

Introduction to Infectious Disease Modelling

Modelling for Pandemic Preparedness and Response Modular
Shortcourse, 2025

James Mba Azam, PhD

2025-09-03

London School of Hygiene and Tropical Medicine, UK

An overview of infectious diseases

What are infectious diseases

disease noun

dis·ease (di-'zēz 

[Synonyms of disease >](#)

- 1 : a condition of the living animal or plant body or of one of its parts that impairs normal functioning and is typically manifested by distinguishing signs and symptoms
: **SICKNESS, MALADY**
 - | *infectious diseases*
 - | *a rare genetic disease*
 - | *heart disease*

Diseases can be classified according to:

- Cause (e.g., infectious, non-infectious)
- Duration (e.g., acute, chronic)
- Mode of transmission (direct or indirect)
- Impact on the host (e.g., fatal, non-fatal)

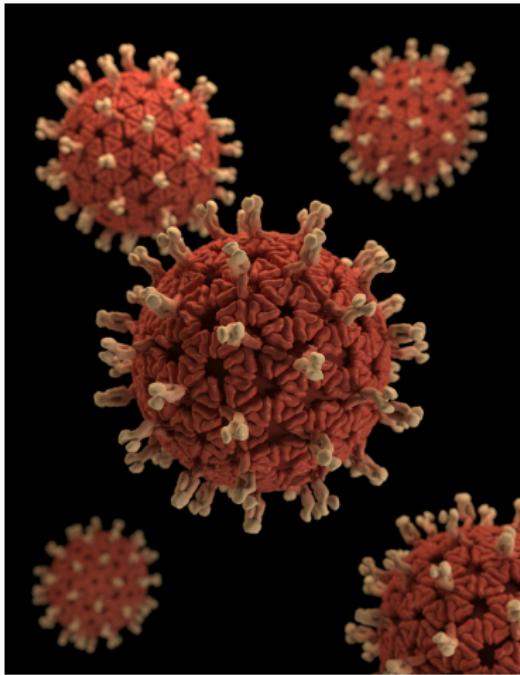


Figure 1: A 3D graphical representation of Rotavirus virions.

Note that these classifications are not mutually exclusive. Hence, a disease can be classified under more than one category at a time.

How are infectious diseases controlled?

- In general, infectious disease control aims to reduce disease transmission.
- The type of control used depends on the disease and its characteristics.
- Broadly, there are two main types of control measures:
 - Pharmaceutical interventions (PIs)
 - Non-pharmaceutical interventions (NPIs)

Pharmaceutical interventions (PIs)

- Pharmaceutical Interventions are medical interventions that target the pathogen or the host.
- Examples:
 - Vaccines,
 - Antiviral drugs, and
 - Antibiotics.

Vaccination



- Activates the host's immune system to produce antibodies against the pathogen.
- Generally applied to reduce the risk of infection and disease.
- The most effective way to prevent infectious diseases.

Challenges with vaccination

- Take time to develop for new pathogens
- Never 100% effective and limited duration of protection
- Adverse side effects
- Some individuals cannot be vaccinated or refuse vaccination
- Some pathogens mutate rapidly (e.g., influenza virus)
- Logistical challenges (e.g., cold chain requirements)

Non-pharmaceutical interventions (NPIs)



- Non-pharmaceutical interventions are measures that do not involve medical interventions.
- Examples:
 - Quarantine,
 - Physical/social distancing, and
 - Mask-wearing.

Quarantine

- *Isolation of individuals who may have been exposed to a contagious disease.*
- Advantage is that it's simple and its effectiveness does not depend on the disease.
- Disadvantages include:
 - Infringement on individual rights
 - Can be difficult to enforce
 - Can be costly
 - Can lead to social stigma

Contact tracing

- Contact tracing is used to identify exposed individuals, i.e., individuals who might have been in contact with an infected/infectious individual.
- It involves identifying, assessing, and managing people who have been exposed to a contagious disease to prevent further transmission.
- It is a critical component of infectious disease surveillance and is often used in combination with other control measures.

Discussion

- How can we quantify the impact of a control measure?

Any Questions?

What are infectious disease models?

- *Models* generally refer to conceptual representations of an object or system.
- *Mathematical models* use mathematics to describe the system. For example, the famous $E = mc^2$ is a mathematical model that describes the relationship between mass and energy.
- *Infectious disease models* use mathematics and/or statistics to represent dynamics/spread of infectious diseases.

- Mathematical models can be used to link the biological process of disease transmission and the emergent dynamics of infection at the population level.
- Models require making some assumptions and abstractions.

- By definition, “all models are wrong, but some are useful” (Box 1979).
 - Good enough models are those that capture the essential features of the system being studied.



Figure 2: George Box

Discussion

- What makes models “wrong” by definition?:

- What makes models “wrong” by definition?:
 - Models simplify reality
 - They cannot fully capture all the complexities of the system being studied.

Factors that influence model formulation/choice

- Accuracy: how well does the model to reproduce observed data and predict future outcomes?
- Transparency: is it easy to understand and interpret the model and its outputs?
 - This is affected by the model's complexity
- Flexibility: the ability of the model to be adapted to different scenarios.
 - For example, can the model be easily modified to account for new data or changing conditions?

What are models used for?

- Generally, models can be used to predict and understand/explain the dynamics of infectious diseases.

Discussion

How are these two uses impacted by accuracy, transparency, and flexibility? That is, which of these factors is more important for each use?

Prediction of the future course

- Must be accurate.
- “But the estimate proved to be off.

Way, way off. Like, 65 times worse than what ended up happening.”

The screenshot shows a news article from AP (Associated Press) titled "CDC's top modeler courts controversy with disease estimate". The article is by Mike Stobbe and was updated at 6:06 PM EST on August 1, 2015. It discusses how Martin Meltzer calculated that 1.4 million people might contract Ebola in West Africa, which received significant attention. The article includes a photo of a person in a white lab coat and a share button.

BY MIKE STOBBE
Updated 6:06 PM EST August 1, 2015

ATLANTA (AP) — Last fall, when Martin Meltzer calculated that 1.4 million people might contract Ebola in West Africa, the world paid attention.

This was, he said, a worst-case scenario. But Meltzer is the most famous disease modeler for the nation's pre-eminent public health agency, the Centers for Disease Control and Prevention. His estimate was promoted at high-level international meetings. It rallied nations to step up their efforts to fight the disease.

Understanding or explaining disease dynamics

- Models can be used to understand how a disease spreads and how its spread can be controlled.
- The insights gained from models can be used to:
 - inform public health policy and interventions.
 - design interventions to control the spread of the disease, for example, randomised controlled trials.
 - collect new data.
 - build predictive models.

Some key milestones in infectious disease modelling

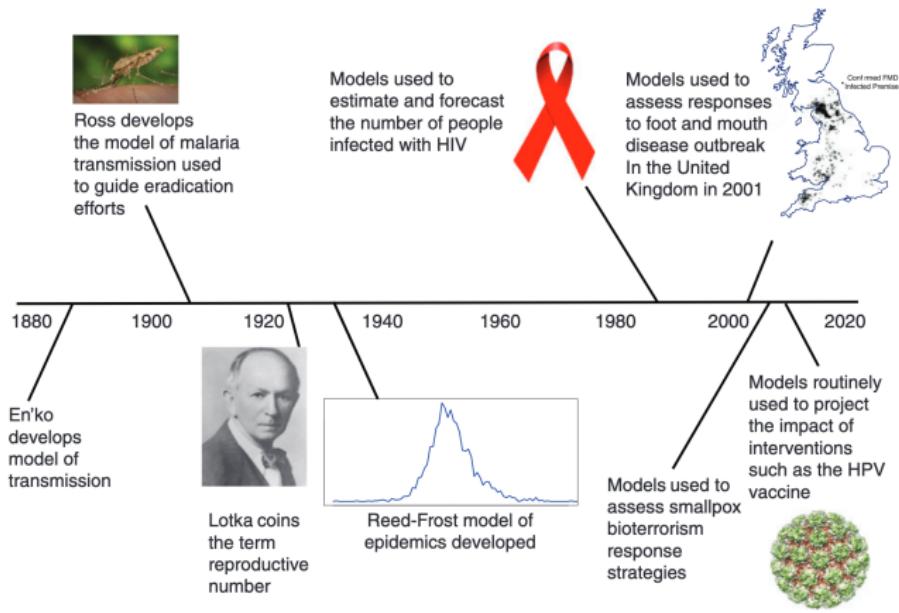


Figure 3: Milestones in mathematical modeling and modeling to inform public policy. Source: Lesslar et al., 2016

Limitations of infectious disease models

- Host behaviour is often difficult to predict.
- The pathogen often has known/unknown characteristics that are difficult to capture.
- Data is often not available or is of poor quality.

Summary

- Models:
 - Simplify reality and allow us to understand complex systems.
 - Simplifications of reality and do not capture all the complexities of the system being studied.
 - Only as good as the data used to calibrate them.
 - Sensitive to the assumptions made during their formulation.
 - Computationally expensive and require a lot of data to run.
 - Difficult to interpret and communicate to non-experts.

Any Questions?

Types of infectious disease models

Overview of infectious disease model types

- Infectious disease models can be classified in several ways:
 - Mathematical approach: mechanistic vs. phenomenological
 - Randomness: deterministic vs. stochastic
 - Time: discrete vs. continuous
 - Space: spatial vs. non-spatial
 - Population structure: homogeneous vs. heterogeneous

- The choice of model type depends on:
 - Research question being addressed
 - Available data and its quality
 - Computational resources and time constraints
 - Expertise of the modelling team
 - Intended audience and use case

Mathematical approach: Mechanistic vs. Phenomenological

Mechanistic models

- Model the underlying biological processes of disease transmission
- Include explicit representations of:
 - Host-pathogen interactions
 - Disease natural history
 - Population mixing patterns

Phenomenological models

- Focus on describing patterns in data
- Do not explicitly model biological mechanisms
- Often statistical in nature

Mathematical approach: Mechanistic vs. Phenomenological

Mechanistic models

- Examples: SIR, SEIR, agent-based models

Phenomenological models

- Examples: curve fitting, time series models, regression models

Mechanistic models: Advantages and disadvantages

Advantages:

- Can provide biological insights
- Allow testing of intervention scenarios
- Can extrapolate beyond observed data
- Parameters often have biological interpretation

Disadvantages:

- Require detailed knowledge of biological processes
- Can be complex to develop and validate
- May require extensive data for parameterization
- Computationally expensive

Phenomenological models: Advantages and disadvantages

Advantages:

- Simpler to develop and implement
- Good for short-term predictions
- Require less detailed data
- Computationally efficient

Disadvantages:

- Limited biological insight
- Poor extrapolation capabilities
- Cannot easily test interventions
- May not capture underlying dynamics

Discussion

Think of a recent epidemic (e.g., COVID-19, Mpox, Ebola). Which type of model would be more appropriate for:

1. Understanding transmission mechanisms?
2. Forecasting case numbers next week?
3. Evaluating vaccination strategies?

Deterministic vs. Stochastic models

Deterministic models

- Same input always produces same output
- Use differential equations or difference equations
- Represent average behavior of the system
- Examples: classical SIR/SEIR models

Stochastic models

- Include random variation
- Same input can produce different outputs
- Account for uncertainty and chance events
- Examples: stochastic SIR, branching processes

When to use deterministic vs. stochastic models?

Deterministic models

Best for:

- Large populations ($>10,000$)
- Understanding general trends
- Parameter estimation
- When computational efficiency is important

Stochastic models

Best for:

- Small populations ($<1,000$)
- Early epidemic phases
- Uncertainty quantification
- Extinction events

When to use deterministic vs. stochastic models?

Deterministic models

Limitations:

- Cannot capture fade-out in small populations
- No uncertainty quantification

Stochastic models

Limitations:

- Computationally intensive
- Require multiple runs
- More complex to implement

Time: Discrete vs. Continuous models

Discrete time models

- Time progresses in fixed steps (hours, days, weeks)
- Use *difference equations*
- Natural for surveillance data

Continuous time models

- Time flows continuously
- Use *differential equations*
- Better for theoretical analysis

Time: Discrete vs. Continuous models

Discrete time models

- Examples: daily case counts,
weekly reports

$$S_{t+1} = S_t - \beta S_t I_t$$

Continuous time models

- Examples: classical epidemic
models

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

Time scale considerations

- Disease natural history:
 - Acute diseases (days) → *discrete time* often sufficient
 - Chronic diseases (years) → *continuous time* may be better

- Data availability:
 - Daily surveillance data → *discrete time* natural
 - Continuous monitoring → either approach possible

- Research question:
 - Policy timing → *discrete time* practical
 - Theoretical insights → *continuous time* elegant

Spatial considerations

Non-spatial (well-mixed) models

- Assume homogeneous mixing
- All individuals have equal contact probability
- Simpler to analyze and implement
- Good for small, homogeneous populations

Spatial models

- Account for geographic structure
- Include movement patterns
- Can model local transmission
- Important for large-scale epidemics

Non-spatial (well-mixed) models

- Simpler to analyze and implement
- Good for small, homogeneous populations

Spatial models

- Can model local transmission
- Important for large-scale epidemics

- **Metapopulation models:**
 - Discrete patches connected by migration
 - Good for cities/regions connected by travel

- **Network models:**

- Individuals as nodes, contacts as edges
- Can represent social or transportation networks

- **Continuous space models:**
 - Individuals distributed in continuous space
 - Include local transmission kernels

Population structure: Homogeneous vs. Heterogeneous

Homogeneous models

- All individuals are identical
- Same contact rates
- Same susceptibility
- Simpler to analyze

Heterogeneous models

- Account for individual differences:
 - Age groups
 - Risk behaviors
 - Vaccination status
 - Geographic location
- More realistic but complex

Common types of heterogeneity

- **Age structure:** Different contact patterns and susceptibility by age
- **Risk groups:** High-risk vs. low-risk populations
- **Behavioral differences:** Compliance with interventions

- **Immunity status:** Vaccinated, previously infected, naive
- **Geographic:** Urban vs. rural, different regions

Network models

- Represent individuals as nodes
- Contacts/relationships as edges
- Can capture:
 - Social networks
 - Sexual networks
 - Transportation networks
 - Hospital networks

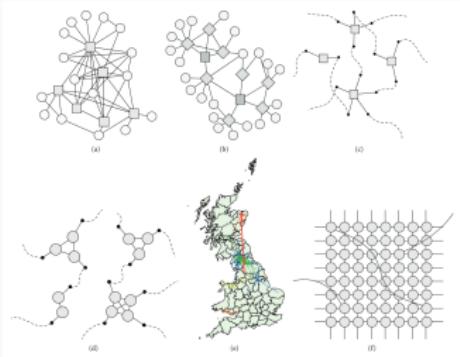


Figure 4: Examples of network models. Source: Danon et al., 2010

Network model advantages

- Capture realistic contact patterns
- Can model targeted interventions (e.g., contact tracing)
- Account for superspreading events
- Useful for outbreak investigation

Network model challenges

- Require detailed contact data
- Computationally intensive
- Network structure may be unknown or time-varying

Agent-based models (ABMs)

- Model individual agents and their interactions
- Each agent has:
 - Individual characteristics
 - Behavioral rules
 - Interaction patterns

Agent-based models (ABMs)

- System behavior emerges from individual interactions
- Very flexible but computationally intensive

Agent-based model applications

- **Household transmission:** Model within-household spread
- **School closures:** Impact on different age groups
- **Behavioral responses:** How people change behavior during epidemics

Agent-based model applications

- **Contact tracing:** Effectiveness of different strategies
- **Vaccination campaigns:** Optimal distribution strategies

Model selection framework

Key questions for model selection:

1. What is your research question?
2. What data do you have available?
3. What is your population size?
4. How important is spatial structure?
5. Do you need to model heterogeneity?
6. What are your computational constraints?
7. What is your timeline?

The modelling for decision-making process

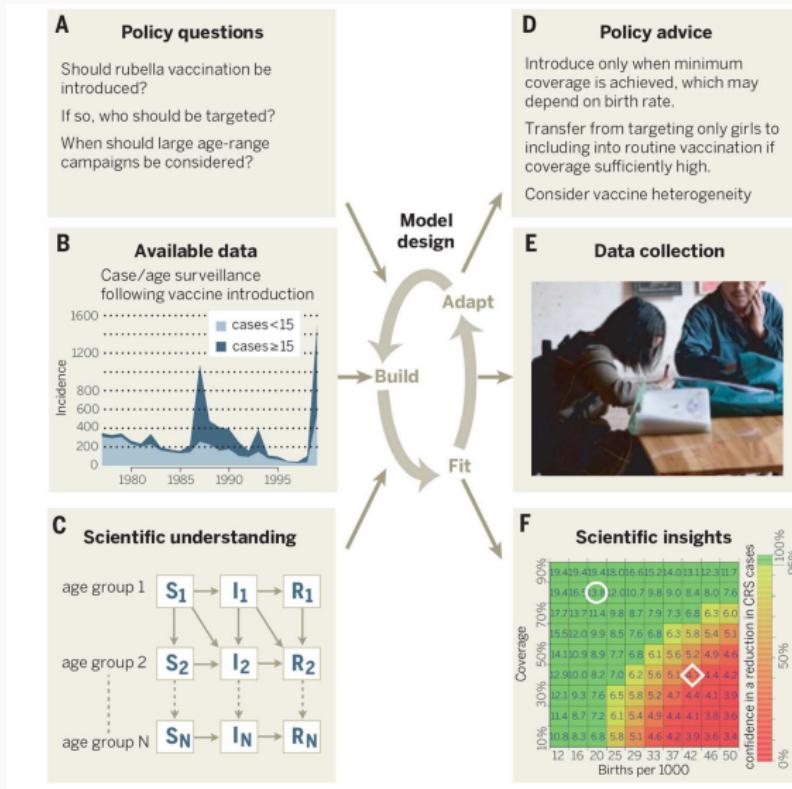


Figure 5: The modelling process involves balancing realism, tractability, and generality. Source: Heesterbeek et al, 2015

Discussion

Consider modeling influenza transmission in a university:

1. What type of model would you choose and why?
2. What heterogeneities would be important to include?
3. Would you use a spatial model?
4. Deterministic or stochastic?

Summary: Choosing the right model type

- **Start simple:** Begin with the simplest model that addresses your question
- **Add complexity gradually:** Only when justified by data or research needs
- **Consider your audience:** Policymakers may prefer simpler, more interpretable models
- **Validate carefully:** More complex models require more validation
- **Document assumptions:** Be explicit about model limitations

Common model progression

1. **Homogeneous compartmental** (SIR/SEIR) → understanding basic dynamics
2. **Add heterogeneity** → age groups, risk groups
3. **Add spatial structure** → metapopulation, networks
4. **Add stochasticity** → uncertainty quantification
5. **Individual-based** → complex behaviors, interventions

Next: Deep dive into compartmental models

- Now that we understand the landscape of infectious disease models...
- We'll focus on **compartmental models** - the most widely used approach
- Starting with the fundamental concepts and building up to specific models (SIR, SEIR)
- These form the foundation for most infectious disease modelling

Any Questions?

Introduction to compartmental models

What are compartmental models?

- Compartmental models:
 - divide populations into compartments (or groups) based on the individual's infection status and track them through time (Blackwood and Childs 2018).
 - are mechanistic, meaning they describe processes such the interaction between hosts, biological processes of pathogen, host immune response, and so forth.

Compartmental models are different from statistical models, which are used to describe the relationship between variables.

- Individuals in a compartment:
 - are assumed to have the same features (disease state, age, location, etc)
 - can only be in one compartment at a time.
 - move between compartments based on defined transition rates.

- Common compartments:
 - Susceptible (S) - hosts are not infected but can be infected
 - Infected (I) - hosts are infected (and can infect others)
 - Removed (R) - hosts are no longer infected and cannot be re-infected

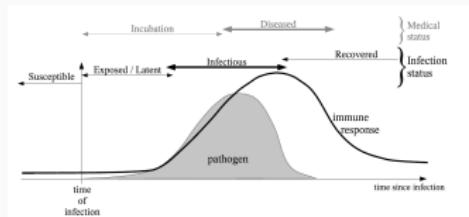


Figure 6: Infection timeline illustrating how a pathogen in a host interacts with the host's immune system
 (Source: Modelling Infectious Diseases of Humans and Animals)

- Other compartments can be added to the model to account for important events or processes (e.g., exposed, recovered, vaccinated, etc.)
- It is, however, important to keep the model simple, less computationally intensive, and interpretable.

- Compartmental models either have *discrete* or *continuous* time scales:
 - Discrete time scales: time is divided into discrete intervals (e.g., days, weeks, months).
 - Continuous time scales: time is continuous and the model is described using differential equations.

- Compartmental models can be *deterministic* or *stochastic*:
 - Deterministic models always return the same output for the same input.
 - Stochastic models account for randomness in the system and model output always varies. Hence, they are often run multiple times to get an average output.

- The choice of model type depends on:
 - the research question,
 - data availability,
 - computational resources,
 - modeller skillset.
- In this introduction, we will focus on deterministic compartmental models with continuous time scales.

- Now, back to the models, we are going to consider infections that either confer immunity after recovery or not.
- The simplest compartmental models for capturing this is the SIR model.

Any Questions?

The Susceptible-Infected-Recovered (SIR) model



This model groups individuals into three *disease states*:

- Susceptible (S): not infected but can be.
- Infected (I): infected & infectious.
- Recovered/removed (R): recovered & immune.

How do individuals move between compartments?

Process 1: Transmission



What drives transmission?

- Transmission is driven by several factors, including:
 - Disease prevalence, I , i.e., number of infected individuals at the time.
 - The number of contacts, C , susceptible individuals have with infected individual.
 - The probability, p , a susceptible individual will become infected when they contact an infected individual.



The transmission term is often defined through the force of infection (FOI), λ .

A tour of the force of infection (FOI)

- FOI, λ , is the per capita rate at which susceptible individuals become infected.

Note

“Per capita” means the rate of an event occurring per individual in the population per unit of time.

- Given the rate per individual per time, FOI, the rate at which new infecteds are generated is given by $\lambda \times S$, where S is the number of susceptible individuals.

- The force of infection is made up of the probabilities/rates that:
 - contacts happen, c ,
 - a given contact is with an infected individual, p , and
 - a contact results in successful transmission, v .

- The FOI can be formulated in two ways, depending on how the contact rate is expected to change with the population size:
 - Frequency-dependent/mass action transmission
 - Density-dependent transmission

Frequency-dependent/mass action transmission

The rate of contact between individuals is constant irrespective of the population density, $\frac{N}{A}$, where N is the population size and A is the area occupied by the population.

- Recall that transmission also depends on the probability of contact with an infected host, p , which is assumed to be $\frac{I}{N}$.
- Hence, the frequency-dependent mass action is given by $\lambda = \beta \times \frac{I}{N}$, where β is the transmission rate.

Question

Why does the frequency-dependent transmission contain $\frac{1}{N}$ if it does not depend on the population density?

- Assume that the rate of new infections is given by
$$\frac{dI}{dt} = S \times c \times p \times v$$
 where S is the number of susceptible hosts, c is the contact rate, and p is the probability of contact with an infected host, and v is the probability of transmission per contact.

- p is usually assumed to be the disease prevalence, $\frac{I}{N}$.
- Hence, the rate of new infections, $\frac{dI}{dt} = S \times c \times \frac{I}{N} \times v$.

- In frequency dependent transmission, the contact rate c is also assumed to be constant, say $c = \eta$ irrespective of population density, $\frac{N}{A}$, where N is the population size and A is the area occupied by the population.
- Hence, $\frac{dI}{dt} = S \times \eta \times v \times \frac{I}{N}$
- Therefore, $\frac{dI}{dt} = \beta S \times \frac{I}{N}$, where $\beta = \eta \times v$, and $\lambda = \beta \times \frac{I}{N}$.

- *Frequency-dependent/mass action transmission* is often used to model sexually-transmitted diseases and diseases with heterogeneity in contact rates.
- Sexual transmission in this case does not depend on how many infected individuals are in the population.

Density dependent transmission

- The rate of contact between individuals depends on the population density, $\frac{N}{A}$.
- Transmission also depends on p - the probability that a given contact is with an infected individual, often taken to be $\frac{I}{N}$.

- The density-dependent transmission is therefore given as
$$\lambda = \beta \times \frac{I}{A}.$$
- Here, because transmission increases with the density of infected individuals, it is called density-dependent transmission.

Note

- Notice that $\lambda = \beta \times \frac{I}{N} \times \frac{N}{A}$ and the N 's cancel out.
- A is often ignored.

- *Density dependent transmission* can be used to model airborne and directly transmitted diseases, for example, measles.

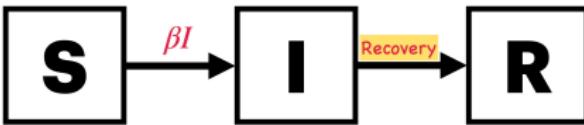
Density-dependent vs frequency dependent transmission

This is one of the most confused and debated concepts in disease modelling. Several studies have attempted to clarify it, including the brilliant work by Begon et al. (2002). Most of the clarifications provided here are based on this paper.

- We will only use the density-dependent formulation in this course.



Process 2: Recovery



Recovery, governed by the recovery rate, γ (rate at which infected individuals recover and become immune).

Some notes on the recovery process

- If the duration of infection is $\frac{1}{\gamma}$, then the rate at which infected individuals recover is γ .
- The average infectious period is often estimated from epidemiological data.

Note

You will learn about parameter estimation in the model fitting and calibration lectures.

Table 1. Epidemiological parameters and the average interepidemic period, T , for various diseases [condensed from (25), where a more extensive compilation of data is given; see also (36)].

Infectious disease	Latent period $1/\sigma$ (days)	Infectious period $1/\gamma$ (days)	Incubation period* (days)	Interepidemic period (years)		Geographical location	Time period
				Average	Range		
Measles	6 to 9	6 to 7	11 to 14	2.2	2 to 4	England and Wales New York City	1855 to 1979
				2.2	2 to 3		1928 to 1968
Whooping cough	6 to 7	21 to 23	7 to 10	3.0	2 to 5	England and Wales Baltimore, Maryland	1855 to 1979
				3.2	2 to 4		1928 to 1954
Poliomyelitis	1 to 3	14 to 20	7 to 12	4.0	3 to 5	England and Wales Finland	1950 to 1966
				4.2	2 to 5		1940 to 1972
Chicken pox	8 to 12	10 to 11	13 to 17	2.5	2 to 4	New York City Glasgow, Scotland	1928 to 1972
				3.0	2 to 4		1929 to 1972
Rubella	7 to 14	11 to 12	16 to 20	3.3	2 to 5	Glasgow, Scotland Baltimore, Maryland	1928 to 1964
				3.4	2 to 7		1928 to 1974
Mumps	12 to 18	4 to 8	12 to 26	3.0	2 to 4	Baltimore, Maryland New York City	1928 to 1973
				3.0	2 to 6		1928 to 1967
Diphtheria	2 to 5	14 to 21	2 to 5	5.1	4 to 6	England and Wales	1897 to 1979
Scarlet fever	1 to 2	14 to 21	2 to 3	4.4	3 to 6	England and Wales	1897 to 1978

*Time to the appearance of the symptoms.

Figure 7: Source: Anderson and May (1982)

Putting it all together



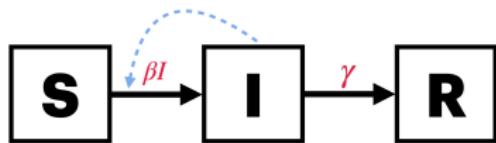
6

Figure 8: An SIR model with transmission rate, β , and recovery rate, γ .

Formulating the model equations

Continuous time compartmental models are formulated using differential equations that describe the change in the number of individuals in each compartment over time.

The SIR model can be formulated as:



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

where β is the transmission rate, and γ is the recovery rate.

The initial conditions are given by:

$$S(0) = N - 1,$$

$$I(0) = 1, \text{ and}$$

$$R(0) = 0.$$

where N is the total population size.

Note

- We represent the compartments as population sizes.:
 - Some modellers often use proportions instead of population sizes as a way to remove the dimensions from the equations.

Model assumptions

- The population is closed: no births, deaths, or migration.
 - Implicitly: the epidemic occurs much faster than the time scale of births, deaths, or migration.
- Individuals are infectious immediately after infection and remain infectious until they recover.

- Mixing is *homogeneous*, i.e., individuals mix randomly:
- Individuals have an equal probability of coming into contact with any other individual in the population.
- Transition rates are constant and do not change over time.
- Individuals acquire “lifelong” immunity after recovery.

Discussion

- What diseases do you think the SIR model is appropriate for?

- The SIR model is appropriate for diseases that confer immunity after recovery. For example, measles and chicken pox.
- Popularised by Kermack and McKendrick in 1927 (Kermack and McKendrick 1927)
 - A must-read paper for budding infectious disease modellers.

What questions can we answer with the SIR model?

- The SIR model can be used to understand the dynamics of an epidemic:
 - How long will the epidemic last?
 - How many individuals will be infected (final epidemic size)?
 - When will the epidemic reach its peak?

Solving the SIR model

- Compartmental models cannot be solved analytically.
- We often perform two types of analyses to understand the long term dynamics of the model:
 - Qualitative: threshold phenomena and analysis of equilibria (disease-free and endemic).
 - Numerical simulations.

i Note

- For this introductory course, we will focus on simulation.
- But first, let's do some qualitative analysis of the SIR.

Threshold phenomena

- Here, we study the conditions under which an epidemic will grow or die out using the model equations.

- Consider the case where $I(0) = 1$ individual is introduced into a population of size N at time $t = 0$.
- That means in a completely susceptible population, we have $S(0) = N - 1$ susceptible individuals.

- At time 0, the disease will not spread if the rate of change of infections is negative, that is $\frac{dI}{dt} < 0$.
- Recall from the SIR model that $\frac{dI}{dt} = \beta SI - \gamma I$.
- Let's solve this equation at $t = 0$ by setting $I = 1$, assuming $\frac{dI}{dt} < 0$.

At $t = 0$, we have

$$\frac{dI}{dt} = \beta SI - \gamma I < 0$$

Factor out I , and we get

$$I(\beta S - \gamma) < 0$$

- Since at $t = 0$, $I > 0$, we have $S < \frac{\gamma}{\beta}$.
- $\frac{\gamma}{\beta}$ is the relative removal rate.

- **Interpretation:**

- At $t = 0$, S must be less than $\frac{\gamma}{\beta}$ for the epidemic to die out.
- If the rate of removal/recovery is greater than the transmission rate, the epidemic will die out.
- Any infection that cannot transmit to more than one host is going to die out.

For the SIR model, the quantity $\frac{\beta}{\gamma}$ is called the reproduction number, R_0 (pronounced “R naught” or “R zero”).

The basic reproduction number, R_0

- The basic reproduction number, R_0 , is the average number of secondary infections generated by a single primary infection in a completely susceptible population.
- The basic reproduction number is a key quantity in infectious disease epidemiology.

- It is often represented as a single number or a range of high-low values.
 - For example, the R_0 for measles is popularly known to be 12 - 18 (Guerra et al. 2017).

- R_0 is often used to express the threshold phenomena in infectious disease epidemiology:
 - If $R_0 > 1$, the epidemic will grow.
 - If $R_0 < 1$, the epidemic will decline.
- A pathogen's R_0 value is determined by biological characteristics of the pathogen and the host's behaviour.

- Conceptually, R_0 is given by

$$R_0 \propto \frac{\text{Infection}}{\text{Contact}} \times \frac{\text{Contact}}{\text{Time}} \times \frac{\text{Time}}{\text{Infection}}$$

- R_0 is unitless and dimensionless.

Numerical simulations

- Numerical simulations can be performed with any programming language.
- This course focuses on the R programming language because:
 - It is a popular language for data analysis and statistical computing.
 - It has a rich ecosystem of packages for solving differential equations.
 - It is free and open-source.

- In R, we can use the `{deSolve}` package to solve the differential equations.
- To solve the model in R with `{deSolve}`, we will always need to define at least three things:
 - The model equations.
 - The initial parameter values.
 - The initial conditions (population sizes).

Practicals

- Let's do a code walk through in R using the script `sir.Rmd`.
 - Link: https://github.com/jamesmbaazam/mppr_r_practicals/

Discussion

- What happens if we increase or decrease the value of R_0 ?
- What happens if we increase or decrease the value of the infectious period?
- How can we flatten the curve?

Any Questions?

The Susceptible-Exposed-Infected-Recovered (SEIR) Model

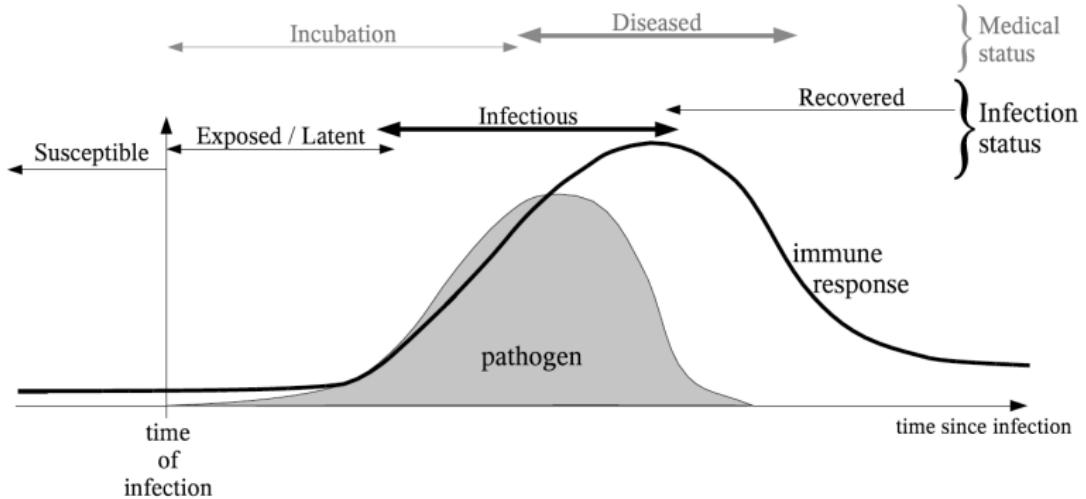


Figure 9: Timeline of infection. Source: Keeling & Rohani, 2008

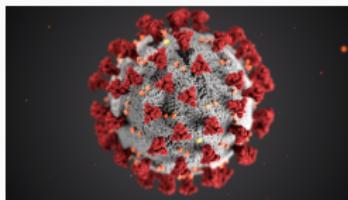
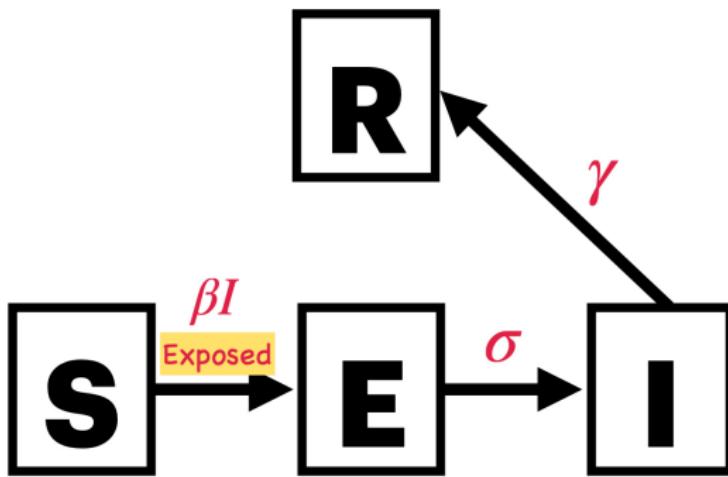


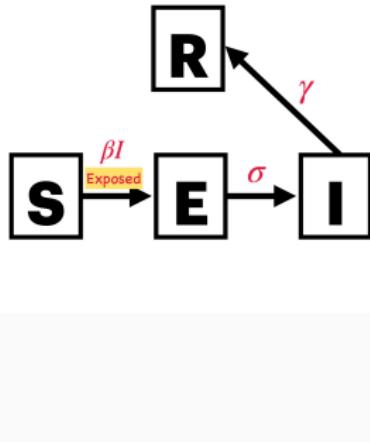
Figure 10: The SARS-CoV-2 virus

- Some diseases have an latent/exposed period during which individuals are infected but not yet infectious. Examples include pertussis, COVID-19, and Ebola.
- Disease transmission does not occur during the latent period because of low levels of the virus in the host.

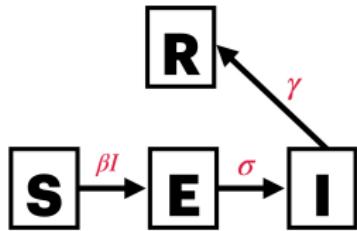


Figure 11: The Ebola virus





- The SEIR model extends the SIR model to include an exposed compartment, E .
- E : infected but are not yet infectious.
- Individuals stay in E for $1/\sigma$ days before moving to I .



Model equations:

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

$$\frac{dE}{dt} = \beta SI - \sigma E$$

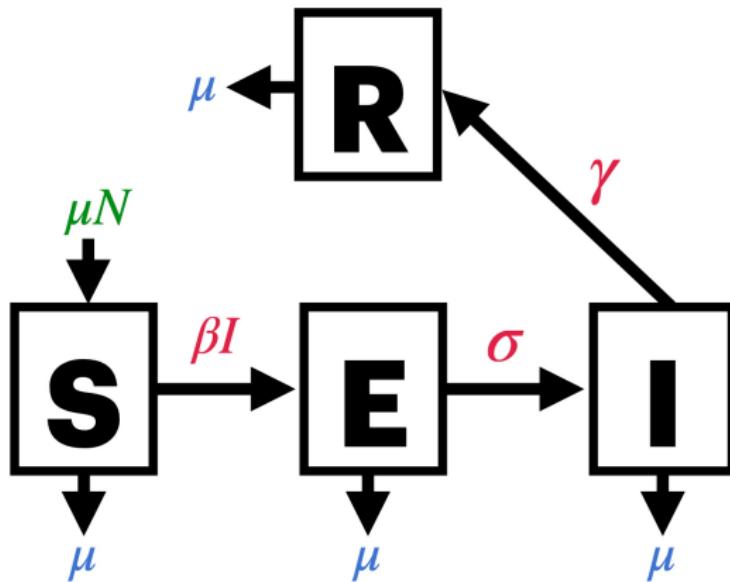
$$\frac{dI}{dt} = \sigma E - \gamma I$$

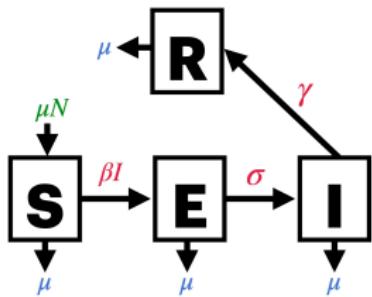
$$\frac{dR}{dt} = \gamma I$$

SEIR model with births and deaths

- Let's relax the assumption about births and deaths in the population.
- We will assume that the susceptible population is replenished with new individuals at a constant rate, μ .
- We will also assume that everyone dies at a constant rate, μ .

Our model schematic now looks like this:





The model equations now become:

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

$$\frac{dE}{dt} = \beta SI - \sigma E - \mu E \quad (2)$$

$$\frac{dI}{dt} = \sigma E - \gamma I - \mu I \quad (3)$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad (4)$$

Discussion

What is the R_0 of the SEIR model?

The R_0 of the SEIR Model

- Beyond the SIR model, calculating R_0 for more complex models can be challenging due to the presence of multiple compartments.
- For complex models, we use the next generation matrix approach (Diekmann, Heesterbeek, and Metz 1990; Diekmann, Heesterbeek, and Roberts 2010).

Note

Using the next generation matrix approach, we can show that the SEIR model with constant births and deaths has

$$R0 = \frac{\beta\sigma}{(\gamma + \mu)(\sigma + \mu)}$$

Numerical simulations

R Practicals

- We can use the same approach as the SIR model to simulate the SEIR model.
- Modify the `sir.Rmd` script to simulate the SEIR model.
 - Link: https://github.com/jamesmbaazam/mppr_r_practicals/

Any Questions?

Modelling epidemic control

i Note

For a recap on the various control measures, refer to Section 5.

Vaccination

For a background on vaccination, refer to Section 7.

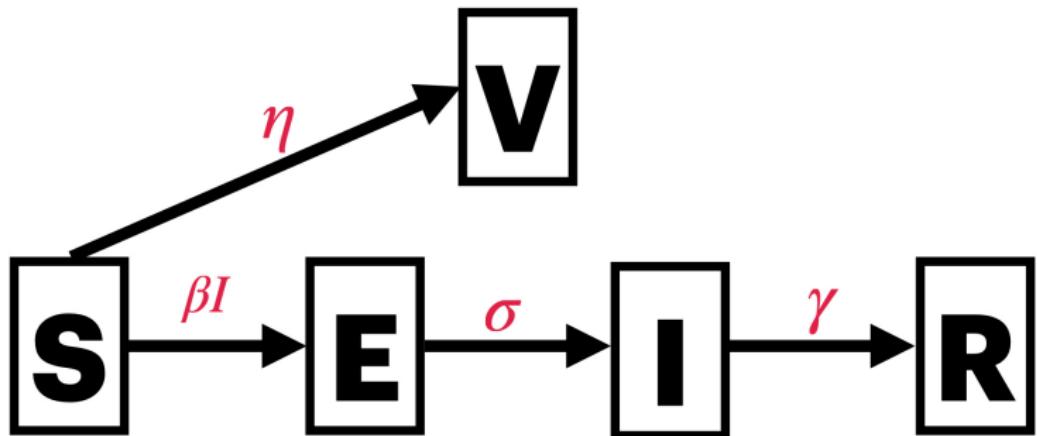


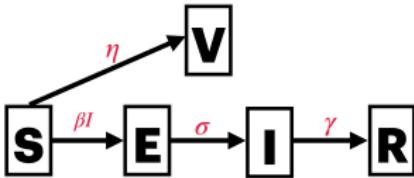
- Vaccination is one of the most effective ways to control infectious diseases.
- Conceptually, vaccination works to reduce the number of susceptible individuals, S .

- There are different types of vaccination strategies, including:
 - pediatric vaccination: vaccinating children to prevent the spread of diseases.
 - mass/random vaccination: vaccinating a large proportion of the population.
 - targeted vaccination: vaccinating specific groups of individuals, example, healthcare workers.
 - pulse vaccination: periodically vaccinating a large number of individuals.

- Let's consider the case of mass/random vaccination.
- Compartmental models can be extended to capture this by adding a new compartment, V .
- Let's consider the SEIR model with vaccination.

The Susceptible-Exposed-Infected-Recovered-Vaccinated (SEIRV) Model





- The SEIRV model is simply the SEIR model with a vaccinated compartment, V .
- The vaccinated compartment represents previously susceptible individuals who have been vaccinated and are immune to the disease.

- The vaccinated compartment is:
 - not infectious and does not move to the exposed or infectious compartments.
 - replenished by the rate of vaccination, η .

The model diagram and equations are as follows:

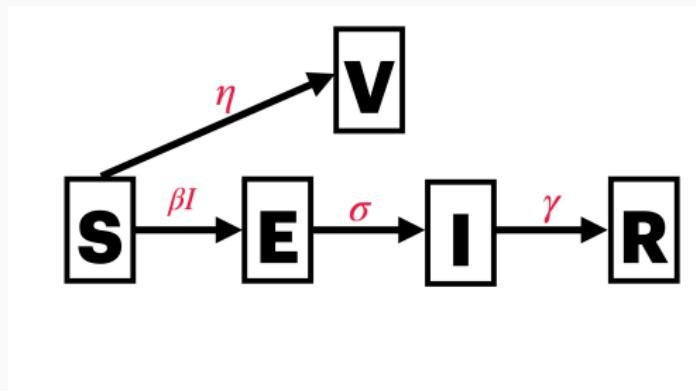
$$\frac{dS}{dt} = -\beta SI - \eta S \quad (5)$$

$$\frac{dE}{dt} = \beta SI - \sigma E \quad (6)$$

$$\frac{dI}{dt} = \sigma E - \gamma I \quad (7)$$

$$\frac{dR}{dt} = \gamma I \quad (8)$$

$$\frac{dV}{dt} = \eta S \quad (9)$$



where η is the rate of vaccination.

Discussion

- What are some of the assumptions of the SEIRV model?
- What are the implications of these assumptions for the model's predictions?

Numerical simulations

R Practicals

- We can use the same approach as the SIR and SEIR models to simulate the SEIRV model.
- Modify the `seir.Rmd` script to simulate the SEIRV model.
 - Link: https://github.com/jamesmbaazam/mppr_r_practicals/

Non-Pharmaceutical Interventions (NPIs)

For a background on non-pharmaceutical interventions, refer to Section 9.

- Conceptually, NPIs usually act to either reduce the transmission rate, β or prevent infected individuals from transmitting.
- NPIs like isolation, social distancing and movement restrictions can reduce the transmission rate, β , by reducing the contact rate between susceptible and infectious individuals.
- Hygiene measures reduce the probability of transmission per contact, thereby reducing the transmission rate, β , since $\beta = c \times p$.

- Let's consider two scenarios that will extend the SIR model to include NPIs:
 1. Modifying the transmission rate, β .
 2. Preventing infected individuals from transmitting through isolation.

Modifying the transmission rate

- NPIs such as social distancing, mask-wearing, and hand hygiene can reduce the transmission rate, β .
- We can model this by making the transmission rate a function of time, $\beta(t)$.

The modified SIR model with a reduced transmission rate is as follows:

$$\frac{dS}{dt} = -\beta(t)SI$$

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

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

- The simplest form is to reduce β by an NPI efficacy, say ϵ .
- Assuming the NPI is implemented between $t_{\text{npi_start}}$ and $t_{\text{npi_end}}$, it means that β remains the same before that period and is modified to $(1 - \epsilon)\beta$ during the period of the NPI, where $0 \leq \epsilon \leq 1$.
- With this knowledge, we can define $\beta(t)$ mathematically as:

$$\beta(t) = \begin{cases} \beta & \text{if } t < t_{\text{npi_start}} \text{ or } t > t_{\text{npi_end}} \\ (1 - \epsilon)\beta & \text{if } t_{\text{npi_start}} \leq t \leq t_{\text{npi_end}} \end{cases}$$

R Practicals

- Let's open the script file `sir_npi.R` and follow along.
 - Link: https://github.com/jamesmbaazam/mppr_r_practicals/

NPIs as compartments

- In the previous example, we retained the SIR model structure and modified the transmission rate.
- We can also model NPIs as compartments in the model. This is useful when we want to treat individuals affected by the NPIs differently.

- For example, isolation is an NPI that prevents infected individuals from transmitting the disease.
- This means that infected individuals in isolation do not contribute to the transmission of the disease and need to be removed from the infected compartment.
- Moving isolated individuals to a separate compartment allows us to track them separately in the model and apply relevant parameters.

- We can model this by introducing a new compartment, Q , for isolating infected individuals.
- Let's assume, infected individuals move to the isolated compartment at a rate, δ .
- Infected individuals in the isolated compartment do not transmit the disease.

The modified SIR model with isolated is as follows:

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

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

$$\frac{dR}{dt} = \gamma I + \tau Q$$

$$\frac{dQ}{dt} = \delta I - \tau Q$$

where δ is the rate at which infected individuals move to the isolated compartment, and τ is the rate at which individuals recover from isolation.

R Practicals

- We can use the same approach as the SIR model to simulate the model with isolation.
- Modify the SIR model in `sir.Rmd` to incorporate the isolated compartment, Q and the relevant parameters.
 - Link: https://github.com/jamesmbaazam/mppr_r_practicals/

Any Questions?

Brief notes on modelling host heterogeneity

- The models we have discussed so far assume that all individuals in the population are identical.
- However, in reality, individuals differ in their susceptibility to infection and their ability to transmit the disease.
- It is essential to capture this heterogeneity in the models in order for the models to be more realistic and useful for decision-making.

- This can be captured by incorporating heterogeneity into the models.
- Heterogeneity is often captured by stratifying the population into different groups.

Age structure

- For many infectious diseases, the risk of infection and the severity of the disease vary by age.
- Hence, it is essential to capture age structure in the models.
- To do this, we divide the population into different age groups and model the disease dynamics within each age group.
- Let's extend the SIR model to include age structure.

- We will divide the population into n age groups.
- Because the homogeneous model has 3 compartments, the age structured one will have $3n$ compartments:
 $S_1, I_1, R_1, S_2, I_2, R_2, \dots, S_n, I_n, R_n$.

- The (compact) model equations are as follows:

$$\frac{dS_i}{dt} = - \sum_{j=1}^n \beta_{ji} S_i I_j$$

$$\frac{dI_i}{dt} = \sum_{j=1}^n \beta_{ji} S_i I_j - \gamma I_i$$

$$\frac{dR_i}{dt} = \gamma I_i$$

- The model can be used to study the impact of age structure on the dynamics of the epidemic.
- For example, we can study the impact of vaccinating different age groups on the dynamics of the epidemic.

Other Heterogeneities

- Other forms of heterogeneity that can be incorporated into the models include:
 - Spatial heterogeneity
 - Temporal heterogeneity
 - Contact heterogeneity
 - Heterogeneity in host behavior

Any Questions?

Final Remarks

Takeaways

In the last two days, we have covered a lot of ground. Here are some key takeaways:

- Infectious diseases are a major public health concern that can have devastating consequences.
- Mathematical models are essential tools for studying the dynamics of infectious diseases and informing public health decision-making.
- Models are simplifications of reality that help us understand complex systems.

- We have discussed several compartmental models, including the SIR, SEIR, and SEIRV models.
- The SIR model is a simple compartmental model that divides the population into three compartments: susceptible, infected, and recovered.
- The SEIR model extends the SIR model by adding an exposed compartment.
- We can model various pharmaceutical and non-pharmaceutical interventions (NPIs) by modifying the transmission rate or adding new compartments.

- We have discussed the basic reproduction number, R_0 , which is a key parameter in infectious disease epidemiology.
- R_0 is the average number of secondary infections produced by a single infected individual in a completely susceptible population.
- If $R_0 > 1$, the disease will spread in the population; if $R_0 < 1$, the disease will die out.

- Deriving the basic reproduction number, R_0 , is an essential step in understanding the dynamics of infectious diseases.
- Deriving R_0 for the simple SIR model is simple as we just need to study the threshold phenomena.
- For more complex models, we can use the next-generation matrix approach to derive R_0 .

- Often, homogeneous models are not sufficient to capture the complexity of infectious diseases.
- Incorporating heterogeneity into the models is essential for capturing the complexity of infectious diseases.
- Age structure is a common form of heterogeneity that can be incorporated into the models.
- Other forms of heterogeneity include spatial, temporal, and contact heterogeneity.

What skills are needed to build and use infectious disease models?

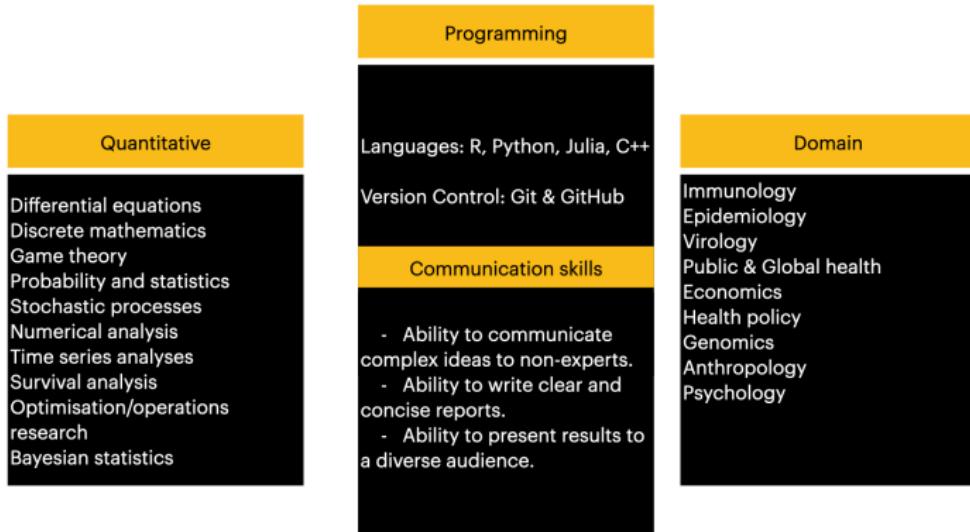


Figure 12: A non-exhaustive list of skills needed for modelling infectious diseases.

Contact Information

James Mba Azam, PhD

Epiverse-TRACE Initiative, London School of Hygiene and Tropical Medicine, UK

Email: james.azam@lshtm.ac.uk

ORCID: 0000-0001-5782-7330

Social Media:

LinkedIn: James Azam

Twitter: james_azam

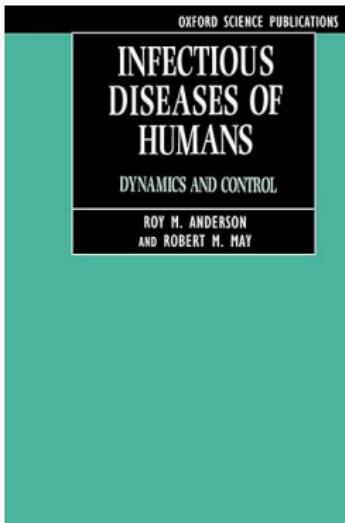
GitHub: jamesmbaazam



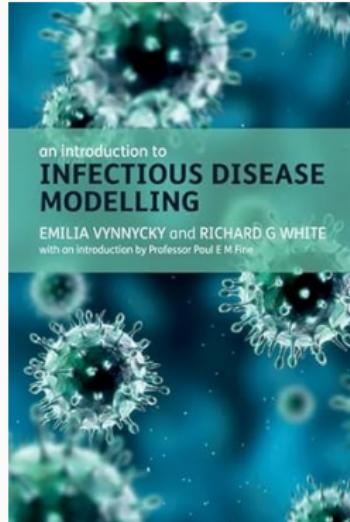
This presentation is made available through a
Creative Commons Attribution 4.0 International License

List of Resources

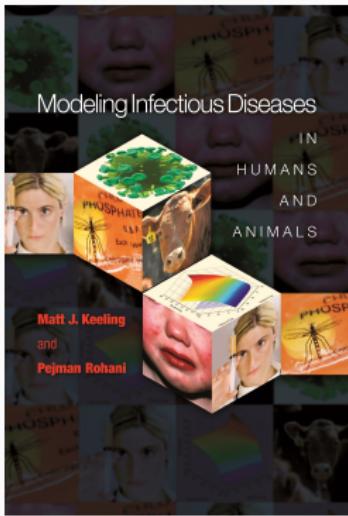
Textbooks



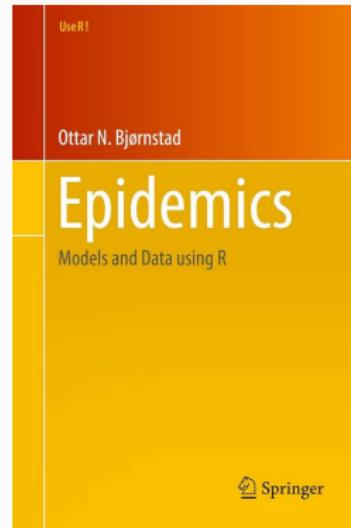
(a) Infectious Diseases of Humans:
Dynamics and Control by Roy M.
Anderson and Robert M. May



(a) Infectious Disease Modelling by
Emilia Vynnycky and Richard White



(a) Modeling Infectious Diseases in Humans and Animals by Matt Keeling and Pejman Rohani



(a) Epidemics: Models and Data Using R by Ottar N. Bjørnstad

Papers and Articles

Modelling infectious disease transmission

- Grassly, N. C., & Fraser, C. (2008). Mathematical models of infectious disease transmission. *Nature Reviews Microbiology*, 6(6), 477–487. <https://doi.org/10.1038/nrmicro1845>
- Kirkeby, C., Brookes, V. J., Ward, M. P., Dürr, S., & Halasa, T. (2021). A practical introduction to mechanistic modeling of disease transmission in veterinary science. *Frontiers in veterinary science*, 7, 546651. <https://www.frontiersin.org/articles/10.3389/fvets.2020.546651/full>

- Blackwood, J. C., & Childs, L. M. (2018). An introduction to compartmental modeling for the budding infectious disease modeler. <https://vttechworks.lib.vt.edu/items/61e9ca00-ef21-4356-bcd7-a9294a1d2f17>

- Cobey, S. (2020). Modeling infectious disease dynamics. *Science*, 368(6492), 713–714.
<https://doi.org/10.1126/science.abb5659>
- Bjørnstad, O. N., Shea, K., Krzywinski, M., & Altman, N. (2020). Modeling infectious epidemics. *Nature Methods*, 17(5), 455–456. <https://doi.org/10.1038/s41592-020-0822-z>

- Bodner, K., Brimacombe, C., Chenery, E. S., Greiner, A., McLeod, A. M., Penk, S. R., & Soto, J. S. V. (2021). Ten simple rules for tackling your first mathematical models: A guide for graduate students by graduate students. *PLOS Computational Biology*, 17(1), e1008539.
<https://doi.org/10.1371/journal.pcbi.1008539>
- Mishra, S., Fisman, D. N., & Boily, M.-C. (2011). The ABC of terms used in mathematical models of infectious diseases. *Journal of Epidemiology & Community Health*, 65(1), 87–94.
<https://jech.bmjjournals.org/content/65/1/87>

- James, L. P., Salomon, J. A., Buckee, C. O., & Menzies, N. A. (2021). The Use and Misuse of Mathematical Modeling for Infectious Disease Policymaking: Lessons for the COVID-19 Pandemic. *41*(4), 379–385.
<https://doi.org/10.1177/0272989X21990391>
- Holmdah, I., & Buckee, C. (2020). Wrong but useful—What COVID-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*.
<https://doi.org/10.1056/nejmp2009027>

- Metcalf, C. J. E. E., Edmunds, W. J., & Lessler, J. (2015). Six challenges in modelling for public health policy. *Epidemics*, 10(2015), 93–96.
<https://doi.org/10.1016/j.epidem.2014.08.008>
- Roberts, M., Andreasen, V., Lloyd, A., & Pellis, L. (2015). Nine challenges for deterministic epidemic models. *Epidemics*, 10(2015), 49–53.
<https://doi.org/10.1016/j.epidem.2014.09.006>

Deriving and Interpreting R₀

- Jones, J. H. (2011). Notes On R₀. Building, 1–19. <https://web.stanford.edu/~jhj1/teachingdocs/Jones-on-R0.pdf>
- Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R₀ in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4), 365–382. <https://doi.org/10.1007/BF00178324>

- Diekmann, O., Heesterbeek, J. A. P., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface*, 7(47), 873–885.
<https://doi.org/10.1098/rsif.2009.0386>

Code repositories

- epirecipes: Code for collate mathematical models of infectious disease transmission, with implementations in R, Python, and Julia.
- Modeling Infectious Diseases in Humans and Animals: Code for the labelled programs in the book “Modeling Infectious Diseases in Humans and Animals”. They are generally available as C++, Fortran and Matlab files.

References

- Anderson, Roy M, and Robert M May. 1982. "Directly Transmitted Infections Diseases: Control by Vaccination." *Science* 215 (4536): 1053–60.
- Begon, Michael, Malcolm Bennett, Roger G Bowers, Nigel P French, SM Hazel, and Joseph Turner. 2002. "A Clarification of Transmission Terms in Host-Microparasite Models: Numbers, Densities and Areas." *Epidemiology & Infection* 129 (1): 147–53.
- Blackwood, Julie C., and Lauren M. Childs. 2018. "An Introduction to Compartmental Modeling for the Budding Infectious Disease Modeler." *Letters in Biomathematics* 5 (1): 195–221. <https://doi.org/10.1080/23737867.2018.1509026>.
- Box, GE. 1979. "All Models Are Wrong, but Some Are Useful." *Robustness in Statistics* 202 (1979): 549.