Epidemic Modelling: An Introduction to the SIR Model

July 6, 2020

1 Introduction

Diseases which affect plants, animals and humans can be categorized on the basis of different criterion. Most often, they are categorized as infectious and non-infectious. Most models focus on studying directly transmitted infectious diseases caused by microparasites like viruses and bacteria. Mathematical models can be used to shape policy decisions and decide methods to control the spread of epidemics. These models are also categorized on the basis of their *accuracy* (which gives a measure of their usefulness, as higher accuracy implies that the model can accurately predict the long-term behaviour of an epidemic), their *transparency*, which determines how easy it is to understand how the different components of the model interact with each other, or their *transparency* (which gives a measure of the flexibility of the model, which is relevant when the model is used to evaluate control policies and adapt to new situations **(1)**).

Since every model is based on certain assumptions, no model can be completely accurate. However, a good model is that which is suited to its purpose, and can be different depending on context. A model is usually built by defining the level of the disease pathogen in the host, on the basis of which the population can be classified into different categories. An individual is said to be *susceptible*, when no pathogen is present. When the level of pathogen in the host is high, and the host can potentially transmit the infection to other susceptible individuals, they are said to be *infected*. Once the pathogen is cleared and the host is no longer infectious, they are said to be *recovered*. This classification forms the basis of the **SIR model**, which is the building block of most epidemiological models.

Introducing additional categories like *exposed* and *carrier*, or accounting for waning immunity, typically adds complexity to the basic SIR model and accounts for multiple real-world scenarios and diseases. These models can also be used to determine parameters such as basic reproductive rate (R_0) and average age of first infection.

2 The SIR Model

Ignoring population demography (like birth rate or death rate), we need to consider the rate at which people move from the $S \to I$ and from the $I \to R$ categories. The rate at which individuals move from I to R is called the recovery

rate γ, the inverse of which is the infectious period. The movement from S to I, which involves disease transmission, is influenced by rates of contact among the population, the proportion of already infected people, and the probability of transferring the pathogen, is called the transmission term.

The transmission term depends on the force of infection, λ, which is the per capita probability of getting the disease. If X is the number of people in S, the number of new infected people produced is λX. **(2)** There are two possible mechanisms of calculating the transmission term. The force of infection can be calculated as $\lambda = \beta Y/N$ where Y is the number of infected individuals, N is the total population size, and β is the product of contact rates and probability of transmitting infection for an individual. This formula is referred to as frequency dependent transmission or mass action. Frequency dependent transmission implies that the per capita probability of catching the infection is dependent on the *proportion* of infected individuals in the population.

The force of infection can also be calculated as $\lambda = \beta Y$, and this mechanism is referred to as density dependent transmission or pseudo-mass action. It implies that the per capita probability of catching the infection increases as the absolute number of infected individuals increase. The distinction between the two becomes relevant when the total population (N) varies. For notation purposes $S = X/N$ is the proportion of susceptible indivduals and $I = Y/N$ is the proportion of infected individuals. $R = Z/N$ is the proportion of recovered individuals, and $S + I + R = 1$

If frequency dependent transmission is considered, the rate at which new infecteds are produced (as a proportion of the total population) is βIS. Thus, the equations describing the model are **(3)**:

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

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

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

These equations have the initial conditions $S(0) > 0$, $I(0) > 0$ and $R(0) = 0$. These SIR equations ignore demography (like birth rate) and cannot be solved analytically. The progression generated from these equations is depicted (figure 1) **(4)**, which depicts a rapid conversion of susceptibles to infected.

3 R_0: Basic Reproductive Ratio

The expression for $\frac{dI}{dt}$ can also be written as

$$\frac{dI}{dt} = I\beta(S - \gamma/\beta)$$

The term β represents the product of number of contacts per unit time and the probability of transmitting the disease for a particular indvidual, and $1/\gamma$ is the infectious period, i.e the amount of time for which an infected person can potentially transmit the disease to susceptibles. Thus their product, β/γ represents the average number of secondary cases arising from an average primary case in

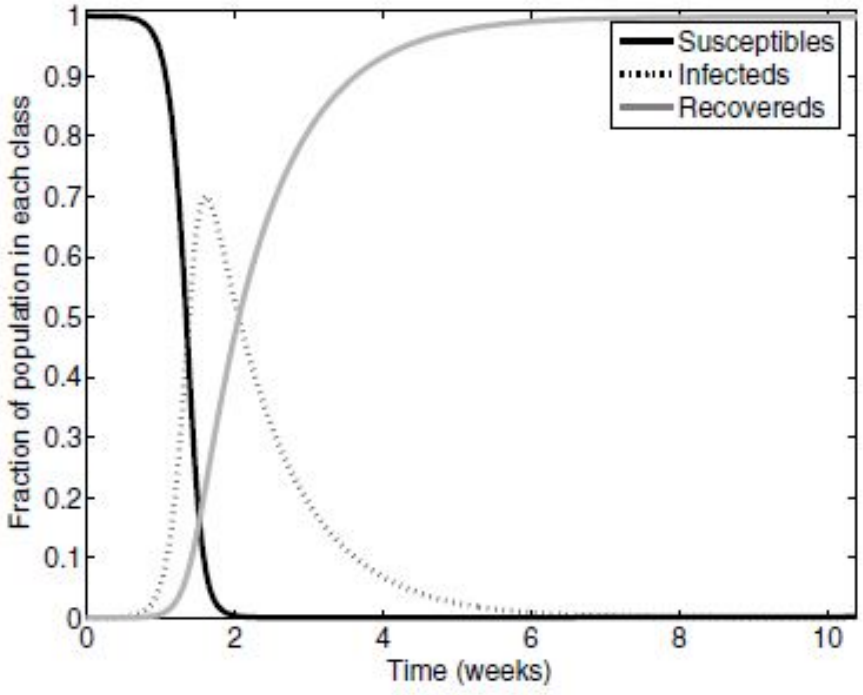

Figure 1: Time evolution of model variables

an entirely susceptible population. This term is called the **basic reproductive ratio** denoted by R_0.

From the above expression for $\frac{dI}{dt}$, it is clear that if $S < 1/R_0$, $\frac{dI}{dt} < 0$ and the infection "dies out". This is known as the threshold phenomenon **(5)**. Thus, if we aassume the entire population is susceptible (S=1), the value of $R_0 > 1$, otherwise the pathogen is unable to invade the population.

Dividing the equations for $\frac{dS}{dt}$ and $\frac{dR}{dt}$, we get,

$$\frac{dS}{dR} = -\beta S/\gamma$$

Replacing β/γ by R_0 and integrating both sides, we get

$$S(t) = S(0)e^{-R_0 R(t)}$$

Assuming R(0)=0, th number of susceptibles declines as the epidemic develops. Since $R \leq 1$, S remains above e^{-R_0}, thus some susceptibles always escape infection. If we assume that the epidemic ends when I=0, we can write **(6)**

$$S(\infty) + R(\infty) = 1$$

$$1 - R(\infty) = S(\infty) = S(0)e^{-R_0 R(\infty)}$$

$$1 - R(\infty) - S(0)e^{-R_0 R(\infty)} = 0$$

The equation is transcendental and we cannot obtain an exact solution, however it can be solved by numerical methods like the Newton-Raphson method.

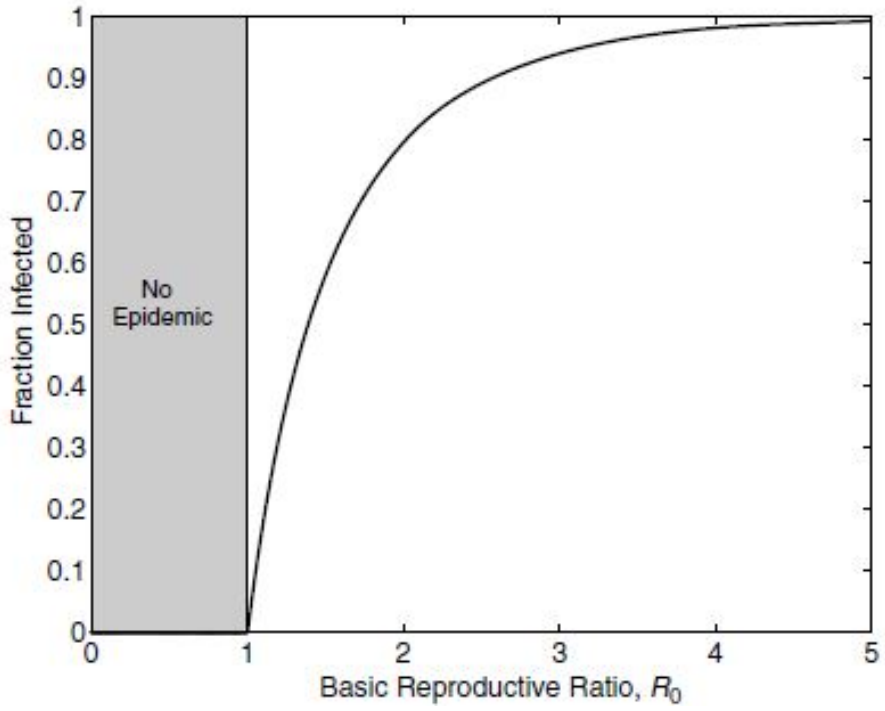

Figure 2: Fraction of population infected vs. R_0

When the equation is solved with the standard assumption S(0)=1, we see that no epidemic is possible when $R_0 < 1$ (figure 2)**(7)**. However, almost everyone (99%) gets infected when R_0 is close to 5, and some susceptibles still escape infection. Thus, the infection dies out due to a decline in the infectious population and not due to a lack of susceptibles **(8)**.

4 Other Models

4.1 SIR Model with demography

If the hosts' natural "lifespan" is assumed to be denoted by $1/\mu$ years, the natural mortality rate of the population becomes μ. By convention, the birth rate of the population is also considered to be μ, so that the total population size remains unchanged with time. Thus, the SIR model (called the generalised SIR model in this case) equations will be: **(9)**

$$\frac{dS}{dt} = \mu - \beta IS - \mu S$$

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

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

Since the period for which an infected individual can transmit the infection is reduced when accounting for natural mortality, R_0 also reduces, and

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

4.2 Mortality due to infection

If the disease is fatal, and the probability of dying during the infection is denoted by ρ, some people are removed from the I class due to death by infection, and the equations for $\frac{dS}{dt}$ and $\frac{dR}{dt}$ remain the same. The removal rate from class I is $(\gamma + \mu)$, from the generalized SIR model. The proportion of people removed due to the disease will be $\frac{\rho(\gamma+\mu)}{(1-\rho)}$. Thus, the expression for $\frac{dI}{dt}$ is: **(10)**

$$\frac{dI}{dt} = \beta IS - \frac{(\gamma + \mu)I}{(1 - \rho)}$$

Only those individuals who do not die during the infectious period will continue transmitting it to others for the entire infectious period, thus

$$R_0 = \frac{\beta(1 - \rho)}{(\gamma + \mu)}$$

4.3 Mortality late in infection

In the above model, if the mortality rate is high, the infectious period is drastically reduced. As ρ approaches one, R_0 drops to 0. Thus mortality can also be considered to take place at the end of the infectious period which gives rise to the following equations: **(11)**

$$\frac{dS}{dt} = \nu - \beta IS - \mu S$$
$$\frac{dI}{dt} = \beta IS - (\mu + \gamma)I$$
$$\frac{dR}{dt} = (1 - \rho)\gamma I - \mu R$$

Here,ν is a fixed birth rate, which is different from the death rate, as the total population size changes due to infection mortality. The expression for R_0 remains the same as the one in the generalised SIR model as the infectious period does not change.

4.4 Fatal Infections

Here, infecteds are assumed to be infectious for $1/\gamma$ time period, after which they succumb to the infection, thus giving rise to the SI models:

$$\frac{dS}{dt} = \nu - \beta IS - \mu S$$
$$\frac{dI}{dt} = \beta IS - (\gamma + \mu)I$$

The infected individual is assumed to be infectious and transmitting the disease for $1/\gamma$ time, thus the expression for R_0 remains the same as the one in the generalised SIR model.

4.5 The SIS Model

In several infections, an individual can be infected many times in their lives, with no apparent immunity. Thus, recovery from the infection is followed by an immediate return to the susceptible pool. This SIS model is described by a pair of coupled differential equations: **(12)**

$$\frac{dS}{dt} = \gamma I - \beta IS$$

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

This simplified model assumes no demogaphy, and $S + I = 1$. Thus, the expression for $R_0 = \beta/\gamma$.

4.6 Waning Immunity in the SIR Model

Some diseases like syphilis give rise to a immunity that diminishes over time, which is an intermediate case of the two extremes (the SIR model with full immunity after infection and the SIS model with no immunity after infection). If we consideer that immunity is lost at the rate "w", individuals move from the recovered class to the sussceptible class at "w" rate. This gives us the following equations:**(13)**

$$\frac{dS}{dt} = \mu + wR - \beta IS - \mu S$$

$$\frac{dI}{dt} = \beta IS - (\gamma + \mu)I$$

$$\frac{dR}{dt} = \gamma I - (w + \mu)R$$

The SIR model with absolute immunity is the case with w=0, where recovered individuals never move to the susceptible class, while the SIS model with no immunity after infection is the case where $w \to \infty$, where there is no R class. The expression for R_0 remains the same as the one in the generalised SIR model, as the infectioous period is not affected.

4.7 The SEIR Model

Adding a latent period where the host is infected with the pathogen, but cannot transmit it to other susceptible individuals gives rise to the SEIR model, and this latent period is accounted for by the addition of the "exposed" class. Individuals are assumed to be in this latent period for $1/\sigma$ time, after which they move into the infectious category. This gives rise to the following equations for the SEIR Model: **(14)**

$$\frac{dS}{dt} = \mu - (\beta I + \mu)S$$

$$\frac{dE}{dt} = \beta IS - (\mu + \sigma)E$$

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

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

We also assume that $S + E + I + R = 1$ and the espression for R_0 is:

$$R_0 = \frac{\beta\sigma}{(\mu + \sigma)(\mu + \gamma)}$$

The additional term in the denominator is due to the natural mortality of some individuals in the exposed class who do not contribute to transmission of the infection.

4.8 Infections with a Carrier state

In some infections, a proportion of infectious individuals become carriers, and transmit the disease at low rates for several years. Thus, a susceptible individual can be infected by either an infectious person or a carrier. If ϵ is the reduced rate of transmission from carriers, the following equations are obtained, with the introduction of the "carrier" class: **(15)**

$$\frac{dS}{dt} = \mu - (\beta I + \epsilon\beta C)S - \mu S$$

$$\frac{dI}{dt} = (\beta I + \epsilon\beta C)S - (\gamma + \mu)I$$

$$\frac{dC}{dt} = \gamma q I - (\Gamma + \mu)C$$

$$\frac{dR}{dt} = \gamma(1 - q)I + \Gamma C - \mu R$$

Here, Γ is the rate at which indviduals leave the carrier class (i.e individuals remain carriers for $1/\Gamma$ time after leaving the infectious class), C is the proportion of carriers in the population and q is the probability that an infectious indvidual becomes a carrier.

R_0 has two contributors, one from the infectious individuals which results in the same expression as the one obtained in the generalised SIR model, and from the carriers, who transmit the infection at low rate. If $\frac{\gamma}{(\gamma+\mu)}$ are the proportion of individuals who leave the infectious class and do not die due to natural mortality, $\frac{q\gamma}{(\gamma+\mu)}$ enter the carrier class, who transmit the infection at the rate of $\epsilon\beta$, and spend $\frac{1}{(\Gamma+\mu)}$ time in the infectious class. Thus, the expression for R_0 becomes:

$$R_0 = \frac{\beta}{(\gamma + \mu)} + \frac{q\gamma}{(\gamma + \mu)}\frac{\epsilon\beta}{(\Gamma + \mu)}$$

5 Epidemic Models and Global Health

Mathematical models have the potential to study complex infectious diseases and shape public health policy at the international level. Since transmission and infection processes can be similar across various pathogens, progress made in one area can build on and contribute to other areas as well.

For instance, the Ebola outbreak of 2014-15 revealed the various challenges involved in developing and using epidemiological models **(16)**.In the early stages

of the infection, despite a lack of data and under-reporting, models were successfully used to develop bed capacity and provide crucial international assistance. However, as the epidemic could not be controlled using traditional interventions, close co-ordination was necessary between modelers and policy makers to develop novel methods of control. Moving ahead, models were also used to inform clinical trials for the development of Ebola vaccines.

Novel pathogens which emerge through zoonosis **(17)** also pose a threat, and have contributed to the current global COVID-19 pandemic. The key areas which are investigated using models are the rates of human to human transmission, and evaluating the risk of the development of more dangerous zoonotic strains. Mathematical models can help design serological surveys, which can help in obtaining in accurate data. They are also used to ascertain which zoonotic diseases pose the most risk to humans.

Genetic sequences of pathogens can also be incorporated into these models as they influence the pathogen's transmission patterns and replication rates. This further influences the spread of the disease**(18)**. Several other factors, such as multiple infections within the same host, and the behaviour of hosts (which influences contact patterns and movement) can also be incorporated into these models, which can make them more accurate and reflective of the complexity of epidemics **(19)**.

6 References

1) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **7-8**

2) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **17**

3) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **19** Equations 2.1-2.3

4) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **20** Figure 2.1

5) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **20**

6) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **22**

7) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **22** Figure 2.2

8) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **22**

9) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **27** Equations 2.13-2.15

10) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **34** Equation 2.26

11) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **37-38** Equations 2.35-2.38

12) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **39** Equation 2.43-2.44

13) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **40** Equations 2.46-2.48

14) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **41** Equations 2.49-2.51

15) Keeling, Matt J., Rohani, Pejman *Modeling Infectious Diseases in Humans and Animals* Princeton University Press (2007) pp. **44** Equations 2.56-2.59

16) Hans Hesterbeek et al. Modeling infectious disease dynamics in the complex landscape of global health *Science* (2015)

17) Hans Hesterbeek et al. Modeling infectious disease dynamics in the complex landscape of global health *Science* (2015)

18) Hans Hesterbeek et al. Modeling infectious disease dynamics in the complex landscape of global health *Science* (2015)