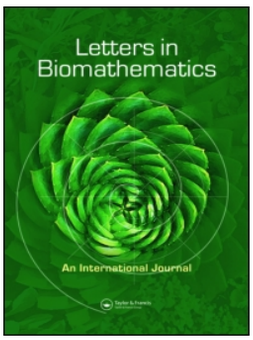


Letters in Biomathematics

ISSN: (Print) 2373-7867 (Online) Journal homepage: https://www.tandfonline.com/loi/tlib20

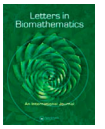
An introduction to compartmental modeling for the budding infectious disease modeler

Julie C. Blackwood & Lauren M. Childs

To cite this article: Julie C. Blackwood & Lauren M. Childs (2018) An introduction to compartmental modeling for the budding infectious disease modeler, Letters in Biomathematics, 5:1, 195-221, DOI: 10.1080/23737867.2018.1509026

To link to this article: https://doi.org/10.1080/23737867.2018.1509026

LETTERS IN BIOMATHEMATICS
2018, VOL. 5, NO. 1, 195–221
https://doi.org/10.1080/23737867.2018.1509026



Taylor & Francis
Taylor & Francis Group

EDUCATION ARTICLE

OPEN ACCESS Check for updates

An introduction to compartmental modeling for the budding infectious disease modeler

Julie C. Blackwood[a]* and Lauren M. Childs [b]*

[a]Department of Mathematics and Statistics, Williams College, Williamstown, Massachusetts; [b]Department of Mathematics, Virginia Tech, Blacksburg, Virginia

ABSTRACT

Mathematical models are ubiquitous in the study of the transmission dynamics of infectious diseases, In particular, the classic 'susceptible-infectious-recovered' (*SIR*) paradigm provides a modeling framework that can be adapted to describe the core transmission dynamics of a range of human and wildlife diseases. These models provide an important tool for uncovering the mechanisms generating observed disease dynamics, evaluating potential control strategies, and predicting future outbreaks. With ongoing advances in computational tools as well as access to disease incidence data, the use of such models continues to increase. Here, we provide a basic introduction to disease modeling that is primarily intended for individuals who are new to developing *SIR*-type models. In particular, we highlight several common issues encountered when structuring and analyzing these models.

ARTICLE HISTORY
Received 23 March 2018
Accepted 19 July 2018

KEYWORDS
Mathematical model; compartmental model; basic reproductive number; transmission

1. Introduction

Recent high profile infectious disease outbreaks around the globe – including SARS, Ebola, and Zika – have reignited a general interest in understanding the spread of disease. The utility of mathematical and computational approaches to simulate and numerically analyze infections within populations has also benefited from major advances in computational power over the past several decades. As a result, public health officials and policy makers have increasingly relied on a quantitative understanding of infectious disease dynamics.

Mathematical models of infectious disease dynamics have a rich history that dates back more than 100 years. Mathematically simple formulations that describe the transition of individuals in a population between 'compartments' that capture the infection status of individuals leads to surprisingly significant insight. Their elegance and simplicity allow the ease of expansion to more complexities through, for example, the addition of compartments. Expanding these models is often straight forward but the apparent simplicity

CONTACT J. C. Blackwood jcb5@williams.edu Department of Mathematics and Statistics, Williams College, Williamstown, Massachusetts
*These authors contributed equally.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

can mask subtle, but important, model structure and parameterization choices. Furthermore, while there are many wonderful texts focused on infectious disease modeling, there exist several complexities that are rarely discussed in sufficient detail for a novice disease modeler.

Here we highlight several of these potential pitfalls, describe how they arise and, when possible, how to address them. We also point to further resources in the literature. We begin with a brief introduction to disease modeling as a means to provide basic context as well as introduce the notation that we will use throughout this paper. We often provide more mathematical detail and discussion in sections in the Appendix.

Following our introduction, we focus on analysis of the compartmental epidemiological model with an emphasis on computing the basic reproductive number, commonly known as R_0. We focus on the following two difficulties: (i) deriving the force of infection and (ii) instances when R_0 does not represent a precise disease extinction threshold. Keep in mind, this is by no means meant to be a all-encompassing discussion of infectious disease modeling but a resource to supplement other more comprehensive texts (e.g. Anderson, May, & Anderson, 1992; Brauer, 2008; Brauer & Castillo-Chavez, 2001; Brauer & Kribs, 2015; Brauer, vanden Driessche, & Wu, 2008; Keeling & Rohani, 2008; Martcheva, 2015; Vynnycky & White, 2010).

2. A brief introduction to disease modeling

Mathematical models that describe the population-level dynamics of infectious diseases are typically based on the classic Susceptible – Infectious – Recovered (*SIR*)-type framework, a structure inspired by the seminal work of Kermack and McKenrick in 1927 (Kermack & McKendrick, 1927). In *SIR*-type models, individuals are separated into compartments based on their infectious status. Thus, any individual can only be in a single compartment at a given time, although they can transition between compartments. In this section, we briefly introduce and provide an analysis of a simple *SIR* model. We then generalize the model analysis so that it can be applied to a large suite of compartmental disease models.

2.1. Constructing compartmental models: the SIR framework

Compartmental disease models divide a population into groups (or compartments) based on each individual's infectious status and track the corresponding population sizes through time. Here, we first consider a model with straight forward properties: susceptible individuals (S) become infected through contact with infectious individuals (I), and infectious individuals recover (R) at a fixed rate γ and confer lifelong immunity. The fixed rate γ has the interpretation that $1/\gamma$ is the average amount of time spent in the infectious class (see Appendix 1 for a detailed explanation). Although there typically is some lag between acquiring infection and becoming infectious, here we do not distinguish between infected and infectious individuals; we instead assume that individuals are infectious immediately upon infection. This assumption will be relaxed later in Section 3.1. Almost all diseases have some delay between infection and infectiousness, however, when that delay is short (hours or days) and the goal is understanding the long term dynamics (years in the future), the assumption that individuals are immediately infectious may be reasonable. Examples of

some infections that are frequently modeled within the *SIR* framework include many common childhood diseases such as chicken pox and measles. An important common feature among these pathogens is that infection is thought to confer life-long immunity so that, from a modeling perspective, individuals move directly into (and remain in) the recovered compartment upon recovery.

Now, to model the infection process we consider a system of ordinary differential equations describing the change in the population size of each compartment. Let λ be the force of infection, or the *per capita* rate at which susceptible individuals acquire infection. It is important to note that the force of infection is not a constant but rather a function of the size of the infectious compartments; that is, $\lambda = \lambda(I)$.

The definition of $\lambda(I)$ can indeed be tricky and depends on assumptions about the nature of interactions between individuals (see Section 4.1). However, complications in the different ways to define the force of infection most often come into play when the total population size is changing, which is not the case at the moment. Thus, we will ignore this complexity for the time being. The force of infection is comprised of a transmission rate β or, more precisely, the product of the contact rate and the probability of transmission given contact, and an interaction term with infectious individuals. In order to preserve the units of the equation, using our definition of β (which has units $time^{-1}$), we will think about the interaction with the proportion of infectious individuals: I/N. Thus,

$$\lambda(I) = \beta \frac{I}{N}.$$

The definition of $\lambda(I)$ will change for each model, although it will always remain a function of the infectious class(es). Here, we have written $\lambda(I)$ as a function of the infectious class, but frequently the dependence is not explicitly included. in model formulation.

For now, we ignore demography, i.e. natural birth and mortality, but relax this assumption later. We can now write down a system of ordinary differential equations (ODEs) of the *SIR* model:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\lambda(I)S, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \lambda(I)S - \gamma I, \quad \frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I. \tag{1}$$

A schematic diagram of this model is provided in Figure 1(A). Here, the nature of our equations show us that the population is 'closed'; that is, the total population size (defined by $N = S+I+R$) is constant over time ($\mathrm{d}N/\mathrm{d}t = 0$). A nice property of this assumption is that the system is entirely determined by two, rather than three, of the equations. For example, we can write the recovered population as $R = N-S-I$ so that only susceptible and infected populations need to be determined explicitly.

A final comment before we move into the model analysis is that throughout this paper S, I, and R always represent population densities (although we will refer to population density and size interchangeably throughout this paper). However, some texts define S, I, and R as proportions of the population in each class, respectively. Non-dimensionalizing the state variables in this way for disease models is a common modeling choice, and the associated derivation is demonstrated in the Appendix (see Appendix 2). Our choice to consistently write the models in terms of densities, rather than proportions, will make it more straight forward to interpret some common pitfalls in disease modeling.

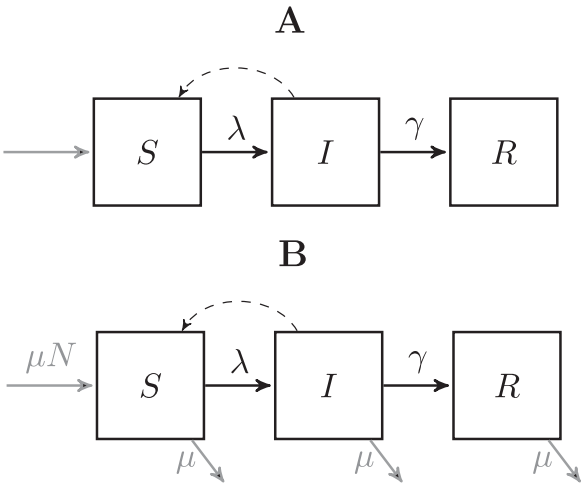


Figure 1. (A) Schematic representation of the standard *SIR* model in the absence of demography, as in Equation (1). *S*, *I*, and *R* represent the total number of susceptible, infected, and recovered individuals in a population, respectively. $\lambda(I)$ is the force of infection, and γ is the recovery rate. The dashed arrow indicates the necessity of interaction with the *I* compartment for transition from *S* to *I*. (B) Same as above, but this schematic includes demography in grey (as in Equation (2)).

2.2. Analysis of the SIR model

Many questions can be asked regarding the dynamics of an infectious disease: How quickly will the disease sweep through a population? How many people will be infected during the outbreak? Will the disease persist in the population? Compartmental ODE models such as the *SIR* framework introduced above can help answer these questions through analysis of the ODE systems.

2.2.1. Finding equilibria

A typical first step in analyzing a system of differential equations is finding the equilibria. At equilibrium we do not mean that individuals are fixed within particular compartments, but rather the rate of individuals entering and leaving a compartment is exactly balanced. This is equivalent to setting the right hand side of our equations, i.e. the rate of change of individuals into and out of each compartment, equal to zero.

In epidemiology, models generally have two important equilibria: (1) the disease-free equilibrium (DFE) and the endemic equilibrium. The DFE requires there to be no infected individuals in the population, or in the case of the *SIR* model $I^* = 0$, where the star designates an equilibrium solution. In contrast, the endemic equilibrium corresponds to the state in which infected individuals persist indefinitely such that $I^* > 0$.

In the model presented in Section 2.1, the canonical DFE is given by $(S^*, I^*, R^*) = (N, 0, 0)$; in other words, the entire population is susceptible to infection. For this model, there is no endemic equilibrium as the infectious population always returns to zero once the epidemic has completed (see Appendix 3). Biologically speaking, an endemic equilibrium requires that there is a continuous supply of susceptible individuals – known as 'susceptible replacement' – in order for infection to persist endemically. This replacement process can, for example, occur through births into the susceptible population or waning of

immunity. Without such replacement, the number of infected individuals returns to zero following an epidemic, as in the *SIR* model without demography (i.e. in Equation (1)). In this section, we therefore introduce a small modification to the *SIR* model of Section 2.1: demography, i.e. the inclusion of natural birth and death.

For mathematical simplicity, we make the following key biological assumptions: (i) the birth rate and death rate are equal and given by μ, (ii) all individuals are capable of reproducing and are equally subject to mortality, and (iii) all individuals are born susceptible to infection. The first two assumptions maintain a constant population size and the third assumption is made for simplicity (and is a valid assumption for many directly transmitted pathogens). It is possible to model infectious disease dynamics without these assumptions. For example, a discussion of models that include demography with variable population size can be found in Brauer (2008), Hethcote (2000), and Ledder (2017). The simple *SIR* model with demography is given by:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \mu N - \beta \frac{I}{N} S - \mu S, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \beta \frac{I}{N} S - (\gamma + \mu) I, \quad \frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - \mu R \tag{2}$$

and a schematic representation of this model is provided in Figure 1(B). For this model, we can find both a DFE and an endemic equilibrium, which are given by:

$$(S^*, I^*, R^*) = (N, 0, 0), \quad \text{and} \tag{3}$$

$$(S^*, I^*, R^*) = \left(\frac{\gamma + \mu}{\beta} N, \frac{\mu}{\beta} N \left(\frac{\beta}{\gamma + \mu} - 1 \right), N - S^* - I^* \right), \tag{4}$$

respectively.

Finding equilibria is only the first step to understanding long-term behaviour in a system. We must also determine which of the behaviours are typically realized. This involves determining the stability of the equilibria and will show us whether we will approach the equilibria or move away from it, assuming we have started nearby. In order to determine the stability of each of these equilibria, we can apply a standard linear stability analysis (more details can be found in any standard ODEs or linear algebra textbook). In brief, we linearize our system around each equilibria using the Jacobian matrix evaluated at the chosen equilibrium of interest. We then find the associated characteristic polynomial, whose roots are the eigenvalues to the Jacobian matrix. For local stability of the equilibrium, we require that all of the eigenvalues have negative real parts; otherwise, the equilibrium is locally unstable. In some cases, the value of the eigenvalues can be calculated directly. However, under more complicated scenarios an alternative method can be used to determine the sign of the eigenvalues (e.g. with a symbolic programming language or analytically with the Routh-Hurwitz criteria, see Brauer & Castillo-Chavez, 2001).

Using linear stability analysis, we can determine the local stability of each equilibrium. However, even for simple systems, this does not always give us a clear biological interpretation on when disease will spread or die out globally. Furthermore, when using more complicated models, there may be multiple stable equilibria present simultaneously. In such cases, it is important to obtain a more global picture of the infection dynamics using, for example, bifurcation analysis to supplement the linear stability analysis (as in Section 4.2.1).

In the standard *SIR* model with demography, however, we can use a standard linear stability analysis to identify conditions under which the DFE and endemic equilibrium are stable as well as unstable. As it turns out, stability is closely related to an important quantity in epidemiology: R_0, often called the basic reproductive number (or ratio). Therefore, we introduce R_0 before discussing the stability of this model.

2.2.2. The basic reproductive number: a biologically motivated introduction

The basic reproductive number, R_0, is arguably the most important quantity in disease modeling and is generally defined as the 'average number of secondary cases arising from a typical primary case in an entirely susceptible population'. There is rich mathematical theory that describes how this quantity can be computed for a range of *SIR*-type models with varying degrees of complexity. We will introduce some of this theory beginning in Section 3. Here, however, we find R_0 for the *SIR* model with demography in an intuitive way motivated by one question central to disease modeling: under what conditions will a disease die out?

To begin answering this question, we assume that the population at the initial time ($t = 0$) has one infected individual and is otherwise susceptible. Therefore, $I(0) = 1$ and $S(0) = N - 1$. The disease will fail to produce additional infections, i.e. fail to invade, if the rate of change in the infectious compartment is decreasing at time $t = 0$, or

$$\left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} < 0.$$

Recalling from Equation (2), the differential equation for I is given by

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta\frac{I}{N}S - (\gamma + \mu)I.$$

Now, we want to evaluate this equation at $t = 0$ by plugging in $S = N-1, I = 1$ and consider when this equation is negative:

$$\left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} = \beta\frac{1}{N}(N-1) - (\gamma + \mu) < 0$$

We rearrange this inequality to a form which provides biological intuition, and now shifts the threshold of interest to one (rather than zero when considering whether the size in the compartment will grow or shrink). Thus,

$$\Rightarrow \frac{\beta}{\gamma + \mu}\frac{N-1}{N} < 1.$$

In other words, the threshold is determined by the product of the transmission rate β with the quantity $1/(\gamma + \mu)$ (which is related to the average time spent in the infectious class and the average lifespan, see Appendix 1) and the initial fraction of susceptible individuals. Combined this must be 'small enough' (i.e. < 1) so that the epidemic cannot take off. As long as we are considering relatively large population sizes, we can look at the approximate

behaviour in the limit $N \to \infty$) and simplify such that

$$\frac{\beta}{\gamma + \mu} < 1$$

implies the disease will die out. A parallel analysis that begins with the question 'under what conditions will a disease persist?' reveals that

$$\frac{\beta}{\gamma + \mu} > 1$$

implies the disease will persist. This quantity, $\beta/(\gamma + \mu)$, is defined as the basic reproductive number for the *SIR* system. It is a threshold quantity where the 'critical value' is at $R_0 = 1$, and we are now better able to interpret our original definition: if a typical infectious individual infects less than one person on average during the time they are infected then a disease will die out. In contrast, if a typical infectious individual infects more than one person on average during the infectious period then the disease will spread. Intuitively, this makes sense: R_0 is the product of the transmission rate and – roughly – the average amount of time spent in the infectious class.

Notice that this definition includes several important assumptions: (i) R_0 is the average number of onward infections produced by a single individual (and therefore excludes individual variation), (ii) the average transmissibility among individuals is used in the definition of R_0 (again ignoring individual variation), and (iii) all individuals in the population are susceptible at time $t = 0$. In other more complicated settings, this definition changes such as when one of these assumptions is broken by the model structure (see Section 4.1.2 on structured populations). We note that the mathematical form of R_0 will differ for each model and numerically will depend on model parameters. However, almost all methods of computing R_0 require examination of when the DFE loses stability.

As you may expect, the way we have derived the basic reproductive number here by examining whether the infectious population grows or shrinks at time zero is only possible when the model framework is simple. However, there are multiple ways to determine R_0 in a more general setting, some of which we discuss later (see Section 3).

2.2.3. R_0 and equilibrium stability

We are now well-suited to determine the stability of both equilibria for the *SIR* model with demography. Now that we have found R_0 for this model, we can actually simplify the complicated endemic equilibrium (Equation (4)) by replacing the appearance of $\beta/(\gamma + \mu)$ with R_0:

$$(S^*, I^*, R^*) = \left(\frac{1}{R_0} N, \frac{\mu}{\beta} N \left(R_0 - 1 \right), N - S^* - I^* \right),$$

It is no coincidence that R_0 keeps appearing. In fact, this threshold is important in evaluating the stability of our equilibrium. As one might expect, as long as $R_0 < 1$ then the DFE is stable. More generally, we can use linear stability analysis, i.e. when the eigenvalues of the

Jacobian matrix evaluated at the DFE have negative real part, to show:

Disease-free equilibrium (DFE) is:
$$\begin{cases} \text{linearly stable} & \text{if } R_0 < 1 \\ \text{unstable} & \text{if } R_0 > 1 \end{cases}$$

Endemic equilibrium is:
$$\begin{cases} \text{unstable} & \text{if } R_0 < 1 \\ \text{linearly stable} & \text{if } R_0 > 1 \end{cases}$$

We emphasize that when using most standard methods for the computation of R_0, which involve linearization around the DFE, we can really only say something definitively about the local stability of the DFE. In other words, the derivation of R_0 typically depends on the stability of the DFE. As R_0 increases past the critical threshold value of one, the DFE is destabilized – thereby often, and indeed in the case of the *SIR* model with demography, this leads to stability of the endemic equilibrium. We will introduce an example later where this is not the case (see Section 4.2.1).

3. Methods for computing R_0

In the previous section, we defined the threshold R_0 for disease invasion simply by examining the sign of the differential equation describing the change in the infectious population for the *SIR* model with demography (see Section 2.2.2). We then described how a linear stability analysis can be used to describe equilibrium stability in terms of R_0 of the equilibria and observed how R_0 provides information on stability properties. See Appendix 4 for a more detailed example of linear stability analysis. This is straightforward for simple models, but as models increase in complexity they often contain multiple compartments with infectious individuals, making this process difficult. In this section we introduce a commonly used method for finding R_0: the next generation matrix method (Diekmann, Heesterbeek, & Johan, 1990; Diekmann, Heesterbeek, & Roberts, 2009). Importantly, this method is centered on considering the dynamics near the DFE.

3.1. Next generation matrix method

We introduce the next generation matrix method by demonstrating the process with a particular example where infected individuals are not immediately infectious. This is a much more biologically valid assumption for most diseases (e.g. Ebola, influenza, and pertussis), as individuals typically have some delay between infection and becoming infectious. However, the *SIR* model, where this delay is ignored, is often a reasonable approximation, especially when the duration of the prior to infectiousness is much shorter than the time spent infectious. Therefore, we now assume that there is a latent period prior to the onset of infectiousness in which individuals have been exposed to disease but are not yet capable of transmission. This class is commonly called the exposed, or E, class and the corresponding model is known as the *SEIR* model (Figure 2).

In particular, we let σ be the rate at which individuals leave the E class (recall that implies that $1/\sigma$ is the average amount of time spent in the E class), and we can extend the model in Equation (2) to:

$$\frac{dS}{dt} = \mu N - \lambda(I)S - \mu S, \tag{5}$$

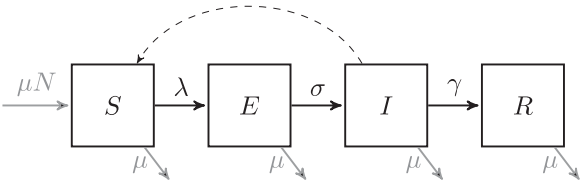


Figure 2. Schematic representation of the *SEIR* model. $\lambda(I)$ is the force of infection, σ is the rate at which individuals transition from the exposed class to the infectious class, γ is the recovery rate, and μ is the natural mortality rate. The dashed arrow indicates the necessity of interaction with the *I* compartment for transition from *S* to *E*.

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \lambda(I)S - (\sigma + \mu)E, \tag{6}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \sigma E - (\gamma + \mu)I, \tag{7}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - \mu R, \tag{8}$$

where $N = S+E+I+R$. In this model, the force of infection $\lambda(I)$ is the same as in the *SIR* model ($\lambda(I) = \beta I/N$). There are no new infections contributed through interactions with the exposed individuals, as they are not infectious. If we try to find R_0 using the same methods as we did for the *SIR* model, we run into trouble. More precisely, when examining the behaviour of I at $t = 0$, in contrast to the *SIR* model, we cannot factor out an I. Fortunately, the next generation matrix method allows us to generalize the process of computing R_0.

A nice feature of the next generation matrix method is that it only requires use of the DFE which is often easy to compute. In the *SEIR* model, the DFE is given by $(S^*, E^*, I^*, R^*) = (N, 0, 0, 0)$. After determining the DFE, we must create a sub-model that only considers the 'disease' compartments, a subset of the equations in the *SEIR* model. The disease compartments are those that include individuals that are in any stage of infection which, for the *SEIR* model, includes both the exposed and infectious individuals. Therefore, this sub-model will only contain the E and I equations (i.e. Equations (6)–(7)), which we write in the form

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \mathcal{F}(\vec{x}) - \mathcal{V}(\vec{x}) \tag{9}$$

where $\vec{x}$ is a vector of the j disease compartments; in the *SEIR* model $j = 2$ because the disease compartments are E and I. The right hand sides of Equations (6)–(7) are therefore contained in the vectors $\mathcal{F}(\vec{x})$ and $\mathcal{V}(\vec{x})$.

Here, $\mathcal{F}(\vec{x})$ contains any terms that directly lead to *new* infections entering each compartment j. Notice that the second element of F is zero because no new infections enter the I compartment, rather they transition from the E compartment into the I compartment. $\mathcal{V}(\vec{x})$, on the other hand, can be further broken down as

$$\mathcal{V}(\vec{x}) = \mathcal{V}^-(\vec{x}) - \mathcal{V}^+(\vec{x}) \tag{10}$$

where $\mathcal{V}^-(\vec{x})$ and $\mathcal{V}^+(\vec{x})$ contain all other outputs and inputs, respectively, from each disease class. This includes terms such as mortality or transition between classes. Although

mathematically it is not necessary to separate $\mathcal{V}(\vec{x})$, we introduce this notation here as it is commonly seen in the literature on the next generation method. For the *SEIR* model, we can now define the vectors in Equations (9)–(10) as

$$\begin{aligned}\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \begin{bmatrix}\frac{\mathrm{d}E}{\mathrm{d}t}\\ \frac{\mathrm{d}I}{\mathrm{d}t}\end{bmatrix} &= \mathcal{F}(\vec{x}) - \left(\mathcal{V}^-(\vec{x}) - \mathcal{V}^+(\vec{x})\right)\\ &= \begin{bmatrix}\beta\frac{I}{N}S\\ 0\end{bmatrix} - \left(\begin{bmatrix}(\sigma+\mu)E\\ (\gamma+\mu)I\end{bmatrix} - \begin{bmatrix}0\\ \sigma E\end{bmatrix}\right)\\ &= \begin{bmatrix}\beta\frac{I}{N}S\\ 0\end{bmatrix} - \begin{bmatrix}(\sigma+\mu)E\\ (\gamma+\mu)I - \sigma E\end{bmatrix} \qquad (11)\end{aligned}$$

Notice, this is exactly equivalent to our sub-model original equations (Equations (6) and (7)).

Next, we must linearize around the DFE, which involves the Jacobian, i.e. the matrix of partial derivatives, evaluated at the DFE. Recalling that the DFE is given by $(S^*, E^*, I^*, R^*) = (N, 0, 0, 0)$, the Jacobian matrix of the sub-model evaluated at the DFE is given by

$$J(S^*, E^*, I^*, R^*) = \begin{bmatrix}\frac{\partial E}{\partial E} & \frac{\partial E}{\partial I}\\ \frac{\partial I}{\partial E} & \frac{\partial I}{\partial I}\end{bmatrix}\Bigg|_{N,0,0,0} \qquad (12)$$

$$= \begin{bmatrix}-(\sigma+\mu) & \beta\frac{S}{N}\\ \sigma & -(\gamma+\mu)\end{bmatrix}\Bigg|_{N,0,0,0} \qquad (13)$$

$$= \begin{bmatrix}-(\sigma+\mu) & \beta\\ \sigma & -(\gamma+\mu)\end{bmatrix}. \qquad (14)$$

As a result, we can factor out the vector $\vec{x} = \left[\begin{smallmatrix}E\\ I\end{smallmatrix}\right]$ on the right hand side, leaving us with

$$\begin{bmatrix}\frac{\mathrm{d}E}{\mathrm{d}t}\\ \frac{\mathrm{d}I}{\mathrm{d}t}\end{bmatrix} = \left(\begin{bmatrix}0 & \beta\\ 0 & 0\end{bmatrix} - \begin{bmatrix}\sigma+\mu & 0\\ -\sigma & \gamma+\mu\end{bmatrix}\right)\begin{bmatrix}E\\ I\end{bmatrix}. \qquad (15)$$

Notice that in the **F** matrix the only non-zero term β arises from the derivative of the first row of $\mathcal{F}$ with respect to I, evaluated at $S = N$. To simplify, we define matrices in Equation (15) as **F** and **V**, respectively, and rewrite Equation (15) as

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = (\mathbf{F} - \mathbf{V})\vec{x}. \qquad (16)$$

Recall from Section 2.2.2 that the ultimate goal is to determine when the number of infectious individuals, I, is increasing or decreasing when the population is near the DFE,

i.e. when a single infected individual is introduced into an otherwise susceptible population. We can now investigate a similar question: when are the compartments containing the infectious classes increasing or decreasing near the DFE? From linear stability analysis we know that stability of the DFE is equivalent to the real part of all eigenvalues of $(\mathbf{F} - \mathbf{V})$ being less than zero. From Equation (16), we can extract what is known as the next generation matrix, or

$$\mathbf{F}\mathbf{V}^{-1}.$$

Here, an analogous threshold (to the eigenvalues of the Jacobian) for when the DFE is stable is given by the spectral radius of the next generation matrix denoted by

$$R_0 = \rho\left(\mathbf{F}\mathbf{V}^{-1}\right) < 1.$$

where the spectral radius (ρ) is the eigenvalue with the largest magnitude. While we omit the proof here, it can be found in other sources, e.g. van den Driessche and Watmough (2002).

For the *SEIR* model, the next generation matrix is given by

$$\mathbf{F}\mathbf{V}^{-1} = \begin{bmatrix} \dfrac{\beta\sigma}{(\sigma+\mu)(\gamma+\mu)} & \dfrac{\beta}{\gamma+\mu} \\ 0 & 0 \end{bmatrix}.$$

We find the eigenvalues of this matrix are $\beta\sigma/(\gamma+\mu)(\sigma+\mu)$ and 0. Here, R_0 (or the spectral radius of $\mathbf{F}\mathbf{V}^{-1}$) is given by

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

For the *SEIR* model, notice that R_0 is very similar to the basic reproductive number in the *SIR* model with demography, $(\beta/(\gamma+\mu))$, multiplied with one additional term, $(\sigma/(\sigma+\mu))$. This term leads to a smaller R_0 for the *SEIR* model because some individuals are lost to natural mortality while in the *E* class and never actually reach the infectious class.

While we provided a brief introduction to the next generation matrix in this section using a particular example (the *SEIR* model), more formal mathematical treatments of this can be found in other references, e.g. Diekmann and Heesterbeek (2000), Martcheva (2015) and van den Driessche and Watmough (2002). We also note that analytically finding the roots of the characteristic polynomial (and thereby finding R_0) can be challenging; it is common to use a symbolic mathematical programme (e.g. Mathematica) or to find it numerically (e.g. using MATLAB). Furthermore, a great deal of tractability, both mathematically and biologically, may be gained by first non-dimensionalizing the model. Such simplifications may reveal biologically relevant quantities that were obscured by relying solely on the mathematical computation. We refer the reader to Ledder (2017) for a more detailed description of non-dimensionalization of models.

3.1.1. A cautionary tale of computing R_0

The next generation matrix method is a widely used method for computing R_0 for a variety of compartmental models that have varying degrees of complexity. There are, however,

instances in which the construction of a model is not amenable to this – or other – standard methods for analytically finding R_0.

One common example in which R_0 cannot be directly computed is when there is 'importation'. Suppose a particular pathogen follows the standard *SEIR* structure with demography as in Equations (5)–(8). Now, assume that there is some external source of infection that does not depend on the number of infected individuals present in the population. For example, an infected individual from another population may briefly encounter a susceptible individual and transmit infection. If we assume the importation rate is given by ϕ, then this modifies the force of infection so that

$$\lambda(I) = \beta \frac{I}{N} + \phi.$$

The S and E equations change to ensure that the total population size remains constant.

While R_0 cannot be computed explicitly for a model containing infectious imports, some modeling studies still attempt to take advantage of the information R_0 provides in similar standard models. For example, R_0 can first be computed for the model in the absence of imports (i.e. when $\phi = 0$). This R_0 then represents an 'intrinsic' basic reproductive number; that is, R_0 in the absence of any spatial processes. The dynamical impact of importation can then be explored numerically through simulations (e.g. Blackwood, Streicker, Altizer, & Rohani, 2013). Other models with importation focus on other aspects of the model all together, such as direct parameterizations that are independent of the precise R_0 value itself (e.g. Shrestha et al., 2015, 2013).

3.2. Other methods for finding R_0

The next generation matrix is the most commonly used method for finding R_0 in standard disease ecology studies. However, there are alternative methods such as the graph theoretic method (de Camino-Beck, Lewis, & van den Driessche, 2009). While we do not discuss them in this paper, a nice discussion can be found in Heffernan, Smith, and Wahl (2005) and Li, Blakeley, and Smith (2011).

4. Treading carefully when formulating a disease model

4.1. Formulating the force of infection

Recall that the force of infection gives the *per capita* rate at which susceptible individuals acquire infection. Generally, the force of infection, denoted by $\lambda(I)$ can be broken apart into three components: (i) the contact rate, (ii) the probability of infection given a contact, and (iii) the size of the infectious population. The product of the first two is the overall transmission rate, standardly denoted by β. Although these quantities are tightly connected we will address each of them independently, beginning with the infectious population.

4.1.1. Infectious population

When the total population size is constant we do not need to worry too much about how the infectious population is used within the system. However, if we expect the total population size to change due to, for example, differing birth and death rates or disease induced mortality, we need to be more careful in its formulation. Throughout our discussion thus far,

the force of infection $\lambda(I)$ depended on the infectious proportion of the total population: I/N.

There are two main ways to incorporate the infectious population into the force of infection $\lambda(I)$. We generally assume that transmission is either 'frequency dependent' or 'density dependent'. In the former, which we have used throughout this paper, $\lambda(I)$ is independent of population size and depends on the fraction of individuals who are infected in a population so that $\lambda(I) = \beta I/N$. In the latter, λ increases proportionally with the number of infected individuals I such that $\lambda(I) = \beta I$. Here, if the density of the infectious population increases, i.e. there are more infectious individuals in the same area, the force of infection also increases.

Frequency dependent transmission, also known as standard incidence, is typically used for sexually transmitted pathogens in human population (Begon et al., 2002; Keeling & Rohani, 2008). As the population size changes, the number of interactions between susceptible and infectious individuals does not change. Hence, the relevancy for sexual transmission, where the number of partners an individual has is likely independent of the number of individuals near by, at least when the population is not very low. In contrast, density dependent transmission allows for reduced interaction between susceptible and infectious individuals as the population size changes (Begon et al., 2002; Keeling & Rohani, 2008). Consider a population that is declining due to disease-induced mortality from a disease with high mortality. As fewer individuals remain, it is more difficult for individuals to come into contact, reducing the force of infection (Keeling & Rohani, 2008). Density dependent transmission is typically used to model airborne and directly transmitted infections such as measles or influenza. We remind the reader that these definitions only differ when the total population size is variable.

Although the vast majority of the literature uses one of these two functions, a variety of phenomenological non-linear formulations have also been introduced (Begon et al., 2002; McCallum, Barlow, & Hone, 2001). When altering the transmission function it is essential that careful consideration is given to the estimation of the transmission rate, β, as the units of the term must be consistent within the equation. For example, a non-linear force of infection such as the following

$$\lambda(I) = \beta I^q$$

where β is the transmission rate, I is the number of infected individuals, and $0 < q < 1$ is a constant defining a power law relationship (Hochberg, 1991; Liu, Levin, & Iwasa, 1986). The exponent on I requires that β take the units $(\sqrt[q]{individuals} \cdot time)^{-1}$ rather than $time^{-1}$ as discussed above in Section 2.

A further complication arises when the population is subdivided by characteristics other than those that are disease-related such as risk status or age. Nothing changes when we consider density-dependent transmission because we are only concerned with the actual number of infectious individuals from a particular group. However, closer attention to the details is required when using frequency dependent transmission: the fraction of individuals infectious from a particular class refers to the number infectious of that class divided by the total number *from that class* rather than the fraction of the total population of all infectious individuals.

To demonstrate this, let us consider the case of an age-structured model, pictured in Figure 3. Although age changes continuously and can be more generally described using

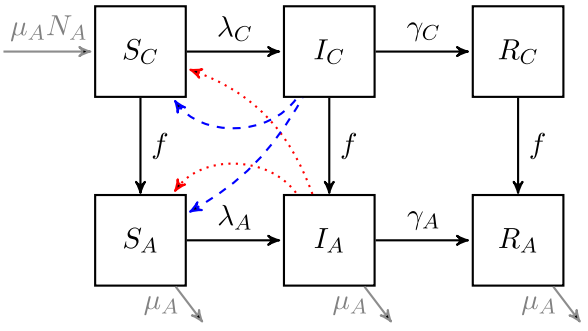


Figure 3. Schematic representation of the age-structured *SIR*. S_C, I_C, and R_C represent the total number of susceptible, infected, and recovered children in a population, and S_A, I_A, and R_A represent the same classes but in the adult population. $\lambda_C(I_C, I_A)$ and $\lambda_A(I_C, I_A)$ are the forces of infection for the child and adult classes, respectively, and γ_C and γ_A are their recovery rates. Dashed lines indicate transmission between classes which contribute to the forces of infection, originating from infectious adults (red lines) or infectious children (blue lines).

a continuous-age model with, for example, partial differential equations (Feng et al., 2016; Martcheva, 2015), we examine an age-structured model with two discrete age classes so that our population is sub-divided into children and adults. Such a model would be relevant for infections that spread much faster among children than adults, whether due to higher susceptibility or greater contact rates. We will use the subscript C to refer to children populations and the subscript A to refer to adult populations. For example, the number of susceptible children will be denoted as S_C. We assume that the total number of children is given by $N_C = S_C + I_C + R_C$ and the total number of adults is given by $N_A = S_A + I_A + R_A$. We also assume that children age into the adult compartment at a rate f and that only adults reproduce and all births enter the S_C class. Finally, only adults are subject to natural mortality so that the total population size remains fixed. Now,

$$\begin{aligned}
\frac{\mathrm{d}S_C}{\mathrm{d}t} &= \mu_A N_A - \lambda_C(I_C, I_A)S_C - fS_C,\\
\frac{\mathrm{d}I_C}{\mathrm{d}t} &= \lambda_C(I_C, I_A)S_C - \gamma_C I_C - fI_C,\\
\frac{\mathrm{d}R_C}{\mathrm{d}t} &= \gamma_C I_C - fR_C,\\
\frac{\mathrm{d}S_A}{\mathrm{d}t} &= -\lambda_A(I_C, I_A)S_A - \mu_A S_A + fS_C,\\
\frac{\mathrm{d}I_A}{\mathrm{d}t} &= \lambda_A(I_C, I_A)S_A - (\gamma_A + \mu_A)I_A + fI_C,\\
\frac{\mathrm{d}R_A}{\mathrm{d}t} &= \gamma_A I_A - \mu_A R_A + fR_C.
\end{aligned}$$

A schematic representation of this model is provided in Figure 3.

Notice that we distinguish between the force of infection to children $\lambda_C(I_C, I_A)$ and the force of infection to adults $\lambda_A(I_C, I_A)$. In the former, the rate at which susceptible children become infected will depend on the rate at which susceptible children contact both infectious adults and children as well as the fraction of the adult and child populations

that are infectious in their respective populations. Additionally, in contrast to the models presented earlier, this model contains multiple infectious classes. Therefore, the forces of infection now depend on both infectious classes. We now assign two subscripts to the transmission rate so that β_{ij} is the transmission rate from individuals in population j to those in population i, where i and j can take on values of C or A. Notice that the proportion of infectious individuals refers to the proportion of each class. Thus, the force of infection for children and adults, respectively, are

$$\lambda_C(I_C, I_A) = \beta_{CC}\frac{I_C}{N_C} + \beta_{CA}\frac{I_A}{N_A},$$
$$\lambda_A(I_C, I_A) = \beta_{AC}\frac{I_C}{N_C} + \beta_{AA}\frac{I_A}{N_A}.$$

4.1.2. Contact rate

A key component of the force of infection is the quantity β, i.e. the transmission rate, which, as described earlier, can be broken down to the product of the probability of transmission per contact and the contacts per unit time. In many unstructured populations (for example, those that do not account for age or sex structure), β can usually be assigned a fixed value without paying direct attention to the two components that make up this quantity. However, there are also a lot of model structures that require more nuanced assumptions on the structure of β. For example, the notion of contacts per time becomes non-trivial when a population has been grouped by factors other than disease status, such as by age.

We follow from the age-structured model introduced in the previous section. The different β_{ij} values are a composition of both the probability of transmission and the contact rate. Contacts of children with other children or of adults with other adults are easiest to account for. However, contacts between adults and children must be symmetric, i.e. an adult contacting a child requires a child contacting an adult. Thus, when recording contacts between individuals as occurs in experimental attempts to estimate contact rates, the resulting data of total reported contacts will be symmetric (see Table 1 for example data). However, as long as the total number of children and adults are different, obtaining the contact rate requires that the total number of reported contacts correct for differences in population size (Table 2 computes the contact rates associated with the data in Table 1).

Table 1. Total reported contacts over a week between 50 children and 100 adults.

	from Children	from Adults
to Children	290	130
to Adults	130	190

Table 2. Calculated contact rates (per week) corresponding to the total reported contact data provided in Table 1. Here, it is assumed that the are 50 children and 100 adults.

	from Children	from Adults
to Children	$\frac{290}{50} = 5.8$	$\frac{130}{100} = 1.3$
to Adults	$\frac{130}{50} = 2.6$	$\frac{190}{100} = 1.9$

4.1.3. Conservation of biting

Models of vector-borne diseases where transmission does not occur directly between humans pose additional complications, as contacts no longer mark the interaction between two individuals. The simplest compartmental model of a vector-borne disease contains susceptible and infectious populations for both human and vector populations. In this model, the subscripts H and V denote variables and parameters that pertain to the human and vector populations, respectively. Here, we will consider standard assumptions for many mosquito-borne illnesses like malaria or dengue that new human infections only originate from bites of the infected vector population and new infected vectors only occur from biting an infected human. Furthermore, we ignore human demography to focus on infection and recovery to a susceptible state at rate γ. Vector demography, as it occurs at a much shorter time scale relative to humans, will be retained such that vector populations have a birth and death rate μ. This gives us the following system of ODEs

$$\begin{aligned}
\frac{\mathrm{d}S_H}{\mathrm{d}t} &= -\lambda_H(I_H, I_V)S_H + \gamma I_H,\\
\frac{\mathrm{d}I_H}{\mathrm{d}t} &= \lambda_H(I_H, I_V)S_H - \gamma I_H,\\
\frac{\mathrm{d}S_V}{\mathrm{d}t} &= \mu N_V - \lambda_V(I_H, I_V)S_V - \mu S_V,\\
\frac{\mathrm{d}I_V}{\mathrm{d}t} &= \lambda_V(I_H, I_V)S_V - \mu I_V,
\end{aligned}$$

where $N_V = S_V + I_V$ is the total vector population size (see schematic representation in Figure 4).

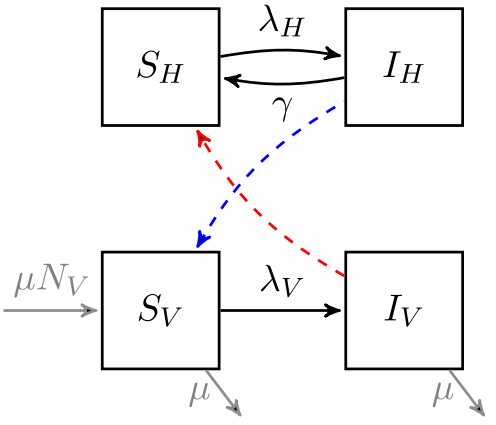


Figure 4. Schematic representation of the vector-borne disease model. S_H and S_V represent the total number of susceptible humans and mosquitoes, respectively, and I_H and I_V represent the total number of infected humans and mosquitoes, respectively. $\lambda_H(I_H, I_V)$ and $\lambda_V(I_H, I_V)$ are the forces of infection for human and mosquito populations, respectively. Here, γ is the recovery rate for infections in humans and this model assumes humans become susceptible to infection again after recovery. Dashed lines indicate transmission between classes which contribute to the forces of infection: humans become infected only through contact with an infected mosquito (red line), and mosquitoes only acquire infection through biting a human (blue line). The only route for mosquitoes to leave the infectious class is by natural mortality.

Given how we defined the force of infection in previous sections, our first instinct may be to define the forces of infection for human and vector populations as follows:

$$\lambda_H(I_H, I_V) = \beta_{HV}\frac{I_V}{N_V}, \quad \lambda_V(I_H, I_V) = \beta_{VH}\frac{I_H}{N_H} \tag{17}$$

such that the infection of humans is dependent upon the transmission rate β_{HV} *from* vectors *to* humans as well as fraction of the vector population that is infected; whereas the infection of vectors is dependent upon the transmission rate β_{VH} *from* humans *to* vectors and the fraction of humans that are infected ($N_H = S_H + I_H$ is the total number of humans). As in the previous models, the transmission rates β_{ij} include both the contact rate and the probability of transmission given contact. Here, contacts that are capable of generating infections occur through bites, b_i, and we denote the probability of infection by α_{ij}. In other words, we require that the transmission rate from j to i (β_{ij}) to be equal to the product of the contact rate b_i and the infection probability α_{ij} so that

$$\beta_{ij} = b_i\alpha_{ij}.$$

Contact through bites has the symmetrical property such that every bite received by a human is a bite taken by a vector, a concept referred to as 'conservation of biting'. We observed a similar idea in contacts between adults and children in Section 4.1.2. Consequently, the number of bites depends on the population sizes of both the human and vector populations. To demonstrate what this means, let us assume that b_V is the biting rate of a single mosquito and b_H is the rate at which a single human gets bitten per unit time. Then the total number of bites received by the human population and total number of bites taken by the mosquitoes should equate as follows:

$$b_H N_H = b_V N_V. \tag{18}$$

Solving for N_V in the equation for conservation of biting (Equation (18)) and substituting in $\lambda_H(I_H, I_V)$ into our original force of infection equations (Equation (17)) leads to

$$\lambda_H(I_H, I_V) = b_V\alpha_{HV}\frac{I_V}{N_H}, \quad \lambda_V(I_H, I_V) = b_V\alpha_{VH}\frac{I_H}{N_H}.$$

As the same biting rate appears in both equations, typically the subscript is dropped. As a note of caution, the canonical model of malaria (known as the Ross-Macdonald model after Ronald Ross and George Macdonald who contributed to its formulation) often refers to biting rate by the parameter a (Anderson et al., 1992; Macdonald, 1957; Ross, 1911). Furthermore, although they use the notion of conservation of biting, the derivation is often omitted. The appearance of the ratio of mosquitoes to humans, commonly denoted by m, results from this assumption.

4.2. Disease can persist when $R_0 < 1$

Up to this point, we have assumed that we can define the threshold of disease spread R_0 by examining the local behaviour near the DFE. In other words, we generate conditions for which the DFE becomes unstable (i.e. when $R_0 > 1$) which, in turn, indicates that the endemic equilibrium exists and is locally stable. However, this scenario is not the only possibility.

4.2.1. Backward bifurcations

As it turns out, there exist cases where both the DFE and an endemic equilibrium stably coexist over a range of parameter values and are separated by an additional unstable equilibrium. In such cases, R_0 is no longer a strict threshold of disease decay and growth. In the infectious disease literature, these scenarios are commonly termed *backward bifurcations*.

In these cases, $R_0 < 1$ remains as a requirement for the DFE to be locally stable. However, the dynamics have changed such that the condition for instability of the disease free equilibrium (that is, $R_0 > 1$) is no longer equivalent to the condition for the stability of a biologically relevant endemic equilibrium. Here, we describe the qualitative structure of the simplest instance where a backward bifurcation is observed. In this case, the system contains three equilibria.

When $R_0 << 1$ only the DFE exists in the system. As R_0 increases, there comes a point $0 < R_0 < 1$ where two endemic equilibria appear. Mathematically this occurs through what is known as a saddle node bifurcation (see for example Strogatz, 2014 or most linear algebra or ODE texts). The appearance of a saddle node bifurcation often arises when the solution for the endemic equilibrium contains a radical (or square root). For $0 < R_0 << 1$, the values of the endemic equilibria are imaginary, i.e. the number under the radical is negative, and are therefore ignored because biological population sizes are real-valued and non-negative. However, as R_0 increases, the term under the radical may switch from negative to positive giving rise to two real-valued, biologically relevant endemic equilibria. This type of equilibrium structure requires that the larger of the newly arisen equilibria is stable and the smaller is unstable. In the majority of backward bifurcations, as R_0 increases the smaller of the two endemic equilibria eventually intersects the DFE and creates a transcritical bifurcation (see Strogatz, 2014) at $R_0 = 1$ such that the DFE becomes unstable. Beyond $R_0 = 1$, the smaller endemic equilibrium is negative and is therefore not biologically realistic. Additionally, the positive, biologically relevant endemic equilibrium is always stable and the DFE is always unstable. Of course, there are more complex equilibrium structures that may not exactly follow this description, but the fact that $R_0 = 1$ is no longer a strict threshold for disease persistence always holds in backwards bifurcations.

One example of a system that exhibits this bifurcation structure is that of an Susceptible-Infectious-Susceptible (*SIS*) model with an imperfect vaccine (Kribs-Zaleta & Velasco-Hernandez, 2000) (see schematic in Figure 5). Assuming that the number of individuals in the vaccinated class is given by V

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= \mu N + \theta V + \gamma I - \lambda(I)S - (\mu + \phi)S, \\
\frac{\mathrm{d}V}{\mathrm{d}t} &= \phi S - \lambda_V(I)V - (\mu + \theta)V, \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \lambda(I)S + \lambda_V(I)V - (\mu + \gamma)I,
\end{aligned}$$

where $N = S + V + I$ is the total population size, $\lambda(I) = \beta(I/N)$ is the force of infection to the susceptible class, $\lambda_V(I) = \sigma\beta(I/N)$ is the force of infection to the susceptible class, β is the transmission rate, μ is the natural birth/mortality rate, θ is the vaccine waning rate, σ is the ability of vaccinated individuals to produce new infections, γ is the recovery rate, and ϕ is the vaccination rate. When the rate of waning of immunity from

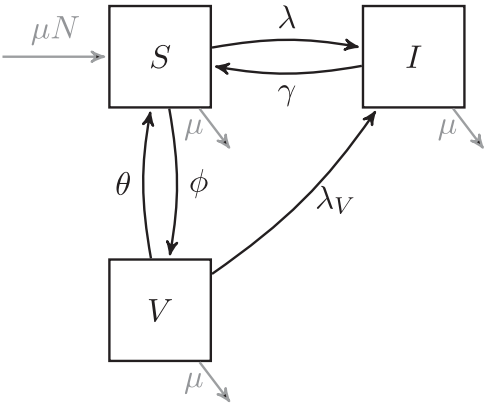


Figure 5. Schematic representation of the *SIS* model with an imperfect vaccine. Here, $\lambda(I) = \beta(I/N)$ and $\lambda_V(I) = \sigma\beta(I/N)$ are the force of infection to the susceptible and vaccinated classes, respectively. β is the transmission rate, θ is the rate of immunity loss, σ represents the ability of vaccinated individuals to produce new infections, γ is the recovery rate, ϕ is the vaccination rate, and μ is the natural mortality rate.

the vaccine exceeds that of recovery rate, it is possible to have a stable endemic equilibrium while $R_0 < 1$. As found in Kribs-Zaleta and Velasco-Hernandez (2000), $R_0(\phi) = (\beta/(\mu + c))((\mu + \theta + \sigma\phi)/(\mu + \theta + \phi))$. Varying R_0 through the variation of the vaccination rate ϕ yields different dynamics dependent on the value of σ, i.e. the ability of vaccinated individuals to produce new infections. Figure 6 shows an example of backward

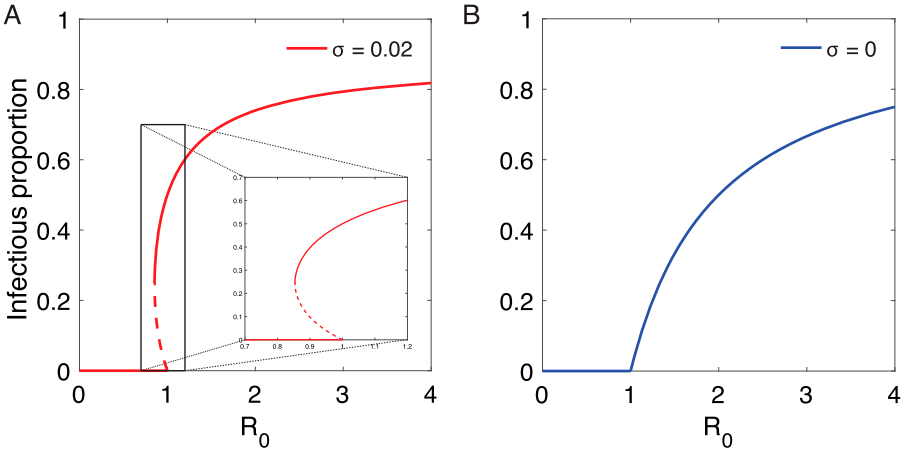


Figure 6. *Backward bifurcation in the presence of an imperfect vaccine.* (A) The presence of a backward bifurcation occurs with $\sigma = 0.02$. Solid line indicates a stable equilibrium while the dashed line represents unstable. The grey shaded area indicates when two stable equilibria are simultaneously present. (B) The backward bifurcation is not observed when when the vaccine is perfect $\sigma = 0$. Parameters are: $\mu = 0.01$; $\theta = 0.01$; $\gamma = 0.5$; $\beta = 3.27$. ϕ is varied between 0 and 100. Initial conditions are: $S(0) = 1 - \eta$; $I(0) = \eta$; $V(0) = 0$ with $\eta = 10^{-4}$ except near $R_0 = 1$ where $\eta = 10^{-10}$ is also used to capture the stable DFE equilibrium. The model and parameters are motivated by Kribs-Zaleta and Velasco-Hernandez (2000).

bifurcation for this system when $\sigma = 0.02$ which disappears when the vaccine is perfect, i.e. $\sigma = 0$.

Backwards bifurcations have arisen in the mathematical analysis of a variety of epidemiological models (e.g. Arino, McCluskey, & van den Driessche, 2003; Dushoff, Huang, & Castillo-Chavez, 1998; Hadeler & van den Driessche, 1997; Martcheva & Thieme, 2003; van den Driessche & Watmough, 2000). However, often the constructions necessary to produce these dynamics require biologically tenuous assumptions such as in the model above where the rate of reinfection from the vaccinated state is greater than the rate of returning to the susceptible state. Another example is a model for Tuberculosis in which a backward bifurcation will occur when infection with re-exposure is more likely than infection with primary exposure (Feng, Castillo-Chavez, & Capurro, 2000; Martcheva, 2015). To our knowledge, however, there exists no real-world experimental evidence of the occurrence of backwards bifurcation in infectious disease systems.

4.2.2. Time varying parameters

The models we have discussed thus far have only considered constant-valued parameters. However, many of these quantities actually vary with time. Take, for example, the transmission rate, β. Some diseases have seasonal patterns, such measles and flu, which may be well represented by a periodically varying transmission rate resulting from, for example, changes in contact rates corresponding to when children are in school or on extended breaks. Over the course of a period, assuming frequency or density dependent transmission, the time varying transmission rate can be replaced, under certain conditions, by its average in the computation of the reproductive number (Wang & Zhao, 2008; Wesley & Allen, 2009). While R_0 for an entire period may be above one, there may be times when an instantaneous R_0 is actually below one. If the epidemic starts at this point in the transmission cycle with a single infectious individual, it can die out.

5. Conclusions

This paper provides an introduction to modeling infectious disease dynamics with a focus on several potential pitfalls. We attempt to show when these issues are likely to arise and how to proceed with models and analyses despite them. This paper provides an introduction to modeling infectious disease dynamics with a focus on several potential pitfalls. We attempt to show when these issues are likely to arise and how to proceed with models and analyses despite them. This paper is intended to serve as a supplement for the many excellent texts on infectious disease dynamics (referenced throughout) and it provides a concise reference that compiles a range of common issues in disease modeling.

Acknowledgments

The authors thank two anonymous reviewers for their helpful feedback throughout the review process. This work was, in part, made possible by the following inspiration:

R_0 is so great
It provides so much info
Except when it fails
– *'An R_0 Haiku'*

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

LMC is supported by Simons Foundation Collaboration Grant 443528.

ORCID

Lauren M. Childs http://orcid.org/0000-0003-3904-3895

References

Anderson, R. M., May, R. M., & Anderson, B. (1992). *Infectious diseases of humans: Dynamics and control*. Vol. 28. Oxford: Wiley Online Library.

Arino, J., McCluskey, C. C., & van den Driessche, P. (2003). Global results for an epidemic model with vaccination that exhibits backward bifurcation. *SIAM Journal on Applied Mathematics*, *64*(1), 260–276.

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., & Turner, J. (2002). A clarification of transmission terms in host-microparasite models: Numbers, densities and areas. *Epidemiology and Infection*, *129*(1), 147–153.

Blackwood, J. C., Streicker, D., Altizer, S. A., & Rohani, P. (2013). Resolving the role of immunity, pathogenesis, and immigration for rabies persistence in vampire bats. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 20837–20842.

Brauer, F. (2008). Compartmental models in epidemiology. In F. Brauer, P. van den Driessche, & J. Wu (Eds.), *Mathematical Epidemiology* (pp. 19–79). Berlin: Springer.

Brauer, F., & Castillo-Chavez, C. (2001). *Mathematical models in population biology and epidemiology*. Vol. 40. New York: Springer.

Brauer, F., & Kribs, C. (2015). *Dynamical systems for biological modeling: An introduction*. Boca Raton: CRC Press.

Brauer, F., vanden Driessche, P. & Wu, J. (Eds.). (2008). *Mathematical epidemiology*. Lecture Notes in Mathematics. Vol. 1945. Berlin: Springer.

de Camino-Beck, T., Lewis, M. A., & van den Driessche, P. (2009). A graph-theoretic method for the basic reproduction number in continuous time epidemiological models. *Journal of Mathematical Biology*, *59*(4), 503–516.

Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Vol. 5. West Sussex: John Wiley and Sons.

Diekmann, O., Heesterbeek, J. A. P., & Johan, A. J. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, *28*(4), 365–382.

Diekmann, O., Heesterbeek, J. A. P., & Roberts, M. G. (2009). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface*, *7*, 873–882.

Dushoff, J., Huang, W., & Castillo-Chavez, C. (1998). Backwards bifurcations and catastrophe in simple models of fatal diseases. *Journal of Mathematical Biology*, *36*(3), 227–248.

Feng, Z., Castillo-Chavez, C., & Capurro, A. F. (2000). A model for tuberculosis with exogenous reinfection. *Theoretical Population Biology*, *57*(3), 235–247.

Feng, Z., Zheng, Y., Hernandez-Ceron, N., Zhao, H., Glasser, J. W., & Hill, A. N. (2016). Mathematical models of Ebola? Consequences of underlying assumptions. *Mathematical Biosciences*, *277*, 89–107.

Hadeler, K. P., & van den Driessche, P. (1997). Backward bifurcation in epidemic control. *Mathematical Biosciences*, *146*(1), 15–35.

Heffernan, J. M., Smith, R. J., & Wahl, L. M. (2005). Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, *2*(4), 281–293.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*(4), 599–653.
Hochberg, M. E. (1991). Non-linear transmission rates and the dynamics of infectious disease. *Journal of Theoretical Biology*, *153*(3), 301–321.
Keeling, M. J., & Rohani, P. (2008). *Modeling infectious diseases in humans and animals*. Princeton: Princeton University Press.
Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A*, *115*, 700–721.
Kribs-Zaleta, C. M., & Velasco-Hernandez, J. X. (2000). A simple vaccination model with multiple endemic states. *Mathematical Biosciences*, *164*(2), 183–201.
Ledder, G. (2017). Scaling for Dynamical Systems in Biology. *Bulletin of Mathematical Biology*, *79*(11), 2747–2772.
Li, J., Blakeley, D., & Smith, R. J. (2011). The failure of R_0. *Computational and Mathematical Methods in Medicine*, *2011*, 1–17.
Liu, W. M., Levin, S. A., & Iwasa, Y. (1986). Influence of nonlinear incidence rates upon the behavior of sirs epidemiological models. *Journal of Mathematical Biology*, *23*(2), 187–204.
Macdonald, G. (1957). *The epidemiology and control of malaria*. London: Oxford University Press.
Martcheva, M. (2015). *Introduction to mathematical epidemiology*. Vol. 61. New York: Springer.
Martcheva, M., & Thieme, H. R. (2003). Progression age enhanced backward bifurcation in an epidemic model with super-infection. *Journal of Mathematical Biology*, *46*(5), 385–424.
McCallum, H., Barlow, N., & Hone, J. (2001). How should pathogen transmission be modelled?. *Trends in Ecology and Evolution*, *16*(6), 295–300.
Ross, R. (1911). *The prevention of malaria*. London: John Murray.
Shrestha, S., Foxman, B., Berus, J., van Panhuis, W. G., Weinberger, D. M., Steiner, C., Viboud, C., & Rohani, P. (2015). The role of influenza in the epidemiology of pneumonia. *Scientific Reports*, *5*, 15314.
Shrestha, S., Foxman, B., Weinberger, D. M., Steiner, C., Viboud, C., & Rohani, P. (2013). Identifying the interaction between influenza and pneumococcal pneumonia using incidence data. *Science Translational Medicine*, *5*, 191ra84.
Strogatz, S. H. (2014). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Boca Raton: Westview press.
van den Driessche, P., & Watmough, J. (2000). A simple sis epidemic model with a backward bifurcation. *Journal of Mathematical Biology*, *40*(6), 525–540.
van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, *180*(1), 29–48.
Vynnycky, E., & White, R. (2010). *An introduction to infectious disease modelling*. Oxford: Oxford University Press.
Wang, W., & Zhao, X-Q. (2008). Threshold dynamics for compartmental epidemic models in periodic environments. *Journal of Dynamics and Differential Equations*, *20*(3), 699–717.
Wesley, C. L., & Allen, L. J. S. (2009). The basic reproduction number in epidemic models with periodic demographics. *Journal of Biological Dynamics*, *3*(2–3), 116–129.

Appendices

Appendix 1. Relating constant rates and duration

The common usage of a constant rate of transition between compartments is frequently made for mathematical simplicity. An underlying assumption for the choice of a constant rate is that average duration spent within a compartment is exactly the reciprocal of the rate of exit from that compartment. Consider a simplified equation for the infectious (I) compartment, such that only recovery occurs and there are no new infections:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \gamma I. \tag{A1}$$

The exit rate from the I compartment, γ, can be found by measuring the average time spent infectious, d. Specifically, it turns out that $\gamma = 1/d$.

To understand the origin of this result, solve Equation (A.1) with an initial condition of $I(0) = I_0$ to obtain the number of infectious individuals at time t

$$I(t) = I_0 \exp^{-\gamma t}. \tag{A2}$$

Dividing both sides by I_0 allows us to rewrite this equation in terms of $I(t)/I_0$, i.e. the fraction of the originally infectious individuals who are still infectious at time t. This is equivalent to the probability of remaining infectious at time t.

The complement of this quantity, which we define as $F(t;\gamma) = 1 - \exp(-\gamma t)$, is then the probability of having left the infectious class by time t. Given that the initial time is assumed to be at $t = 0$ (that is, there are no infections prior to the initial time), $F(t;\gamma)$ is exactly the cumulative distribution function for a probability distribution known as the exponential distribution where γ is the rate parameter. Any standard statistical text will demonstrate that the mean of the exponential distribution with parameter γ is exactly $1/\gamma$, i.e. the average time an individual is infectious.

Appendix 2. SIR model in terms of population proportions

As discussed in the main text, the state variables in compartmental models can be written in terms of proportions of the population instead of the population density. This requires a change of variables, which we demonstrate here for the *SIR* model without demography (Equation (1) in main text). We first define the new variables:

$$\tilde{S} = \frac{S}{N}, \quad \tilde{I} = \frac{I}{N}, \quad \tilde{R} = \frac{R}{N}$$

so that $\tilde{S}$, $\tilde{I}$, and $\tilde{R}$ are the proportions of the population that are susceptible, infectious, and recovered, respectively. For each of these, we can solve for the state variables as defined in the original model (Equation (1)) and obtain

$$S = \tilde{S}N, \quad I = \tilde{I}N, \quad R = \tilde{R}N.$$

Noting that the total population size, N, is a constant we now substitute the new definition for S into Equation (1) and

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}(\tilde{S}N)}{\mathrm{d}t} = N\frac{\mathrm{d}\tilde{S}}{\mathrm{d}t}$$

and thus replacing $\mathrm{d}S/\mathrm{d}t$ and plugging in for S and I we obtain

$$-\beta\frac{(\tilde{I}N)}{N}(\tilde{S}N) = N\frac{\mathrm{d}\tilde{S}}{\mathrm{d}t}$$

which can be simplified by canceling the Ns on both sides of the equation so that

$$\frac{\mathrm{d}\tilde{S}}{\mathrm{d}t} = -\beta\tilde{I}\tilde{S}.$$

A similar process applies for the I and R equations, and it can be shown that

$$\frac{\mathrm{d}\tilde{I}}{\mathrm{d}t} = \beta\tilde{I}\tilde{S} - \gamma\tilde{I}, \quad \frac{\mathrm{d}\tilde{R}}{\mathrm{d}t} = \gamma\tilde{I}.$$

In more complex models, this change of variables can be applied in a similar way. We refer interested readers to Ledder (2017) to read more about non-dimensionalization of models in epidemiology.

Appendix 3. SIR model without demography

In Section 2.2, we introduced and analyzed the *SIR* model with demography. Interestingly, it turns out that if we eliminate demography by setting $\mu = 0$ then the standard analysis of finding equilibria

and R_0 no longer applies in the same way. First, recall the *SIR* model without demography:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta\frac{I}{N}S, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \beta\frac{I}{N}S - \gamma I, \quad \frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I.$$

In this model, there is only a DFE. In fact, there are *infinitely* many DFEs because as long as $I = 0$ then the value of S and R does not matter.

Suppose we try to find R_0 using, for example, the next generation matrix. In this case, the $\mathbf{F}$ matrix is simply a scalar, $\beta(1/N)S$ and the $\mathbf{V}$ matrix is also a scalar given by γ (note that they are scalars because there is only one infectious class, I). Evaluated at the DFE,

$$\mathbf{FV}^{-1} = \frac{\beta}{\gamma}$$

so that $R_0 = \beta/\gamma$. The magnitude of R_0 relative to one determines whether the disease initially grows ($R_0 > 1$) or declines ($R_0 < 1$). In contrast to the *SIR* model with demography, you always return to a disease free state.

Although R_0 does not directly relate to the global asymptotic stability of the DFE, it does provide information on the shape of the epidemic curve. Namely, if $R_0 < 1$ then I is always decreasing. Consequently, there is no epidemic. If $R_0 > 1$, then I increases until S is equal to $\gamma N/\beta$. At that point, I decreases back towards zero. In this case, the dynamics follow a standard epidemic curve – increasing cases followed by a peak and eventually extinction of the outbreak. Therefore, although R_0 does not directly impact the global stability of the DFE it does play a substantial role in the short term epidemic dynamics.

Appendix 4. Linear stability analysis

We show how to perform a linear stability analysis for any two species model through the example of the *SIR* model with demography. Notice that since the total population size is constant we only need to use two equations to describe the system. Thus, our populations are represented by S and I so that

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \mu N - \beta\frac{I}{N}S - \mu S = F_1(S, I)$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta\frac{I}{N}S - (\gamma + \mu)I = F_2(S, I)$$

with an equilibrium point denoted by (S^*, I^*). Now, suppose that $(x(t), y(t))$ is a small perturbation away from the equilibrium such that

$$x(t) = S(t) - I^* \Longrightarrow S(t) = x(t) + S^*$$
$$y(t) = I(t) - I^* \Longrightarrow I(t) = y(t) + I^*$$

Our goal is to determine how the perturbation (x, y) behaves as time increases, so we want to find the derivatives of x and y. Using the equalities above, this leaves us with

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}S}{\mathrm{d}t} = F_1(S, I) = F_1(x + S^*, y + I^*) \tag{A3}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}I}{\mathrm{d}t} = F_2(S, I) = F_2(x + S^*, y + I^*) \tag{A4}$$

We use the Taylor expansion to approximate these equations near the equilibrium. We now have two variables, so we need to find the Taylor expansion in two dimensions. Recall that we can approximate a function $f(w, z)$ at a point (w^*, z^*) using the following:

$$f(w, z) \approx f(w^*, z^*) + \frac{1}{1!}\left[(w - w^*)\frac{\partial f(w^*, z^*)}{\partial w} + (z - z^*)\frac{\partial f(w^*, z^*)}{\partial z}\right] + h.o.t.$$

where *h.o.t.* are higher order terms. We can apply this to Equations (A3) and (A4). Because the perturbations are assumed to be small, we ignore higher order terms and after some simplification

we are left with the following linear approximations:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial F_1(S^*, I^*)}{\partial S}x + \frac{\partial F_1(S^*, I^*)}{\partial I}y$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial F_2(S^*, I^*)}{\partial S}x + \frac{\partial F_2(S^*, I^*)}{\partial I}y$$

We can alternatively write these approximations using slightly different notation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \left.\frac{\partial F_1}{\partial S}\right|_{(S^*,I^*)} x + \left.\frac{\partial F_1}{\partial I}\right|_{(S^*,I^*)} y$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \left.\frac{\partial F_2}{\partial S}\right|_{(S^*,I^*)} x + \left.\frac{\partial F_2}{\partial I}\right|_{(S^*,I^*)} y$$

where the vertical bar indicates that the partial derivatives are being evaluated at the equilibrium. It is more convenient to write this in matrix form, i.e. the Jacobian. We can therefore rewrite the linear approximation as:

$$\begin{bmatrix} \frac{\mathrm{d}x}{\mathrm{d}t} \\ \frac{\mathrm{d}y}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial I} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial I} \end{bmatrix}_{(S^*,I^*)} \begin{bmatrix} x \\ y \end{bmatrix} \tag{A5}$$

To simplify this even further, we make the following definitions:

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \begin{bmatrix} \frac{\mathrm{d}x}{\mathrm{d}t} \\ \frac{\mathrm{d}y}{\mathrm{d}t} \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} \frac{\partial F_1}{\partial S} & \frac{\partial F_1}{\partial I} \\ \frac{\partial F_2}{\partial S} & \frac{\partial F_2}{\partial I} \end{bmatrix}_{(S^*,I^*)}, \quad \vec{n} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Note that matrices will be indicated by bold font and vectors will have an arrow above them. We can now simply write Equation (A.5) as

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \mathbf{J}\vec{n}$$

where **J** is called the ***Jacobian*** matrix and here it is evaluated at the equilibrium. In the case of the *SIR* model the Jacobian is

$$\begin{bmatrix} -\beta\frac{I^*}{N} - \mu & -\beta\frac{S^*}{N} \\ \beta\frac{I^*}{N} & \beta\frac{S^*}{N} - (\mu + \gamma) \end{bmatrix}$$

evaluated at the DFE $(S^*, I^*) = (0, 0)$ makes the linearized system

$$\begin{bmatrix} \frac{\mathrm{d}S}{\mathrm{d}t} \\ \frac{\mathrm{d}I}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} -\mu & -\beta \\ 0 & \beta - (\mu + \gamma) \end{bmatrix} \begin{bmatrix} S \\ I \end{bmatrix}$$

Note that the Jacobian can be extended to higher dimensions.

As our goal is to determine the long term behaviour of the perturbation from the equilibrium, we need to understand how $\vec{n}$ will change. From calculus we know that in a one dimensional setting

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} = \alpha n(t) \quad \Longrightarrow \quad n(t) = n_0\,\mathrm{e}^{\alpha}$$

The dynamics of this equation can then be determined by the sign of α. If it is positive, the perturbation grows with time. In contrast if it is negative, the perturbation shrinks with time. Now, if α represents the Jacobian evaluated around an equilibrium, these dynamics tell us about the stability of the equilibrium, i.e. when $\alpha < 0$ the equilibrium is stable and when $\alpha > 0$ the equilibrium is unstable.

A similar idea applies to higher order systems. Let's return to our two dimensional system

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \mathbf{J}\vec{n}. \tag{A6}$$

It turns out it *is* possible to analyze equilibria in a similar manner and a solution of Equation (A.6) exists and is of the form

$$\vec{n} = \mathrm{e}^{\lambda t}\vec{n}_0 \tag{A7}$$

where $\vec{n}_0$ is a non-zero vector and λ is a constant. Differentiating Equation (A.7) leaves us with

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \lambda\,\mathrm{e}^{\lambda t}\vec{n}_0.$$

Substituting this into Equation (A.6) gives us that

$$\lambda\,\mathrm{e}^{\lambda t}\vec{n}_0 = \mathbf{J}\vec{n}$$

but from Equation (A.7), this means that

$$\lambda\,\mathrm{e}^{\lambda t}\vec{n}_0 = \mathbf{J}\,\mathrm{e}^{\lambda t}.\vec{n}_0$$

Here, $\mathrm{e}^{\lambda t}$ is a non-zero constant so we can divide it out and we are left with

$$\lambda\vec{n}_0 = \mathbf{J}\vec{n}_0 \tag{A8}$$

Here, $\vec{n}_0$ is called an ***eigenvector*** and λ is an ***eigenvalue*** of the matrix $\mathbf{J}$. Details and further explanation on eigenvalues and eigenvectors can be found in any standard linear algebra or ordinary differential equation textbook.

Keep in mind that our end goal is to determine how the dynamics change, which given Equation (A.6) depend on λ. To find λ, we subtract $\lambda\vec{n}_0$ from both sides of Equation (A.8), leaving us with and we are left with

$$\begin{aligned} 0 &= \mathbf{J}\vec{n}_0 - \lambda\mathbf{I}\vec{n}_0 \\ &= (\mathbf{J} - \lambda\mathbf{I})\,\vec{n}_0 \end{aligned}$$

Notice that we cannot simply pull out $\vec{n}_0$ from the above equation because $\mathbf{J}$ and λ have different dimensions. Therefore, we needed to introduce a 2x2 identity matrix $\mathbf{I}$.

To have a non-zero solution of this equation the ***determinant*** of $(\mathbf{J} - \lambda\mathbf{I})$ must equal zero. We further you to a standard ordinary differential equation text to find the determinant.

Thus, a general solution the two-dimensional system

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \mathbf{J}\vec{n}$$

is given by

$$\vec{n} = c_1\,\mathrm{e}^{\lambda_1 t}\vec{n}_1 + c_2\,\mathrm{e}^{\lambda_2 t}\vec{n}_2$$

where $\vec{n}_1$ and $\vec{n}_2$ are the eigenvectors corresponding to the eigenvalues λ_1 and λ_2, respectively, and c_1 and c_2 are arbitrary constants.

In the case of the *SIR* model we find that the two eigenvalues of $\mathbf{J}$ are $\lambda_1 = -\mu$ and $\lambda_2 = \beta - (\mu + \gamma)$. As our parameters are all positive to ensure biological realism, $\lambda_1 < 0$ always. However, λ_2 may be positive or negative dependent upon the relative sizes of β and $\mu + \gamma$. The DFE equilibrium will be linearly stable when

$$\beta - (\mu + \gamma) < 0.$$

We can rearrange this quantity to find an equivalent comparison, i.e.

$$\frac{\beta}{(\mu + \gamma)} < 1$$

Notice, this is in fact our definition of R_0. In fact, it can be shown that when $R_0 < 1$ then our eigenvalues are negative and the DFE is linearly stable and when $R_0 > 1$ then at least one eigenvalue is positive and the DFE is unstable.

We can now move on to determine what the values of λ mean in terms of linear stability of the equilibria. Here, we will list the different cases; there are two main cases (each with three subcases each):

(1) Both eigenvalues λ_1 and λ_2 are real.
   (a) Both eigenvalues are positive. In this case, $e^{\lambda_i t} \longrightarrow \infty$ for both eigenvalues. In other words, perturbations from equilibrium will always move away from the equilibrium. Therefore, the equilibrium is *unstable*.
   (b) Both eigenvalues are negative. In this case, $e^{\lambda_i t} \longrightarrow 0$ for both eigenvalues. In other words, perturbations from equilibrium will always move back to the equilibrium. Therefore, this equilibrium is *linearly stable*.
   (c) One eigenvalue is positive and the other is negative. In this case, the equilibrium is seemingly stable in 'one direction' and unstable in the other. Therefore, the equilibrium is *unstable*. In higher, dimensions, a single positive eigenvalue will be sufficient to make the equilibrium unstable.

(2) Both eigenvalues are imaginary and given by $\lambda_{1,2} = a_1 \pm b_1 i$.
   (a) The eigenvalues are purely imaginary (i.e. $a_1 = 0$). Here, we can use Euler's formula to examine the long-term behaviour of perturbations equilibria. Without loss of generality, let's look at $\lambda_1 = b_1 i$. Then

$$e^{b_1 i t} = \cos(b_1 t) + i \sin(b_1 t). \tag{A9}$$

   The presence of sines and cosines indicates that the dynamics near an equilibrium will oscillate. Specifically, in this case the solutions to $d\vec{n}/dt = \mathbf{J}\vec{n}$ will neither grow nor decay and instead oscillate with no *damping*.
   (b) The real part of the eigenvalues is positive ($a_1 > 0$). Here, to determine stability we care about (again without loss of generality) the long-term behaviour of $e^{(a_1+b_1 i)t}$. Using Euler's formula again, the long-term behaviour of a perturbation will depend on

$$e^{(a_1+b_1 i)t} = e^{a_1 t}\left(\cos(b_1 t) + i \sin(b_1 t)\right). \tag{A10}$$

   Since $a_1 > 0$ the exponential term grows and we therefore get oscillations of increasing magnitude. Therefore, the equilibrium is *unstable*.
   (c) The real part of the eigenvalues is negative ($a_1 < 0$). Again, we are concerned with the long-term behaviour of $e^{(a_1+b_1 i)t}$. We will again arrive at Equation (A.10) and since $a_1 < 0$ the exponential term decays and we therefore get oscillations of decreasing magnitude. Therefore, the equilibrium is *linearly stable*.

To summarize, an equilibrium is *linearly stable* if the real part of all eigenvalues is less than zero, i.e. $\mathrm{Re}(\lambda_i) < 0$ for all i. Otherwise we say the equilibrium is *unstable*.