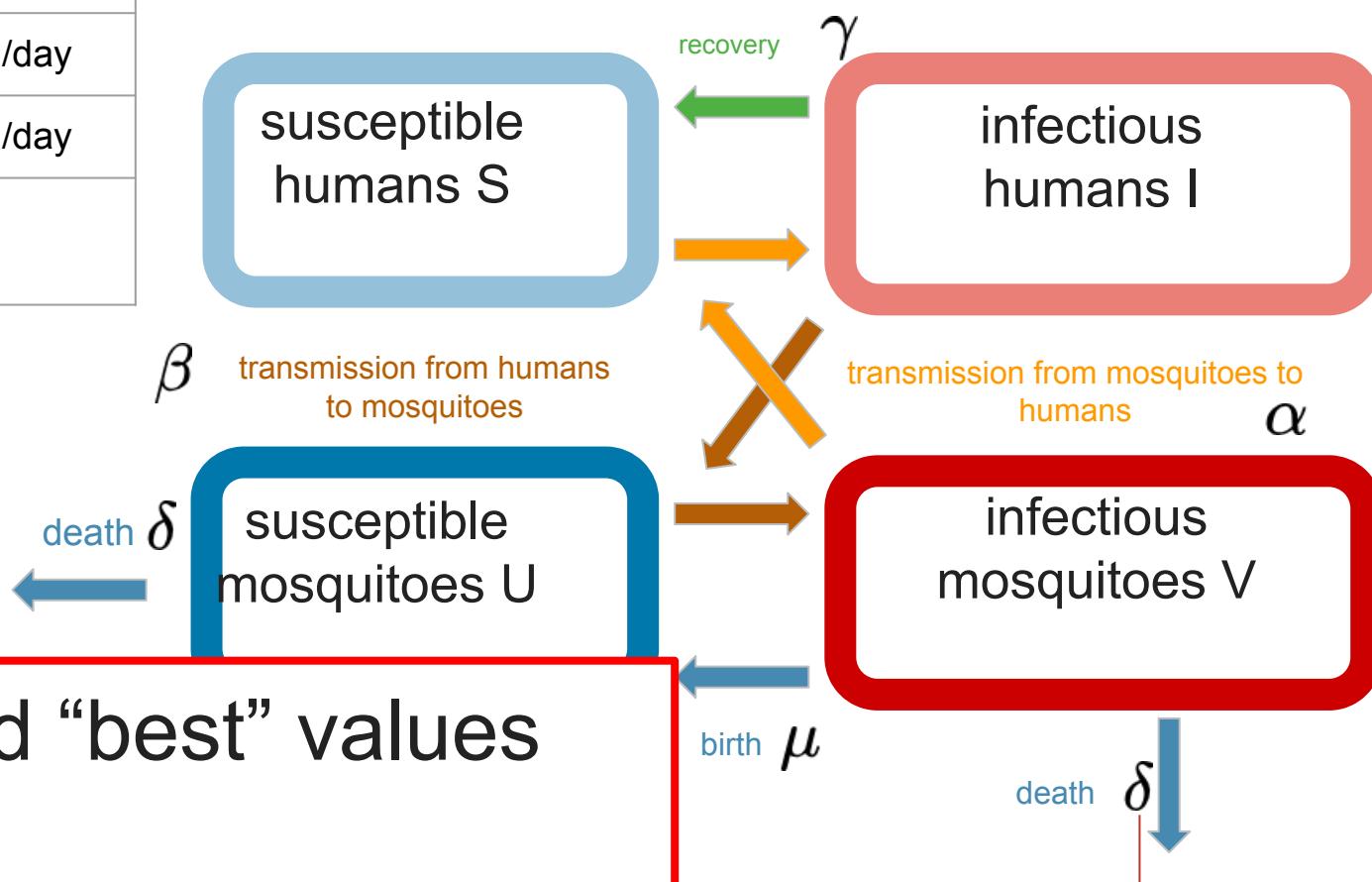
# Fitting the curve for the Ross-McDonald model

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	?	1
delta	mosquito mortality rate	0.13	1/day
mu	per capita mosquito birth rate	0.13	1/day
beta	probability of transmission to mosquitoes	?	1

prevalence at endemic equilibrium=0.35

alpha\_2=?

beta=?



Sw  
Use optim in R and find “best” values for alpha\_2 & beta!

# Fitting the curve for the Ross-McDonald model

parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	?	1

```
c H=5000; I0=1; V0=8; VectorHumanRatio=5; finalT=8*365
r x0 <-c(S=H-I0, I=I0, U=H*VectorHumanRatio-V0, V=V0)
b time.points<-seq(0,finalT,1)

RSS.RM<-function(parameters) {
  parms<-c(alpha=0.5*parameters[1], gamma=1/285, beta=parameters[2],
  mu=0.13, delta=0.13)
  ...
  return(rss)
}
```

# Fitting the curve for the Ross-McDonald model

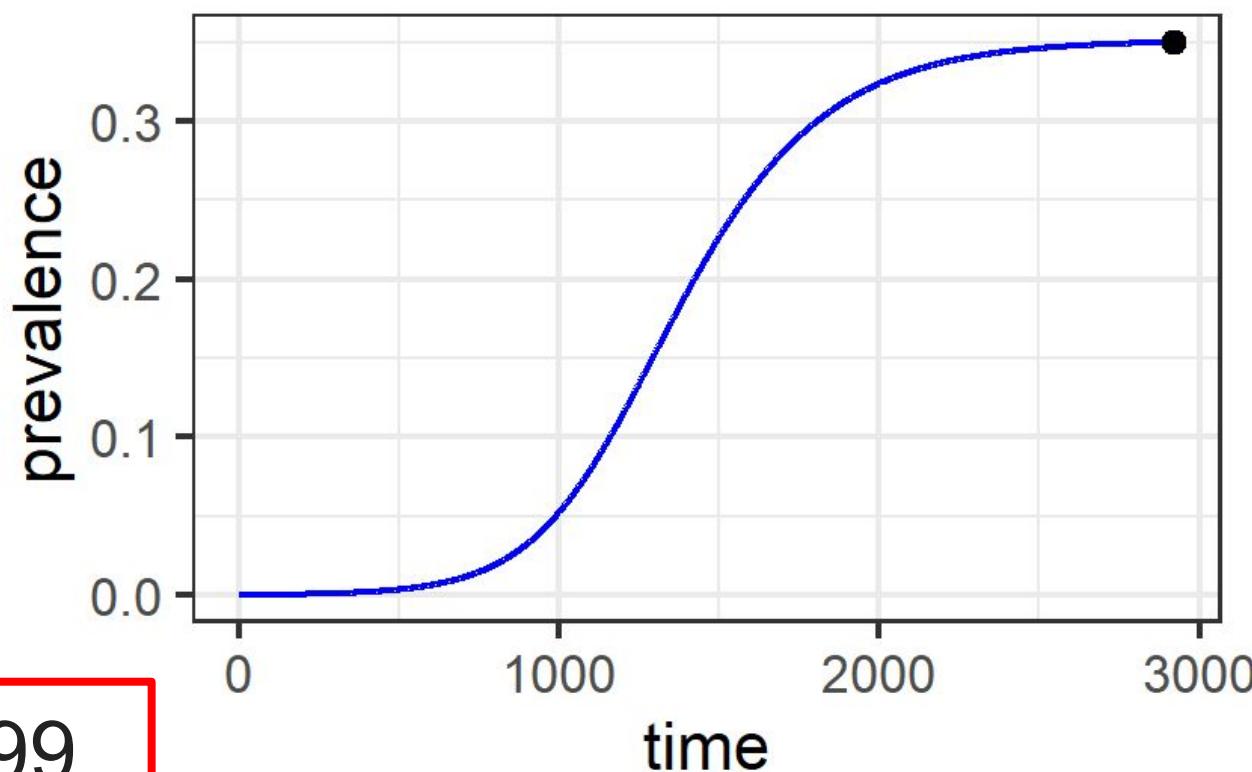
parameter	description	value	unit
gamma	reciprocal of untreated infection duration	1/285	1/day
alpha_1	biting rate within gonotrophic cycle	0.5	1/day
alpha_2	probability of transmission to humans	?	1
delta	mosquito mortality rate	0.13	1/day
mu	per capita mosquito birth rate	0.13	1/day
beta	probability of transmission to mosquitoes	?	1

prevalence at endemic equilibrium=0.35

alpha\_2=?

beta=?

alpha\_2=0.001697, beta=0.2999



## Key takeaway points:

---

- curve fitting assumes that **observations are without measurement error**, and that consecutive observations are **independent**
- use optimization to identify **optimal parameter sets** such that the residual **error** between observations and model outputs **minimal**
- **parsimony** is an epistemological method: “of two competing theories, the simpler one is preferred”
- **parsimony** is **quantified** by information criteria: how much better is the data explained by a model if we add additional parameters to it?



## 5 - Statistical inference

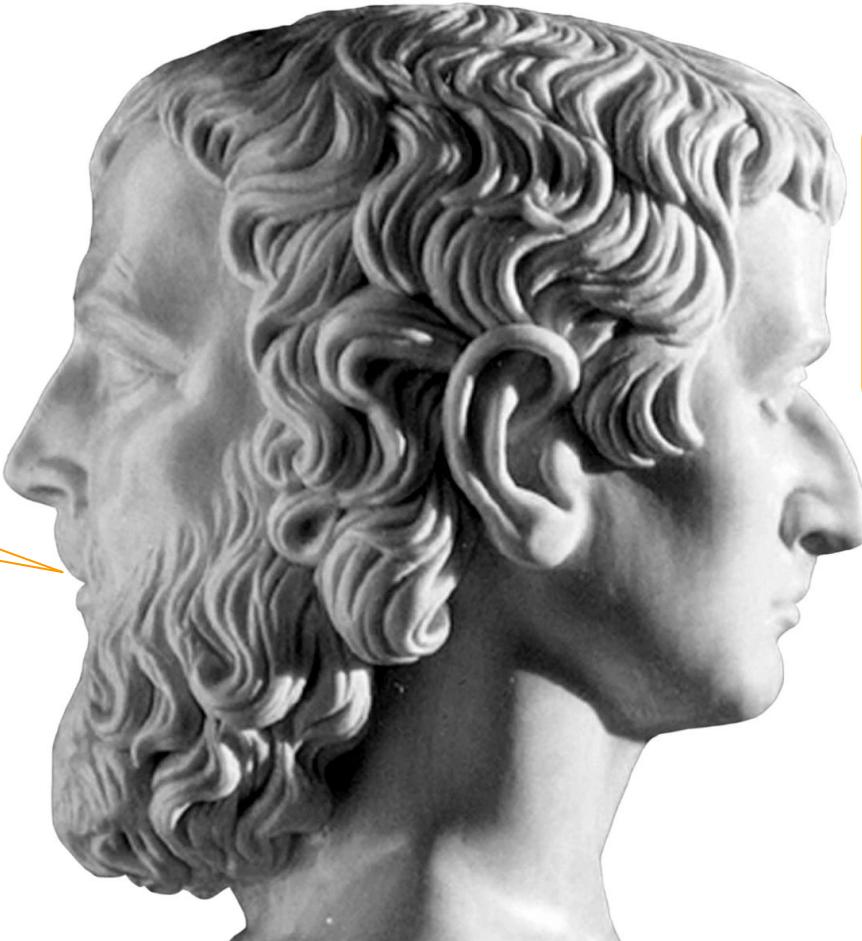
# Three paradigms of statistical inference

the model is fixed,  
the data is random

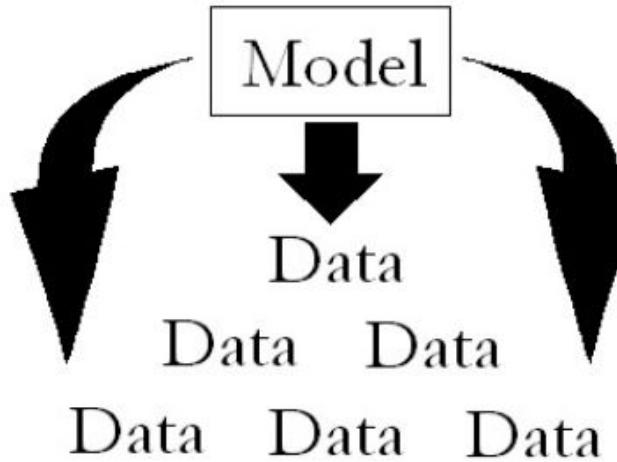
**frequentist**

the model is random,  
the data is fixed

**Bayesian**

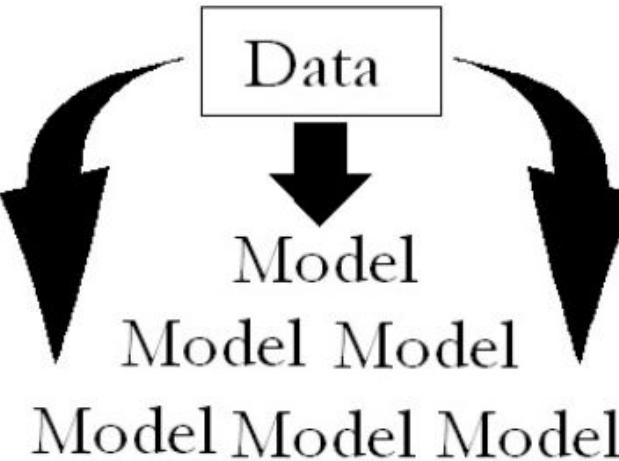


# Three paradigms of statistical inference



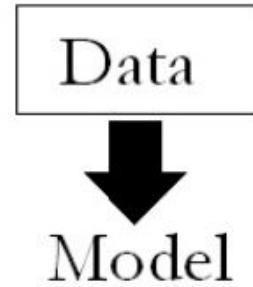
## frequentist

Can we **reproduce**  
**data** by sampling from  
statistical population for  
a **given model**  
**parameter**?



## Bayesian

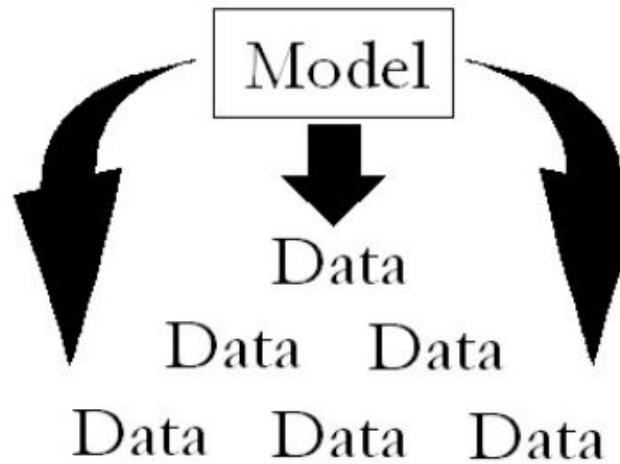
Given the data, what is  
the **most plausible**  
probability distribution  
of **model parameters**?



## likelihoodist

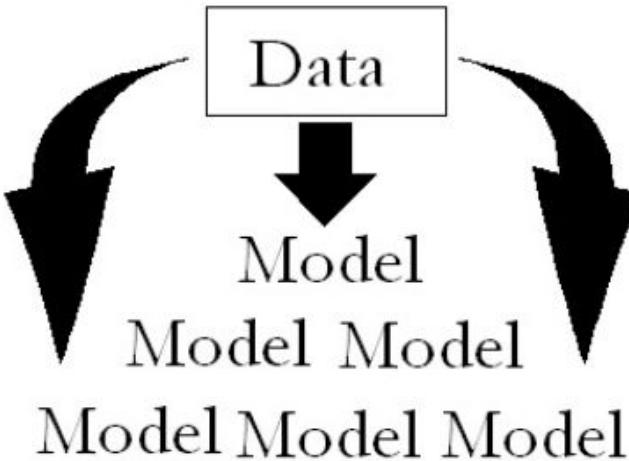
Measure **evidence**  
from data by likelihood  
function!

# Three paradigms of statistical inference



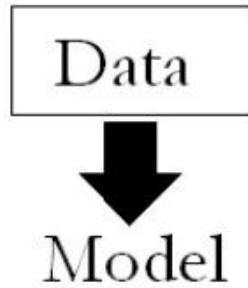
## frequentist

Can we **reproduce**  
**data** by sampling from  
statistical population for  
a **given model**  
**parameter**?



## Bayesian

Given the data, what is  
the **most plausible**  
probability distribution  
of **model parameters**?

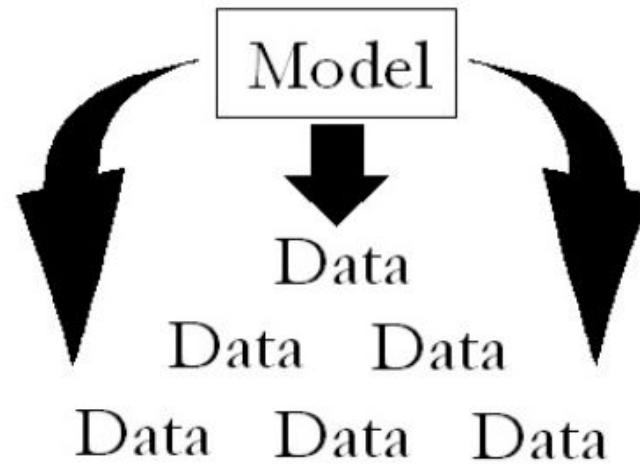


## likelihoodist

Measure **evidence**  
from data by likelihood  
function!

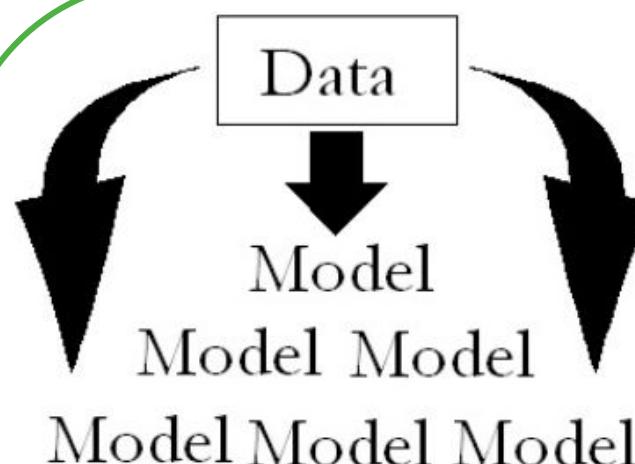
# Three paradigms of statistical inference

parameter inference for  
disease transmission models



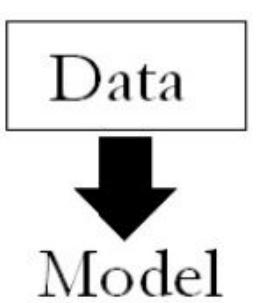
## frequentist

Can we **reproduce**  
**data** by sampling from  
statistical population for  
a **given model**  
**parameter**?



## Bayesian

Given the data, what is  
the **most plausible**  
probability distribution  
of **model parameters**?



## likelihoodist

Measure **evidence**  
from data by likelihood  
function!

*posterior distribution*

*maximum likelihood*

# Maximum likelihood

- **data** is considered **random sample** from unknown statistical population
- **maximum likelihood**: find probability distributions which is most likely to generate such samples
- parameterized distribution  $f_\theta(\cdot)$  for  $\theta$  in parameter space  $\Theta$
- observed data sample  $X \in \mathbb{R}^d$ : **likelihood**  $L : \theta \mapsto f_\theta(X)$
- **goal: find**  $\hat{\theta}(X) = \operatorname{argmax}_\theta L(\theta)$
- $\hat{\theta}$  maximum likelihood estimator which is random variable  $\hat{\theta} : \mathbb{R}^d \rightarrow \Theta$

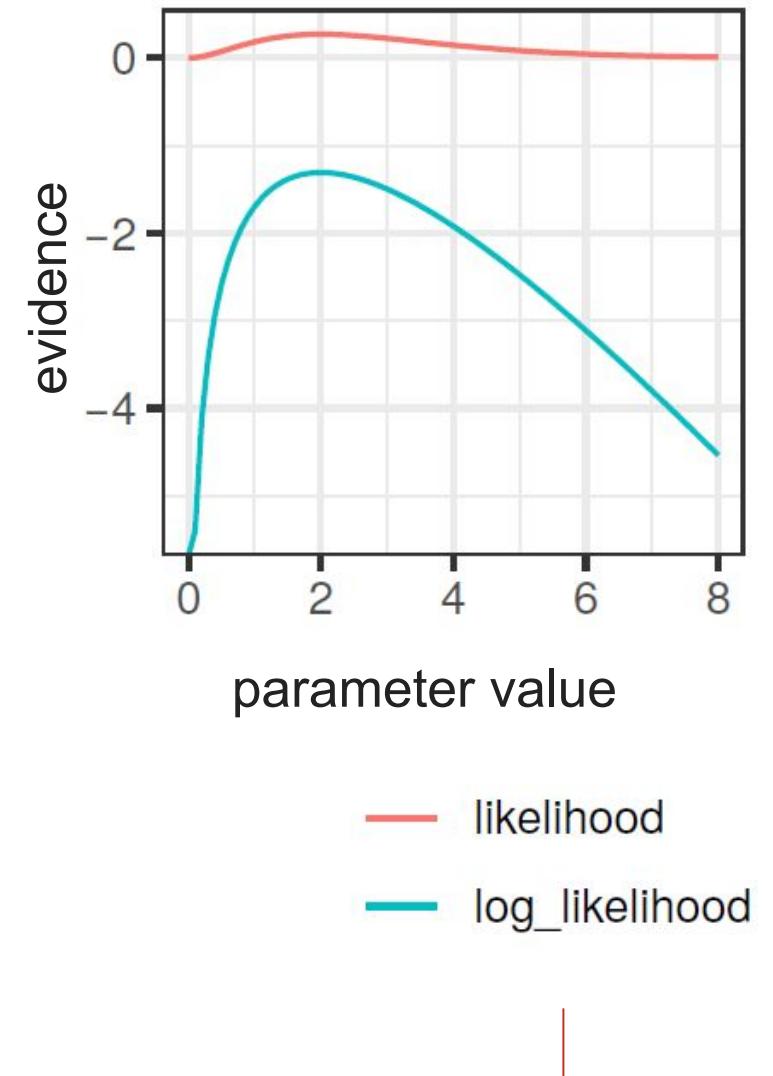
# Maximum likelihood

- likelihood is often log-transformed

$$\ell = \log L$$

- solve  $\frac{\partial \ell}{\partial \theta} = 0$

- check whether Hessian  $\frac{\partial^2 \ell}{\partial \theta^2}$  is negative definite, concave



# Maximum likelihood

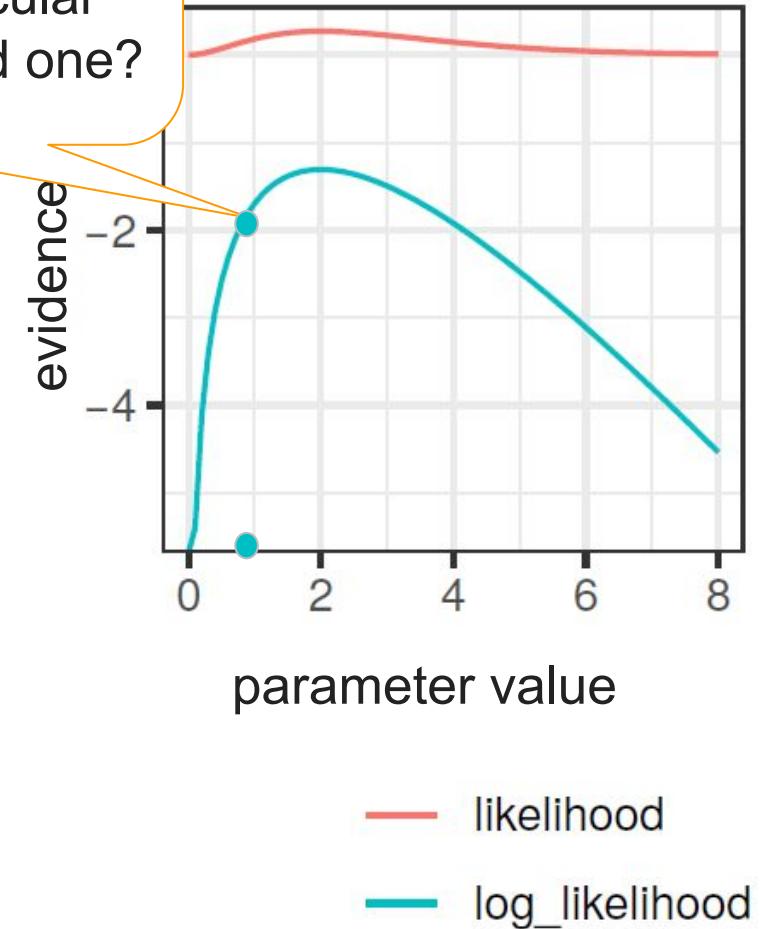
- likelihood is often log-transformed

$$\ell = \log L$$

- solve  $\frac{\partial \ell}{\partial \theta} = 0$

- check whether Hessian  $\frac{\partial^2 \ell}{\partial \theta^2}$  is negative definite, concave

Given data, what is the evidence that our model evaluated for this particular parameter to be the good one?



# Maximum likelihood

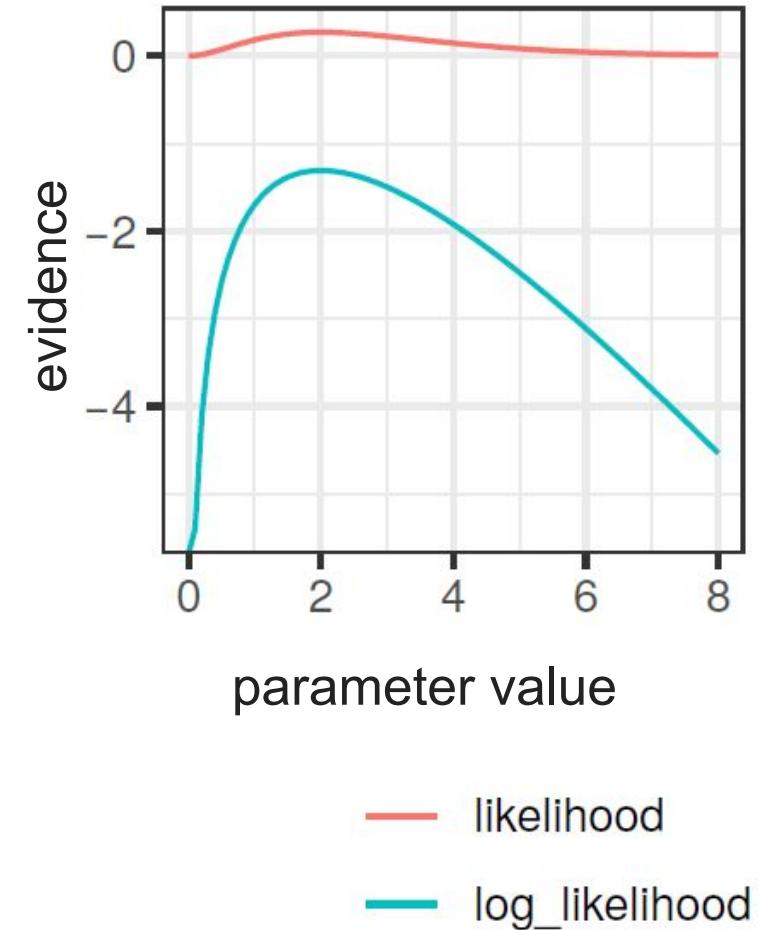
- likelihood is often log-transformed

$$\ell = \log L$$

- solve  $\frac{\partial \ell}{\partial \theta} = 0$

- check whether Hessian  $\frac{\partial^2 \ell}{\partial \theta^2}$  is negative definite, concave

- When does **minimizing the residual sum of squares** correspond to **maximum likelihood**?



# Maximum likelihood

- likelihood is often log-transformed

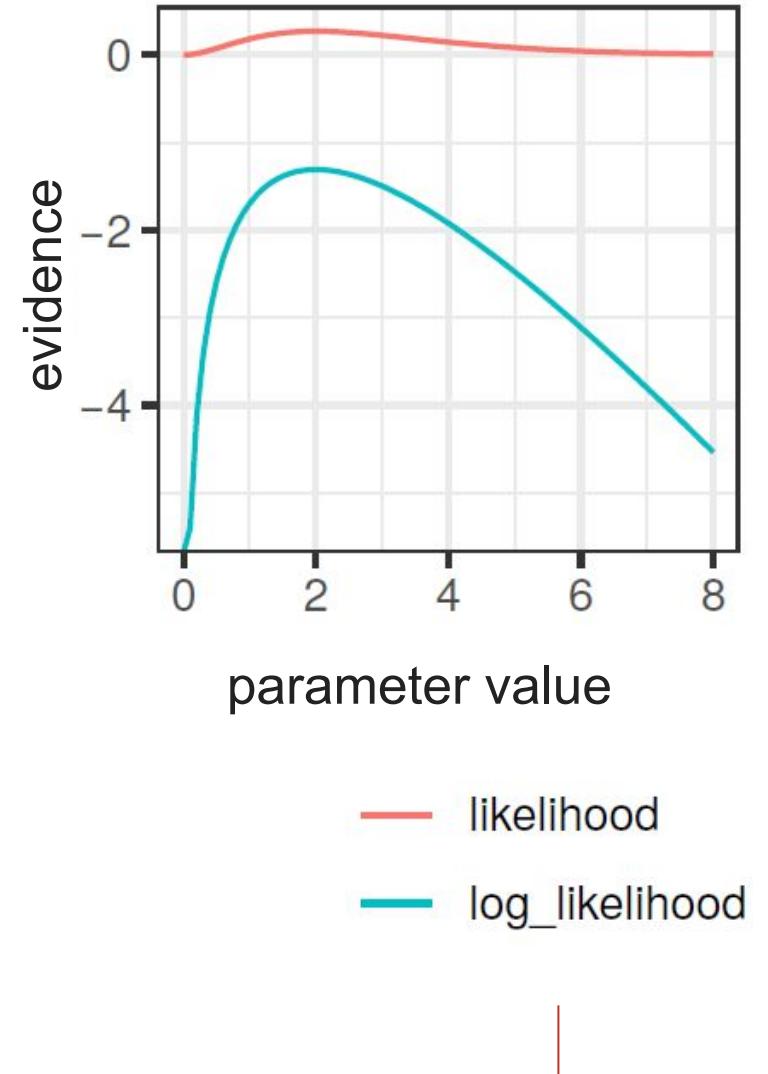
$$\ell = \log L$$

- solve  $\frac{\partial \ell}{\partial \theta} = 0$

- check whether Hessian  $\frac{\partial^2 \ell}{\partial \theta^2}$  is negative definite, concave

- When does **minimizing the residual sum of squares** correspond to **maximum likelihood**?

Only if  $f_\theta(\cdot)$  is Gaussian!



# Maximum log-likelihood estimator

- maximum log likelihood estimator  $\hat{\theta}$  depends on sample of size d

- **consistent** estimator:

with more samples, estimator converges towards true value:

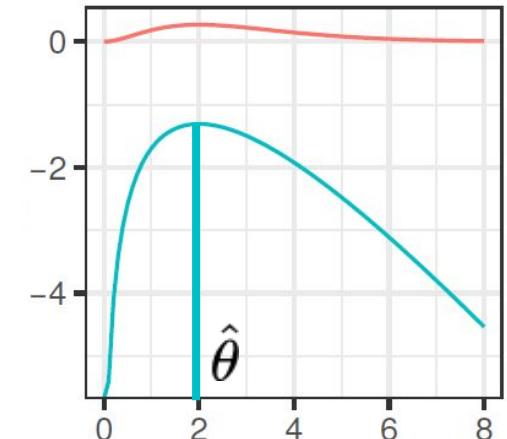
$$\forall \epsilon > 0 : \lim_{d \rightarrow \infty} \mathbb{P}(|\hat{\theta}^d - \theta_0| > \epsilon) = 0$$

- **efficient** estimator:  $\text{Var}(\hat{\theta}(X))$  as small as possible

- **Cramér-Rao bound**:  $\text{Var}(\hat{\theta}(X)) \geq \left[ -d \mathbb{E} \left( \frac{\partial^2 \ell}{\partial \theta^2}(\theta, X|\theta) \right) \right]^{-1}$

- **curvature**:  $\text{Var} \left( \frac{\partial}{\partial \theta} \log f_\theta \right) = -\mathbb{E} \frac{\partial^2}{\partial \theta^2} \log f_\theta$

Fisher information



# Maximum log-likelihood estimator

- maximum log likelihood estimator  $\hat{\theta}$  depends on sample of size d
- **consistent** estimator:

with more samples, estimator converges towards true value:

$$\forall \epsilon > 0 : \lim_{d \rightarrow \infty} \mathbb{P}(|\hat{\theta}^d - \theta_0| > \epsilon) = 0$$

Fisher information

- **efficient** estimator:  $\text{Var}(\hat{\theta}(X))$  as small as possible

- **Cramér-Rao bound**:  $\text{Var}(\hat{\theta}(X)) \geq \left[ -d \mathbb{E} \left( \frac{\partial^2 \ell}{\partial \theta^2}(\theta, X|\theta) \right) \right]^{-1}$

- **curvature**:  $\text{Var} \left( \frac{\partial}{\partial \theta} \log f_\theta \right) = -\mathbb{E} \frac{\partial^2}{\partial \theta^2} \log f_\theta$

“variance of the score is the curvature of the log-likelihood function”



# Maximum log-likelihood estimator - example

- Gaussian likelihood with parameter  $\theta$

$$f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) \quad \sigma \text{ fixed}$$

- observations:  $X_1=7, X_2=8$
- calculate maximum likelihood estimator  $\hat{\theta}$



# Maximum log-likelihood estimator - example

- Gaussian likelihood with parameter  $\theta$

$$f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) \quad \sigma \text{ fixed}$$

- observations:  $X_1=7, X_2=8$
- calculate maximum likelihood estimator  $\hat{\theta}$  based on the observation

$$f_{\theta}((7, 8)) = \mathbb{P}(7, 8|\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(7-\theta)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(8-\theta)^2}{2\sigma^2}\right)$$

$$\ell(\theta) = 2 \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(7-\theta)^2}{2\sigma^2} - \frac{(8-\theta)^2}{2\sigma^2} \quad \frac{\partial\ell}{\partial\theta} = \frac{1}{2\sigma^2}(7 + 8 - 2\theta) = 0$$
$$\hat{\theta} = \frac{7+8}{2}$$



# Maximum log-likelihood estimator - example

- Gaussian likelihood with parameter  $\theta$

$$f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) \quad \sigma \text{ fixed}$$

- observations:  $X_1=7, X_2=8$

maximum log-likelihood estimator

- calculate maximum likelihood estimator  $\hat{\theta}$  based on the observation

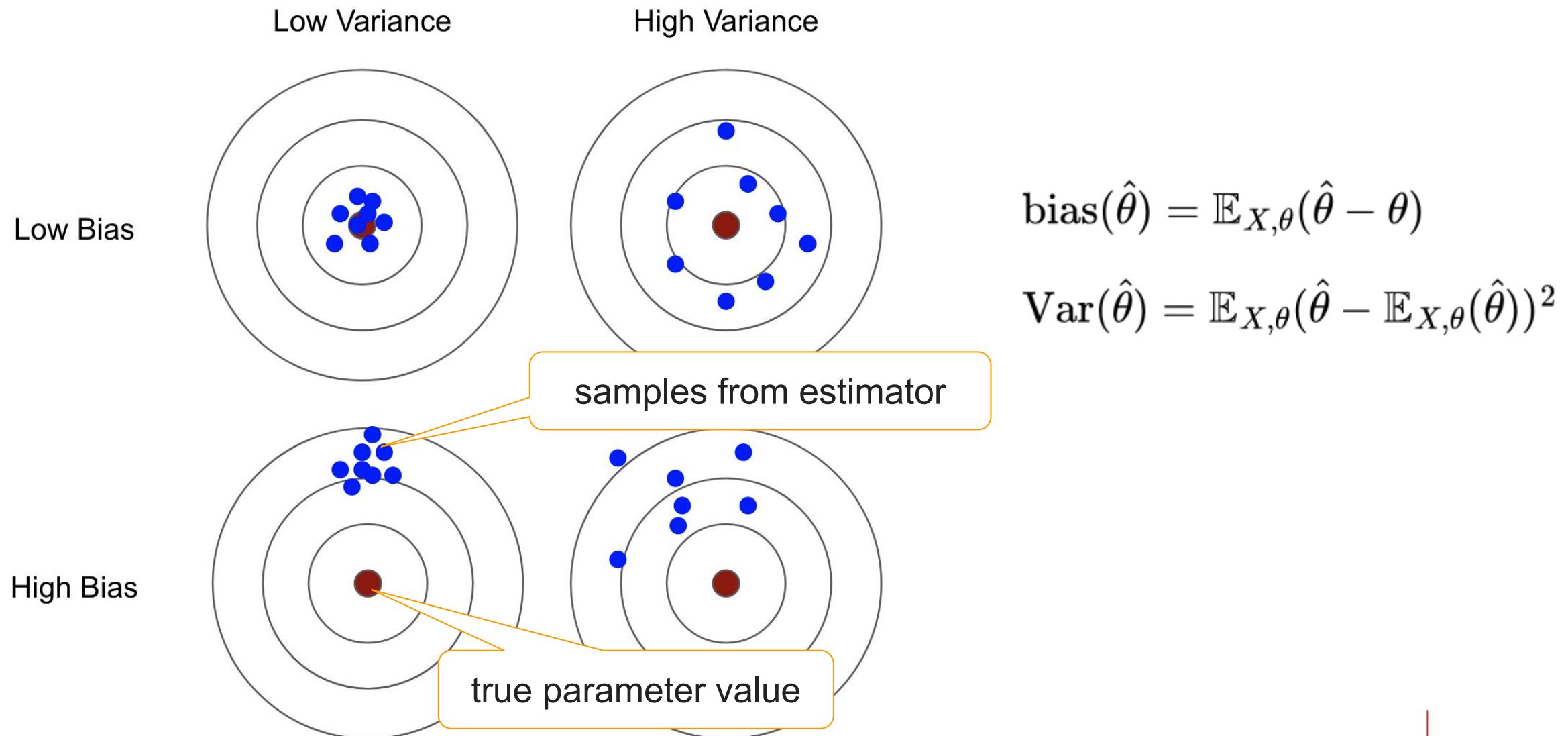
$$f_{\theta}((7, 8)) = \mathbb{P}(7, 8|\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(7-\theta)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(8-\theta)^2}{2\sigma^2}\right)$$

$$\ell(\theta) = 2 \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(7-\theta)^2}{2\sigma^2} - \frac{(8-\theta)^2}{2\sigma^2}$$

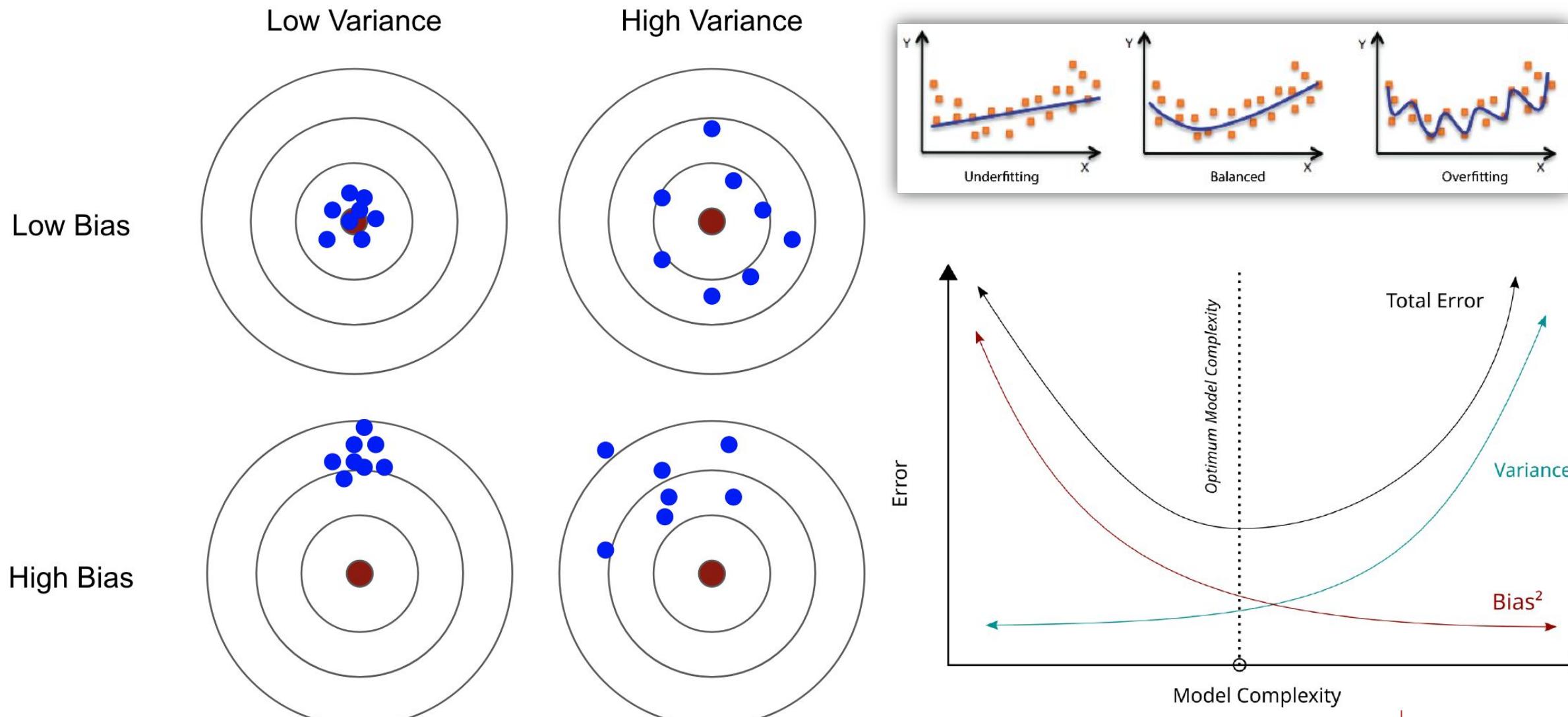
$$\frac{\partial \ell}{\partial \theta} = \frac{1}{2\sigma^2}(7 + 8 - 2\theta) = 0$$

$$\hat{\theta} = \frac{7+8}{2}$$

# Variance-bias trade-off for estimators



# Variance-bias trade-off for estimators



# Maximum log-likelihood for prevalence

Let's go back to our SIR model

$$X = [X_1, \dots, X_{14}]$$

$$\theta = (\beta, \gamma)$$

observations of hosts in bed

model parameters

$$\frac{dS}{dt} = -\beta I \frac{S}{N}$$

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

$$\frac{dR}{dt} = \gamma I$$

$$I(1, \theta), \dots, I(14, \theta)$$

numerical solutions  
for given parameter

# Maximum log-likelihood for prevalence

Let's go back to our SIR model

$$X = [X_1, \dots, X_{14}]$$

$$\theta = (\beta, \gamma)$$

observations of hosts in bed

model parameters

$$f_\theta(X) = \sum_{i=1}^{14} \binom{N}{X_i} \left( \frac{I(i, \theta)}{N} \right)^{X_i} \left( 1 - \frac{I(i, \theta)}{N} \right)^{N-X_i}$$

binomial likelihood function  
with N=763

$$\frac{dS}{dt} = -\beta I \frac{S}{N}$$

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

$$\frac{dR}{dt} = \gamma I$$

$$I(1, \theta), \dots, I(14, \theta)$$

numerical solutions  
for given parameter



AIDM\_04.R

Use `optim` in R to find **maximum likelihood estimator** for this particular data!

# Maximum log-likelihood for prevalence

$$f_{\theta}(X) = \sum_{i=1}^{14} \binom{N}{X_i} \left( \frac{I(i,\theta)}{N} \right)^{X_i} \left( 1 - \frac{I(i,\theta)}{N} \right)^{N-X_i}$$

```
####log likelihood for prevalence in SIR model
loglikelihood_binom <- function(parameters =c(beta =1.1, gamma=1/3)) {
  ## simulate prevalence , per time point , calculate binomial log likelihood

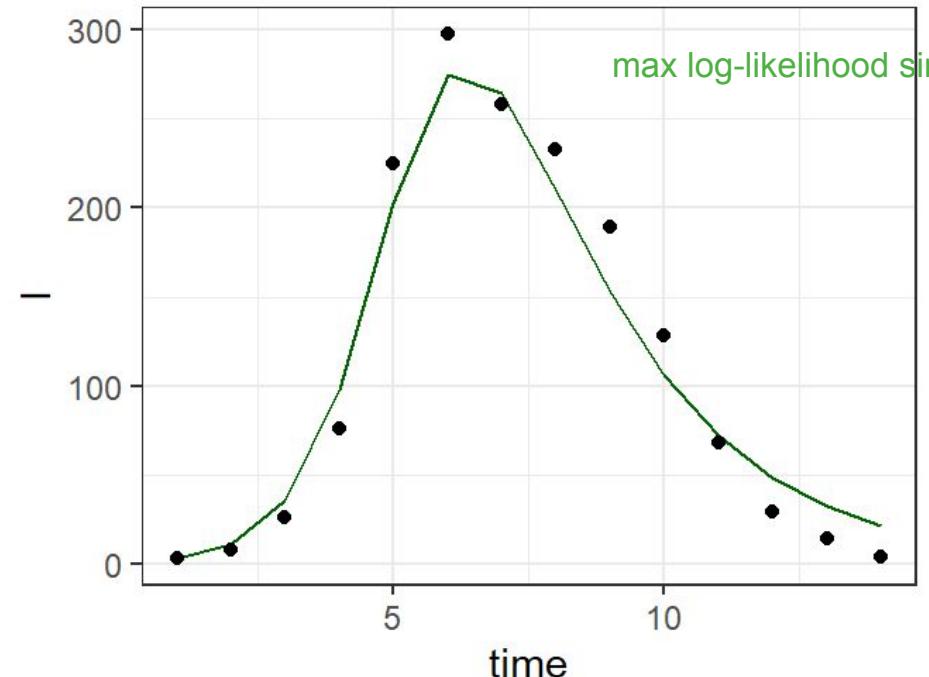
  as.data.frame(ode(initial.condition,time.points,SIR.model,parameters))->solution_SIR
  merge(solution_SIR,
        influenza_england_1978_school) %>%
    mutate(beta=parameters[1],gamma=parameters[2]) %>%
    mutate(loglikelihood=dbinom(in_bed,size=763,prob=I/763,log=T)) %>%
    summarize(loglikelihood=sum(loglikelihood))->df
  return(df$loglikelihood)
}
```

# Maximum log-likelihood for prevalence

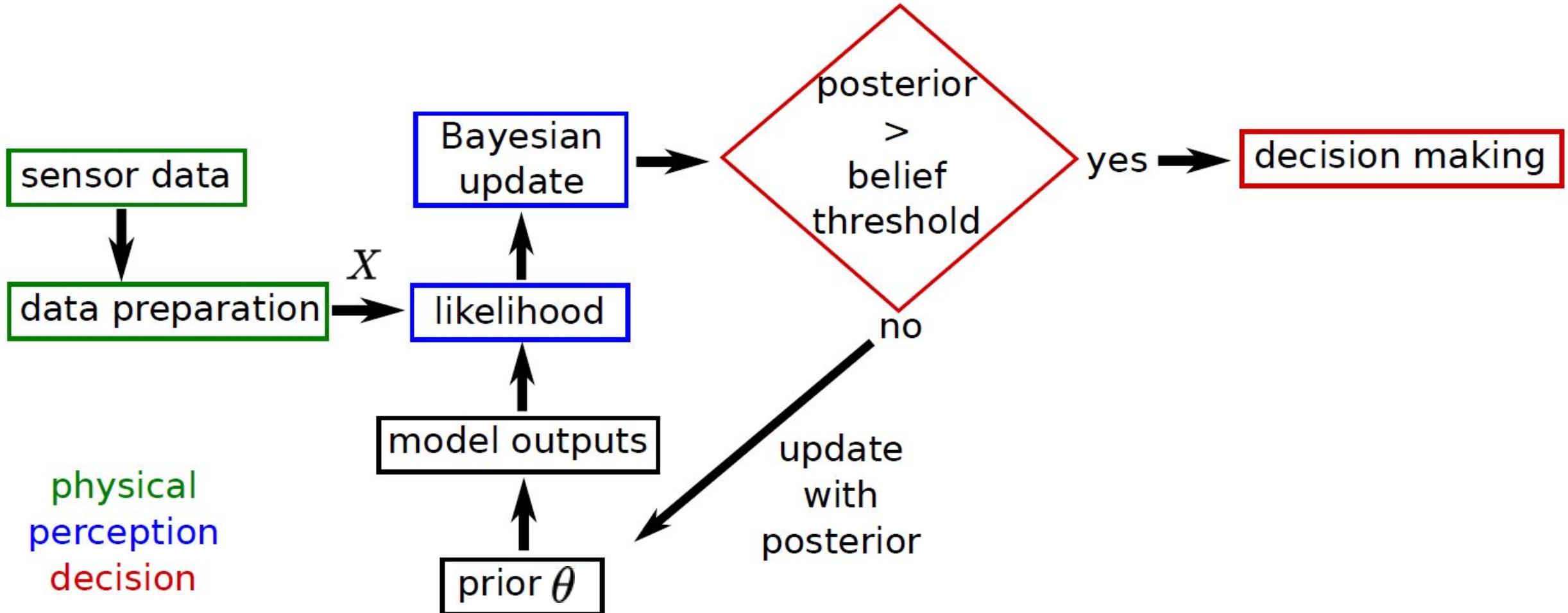
$$f_{\theta}(X) = \sum_{i=1}^{14} \binom{N}{X_i} \left( \frac{I(i, \theta)}{N} \right)^{X_i} \left( 1 - \frac{I(i, \theta)}{N} \right)^{N-X_i}$$

```
####log likelihood for prevalence in SIR model
loglikelihood_binom <- function(parameters = c(beta =1.1, gamma = 0.1), N = 763, in_bed = 14)
  ## simulate prevalence , per time point , calculate log likelihood
```

```
as.data.frame(ode(initial.condition, time.points, SIR.model, parameters)) -> solution_SIR
merge(solution_SIR,
      influenza_england_1978_school) %>%
  mutate(beta=parameters[1], gamma=parameters[2]) %>%
  mutate(loglikelihood=dbinom(in_bed, size=N, prob=beta/N, log=T)) %>%
  summarize(loglikelihood=sum(loglikelihood)) -> df
return(df$loglikelihood)
}
```



# Bayesian inference system



# Bayesian updating



the model is random,  
the data is fixed

knowledge about the  
parameter **before** having  
seen the data

prior

$\theta$   
 $X$

$\mathbb{P}(\theta)$

$\mathbb{P}(X|\theta)$   
likelihood

knowledge about the  
parameter **after** having  
seen the data

posterior

$\mathbb{P}(\theta|X)$

$$\mathbb{P}(\theta|X) > \mathbb{P}(\theta)$$

data evidence increase our  
belief in parameter hypothesis

# Bayesian updating



the model is random,  
the data is fixed

knowledge about the  
parameter **before** having  
seen the data

prior

$\theta$   
 $X$

$\mathbb{P}(\theta)$

$\mathbb{P}(X|\theta)$   
likelihood

knowledge about the  
parameter **after** having  
seen the data

posterior

$\mathbb{P}(\theta|X)$

Bayesian theorem

$$\mathbb{P}(\theta|X) = \frac{1}{\mathbb{P}(X)} \mathbb{P}(X|\theta) \mathbb{P}(\theta) = C \mathbb{P}(X|\theta) \mathbb{P}(\theta)$$

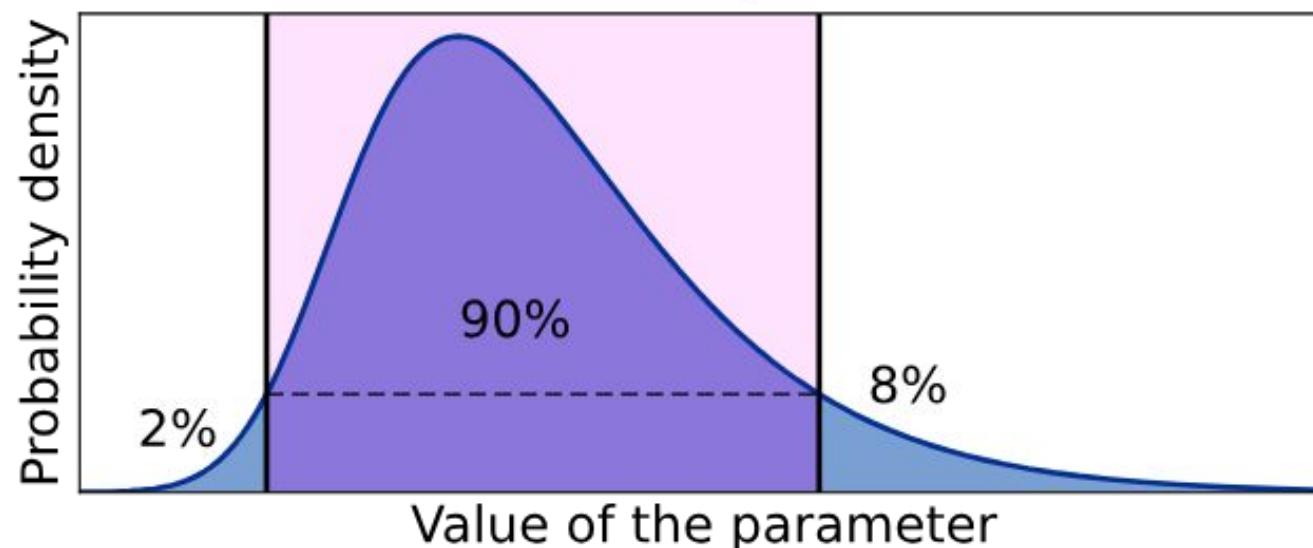
marginal likelihood

$$\mathbb{P}(X) = \int \mathbb{P}(X|\theta) \mathbb{P}(\theta) d\theta$$

integration might  
be costly!

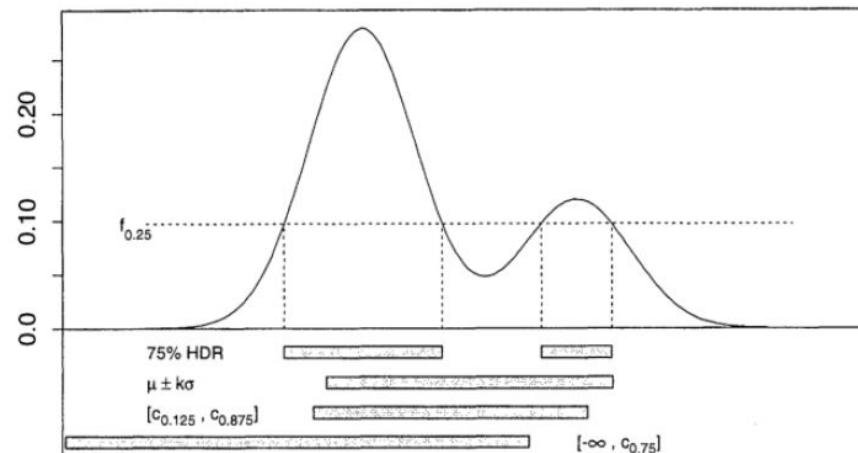
# Uncertainty from Bayesian posterior: credible intervals

- ▶ lower and upper bound such that posterior probability that unknown parameter falls within is  $X = 95, 90, 89, 75\%$
- ▶ depends on prior, not unique...
- ▶ **highest density**: region such that density function is highest with probability  $X\%$
- ▶ **equal-tailed**: e.g. 75% CI: 12.5th and 87.5th percentile, transformation invariant



# Uncertainty from Bayesian posterior: credible intervals

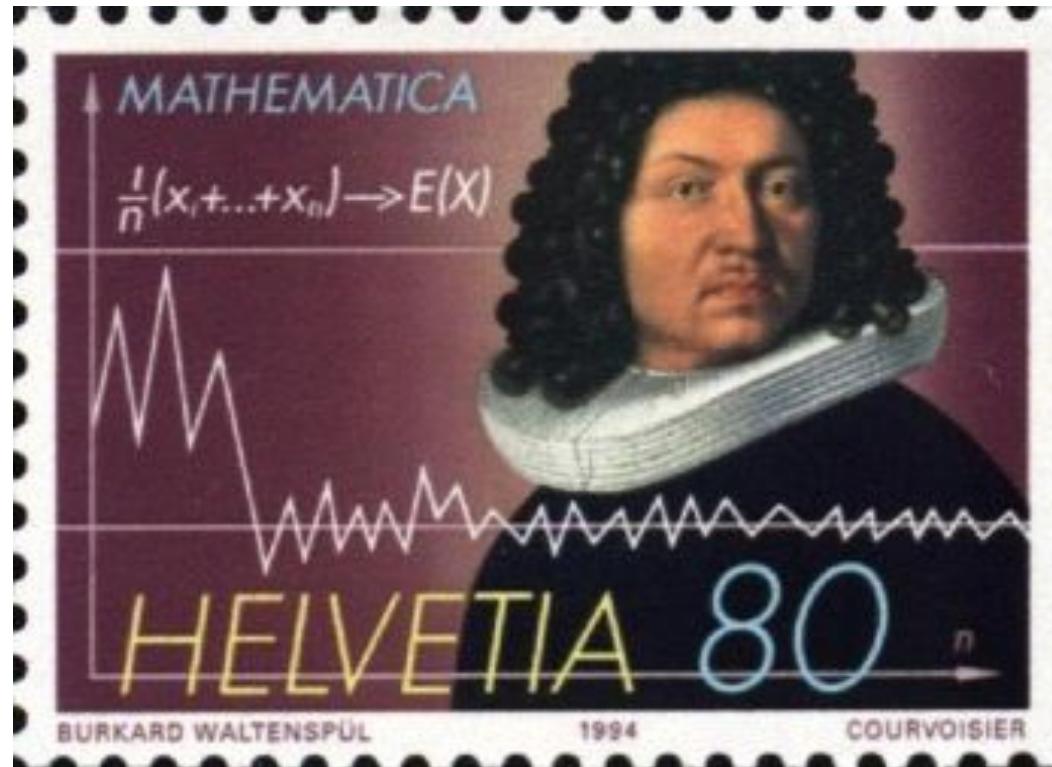
- ▶ lower and upper bound such that posterior probability that unknown parameter falls within is  $X = 95, 90, 89, 75\%$
- ▶ depends on prior, not unique...
- ▶ **highest density**: region such that density function is highest with probability  $X\%$
- ▶ **equal-tailed**: e.g. 75% CI: 12.5th and 87.5th percentile, transformation invariant



# Law of large numbers

- Given independent, identically distributed random variables  $X_1, \dots, X_d$ , then

$$\lim_{d \rightarrow \infty} \frac{1}{d} \sum_i X_i = \mathbb{E}(X_1)$$



# Law of large numbers

- Given independent, identically distributed random variables  $X_1, \dots, X_d$ , then

$$\lim_{d \rightarrow \infty} \frac{1}{d} \sum_i X_i = \mathbb{E}(X_1)$$

- weak law:** convergence in probability;  $\lim_{d \rightarrow \infty} \mathbb{P}\left(\left|\frac{1}{d} \sum_i X_i - \mathbb{E}(X_1)\right| < \epsilon\right) = 1$
- strong law:** almost sure convergence
- Proof: assume finite variance, we know  $\mathbb{E}\left(\frac{1}{d} \sum_i X_i\right) = \mathbb{E}(X_1)$  and  $\text{Var}\left(\frac{1}{d} \sum_i X_i\right) = \frac{1}{d^2} \sum_i \text{Var}(X_i)$   
Chebyshev's inequality  $\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}$  used for  $X = \frac{1}{d} \sum_i X_i$ :

$$\mathbb{P}\left(\left|\frac{1}{d} \sum_i X_i - \mathbb{E}(X_1)\right| < \epsilon\right) \geq 1 - \frac{\sigma^2}{d\epsilon^2}$$

# Monte Carlo integration

- Bayesian posterior: need to calculate integral for marginal likelihood function:

$$f(\theta) = \mathbb{P}(X|\theta)\mathbb{P}(\theta)$$

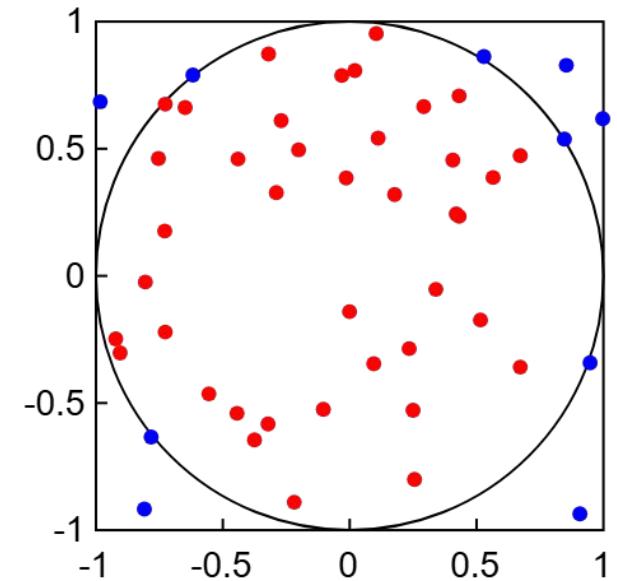
- **numeric** integration: difficult in higher dimensions
- **Monte Carlo**: sample function values **randomly**:  $f(\omega_i)$

- **Monte Carlo estimator**:

$$I_N = \frac{1}{N} \sum_{i=1}^N f(\omega_i)$$

- strong **Law of Large Numbers**:

$$\lim_{N \rightarrow \infty} I_N = \int f(\theta) d\theta, a.s.$$



$$f(\omega_i) = \mathbb{1}_{|x|<1}(\omega_i)$$

$$\lim_{N \rightarrow \infty} I_N = \pi$$

# Monte Carlo integration

- Bayesian posterior: need to calculate integral for marginal likelihood function:

$$f(\theta) = \mathbb{P}(X|\theta)\mathbb{P}(\theta)$$

- **numeric** integration: difficult in higher dimensions

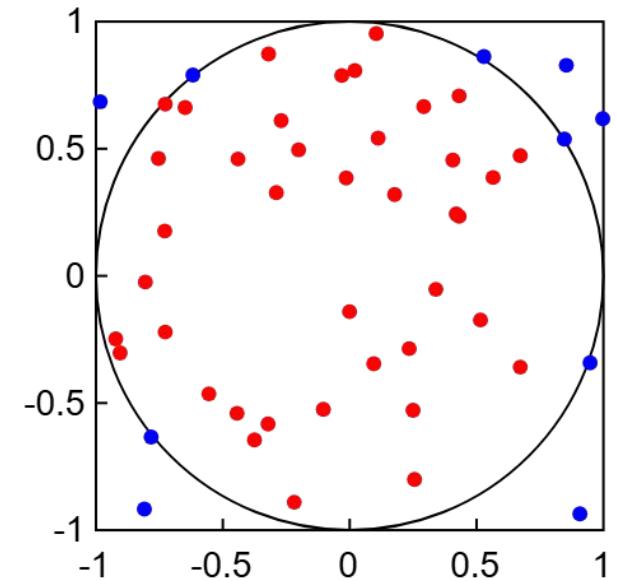
- **Monte Carlo estimator** depends on **prior**  $\mathbb{P}(\theta)$

- **Monte Carlo estimator:**

$$I_N = \frac{1}{N} \sum_{i=1}^N f(\omega_i)$$

- strong **Law of Large Numbers**:

$$\lim_{N \rightarrow \infty} I_N = \int f(\theta) d\theta, a.s.$$



$$f(\omega_i) = \mathbb{1}_{|x|<1}(\omega_i)$$

$$\lim_{N \rightarrow \infty} I_N = \pi$$

# Monte Carlo integration

- Bayesian posterior: need to calculate integral for marginal likelihood function:

$$f(\theta) = \mathbb{P}(X|\theta)\mathbb{P}(\theta)$$

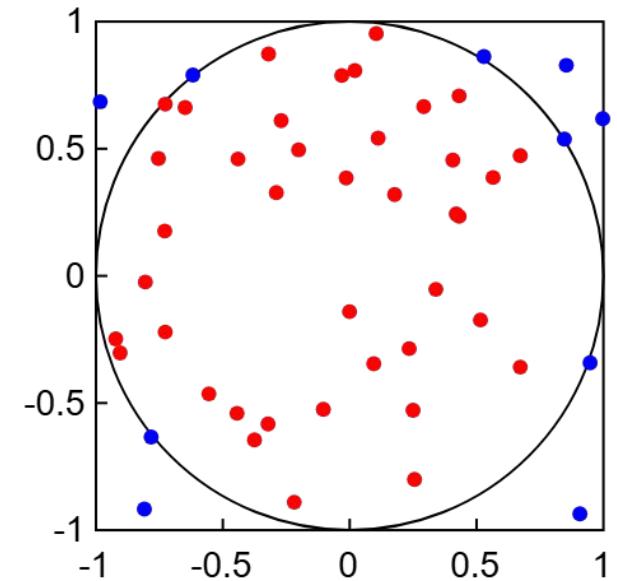
- **numeric** integration: difficult in higher dimensions

- **Monte Carlo estimator** depends on **prior**  $\mathbb{P}(\theta)$

- **Monte Carlo estimator:**

$$I_N = \frac{1}{N} \sum_{i=1}^N f(\omega_i)$$

- strong **Law of Large Numbers**:  
Monte Carlo estimator is independent from dimension or parameter space



$$f(\omega_i) = \mathbb{1}_{|x|<1}(\omega_i)$$

$$\lim_{N \rightarrow \infty} I_N = \pi$$

# Monte Carlo integration: variance and error

- ▶ sample variance  $\text{Var}(f) = \frac{1}{N-1} \sum_i (f(\omega_i) - I_N)^2 = \sigma_N^2$
- ▶ sample variance of integral:  $\text{Var}(I_N) = \frac{1}{N} \sigma_N^2$
- ▶ if  $\{\sigma_N^2\}$  bounded, then convergence: Law of Large Numbers
- ▶ Central Limit Theorem:

$$\frac{I_N - \int f(\theta) d\theta}{\sqrt{\sigma_N^2/N}} \xrightarrow{\text{proba}} \mathcal{N}(0, 1)$$

- ▶ Monte Carlo error estimate:  $I_N \pm 1.96 \sqrt{\sigma_N^2/N}$

# Monte Carlo integration: variance and error

**variance** of Monte Carlo  
integral decrease with  
increasing sample size

- ▶ sample variance  $\text{Var}(f) = \frac{1}{N-1} \sum_i (f(\omega_i) - I_N)^2$
- ▶ sample variance of integral:  $\text{Var}(I_N) = \frac{1}{N} \sigma_N^2$
- ▶ if  $\{\sigma_N^2\}$  bounded, then convergence: Law of Large Numbers
- ▶ Central Limit Theorem:

$$\frac{I_N - \int f(\theta) d\theta}{\sqrt{\sigma_N^2/N}} \xrightarrow{\text{proba}} \mathcal{N}(0, 1)$$

- ▶ Monte Carlo error estimate:  $I_N \pm 1.96 \sqrt{\sigma_N^2/N}$

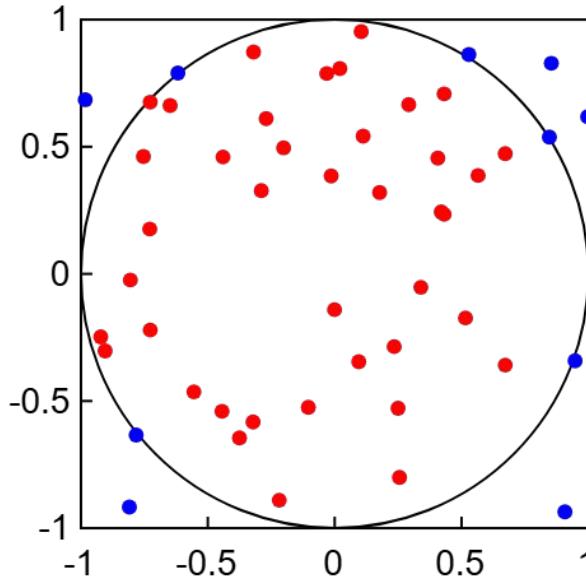
# Monte Carlo integration: variance and error

- ▶ sample variance  $\text{Var}(f) = \frac{1}{N-1} \sum_i (f(\omega_i) - I_N)^2 = \sigma_N^2$
- ▶ sample variance of integral:  $\text{Var}(I_N) = \frac{1}{N} \sigma_N^2$
- ▶ if  $\{\sigma_N^2\}$  bounded, then convergence: Law of Large Numbers
- ▶ Central Limit Theorem:

In order to decrease Monte Carlo error by factor 10, you need to increase sample number by factor 100!

- ▶ Monte Carlo error estimate:  $I_N \pm 1.96 \sqrt{\sigma_N^2/N}$

# Monte Carlo integration: example



- sample randomly N points in the square
- calculate the fraction of points that fall within the circle
- use  $\text{Ratio} = \frac{\text{Area of circle}}{\text{Area of square}} = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}$
- calculate area of circle by multiplying fraction of points within circle by 4

Implement in R!

AIDM\_05.R

# Monte Carlo integration: example for marginal likelihood

- sample N points uniformly distributed in the interval [1.2,2] for **beta** and [0.2,0.6] for **gamma**
- calculate Monte Carlo estimator for marginal likelihood of the SIR model with
- use  $\theta = (\beta, \gamma)$

$$f(\theta) = \mathbb{P}(X|\theta)\mathbb{P}(\theta)$$

posterior probability distribution

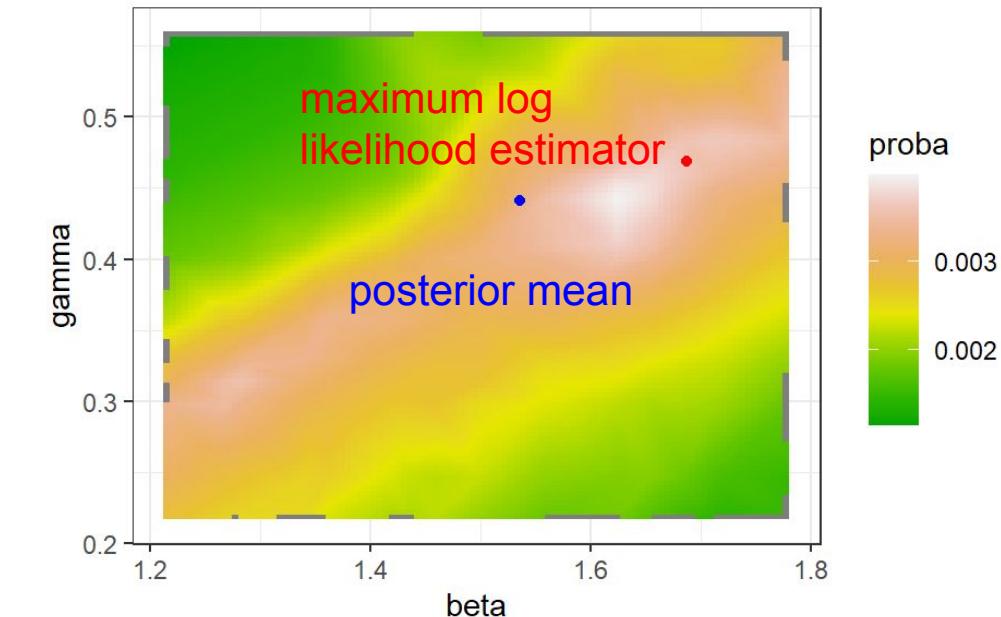
posterior  
distribution

$$\mathbb{P}(\theta|X) = \frac{1}{\mathbb{P}(X)} \mathbb{P}(X|\theta)\mathbb{P}(\theta)$$

Monte Carlo integral



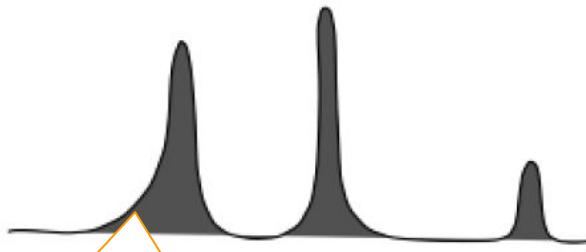
Implement in R!



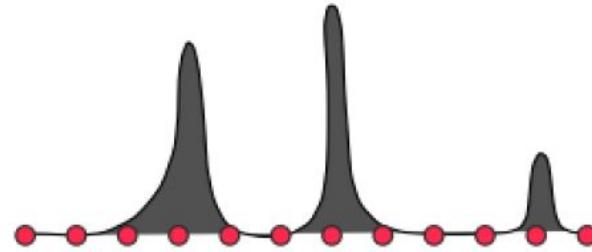
# Monte Carlo integration: importance sampling

importance sampling uses proposal distribution  $q$  to sample more important parameter regions more often

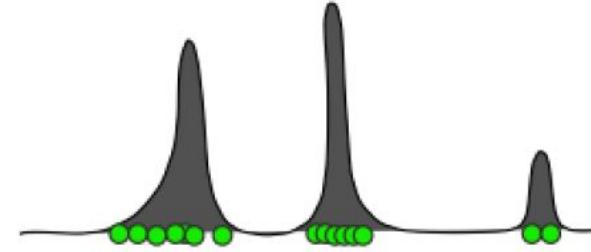
$$\int f(\theta) d\theta = \int \mathbb{P}(X|\theta) \mathbb{P}(\theta) d\theta = \int \mathbb{P}(X|\theta) \underbrace{\frac{\mathbb{P}(\theta)}{q(\theta)}}_{\text{importance weight}} q(\theta) d\theta \underbrace{q(\theta)}_{\text{sampling probability}}$$



target distribution =  
prior



Uniform  
distribution



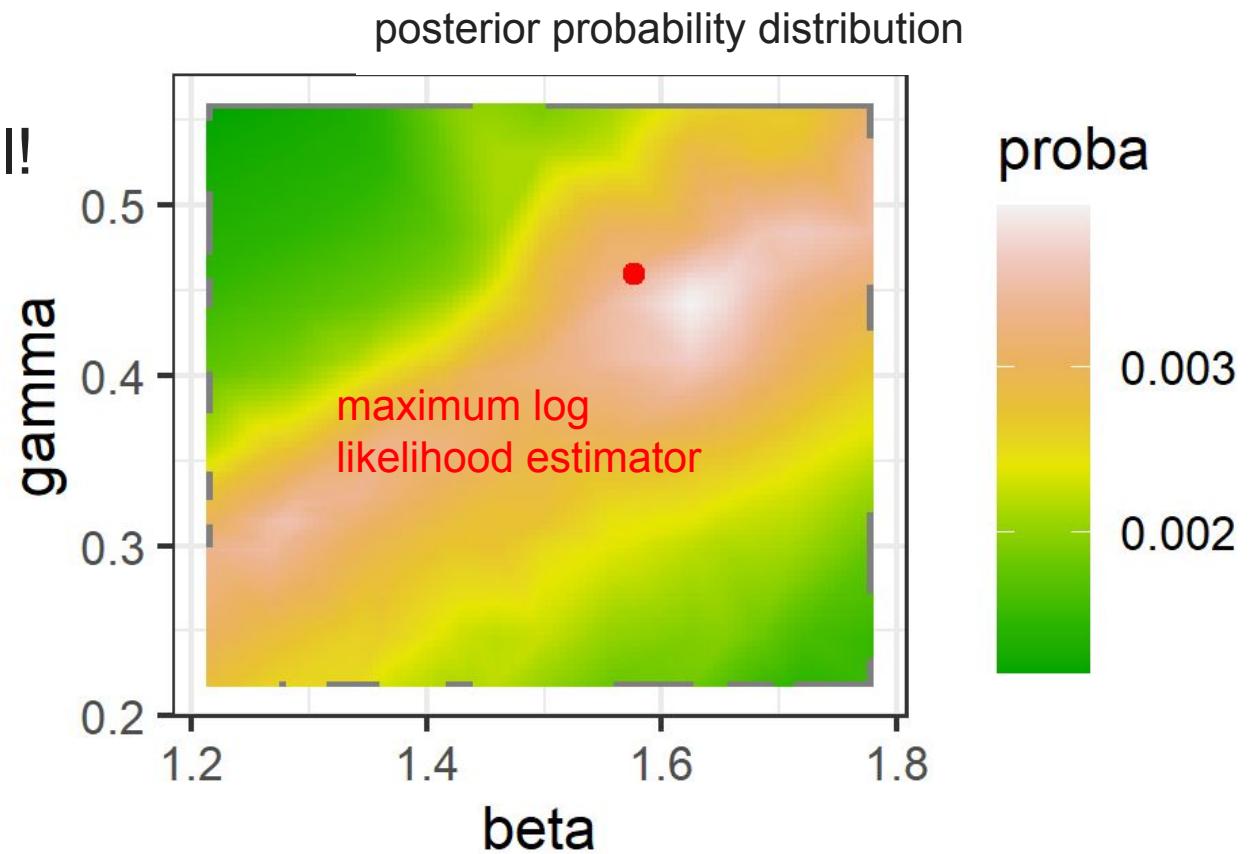
Importance  
sampling

# Monte Carlo integration: importance sampling

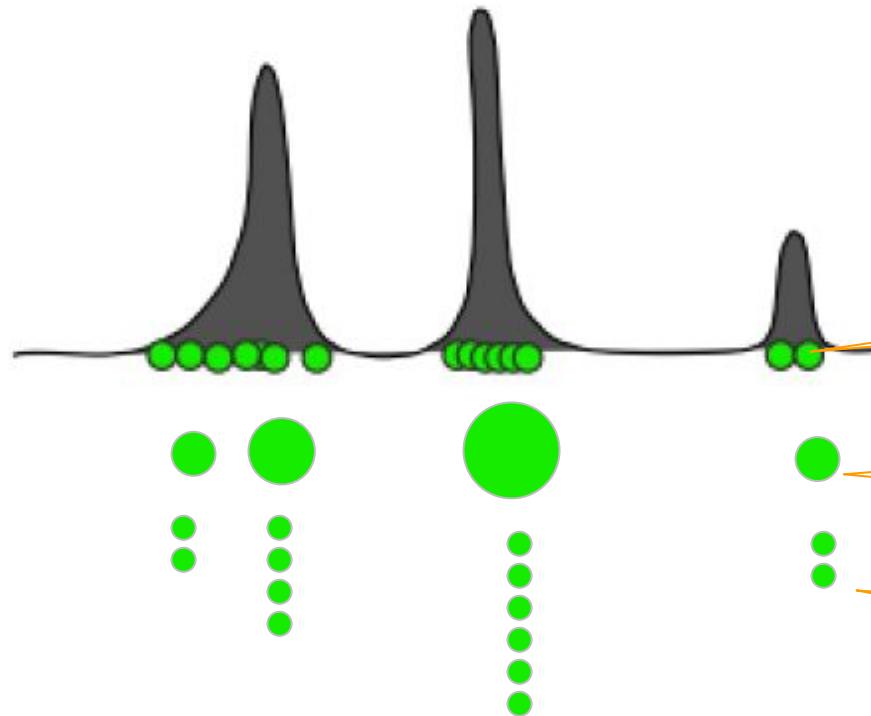
importance sampling uses proposal distribution  $q$  to sample more important parameter regions more often

  
AIDM\_05.R

Implement in R with truncated normal distribution as proposal!



# Monte Carlo integration: sampling importance resampling



**1. Sampling (Importance Sampling):** draw samples from proposal distribution.

$$\omega_1, \dots, \omega_N \sim q$$

**2. Weights:** proportional to how likely sample is under the target distribution

$$W_i = \frac{\mathbb{P}(\omega_i)}{q(\omega_i)}$$

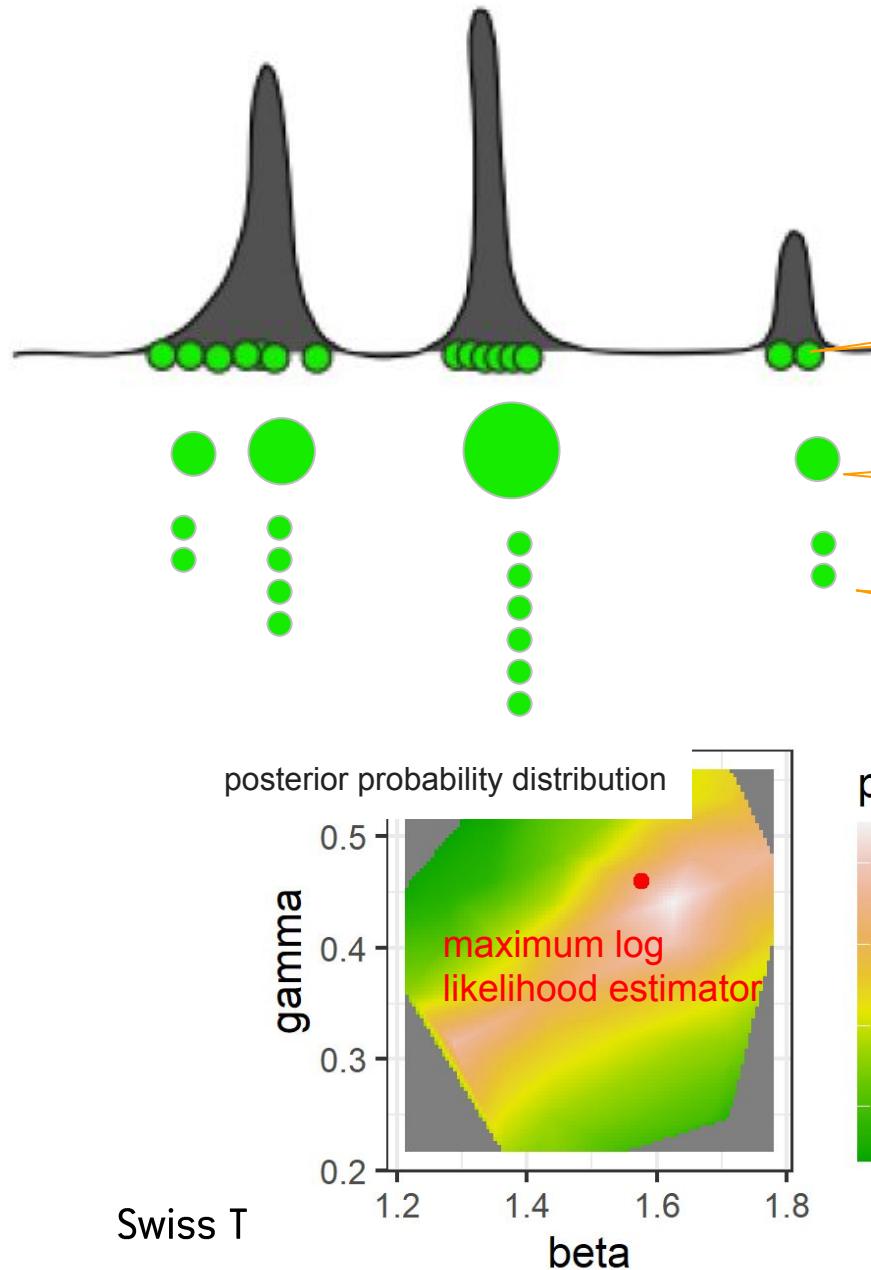
**3. Resampling:** resample with replacement using the weights

$$\hat{\omega}_1, \dots, \hat{\omega}_k \sim \{\omega_1, \dots, \omega_N\} \quad \tilde{W}_i = \frac{W_i}{\sum_i W_i}$$

$$\int f(\theta) d\theta = \int \mathbb{P}(X|\theta) \mathbb{P}(\theta) d\theta$$

Use samples from resampling step to calculate Monte Carlo integral estimator!

# Monte Carlo integration: sampling importance resampling



**1. Sampling (Importance Sampling):** draw samples from proposal distribution.

$$\omega_1, \dots, \omega_N \sim q$$

**2. Weights:** proportional to how likely sample is under the target distribution

$$W_i = \frac{\mathbb{P}(\omega_i)}{q(\omega_i)}$$

**3. Resampling:** resample with replacement using the weights

$$\hat{\omega}_1, \dots, \hat{\omega}_k \sim \{\omega_1, \dots, \omega_N\} \quad \tilde{W}_i = \frac{W_i}{\sum_i W_i}$$



Implement in R with truncated normal distribution as proposal and resample 50 times!

## Key takeaway points:

---

- **paradigms of inference** distinguish whether model parameter is fixed, and data random, or parameter random and data fixed
- **maximum log likelihood** approach is an alternative to curve fitting, taking also into consideration parameter uncertainty
- **Bayesian inference** allows to improve prior knowledge on parameters by comparing model outputs with data
- **Monte Carlo integration** is a (not always) efficient way to calculate high-dimensional integrals for Bayesian posterior



**Modeling malaria transmission and drug resistance at the population level**

**Christian Selinger & Monica Golumbeanu**

# How antimalarial combination therapy works

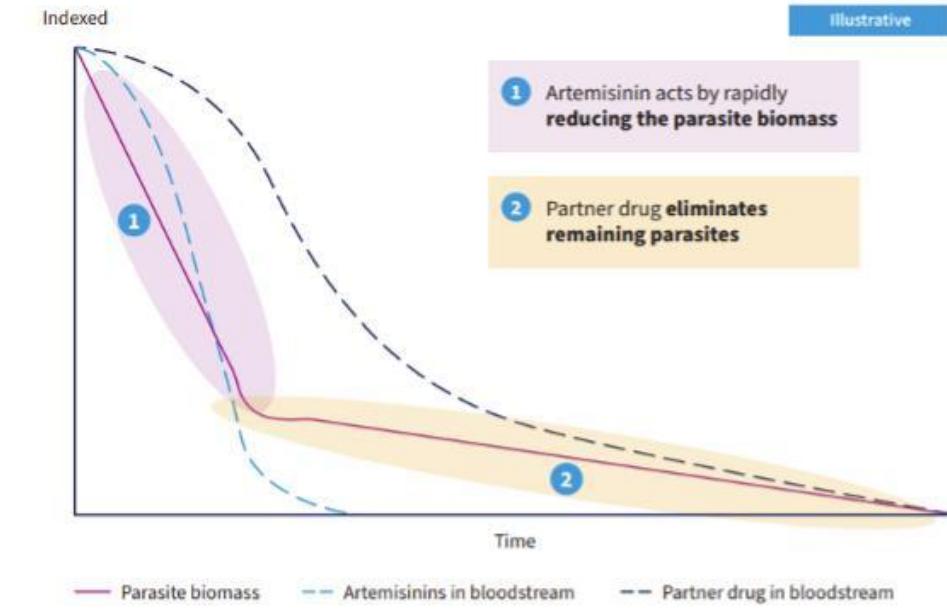
## Rapid Parasite Clearance (Artemisinin Component)

- The artemisinin derivative (e.g., artesunate, artemether, or dihydroartemisinin) acts quickly, significantly reducing the parasite load within the first 24-48 hours.

## Sustained Killing Effect (Partner Drug)

- Since **artemisinin clears most but not all parasites**, a longer-acting partner drug (e.g., lumefantrine, amodiaquine, mefloquine, piperaquine) eliminates the remaining parasites over several days.
- The **partner drug has a longer half-life**, preventing reinfection and reducing the risk of **recrudescence (treatment failure due to surviving parasites)**.

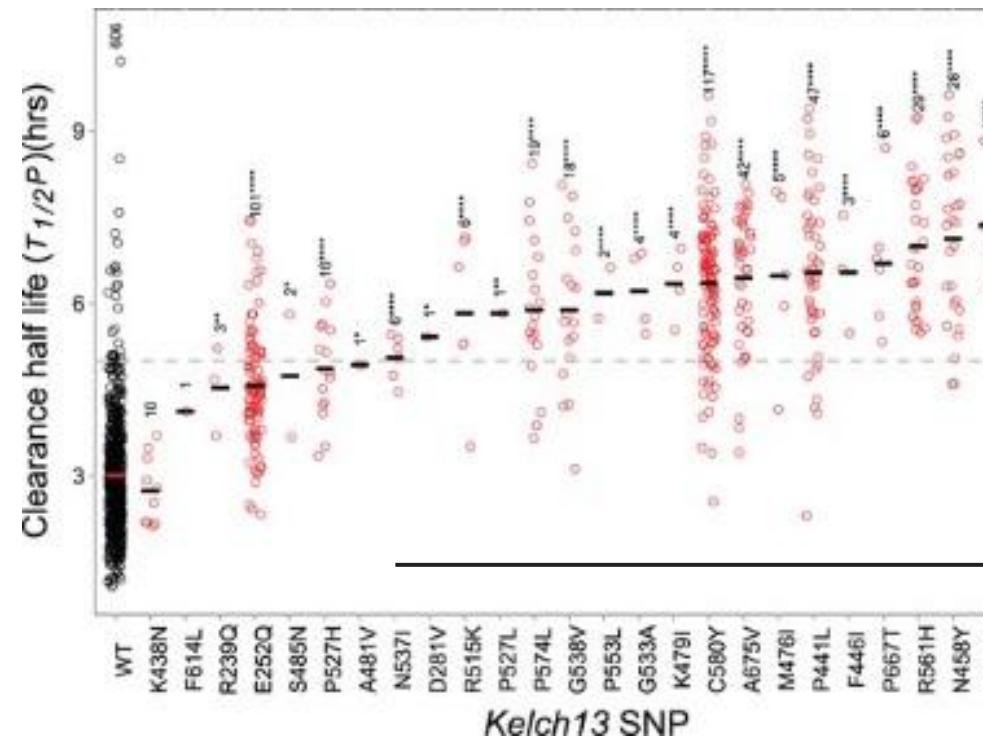
Figure 1. Evolution of parasite biomass in the body following ACT administration



<https://www.mmv.org/malaria/symptoms-and-treatments/about-arte-misinin-and-ACTs>

# Different levels of antimalarial resistance are present in a population

- Over time, parasites naturally undergo mutations which allow them to **evolve** and, under right conditions (selective pressure), they **evade the effects of antimalarial treatments**, making them less effective.
- Antimalarial resistance** is when *malaria parasites* (like *Plasmodium falciparum*) **no longer respond effectively to antimalarial drugs** that previously worked to treat the infection.



Delay in parasite clearance across parasite genotypes

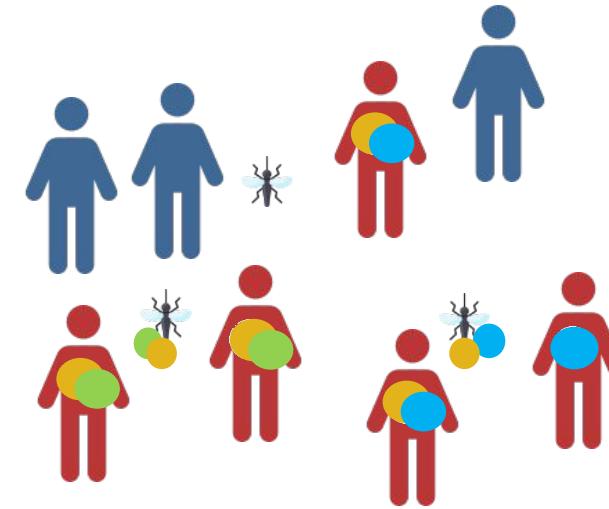
# What drives the spread of antimalarial resistance in a population?

**Drug pressure:** treating infections kills sensitive parasites but allows resistant ones to survive and multiply.

**Fitness cost/advantage:** in the absence of drugs, resistant parasites may spread more slowly (or even be outcompeted)

**Transmission dynamics:** infected humans pass parasites to mosquitoes, which infect other humans

**Human movement & mosquito migration:** Spread of resistant parasites geographically



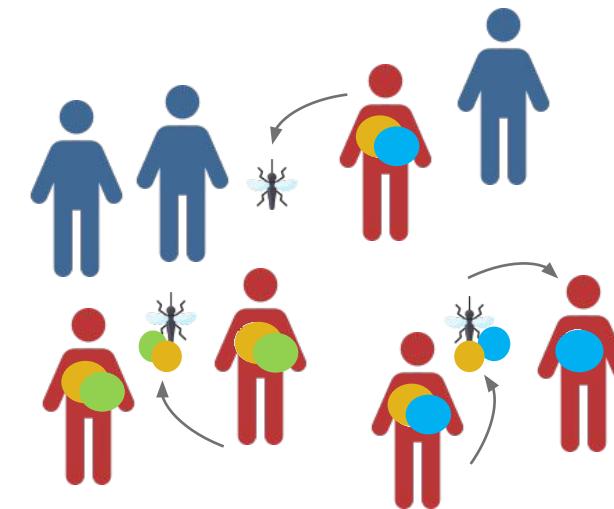
# What drives the spread of antimalarial resistance in a population?

**Drug pressure:** treating infections kills sensitive parasites but allows resistant ones to survive and multiply

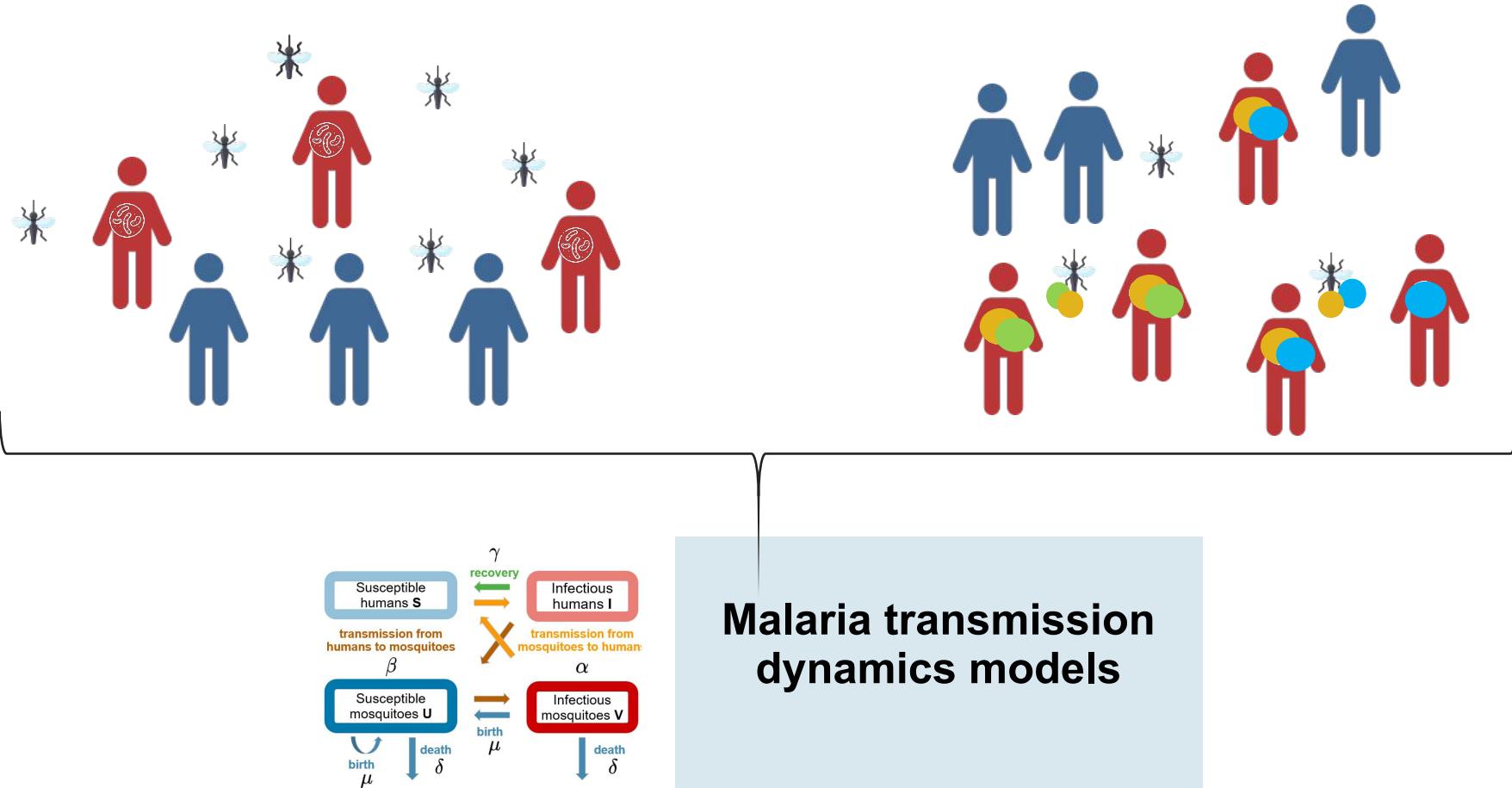
**Fitness cost/advantage:** in the absence of drugs, resistant parasites may spread more slowly (or even be outcompeted)

**Transmission dynamics:** infected humans pass parasites to mosquitoes, which infect other humans

**Human movement & mosquito migration:** Spread of resistant parasites geographically

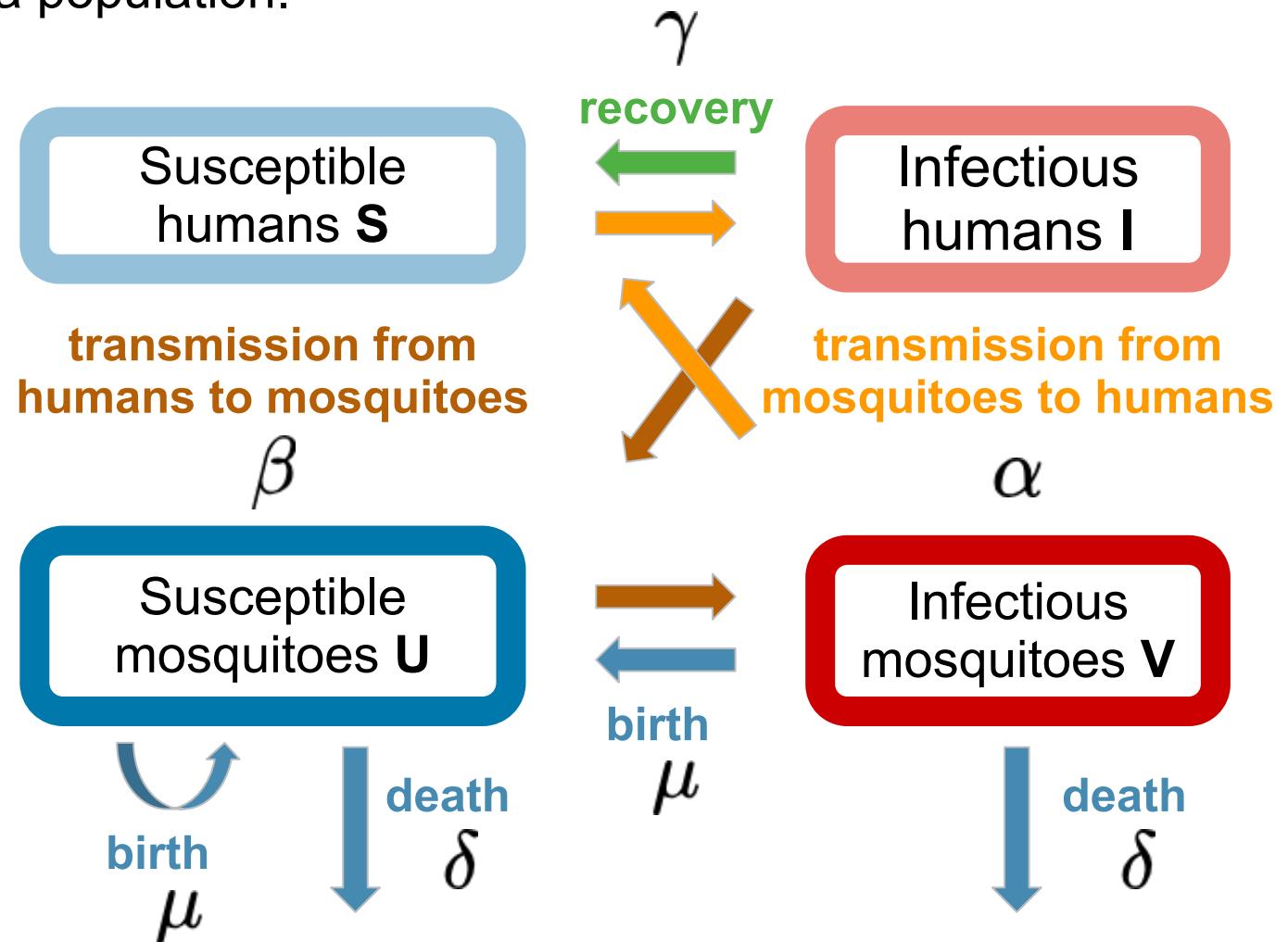
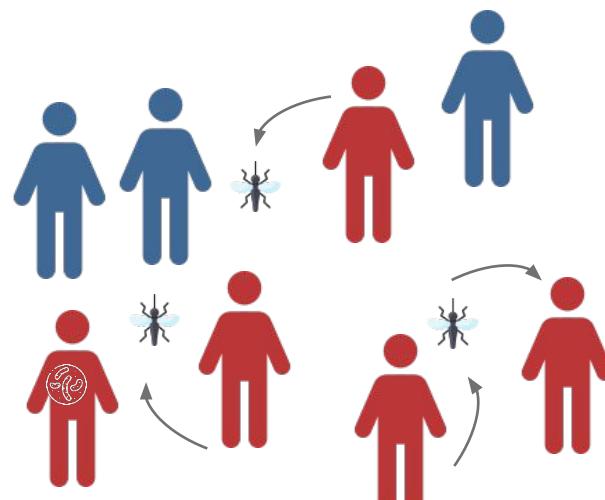


# Integrating malaria transmission dynamics with parasite genetics



# Malaria transmission dynamics model

A **malaria transmission model** is a mathematical framework used to simulate and analyze how malaria spreads in a population.



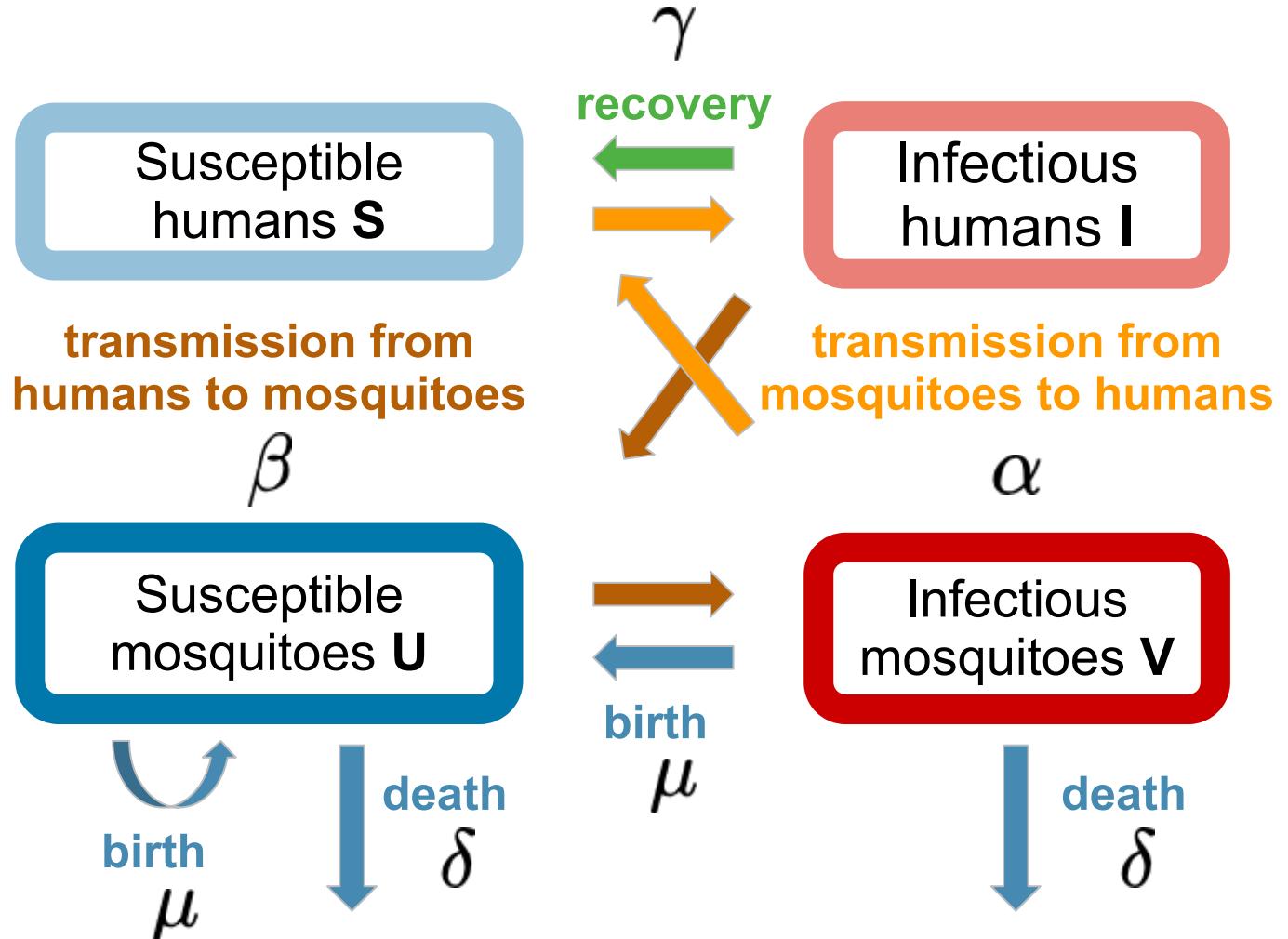
# Malaria transmission dynamics model

Model equations:

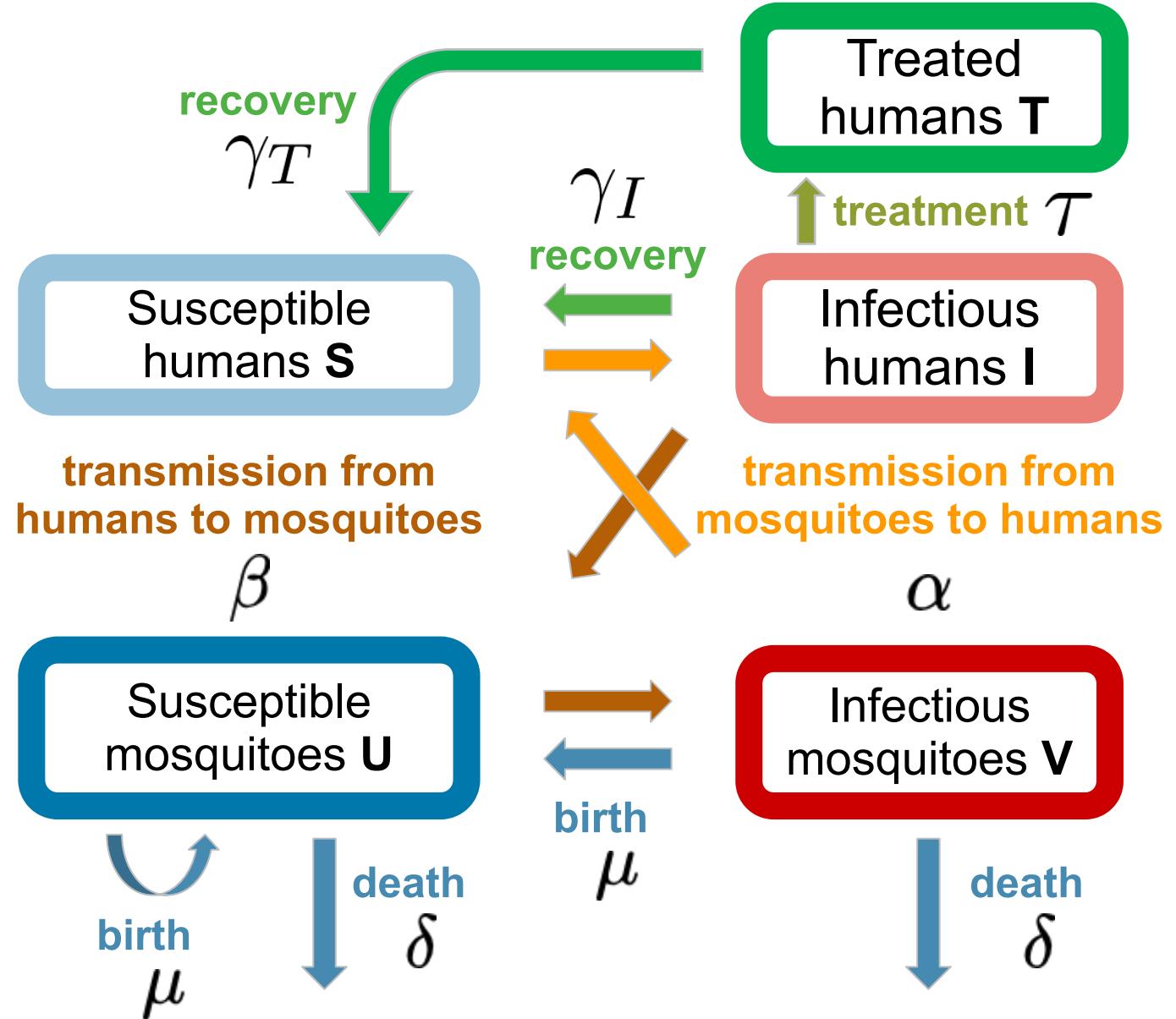
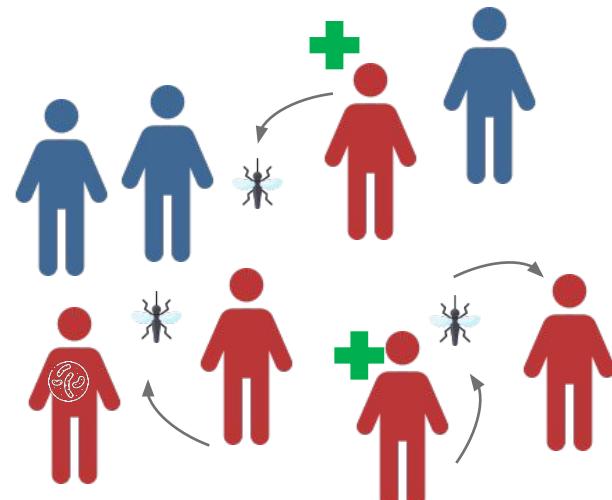
$$\begin{cases} \frac{dS}{dt}(t) = -\alpha \frac{V(t)}{H} S(t) + \gamma I(t) \\ \frac{dI}{dt}(t) = \alpha \frac{V(t)}{H} S(t) - \gamma I(t) \\ \frac{dU}{dt}(t) = -\beta \frac{I(t)}{H} U(t) + \mu M - \delta U(t) \\ \frac{dV}{dt}(t) = \beta \frac{I(t)}{H} U(t) - \delta V(t) \end{cases}$$

Assumptions:

- Constant mosquito population size:  
 $M = U(0) + V(0)$
- Birth and death rates are equal:  
 $\mu = \delta$
- Constant human population size  
 $H = S(0) + I(0)$



# Malaria transmission dynamics model with treatment

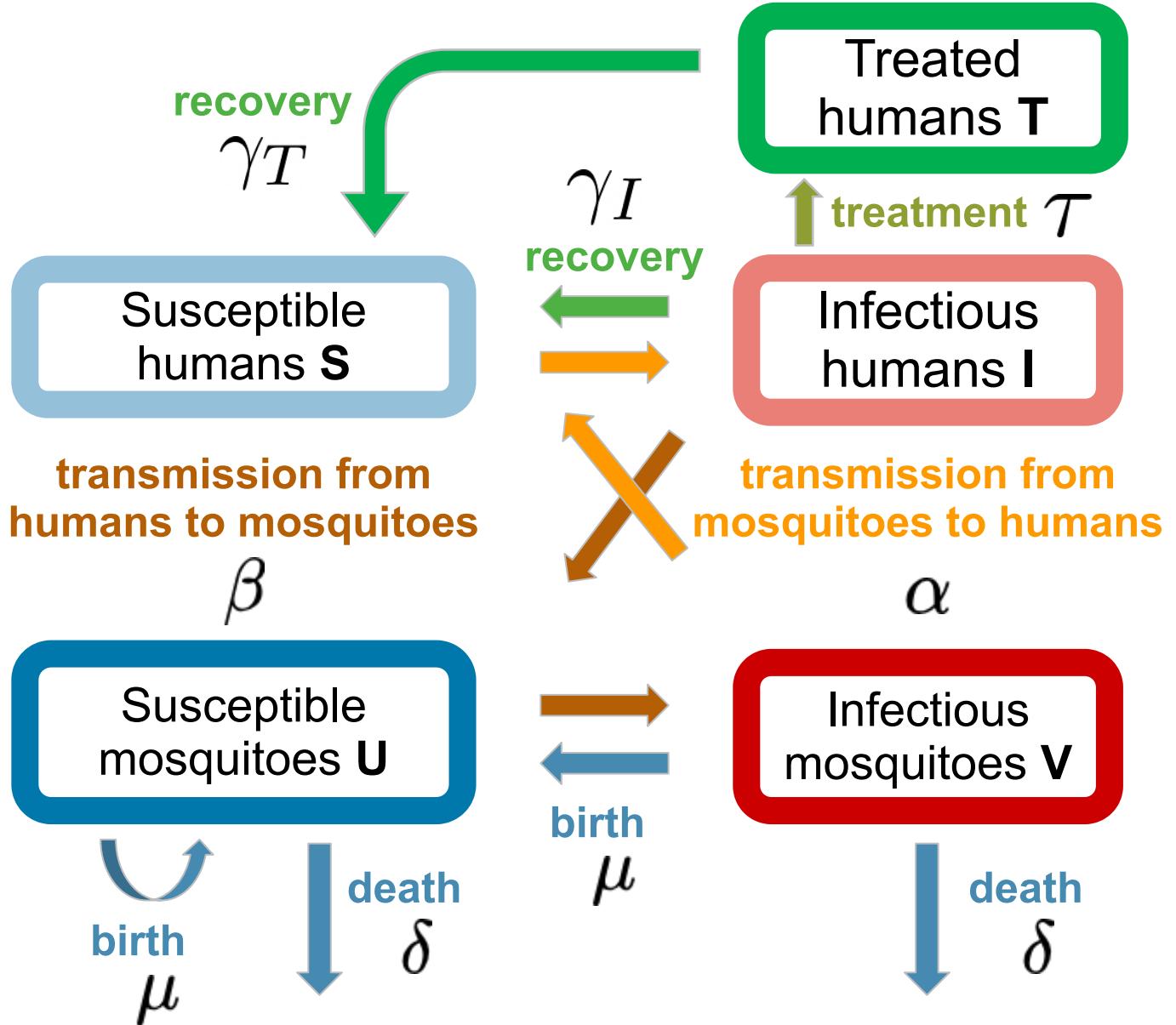


# Malaria transmission dynamics model with treatment

$$\begin{aligned}\frac{dS}{dt} &= -\alpha S \frac{V}{H} + \gamma_I I + \gamma_T T \\ \frac{dI}{dt} &= \alpha S \frac{V}{H} - \tau I \\ \frac{dT}{dt} &= \tau I - \gamma_T T \\ \frac{dU}{dt} &= -\beta U \frac{I}{H} + \mu(U + V) - \delta U \\ \frac{dV}{dt} &= \beta U \frac{I}{H} - \delta V\end{aligned}$$

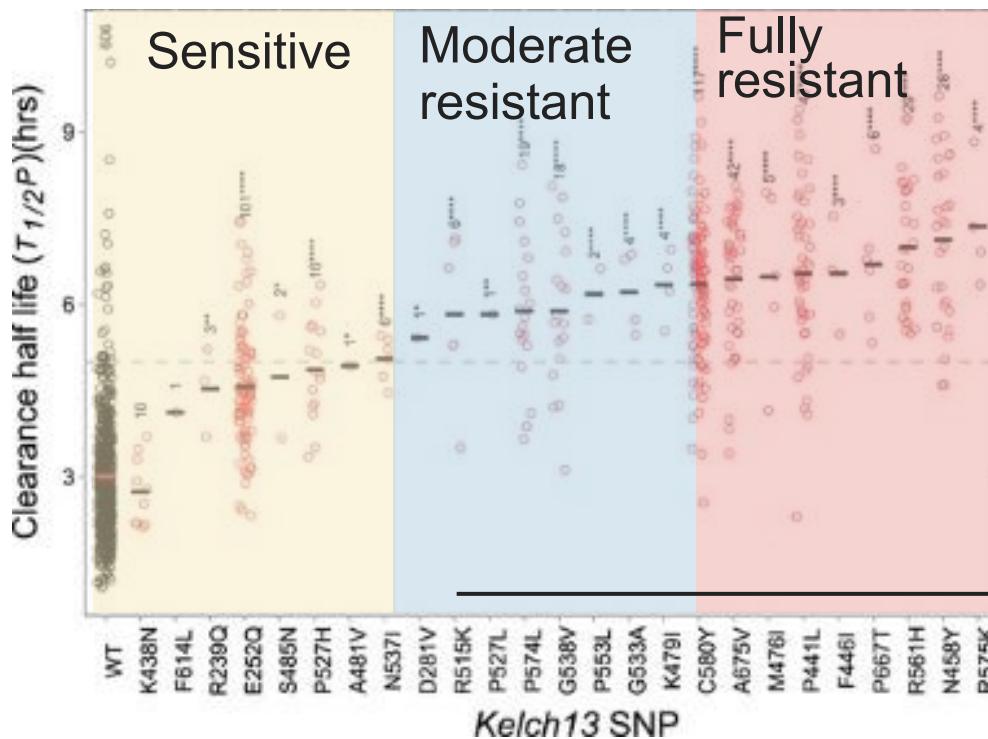
## Assumptions:

- Constant mosquito population size:  
 $M = U(0) + V(0)$
- Birth and death rates are equal:  
 $\mu = \delta$
- Constant human population size  
 $H = S(0) + I(0)$



# Different levels of antimalarial resistance are present in a population

- Over time, parasites naturally undergo mutations which allow them to **evolve** and, under right conditions (selective pressure), they **evade the effects of antimalarial treatments**, making them less effective.
- Antimalarial resistance** is when *malaria parasites* (like *Plasmodium falciparum*) **no longer respond effectively to antimalarial drugs** that previously worked to treat the infection.

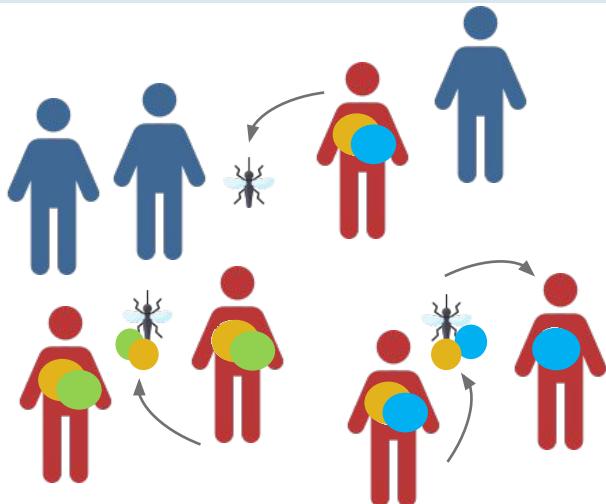


Delay in parasite clearance across parasite genotypes

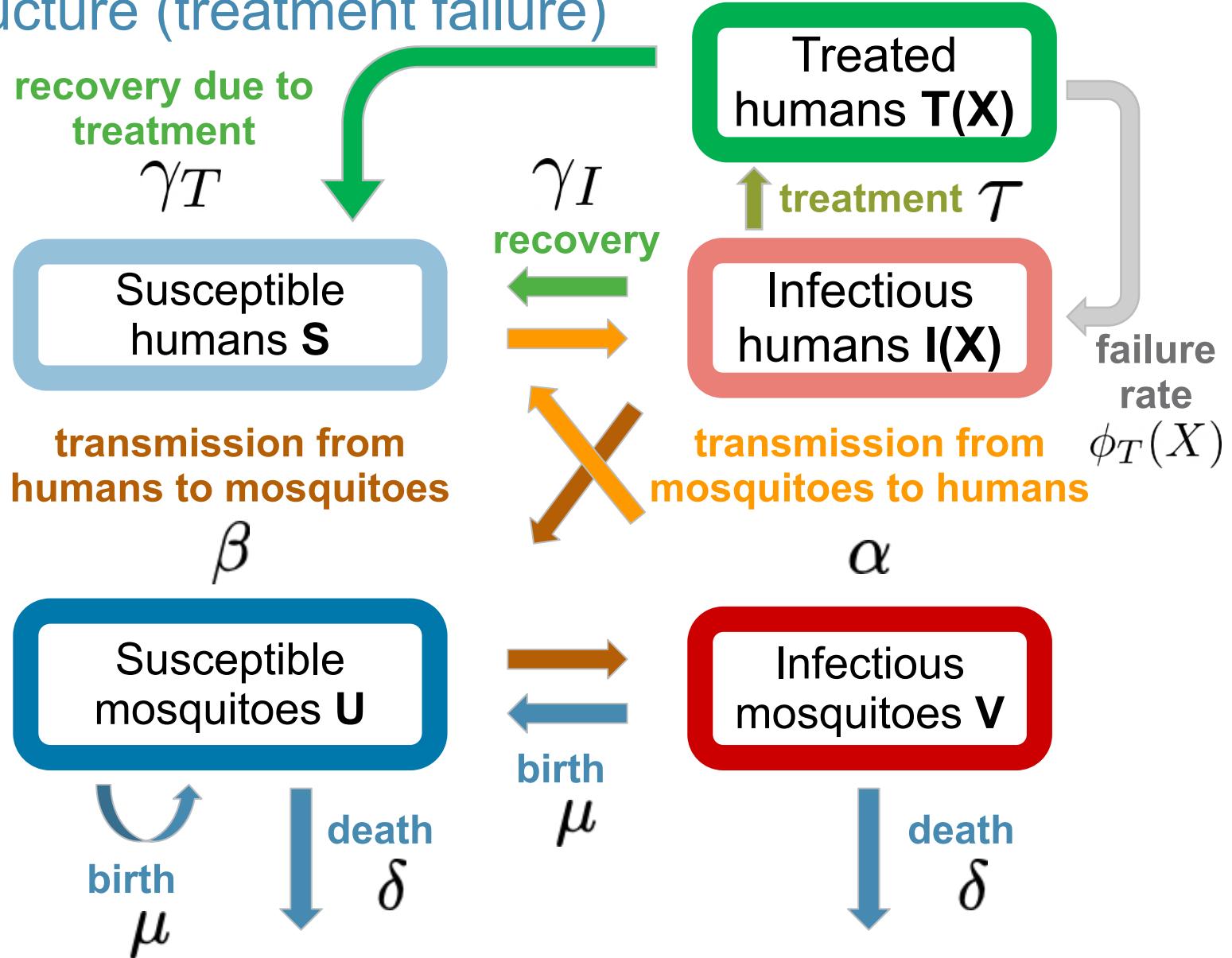
Varying drug failure rates across parasite genotypes

# Malaria transmission dynamics model with treatment and parasite genotype structure (treatment failure)

Humans are infected with parasites with **different genotypes** at frequencies  $f_X$  (different levels of treatment failure) → infected humans  $I(X)$



Each parasite genotype has associated a **treatment failure rate**  $\tau_X$

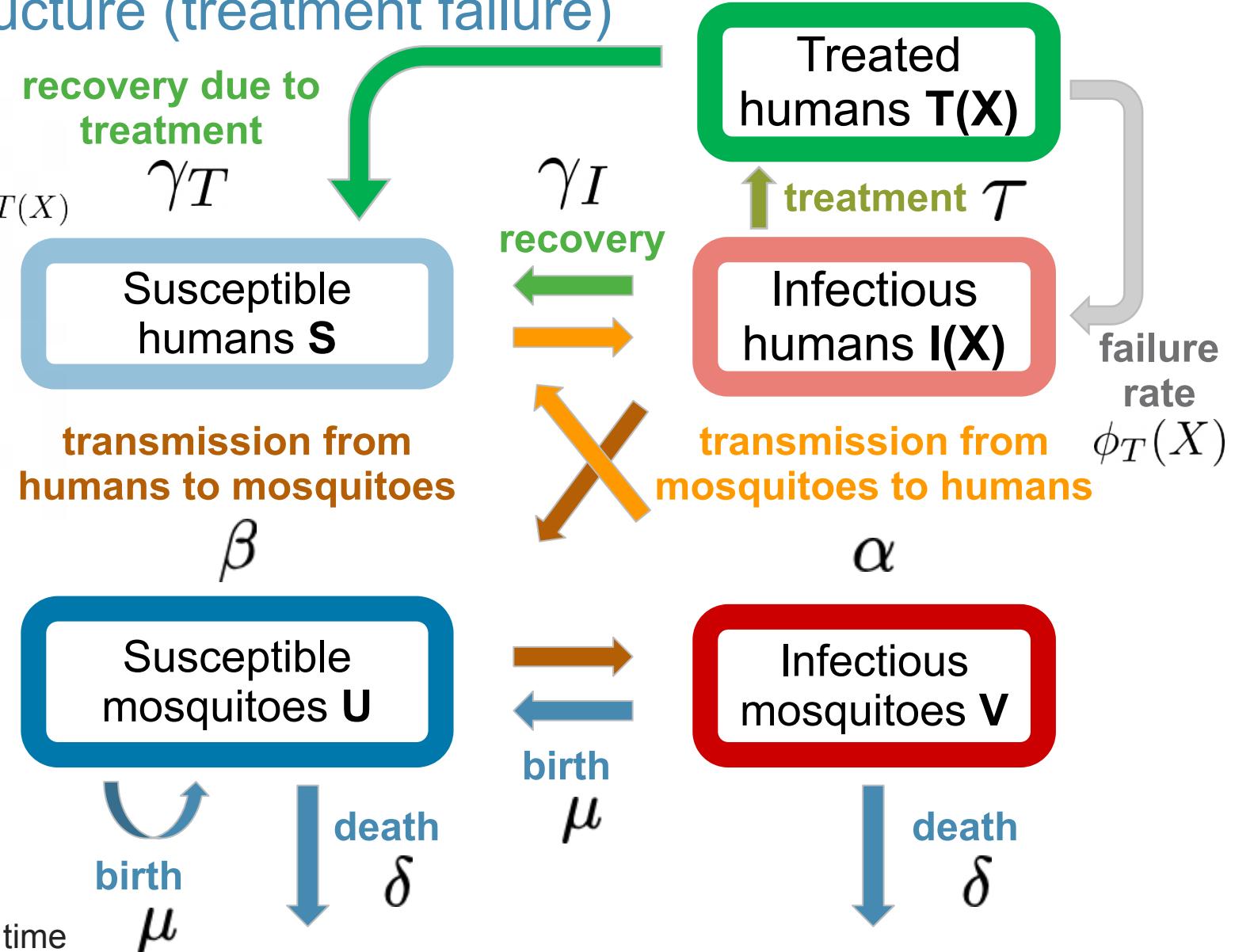


# Malaria transmission dynamics model with treatment and parasite genotype structure (treatment failure)

$$\begin{cases} \frac{dS}{dt} = -\alpha S \frac{V}{H} + \gamma_I \sum_X I(X) + \gamma_T \sum_X T(X) \\ \frac{dI(X)}{dt} = \alpha S \frac{V}{H} f_X - \tau I(X) - \gamma_I I(X) + \phi_T(X)T(X) \\ \frac{dT(X)}{dt} = \tau I(X) - \gamma_T T(X) - \phi_T(X)T(X) \\ \frac{dU}{dt} = -\beta U \sum_X \frac{I(X)}{H} + \mu(U + V) - \delta U \\ \frac{dV}{dt} = \beta U \sum_X \frac{I(X)}{H} - \delta V \end{cases}$$

## Assumptions:

- Constant mosquito population size:  
 $M = U(0) + V(0)$
- Birth and death rates are equal:  
 $\mu = \delta$
- Constant human population size  
 $H = S(0) + I(0)$
- Genotype frequencies  $f_X$  are constant in time



# Malaria transmission dynamics model with treatment and parasite genotype structure (treatment failure)

AIDM\_RMresist.R

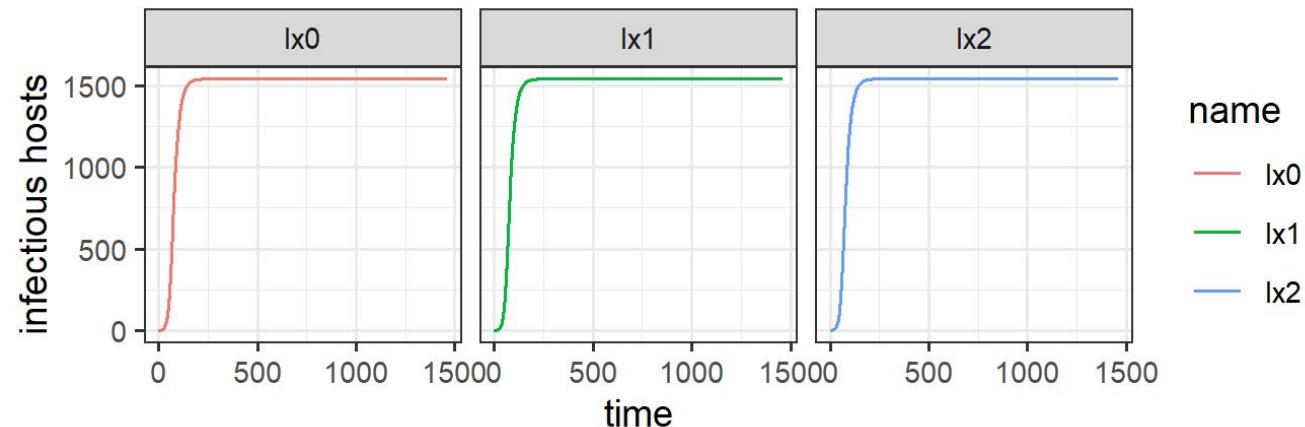
```
H=5000;V0=8; VectorHumanRatio=5; finalT=4*365
x0 <- c(S=H-3, Ix0=1, Ix1=1, Ix2=1, Treatx0=0, Treatx1=0, Treatx2=0,
         U=H*VectorHumanRatio-V0, V=V0) #initial condition
time.points<-seq(0,finalT,1)##time unit in days

parms<-c(
alpha=0.5*0.022,
gammaI=1/285,
gammaT=1/30,
phiTx0=0,phiTx1=1/15,phiTx2=1/5,##failure rate
tau=0,
beta=0.48, mu=0.13, delta=0.13,
fx0=1/3,fx1=1/3,fx2=1/3
)
```

Initial condition with one infection per genotype.

Failure rates for different genotype = phenotype!

Same probability to acquire any genotype during transmission



treatment rate = 0  
coexistence of 3 genotypes (no fitness cost for resistant genotypes x1 & x2!)

# Malaria transmission dynamics model with treatment and parasite genotype structure (treatment failure)

AIDM\_RMresist.R

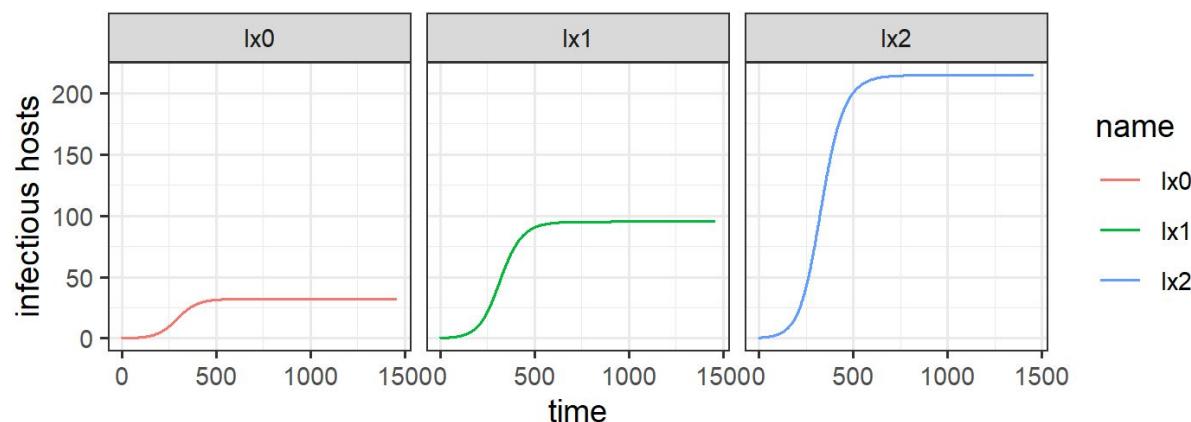
```
H=5000;V0=8; VectorHumanRatio=5; finalT=4*365
x0 <- c(S=H-3,Ix0=1,Ix1=1,Ix2=1,Treatx0=0,Treatx1=0,Treatx2=0,
         U=H*VectorHumanRatio-V0,V=V0)##initial condition
time.points<-seq(0,finalT,1)##time unit in days

parms<-c(
alpha=0.5*0.022,
gammaI=1/285,
gammaT=1/30,
phiTx0=0,phiTx1=1/15,phiTx2=1/5,##failure rate
tau=0,
beta=0.48, mu=0.13, delta=0.13,
fx0=1/3,fx1=1/3,fx2=1/3
)
```

Initial condition with one infection per genotype.

Failure rates for different genotype = phenotype!

Same probability to acquire any genotype during transmission



name  
— Ix0  
— Ix1  
— Ix2

treatment rate = 0.4  
coexistence of 3 genotypes, fitness advantage for resistant types x1, x2

# Malaria transmission dynamics model with treatment and parasite genotype structure (treatment failure)

AIDM\_RMresist.R

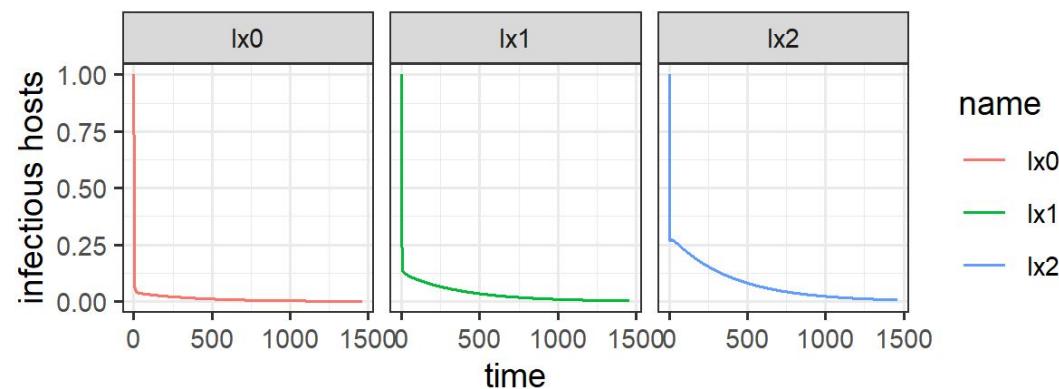
```
H=5000;V0=8; VectorHumanRatio=5; finalT=4*365
x0 <- c(S=H-3,Ix0=1,Ix1=1,Ix2=1,Treatx0=0,Treatx1=0,Treatx2=0,
         U=H*VectorHumanRatio-V0,V=V0)##initial condition
time.points<-seq(0,finalT,1)##time unit in days

parms<-c(
alpha=0.5*0.022,
gammaI=1/285,
gammaT=1/30,
phiTx0=0,phiTx1=1/15,phiTx2=1/5,##failure rate
tau=0,
beta=0.48, mu=0.13, delta=0.13,
fx0=1/3,fx1=1/3,fx2=1/3
)
```

Initial condition with one infection per genotype.

Failure rates for different genotype = phenotype!

Same probability to acquire any genotype during transmission



treatment rate = 0.8  
all three genotypes are eliminated