**Introduction to infectious disease modelling: a beginner’s guide**

This guide aims to introduce infectious disease modelling, drawing on the principles, relevance, and basic mechanics discussed in the provided sources. It is intended for absolute beginners in the field.

**Section 1: What is Infectious Disease Modelling?**

**Definitions and goals of modelling:**

Infectious disease modelling involves using mathematical and statistical methods to understand the spread and transmission of infectious diseases (1). Disease spread models are simplified representations of real-life systems of disease transmission (2). They are also known as mechanistic models because they include explicit hypotheses about the biological mechanisms driving infection dynamics. Models are used to understand the dynamics of disease spread, predict disease transmission and its effects, identify outbreak scenarios, and evaluate the effectiveness of interventions (1). They help in guiding public health decisions and policies.

**Historical context:**

The classic SIR model is a foundational framework in epidemiological modelling. It was first introduced by Kermack and McKendrick in 1927 (3). Their work is described as seminal in compartmental modelling (4). The model proposed by Kermack and McKendrick is the starting point for the study of epidemic models described in some sources (5).

**Why we model diseases: understanding, forecasting, planning:**

Models help in understanding the spread and impact of infectious diseases (1). They allow for the prediction of disease transmission, forecast numbers of cases and deaths, and estimate final outbreak size (6). Modelling is essential for ecologists, veterinarians interested in infectious diseases, public health professionals, epidemiologists, and policymakers (3). Mathematical models are instrumental in simulating disease dynamics, allowing researchers to predict outbreak patterns and assess the impact of interventions. They can guide effective public health responses before real-world interventions are tested (3). Models can help policymakers make data-informed decisions that can save lives.

**Section 2: The Building Blocks of a Model**

**Key concepts:**

***Compartmental models:*** This is a common type of infectious disease model. They are called compartmental models because they divide the population into compartments based on the individual's status with respect to the disease (1). These compartments are defined based on the epidemiological characteristics of the disease.

***SIR, SEIR:***

The SIR model divides the population into three compartments: Susceptible, Infected, and Recovered. Diseases that confer immunity against re-infection can use the SIR structure, indicating passage from susceptible (S) to infective (I) to removed (R) (5). The recovered group (R) can include recovered individuals and deaths caused by the disease.

The SEIR model is an extension of the SIR model. It includes an additional compartment for Exposed individuals who have been infected but are not yet infectious (2). Exposed individuals are in the incubation period and may not show symptoms yet. This model is useful for understanding diseases with an incubation period. Other variations like SIS (no immunity) and SIRS (temporary immunity) also exist.

**Parameters:**

*Beta (β):* This represents the transmission rate or the probability of infection upon contact. It is the per capita rate at which two specific individuals come into effective contact per unit time. When the infected or susceptible populations are considered, beta causes the susceptible population to decrease and the infected population to increase (3).

*Gamma (γ):* represents the recovery rate, or the rate at which infectious individuals move to the recovered category. When the removed population is considered, gamma causes a portion of the infected population to move to the removed population. The recovery rate quantifies the transition from the infectious to the recovered state (1).

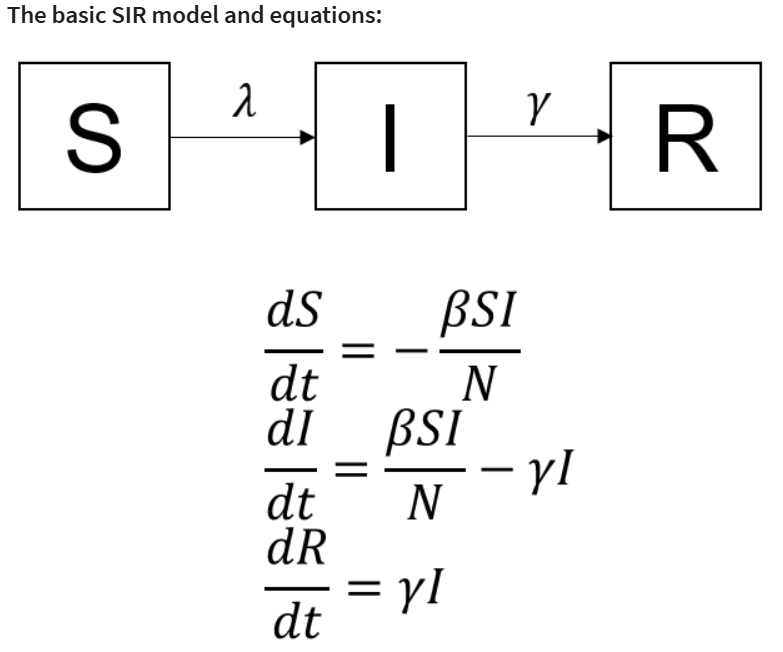
*Contact rate:* This is a component of the transmission rate. Assumptions about contact rates are crucial for modelling. Contact patterns vary significantly across different populations and affect disease spread (2).

*R₀ (Basic Reproduction Number):* This parameter is derived from the model. R₀ measures the average number of secondary cases generated by a single infected individual in a completely susceptible population (1). It is calculated from the SIR model as R₀ = beta / gamma. R₀ is used to determine the transmissibility of a disease and predict the potential for an outbreak. If R₀ > 1, the disease is likely to spread; if R₀ < 1, it may naturally decline and disappear. Reducing R₀ below one is necessary to prevent a disease from becoming endemic (3).

**Assumptions:**

Compartmental models often assume the mixing of members is homogeneous, sometimes called *mass action incidence* (5). This means each individual has the same susceptibility, infection, and removal chance. A basic SIR model assumes a closed population, meaning there are no births or immigration introduced to the population. It also assumes individuals only leave the population through death caused by the infection. A key assumption in cohort models, which can be related to compartmental models, is that future behaviour depends only on the present health state and transition probabilities, and past events are not remembered.

The flow chart structure below represents the basic SIR model. Individuals move from the Susceptible (S) compartment to the Infected (I) compartment, and then to the Recovered (R) compartment. Arrows indicate the transitions between states.

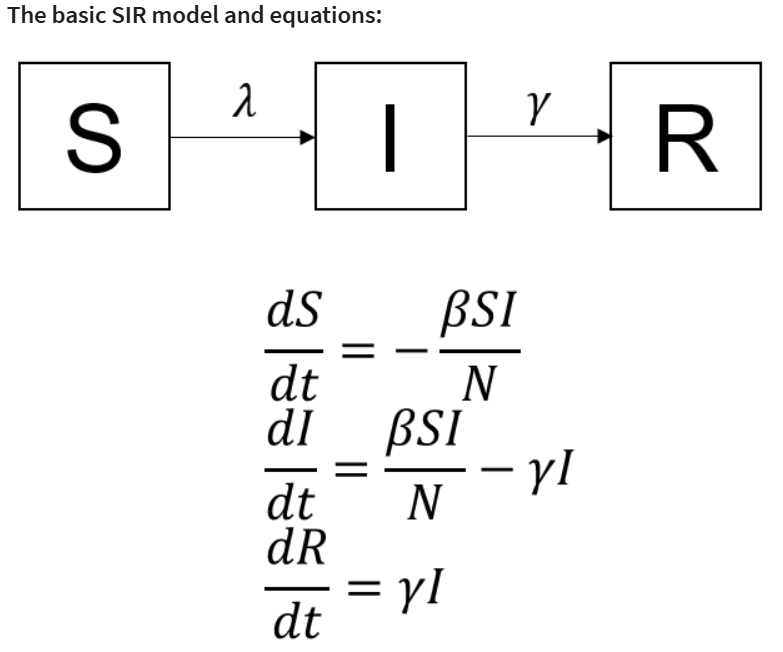


**Section 3: The SIR Model – A Gentle Walkthrough**

What does SIR stand for? SIR stands for Susceptible, Infected, and Recovered. These are the three compartments the population is divided into (7).

*Model structure (basic ODEs or flow diagram):*

The SIR model is a compartmental model. It is often represented by a system of ordinary differential equations (ODEs). The system of ODEs that represent the changes across the population is given by:



**Equations explained:**

*dS/dt:* This equation describes the rate of change of susceptible individuals over time. The number of susceptible individuals decreases because they become infected through contact with infectious individuals. The rate of decrease is proportional to the transmission rate (β) and the number of susceptible (S) and infected (I) individuals. The minus sign indicates a decrease.

*dI/dt:* This equation describes the rate of change of infected individuals over time. The number of infected individuals increases due to new infections (βSI, the same rate as susceptible individuals are leaving their compartment). The number of infected individuals decreases as they recover (at rate γ). The term −γI() represents the removal rate from the infected compartment.

dR/dt: This equation describes the rate of change of recovered individuals over time. The number of recovered individuals increases as infected individuals recover (at rate γ). The rate of increase is proportional to the recovery rate (γ) and the number of infected individuals (I).

**What does it simulate? (for instance, peak infections, outbreak duration):**

* + The SIR model describes the dynamics of a disease outbreak.
  + It can estimate the basic reproductive number (R₀).
  + It predicts the number of individuals in each compartment over time.
  + Simulation results can illustrate scenarios of disease spread and show the time series of susceptible, infectious, and recovered individuals. This implicitly simulates the rise and fall of cases (the epidemic curve, represented by the infected compartment I(t)) and the eventual number of removed individuals. It provides insights into the potential impact of a disease outbreak.

**A Simple Simulation**

*Build or run a basic SIR simulation:*

* + The sources provide examples of implementing infectious disease models, including the SIR model, in R.
  + Implementing the model involves creating code to represent the system of ODEs. This includes defining a function for the model equations that takes time, variables (compartment sizes), and parameters (β and γ) as inputs and returns the rate of change for each compartment.
  + The code then defines the initial conditions for the compartments and the parameter values (β, γ). A sequence of time points for the simulation is also created.
  + An ODE solver is used to solve the differential equations over the specified time period. The deSolve library in R is mentioned for solving ODEs.
  + The output from the ODE solver, which contains the compartment sizes at each time point, is typically stored in a data structure, such as a data frame.
  + Let learners modify β and γ: The parameter values for β and γ are defined at the beginning of the code. Changing these values before running the simulation would allow learners to explore the impact of different transmission and recovery rates on the epidemic dynamics. A specific example provides parameter values for COVID-19 (β=0.3, γ=0.4) for a case study.
* *Show graphs:*
  + The results of the simulation are typically plotted as time series graphs showing the number of individuals in each compartment over time. The ggplot2 library in R is mentioned for creating high-quality graphics and visualizations.
  + The plot function from the base R graphics library can also be used.
  + Legends are added to plots to identify the lines corresponding to susceptible, infectious, and recovered individuals.

References

1. Ali A. A Beginner’s Guide to Infectious Disease Modeling in R for Ecologists [Internet]. Available from: https://www.researchgate.net/publication/368141421

2. Kirkeby C, Brookes VJ, Ward MP, Dürr S, Halasa T. A Practical Introduction to Mechanistic Modeling of Disease Transmission in Veterinary Science. Front Vet Sci. 2021 Jan 26;7.

3. Misra S. Mathematical Modeling of Infectious Disease Spread Using the SIR Model. Biomed J Sci Tech Res [Internet]. 2024 Nov 12;59(3). Available from: https://biomedres.us/fulltexts/BJSTR.MS.ID.009302.php

4. Blackwood JC, Childs LM. An introduction to compartmental modeling for the budding infectious disease modeler. Lett Biomath. 2018 Dec 14;5(1):195–221.

5. Brauer F. Compartmental Models in Epidemiology.

6. Mac S, Mishra S, Ximenes R, Barrett K, Khan YA, Naimark DMJ, et al. Modeling the coronavirus disease 2019 pandemic: A comprehensive guide of infectious disease and decision-analytic models. J Clin Epidemiol. 2021 Apr 1;132:133–41.

7. Lord AG. Modeling COVID-19 Data using an SIR Model. Vol. 18, Citations Journal of Undergraduate Research. 2021.