

Introduction to Infectious Disease Modelling

James Mba Azam, PhD

2024-06-11

Epiverse-TRACE Initiative, 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*

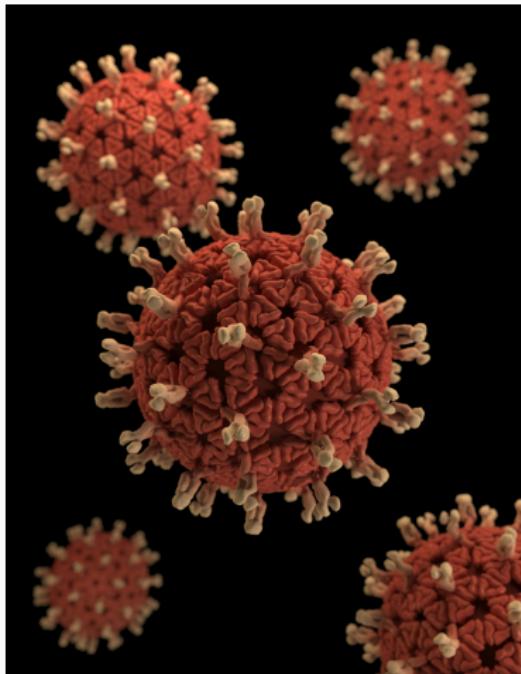


Figure 1: A 3D graphical representation of Rotavirus virions.

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)

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?

Infectious disease models

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/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.

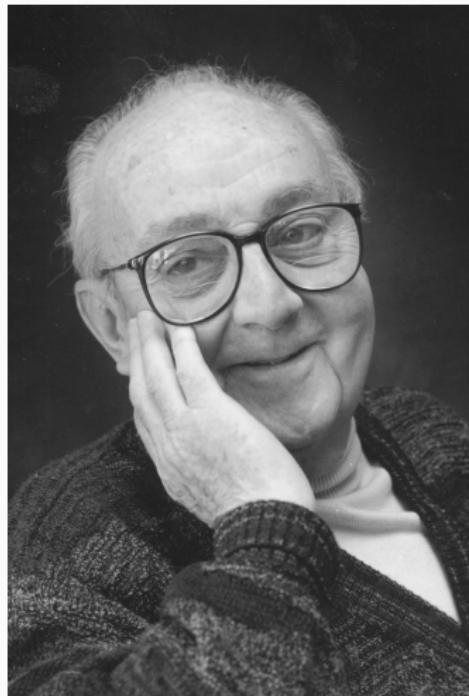


Figure 2: George Box

- What makes models “wrong” by definition?:
 - Simplifications of reality; not capturing 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.

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?

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, a disease modeler for the Centers for Disease Control and Prevention, 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 sidebar with links to other news stories like the Israel-Hamas war, Trump trial, France shooting, Pelosi in China, and Meghalaya's premieres. There is also a call to action to support independent fact-based journalism and a "DONATE" button.

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.

Limitations of infectious disease models

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

Summary

- Models:
 - Simplifications of reality and do not capture all the complexities of the system being studied.
 - Only as good as the data used to parameterize 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.

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 as 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

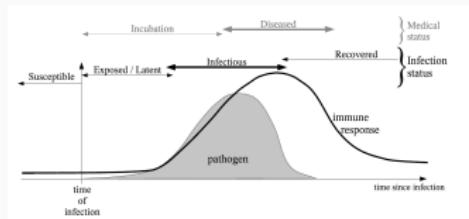


Figure 3: Infection timeline illustrating how a pathogen in a host interacts with the host's immune system
 (Source: Modelling Infectious Diseases of Humans and Animals)

- 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.

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	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	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	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	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	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	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 4: Source: Anderson and May (1982)

Putting it all together



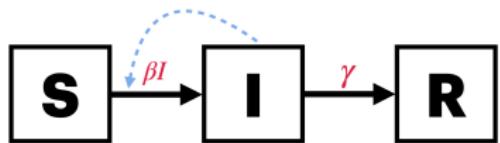
6

Figure 5: 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:



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

$$\frac{dI}{dt} = \dot{I} = \beta SI - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \dot{R} = \gamma I \quad (3)$$

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

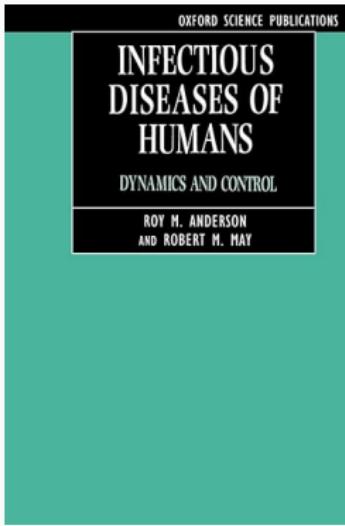
- Let's do a code walk through in R using the script
01_sir.Rmd.

Discussion

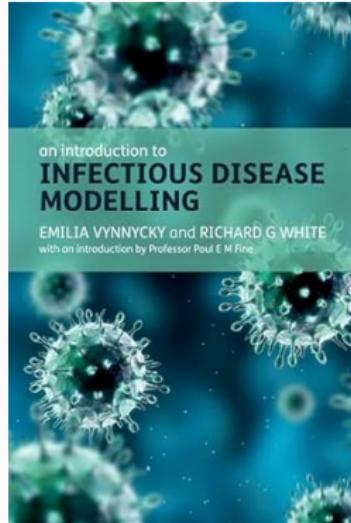
- What happens if we increase or decrease the value of R_0 ?
- What happens if we increase or decrease the value of the infectious period?
- How can we flatten the curve?

List of Resources

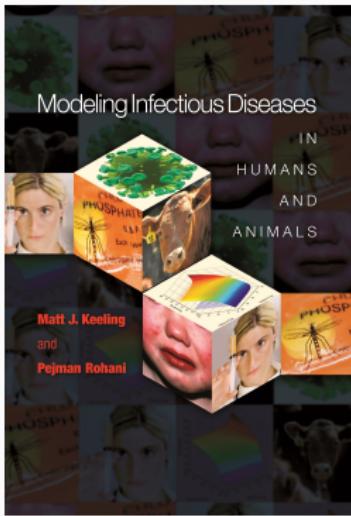
Textbooks



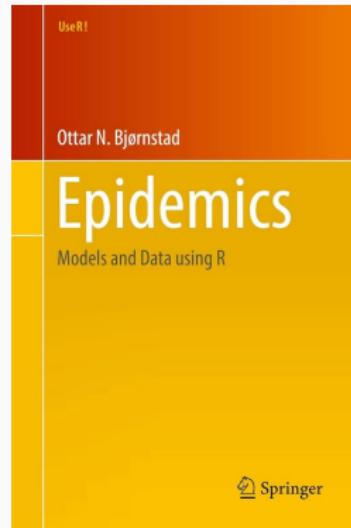
(a) Infectious Diseases of Humans:
Dynamics and Control by Roy M.
Anderson and Robert M. May



(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>

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.
- Guerra, Fiona M, Shelly Bolotin, Gillian Lim, Jane Heffernan, Shelley L Deeks, Ye Li, and Natasha S Crowcroft. 2017. "The Basic



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