Unravelling Transmission in Epidemiological Models and its Role in the Disease-Diversity Relationship.

Marjolein E.M. Toorians[1,*], Ailene MacPherson[2], and T. Jonathan Davies[1,3,4]

[1]Department of Botany, Biodiversity Research Centre, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada
[2]Department of Mathematics, Simon Fraser University, Burnaby, Canada
[3]African Centre for DNA Barcoding, University of Johannesburg, Johannesburg 2092, South Africa
[4]Department Forest & Conservation Sciences, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada
[*]Corresponding author

October 20, 2021

Abstract

With the decrease of biodiversity worldwide coinciding with an increase in disease outbreaks, investigating this link is more important then ever before. This review outlines the different modelling methods commonly used for pathogen transmission in animal host systems. There are a multitude of ways a pathogen can invade and spread through a host population. The assumptions of the transmission model used to capture disease propagation determines the outbreak potential, the net reproductive success (R_0). This review offers an insight into the assumptions and motivation behind common transmission mechanisms and introduces a general framework with which contact rates, the most important parameter in disease dynamics, determines the transmission method. By using a general function introduced here and this general transmission model framework, we provide a guide for future disease ecologists for how to pick the contact function that best suites their system. Additionally, this manuscript attempts to bridge the gap between mathematical disease modelling and the controversially and heavily debated disease-diversity relationship, by expanding the summarized models to multiple hosts systems and explaining the role of host diversity in disease transmission. By outlining the mechanisms of transmission into a stepwise process, this review will serve as a guide to model pathogens in multi-host systems. We will further describe these models it in the greater context of host diversity and its effect on disease outbreaks, by introducing a novel method to include host species' evolutionary history into the framework.

Keywords: Epidemiological Models, Transmission, Biodiversity, Dilution Effect

* Marjolein.toorians@botany.ubc.ca

1 Introduction

Worldwide, many natural populations are declining rapidly due to anthropogenic influences (IPBES, 2020), resulting in a shift in the abundance of humans and livestock relative to wildlife species (Bar-On et al., 2018). In some cases, these demographic changes have led to increasing contact between humans and livestock with wildlife, providing new opportunities for disease spillover (Graham et al., 2008). This spillover can occur when a new pathway between two species is established, for example, due to travel, invasion or other forms of translocation of the pathogen. Most emerging pathogens in humans are directly transmitted viruses or bacteria that have crossed the species barrier, perhaps multiple times, likely facilitated by changes to the environment made by humans (Dobson and Foufopoulos, 2001). However, to make predictions of how biodiversity losses and the increasing abundance of humans and domesticated species will reshape the disease landscape, we need to understand how the densities of species in ecosystems affect the maintenance and onward transmission of diseases (Keesing et al., 2010).

Transmission describes the process by which an infected individual transmits a pathogen to an uninfected individual, and is a critical step in disease outbreak and spillover (Park et al., 2018). When interspecific transmission is sufficiently high, a pathogen may invade the novel host population (Daszak et al., 2000; Fenton and Pedersen, 2005; Wolfe et al., 2007). Thus a key strategy for preventing spillover events is reducing transmission from the reservoir (a population of a single species that can maintain the disease) to a susceptible, novel target population (Haydon et al., 2002). Understanding the process of transmission within and between species is essential to reduce the risk of future disease spillover and emergence events.

Transmission is a complex, stepwise process (see Box 1 for more detail on the biological process of transmission). The assumption made about this biologically complex process are, however, often opaque, hidden within a single composite parameter. In the susceptible-infected-recovered (SIR) compartmental model (Kermack and McKendrick, 1927; Anderson and May, 1982), and the focal model for this manucript, the transmission process is often captured by simplifying assumptions as being *density-dependent*, *density-independent* or some intermediate between these two extremes. Transmission is determined not only by the inherent behaviour of the pathogen, but also by that of the host or vector, as diseases rely on their host for reproduction (Rohani et al., 2003; Han et al., 2015). Properties of the host population that affect disease transmission include contact rates, population density, and individual birth and death rates, whereas other disease parameters, such as recovery and transmission rates are considered pathogen dependent parameters (Han et al., 2020). In addition, transmission is unlikely a homogeneous process, but integrates across many heterogeneities at different stages of transmission, including host contact rates and susceptibility McCallum (2016); Smith et al. (2009). However, in many epidemiological models transmission is still represented as a simple, singular constant.

The choice of transmission model affects the projected disease dynamics, and thus the potential management decisions that follow. It is critical, therefore, to choose the model that best represents the pathogen life-history and host population behaviour. To make informed mathematical assumptions, we need to understand the mechanisms of transmission and the assumptions that underlie the models that represent them. In this paper, we synthesize common transmission model assumptions, their biological implications and how these affect the dynamics of the simple compartmental (SIR) model of infectious disease spread. While limited in epidemiological complexity, focusing on this subset of models allows us to examine details of the transmission process relevant for inter-specific disease transmission. In particular, we will explore the effect of host contact rate and structure on disease transmission and show it is the most important determinant of these variations. We will further assess how disease transmission is involved in spillovers, tying it to the heavily debated *disease-diversity relationship* (Halliday and Rohr, 2019).

Box 1: The Biology of Transmission

Transmission is a complex process, and with its many steps many heterogeneities can arise. Transmission is the transfer of infectious particles, or propagules, from the infected (donor) host, to the naive, receiving host. Transmission potential is then dependent on a threshold propagule load within the donor host (Wilber et al., 2016; Blaser et al., 2014). As pathogen particles often have to travel through space between leaving the donor host and establishing in the recipient host, the environment and the time a particle is *free-roaming* can play a role (Tien and Earn, 2010). McCallum (2016) deconstructed transmission into five discrete stages:

- Stage 1: Dynamics of propagules within- donor host.
- Stage 2: Production of pathogen-infective stages in donor host.
- Stage 3: Pathogen survival and growth in the environment (including the environment of an intermediate host).
- Stage 4: Exposure of the recipient host.
- Stage 5: Reception and growth of the pathogen load in the recipient host.

Some diseases (e.g. HIV) are never exposed to the environment, and are directly transmitted, these pathogens skip Stage 3. Between these stages of transmission, heterogeneities can arise and threshold-like behaviours may emerge which can affect the overall transmission. Conventional, linear frameworks of transmission may overlook *super-spreaders* resulting from high infection loads (McCallum et al., 2017). For instance, nematode crowding within hosts causes an increased immune response, saturates the relationship between transmission Stages 1 and 2, such that at high pathogen levels the production of the dispersal pathogen life-stage plateaus (McCallum et al., 2017), e.g. Malaria. Age of infection and pathogen competition (in case of co-infection) can also have an effect on the production of pathogen dispersal stages. For Stage 3, the environment determines the subsequent received load. Some dispersal stages can linger outside hosts longer then others. For example, the decay rate of *Mycobacteria*, causing tuberculosis, is dependent on the temperature of the medium and likely also to light-exposure, therefore differing among seasons (Fine et al., 2011). The received load (Stage 4) is impacted by multiple factors including: the relative density of the donor and recipient hosts, the survival rate of the free-living pathogen stage, and, for trophically transmitted pathogens, the Holling type II function response. The final load that establishes in the recipient host (Stage 5) depends on the quality of the received particle, the hosts' immune response (which may depend on its genotype) and the presence of other pathogens (co-infection). Finally, feeding back from stage 5 to 1, while a host typically does not become infective immediately following exposure, the host response to the pathogen can modify subsequent recipient host's received pathogen load.

As transmission is a step-wise process, there are many opportunities for heterogeneity to arise in this pathway (McCallum et al., 2017). Within a host population, not all individuals are the same. Individuals may vary in susceptibility and infectivity, and hosts likely do not have a homogeneous contact structure. Cross et al. (2004, 2007) showed in models of bovine Tuberculosis (*Mycobacterium bovis*, or bTB) in buffalo, that allowing for social groupings can break down standard SIR disease dynamics. An illustration of how contact structure and animal behaviour can effect disease dynamics is provided by the European badger, also a reservoir for bTB. In this example, culling intended to reduce number of infected individuals instead resulted in an increased contact rate (as aggregation increased in lower-density groups), and thus increased disease prevalence (Vicente et al., 2007). Additionally, individual variation, well captured by the concept of *superspreaders*, dramatically influences outbreak potential and affects disease dynamics (Lloyd-Smith et al., 2005). The concept can be extended to *super-movers*, *super-recipients*, *super-shedders*, and *super-susceptibles* (Streicker et al., 2013; Craft, 2015; White et al., 2017).

A key challenge of modeling transmission, and the focus of this manuscript, is capturing the inherent nature of host density and contact rates. For example, contacts cannot indefinitely increase with population density, as is assumed in some models (Smith et al., 2009) and contact structure has a great effect on the outcome as well (see examples in Box 1). The epidemiological contact process and the emergent relationship between host density and contact rate can be modeled both explicitly, in the form of density-dependent or density-independent compartmental models, and

implicitly, e.g., as a network or individual-based model. Here, we will focus on the representation of transmission in the the SIR compartmental model. This paper will serve as a step-by-step guide to decomposing transmission and introduces a novel general contact rate-function that we then use to unify existing model assumptions by imposing differing assumptions for the contact rate. This review concludes by examining how the biology of transmission and the mathematical assumptions made about it are involved in the disease-diversity relationship our theoretical understanding of this relationship. To clarify the disease-diversity relationship we propose a novel method for including host species' identities (evolutionary history) into the framework, such that we can investigate the effect of biodiversity on disease outbreaks. This review and its novel functions will serve as a helpful tool for modellers to accurately describe transmission and therefore their specific system's disease dynamics.

Box 2: Transmission-virulence trade-off:
Pathogen load is often used as a proxy for *virulence*, the pathogen-induced host mortality. A higher pathogen load also translates to a higher rate of transmission (transmission stage 2), as there are more dispersal propagules (McCallum, 2016; Wilber et al., 2016). For successful transmission, a sufficient number of propagules need to be produced within the donor host, survive the intermediate stages, and establish in the recipient host. When virulence is high, transmission may be reduced due to the negative effect of high pathogen load on host behaviour (e.g. high fevers in influenza immobilizes patients which prevents the disease from further spreading (Dieckmann et al., 2005), including death of the host - a pathogen that kills its host quickly reduces its window of opportunity for infection. This is known as the *transmission - virulence trade-off* (Cressler et al., 2016). However, pathogen load is not always a good predictor of transmission success. For example, HIV has highest transmission potential at intermediate viral loads (McKay et al., 2020).

2 Decomposing the transmission function, $F(S, I)$

In the classic SIR model, the transmission rate, β, describes the rate of new infections per unit of time, given an interaction event between a S and an I individual. It is often modeled as a single constant but can represent a composite of a lot of different underlying processes. These dynamics are described by the *transmission function*, $F(S, I)$, encompassing S and I individuals and the rate of transmission, β. A great number of variations exist of this transmission function, a number of them are shown in Table 1. The transmission term can also be described as the product of the force of infection and the density of susceptible individuals (de Jong et al., 1995). The transmission function is represented in the SIR model as in the derivative for S and I as:

$$\frac{dS}{dt} = -F(S, I) \tag{1a}$$

$$\frac{dI}{dt} = F(S, I) - \Gamma I \tag{1b}$$

where t is time and γ the removal rate (think of this as the recovery and mortality rates, combined). S and I are the number of susceptible and infected individuals, respectively. Throughout this paper the total number of hosts is defined as $N = S + I + R$, unless specified otherwise. Assuming infection is a probability, the rate of infection can be described as (Begon et al., 2002):

$$F(S, I) = \beta S \frac{I}{N} \tag{2}$$

where β is the transmission rate, determining the speed with which a susceptible individual (S) is infected by an infected individual (I), which is a probability event, $\frac{I}{N}$, with N being the total population size. There are several equally valid models, this one being the most common in ecology. Here we decompose the various commonly used transmission functions, $F(S, I)$, of compartmental SIR models (table 1). We begin with the different relationships between susceptible and infected individuals for homogeneously and heterogeneously distributed populations (not including explicit contact networks) (table 1), and finish with explaining how the contact rate (κ) determines the density or frequency-dependent nature of transmission, together with the probability of successful transmission (c). We provide a schematic illustrating the framework of our decomposition throughout the paper in Figure 1, starting

with standard homogeneity where we decompose β (left column) and finishing with some examples of heterogeneity (right column). Table 2 summarizes the parameters that are used here and in the equations throughout this review.

In this model the criterion for a disease outbreak is directly determined by the transmission rate. The basic reproductive rate, R_0, describes the number of secondary infections arising from a single infected host in a fully susceptible host population. This quantity is also known as the *Lifetime Reproductive Success* or *Fitness* of the pathogen. In the SIR model, assuming density-dependent transmission and setting Equation (1) to zero, R_0 can be defined as:

$$R_0 = \frac{\beta}{\Gamma} \tag{3}$$

Emergence of disease, or pathogen persistence in a community, will occur when $R_0 > 1$. If $R_0 < 1$, the disease will go extinct, whereas an R_0 equal to 1 represents an endemic disease. Equation (3) shows that increasing β will increase R_0, and thus will determine the endemic size and the intrinsic rate of increase of the pathogen.

$F(S,I)$	**Description**	**units β_x**	**Source**
	Homogeneous population		
$\frac{\beta_A SI}{A}$	Simple density-dependence (DD)	$\text{ind}^{-1}\ \text{m}^{-2}\ \text{t}^{-1}$	Anderson and May (1982)
$\beta_D SI$	Simple DD with constant area	$\text{ind}^{-1}\ \text{t}^{-1}$	Anderson et al. (1992)
$\frac{\beta_F SI}{N}$	Simple frequency-dependence (FD)	t^{-1}	Anderson et al. (1992)
$\frac{\beta_q SI}{N^q}$	Intermediate	$\text{ind}^{(q-1)}\ \text{t}^{-1}$	Smith et al. (2009)
$\frac{\beta_K SI}{K}$	Carrying capacity dependence	t^{-1}	This manuscript
$\frac{\beta_u N}{\omega_0+S+I}$	Asymptotic transmission	t^{-1}	Diekmann and Kretzschmar (1991)
	Individual heterogeneity		
$\beta_{PL} S^m I^n$	Power law	$\text{ind}^{1-m-n}\ \text{t}^{-1}$	Hochberg (1991); Novozhilov (2008)
	Spatial heterogeneity		
$\beta_r I(N - \frac{I}{p})$	Refuge effect	$\text{ind}^{-1}\ \text{t}^{-1}$	Barlow (1991)
$\theta S \cdot ln(1 + \frac{\beta_b I}{\theta})$	Negative binomial	$\text{ind}^{-1}\ \text{t}^{-1}$	Barlow (2000); Greer et al. (2008)

Table 1: Different compositions of the variables S and I and the transmission rate, β, in the transmission function ($F(S,I)$) in traditional SIR models. Note that for each model β assumes different values, depending on the units. See table 2 for parameter definitions.

Parameter	**Description**	**Unit**
β_x	Transmission rate	see table 2
κ	Contact rate	t^{-1}
c	Probability of successful transmission	-
ξ	Scale for host-pathogen combination	-
A	Area	m^2
q	Contribution of 1 added individual for average κ	-
m	Contribution of 1 S individual	-
n	Contribution of 1 I individual	-
$\omega_{S,I,R}$	Contact structure	-
ω_0	Critical population size	ind
p	Aggregation	-
θ	Heterogeneous mixing term	$\text{ind}^{-1}\ \text{t}^{-1}$

Table 2: Parameters of transmission functions described in table 1.

2.1 Homogeneity

We start our decomposition of the transmission function with homogeneous populations (Figure 1, left column). There are various different options for modelling the interaction between the susceptible and infected individuals, and a pathogen's transmission rate, β (McCallum et al., 2001; Begon et al., 2002; Hoch et al., 2008). However, all formulations of β depend on the contact (κ) between hosts, and thus the transmission function can be defined by the contact rate. In this paper, we introduce a novel general function for the contact rate from which we will show we can derive all βs described in table 1.

Decomposing transmission rate, β

In homogeneous populations, the rate at which new infections occur, β, is the product of the contact rate (κ) and the proportion of those contacts resulting in infection (probability c) (McCallum et al., 2001, 2017). Most commonly, β is described as:

$$\beta = \kappa \cdot c \tag{4}$$

Using this linear relationship between κ and c, we can then determine the R_0 of a disease as:

$$R_0 = \kappa \cdot c \cdot \tau \tag{5}$$

Where τ is the duration of the infection. This formulation is commonly adopted to approximate R_0 from empirical data (Anderson et al., 1992).

The dynamics of the transmission, $F(S, I)$, are dependent on the type of contacts between hosts, as described in the rate κ, such that, combining standard equation (2), with equation (4):

$$F(S, I) = \beta S \frac{I}{N} = \kappa c S \frac{I}{N} \tag{6}$$

Keeling and Rohani (2008) describe an alternative formulation to model frequency-dependent transmission:

$$\beta = -\kappa \cdot ln(1 - c) \tag{7}$$

Here, contacts are averaged per time unit and constant over density, this definition can therefore only be used for diseases where transmission depends on density-independent contact rates among hosts, such as is more often the case for sexually transmitted diseases.

General contact rate

The dynamics of κ determine the density-dependent (DD) or frequency-dependent (FD) nature of $F(S, I)$. Here we present a generalized version of this contact rate that can be adapted to reveal the different natures of the transmission functions, $F(S, I)$, described in table 1:

$$\kappa(N) = \frac{\xi N^{(1-q)}}{(\omega_0 + \omega_S S + \omega_I I + \omega_R R)} \tag{8}$$

The denominator shows the relative contribution of S, I and R on the disease spread managed by parameters ω_S, ω_I and ω_R, respectively. when $\omega_0 > 1$, the contact rate can shift from DD to FD with increasing total population size, N (the sum of S, I and R). Parameter q fixes the strength of DD versus FD contacts (Smith et al., 2009), and can generate the four main transmission functions depicted in Figure 2, as we will expand on in the next sections.

2.1.1 Density-dependence

The DD-transmission function is often referred to as *(pseudo-) mass action*, where population size can change but density remains constant (Kermack and McKendrick, 1927). When $\omega_S = \omega_I = \omega_R = \frac{A}{N}$, $q = 0$ and $\omega_0 = 0$ in equation (8), transmission is DD, and, ω_x in the denominator represents the relative space an individual occupies in area, A. Here, encounters between individuals are assumed to be brief, and thus scale linearly with population density. In this model, κ is thus host density-dependent, and can be simply expressed as:

$$\kappa(N) = \frac{\xi N}{A} \tag{9}$$

Where ξ is a scaling constant, which varies for each host and its specific pathogen, and determines the contacts between N hosts given area A (Begon et al., 2002; Smith et al., 2009). Here, the number of transmission events grows linearly with the density of hosts, and the number of cases can grow indefinitely (Anderson and May, 1982; Antonovics, 2017). Combining equation (2) and equation (9), the term N is canceled:

$$F(S,I) = \kappa c S \frac{I}{N}$$

$$= \frac{\xi N}{A} c S \frac{I}{N} \tag{10}$$

$$= \xi \cdot c \frac{SI}{A} \tag{11}$$

$$= \beta_A \frac{SI}{A} \tag{12}$$

This brings us to the first equation in table 1. Here, β_A includes area in its units. This can be useful for modelling wildlife populations, as area is important in moderating population behaviour (e.g. aggregation and dispersal) in response to density changes (Civitello et al., 2018).

Often times area is simply considered as a constant ($A = 1$), implicitly placing dependence on space in β, which further simplifies the the transmission function to:

$$F(S,I) = \beta_D SI \tag{13}$$

This returns our second equation in table 1.

It is assumed that with constant area, an increase in number of individuals causes an increase in density, hence *Density-dependence* (Kermack and McKendrick, 1927). This transmission model (Figure 3A) can fit well to directly transmitted diseases, such as the common cold or influenza, but is most commonly used for modeling wildlife diseases, such as rabies virus in reservoir dogs or tuberculosis in possums (White et al., 2017; Rhodes et al., 1998; Barlow, 1991).

2.1.2 Frequency-dependence

The mass-action function above, takes the assumption that the density of contacts per unit of time is proportional to the density of both susceptible and infected individuals (Kermack and McKendrick, 1927). When $\omega_S = \omega_I = \omega_R = 1$, $q = 1$ and $\omega_0 = 0$ in equation (8), contact rate is constant ($\kappa = \xi$), and independent of host population density, which returns Frequency-Dependent (FD) transmission:

$$\beta = \kappa \cdot c = \xi \cdot c \tag{14}$$

Plugging this into equation (2), we return the familiar FD transmission function (row 3 of table 1):

$$F(S,I) = \xi c S \frac{I}{N} = \beta_F \frac{SI}{N} \tag{15}$$

Here, infections rise with the probability ($\frac{I}{N}$), in contrast to assuming random interaction depending on densities, as in DD. Now, we are working with $\frac{I}{N}$, or the *density* of infected hosts, instead of *number* of hosts, as in DD. A key assumption here is the limited transmission due to the constant number of contacts per time unit (note that here β_F is an order of magnitude lower than β_D). An example of FD transmission includes sexually transmitted diseases for which transmission relies on the limited number of contacts or 'transmission events' between sexual partners. Vectored diseases, such as Lyme's disease, have also been approximated with FD, with transmission dependent on contact with an intermediate host. This vector determines the speed of transmission, which is limited by, for example, in the case of Lyme's disease, the number of meals the vector has per time unit, making the transmission host density independent. This FD force of infection dynamic is shown in Figure 3B.

An interesting approximation (and simplification) of FD transmission, can be derived by using carrying capacity (K) instead of population size: $F(S,I) = \frac{\beta_K SI}{K}$ (row 5, table 1. At low population densities transmission is FD; however, transmission does not asymptote at higher population densities, but rather resembles DD dynamics, thus combining the behaviour of DD with the infection speed of FD. This dynamic, hereby termed Carrying Capacity Dependent transmission, is shown in Figure 3D.

Comparison DD and FD

The per capita transmission in FD transmission ($\hat{\beta} = \frac{1}{SI}\frac{dI}{dt}$) declines with increasing population size (Ferrari et al., 2011), as evident from the measles outbreak in England (Bjørnstad et al., 2002), where increasing city size decreased the mean β, therefore following FD dynamics. In contrast, $\hat{\beta}$ for DD is constant and independent of population size. Therefore, in FD R_0 remains constant across varying N, but increases with N in DD (Ferrari et al., 2011). One consequence of this difference is that DD models predict a critical population density as an invasion threshold, which is not the case for FD.

2.1.3 Intermediate DD-FD

Often, host contact behaviour can be found somewhere on the DD-FD continuum (Anderson et al., 1992; Smith et al., 2009; Ferrari et al., 2011). For example, it is unlikely that contact rates increase indefinitely with N, as is assumed under DD, or are completely density independent, as assumed in FD (McCallum et al., 2001). The general contact rate function, defined above, can be adapted to model intermediate DD-FD transmission, by means of parameter q ($0 < q < 1$) and setting $\omega_S = \omega_I = \omega_R = \frac{A}{N}$, and $\omega_0 = 0$ in equation (8) such that:

$$\kappa(N) = \xi \frac{N^{(1-q)}}{A} \tag{16}$$

Where the dimensions for ξ are individuals$^{(q-1)}t^{-1}$, ξ & A are constants as in equation (9). Combining equation (16) with equation (2), brings us:

$$F(S,I) = \kappa c S \frac{I}{N}$$

$$= \xi \frac{N^{(1-q)}}{A} cS\frac{I}{N} \tag{17}$$

$$= c\xi \frac{SI}{A}\frac{1}{N^q} \tag{18}$$

$$= \frac{c\xi SI}{N^q} \tag{19}$$

$$\tag{20}$$

Substituting in β_q for $c \cdot \xi$, we get $\frac{\beta_q SI}{N^q}$. Units of β are also on the DD-FD spectrum, determined by q (table 1). With $0< q <1$, dynamics range from fully density-dependent to frequency-dependent. When $q = 1$, N drops and thus contact rate is again independent of N, thus $q = 1$ is FD, and $q = 0$ is DD. Therefore, we can interpret q as the relative importance of a single added host within a population to the average contact rate. This formulation was presented by Smith et al. (2009) to describe cowpox in voles, where q could also vary seasonally (more FD in late summer and more DD in late winter). These intermediate FD-DD dynamics are displayed in Figure 3C.

2.1.4 Asymptotic transmission

By including an additional term for the saturation of contacts at high population sizes, we can also allow transmission to vary between DD and FD depending on the total population size (Antonovics, 2017). When $\omega_S = \omega_I = \omega_R = 1$, q=0 and $\omega_0 > 0$ in equation (8), we can derive an asymptotic contact rate function:

$$\kappa(N) = \frac{\xi N}{\omega_0 + N} \tag{21}$$

where ω_0 is a constant which determines the behaviour of the contact function in relation to population size. When $\omega_0 << N$, transmission is FD, and when $\omega_0 >> N$, transmission is DD. The transmission function is then:

$$F(S,I) = \kappa cS\frac{I}{N}$$

$$= \frac{\xi N}{\omega_0 + N} cS\frac{I}{N} \tag{22}$$

$$= \frac{\xi}{\omega_0 + N} cSI \tag{23}$$

$$= \beta_u \frac{SI}{\omega_0 + N} \tag{24}$$

where $\beta_u = \xi c$, as in FD. Some diseases are more likely to spread asymptotically, such that at low host densities contacts are directly proportional to host density (DD), but a maximum rate of contact is obtained at high host densities (e.g. due to spatial or social distribution or limitation of time) (Diekmann and Kretzschmar, 1991; McCallum et al., 2001). We can thus define $\omega_0 = N^*$ (with $1 < \omega_0 < \infty$), as the half saturation constant in Michealis-Menten kinetics, which is the critical level after which contacts start to saturate. The lower ω_0, the higher the affinity of individual contacts and the quicker saturation of contacts is reached. Figure 3E shows the asymptotic transmission dynamics.

2.2 Heterogeneity

The models described above all assumed individuals were homogeneous for their susceptibility and that epidemiologically relevant contacts between hosts occur at random, known as the assumption of *random mixing*. However, this is not always the case, for example, a host's behaviour can change the probability of encountering an infected individual and the geographic distribution and movement of hosts plays a significant role as well (Pope et al., 2007; Vicente et al., 2007). This heterogeneity can affect the overall disease spread (Ferrari et al., 2011), which is why we decompose $F(S,I)$ seperately for these circumstances (Figure 1, right column).

Behavioural heterogeneity

There are various approaches that allow us to include heterogeneity in behaviour of S and I, most accurate of which are network or individual-based models (Ferrari et al., 2011). Here, we review some simplified and commonly adopted formulations for compartmental SIR models.

2.2.1 Power-law transmission

This non-linear density-dependent transmission function simulates heterogeneous interactions between S and I individuals with accelerating and decelerating disease spread by raising S and I to the powers m and n, respectively (Hochberg, 1991):

$$F(S,I) = \beta_{PL} S^m I^n \tag{25}$$

Parameters m and n determine the speed of infection, and represent how the densities S and I affect the *per capita* transmission efficiency of the pathogen, such as the susceptibility (by m) or the infectivity (by n) of an individual (Hochberg, 1991; Novozhilov, 2008). When both parameters are unity, the model describes simple DD (with constant area). There are many possible response types, as theoretically $-\infty < m, n < \infty$. Some biologically relevant ones are considered further, below.

Assuming $n = 1$, when $m > 1$, $F(S,I)$ accelerates, and new infections arise exponentially with linearly increasing number of susceptible individuals. When $0 < m, n < 1$ rate of new infections decelerates, and this decelerating response behaves similarly to standard DD transmission Gubbins et al. (2000). Figure 3F shows the power law transmission dynamics with varying n. We could find no real world examples of exponents $0< m <1$, but it would represent a case for which increasing the number of S decreases the rate of transmission (Novozhilov, 2008). Greer et al. (2008) found that *Ambystoma tigrinum* virus in its host was best described with a Power-law function with $m = 1$ and $n = 0.255$, suggesting that infection scales positively but non-linearly with I. When $m < 0$, S is inversely related to $F(S,I)$, and when $n <0$, I is inversely related to $F(S,I)$. An example of the latter is seen in *Bacillus thuringiensis* infections of *Plodia interpunctella* moths, transmission requires cannibalism of infected individuals, and thus transmission increases with S, and a decrease with I (Knell et al., 1996). When $n = 0$, the infection risk is a constant, and independent of the density of infected individuals, as for example where there is constant re-inoculation of the disease from another host, e.g. rabies from a bat reservoir host (Greer et al., 2008; Mollentze et al., 2020). By sampling from a distributions of m and n, the power-law function can also approximate heterogeneity in transmission rate among S and I individuals (White et al., 2017).

Spatial heterogeneity

The assumption in SIR models of spatial homogeneous populations is an oversimplification of realistic population dynamics, especially for species exhibiting territoriality, sociality and other complex group or individual behaviours (Cross et al., 2004; Viana et al., 2014; White et al., 2017). For example, all else being equal, spatial proximity to an infected individual would be predicted to increase transmission potential. Contact networks provide a useful approach to model population behaviour (Craft, 2015), but can rapidly become analytically intractable. Here we illustrate how heterogeneity in contacts can be included into our compartmental SIR models, making a few simplifying assumptions, and holding transmission rate, β, constant.

2.2.2 Refuge effect

The refuge model, first applied to TB in possums, attempts to include spatial aggregation of hosts, with the transmission function incorporating host clustering (Barlow, 1991). In this model, the population consists of patches, with and without disease, and p is the proportion of the total area that is occupied by diseased individuals, and represents how aggregated the disease is in space. The more the disease is aggregated (the smaller p), the lower the

probability that an infectious individual will encounter a susceptible individual, and thus:

$$F(S,I) = \beta_r I(N - \frac{I}{p}) \tag{26}$$

Assuming $N = S + I$, so in this case a simple SI model excluding the R compartment, the number of susceptible individuals is $S = Np - I$. If $p = 1$, the diseased patch equals the total area, there is no aggregation of the disease, and the transmission function simplifies to $F(S,I) = \beta_r I(N - I) = \beta_r SI$, equation (13), DD dynamics. As p decreases ($0< p <1$), disease aggregation increases as a proportion p of the total area, as increasing I/p, decreasing S, indicates a greater proportion of infected individuals occupying the patch. With greater aggregation, transmission $F(S,I)$ decreases, as individuals are now less homogeneously distributed. For $I = pN$, which is the maximum density of individuals in the diseased patch, $F(S,I) = 0$.

Possums in New Zeland revealed an aggregated spatial prevalence pattern due to variation in carrying capacity across habitat patches (Barlow, 1991; May and Anderson, 1984), returning a maximum contact rate per patch, $\beta_r K$. Figure 4A shows this transmission dynamic for three different values of p.

2.2.3 Negative binomial

The negative binomial model builds upon the Refuge effect (Barlow, 1991) and was also fit to model TB in possums in a heterogeneous environment, but with varying carrying capacities (different resource availability) per patch, and therefore allowed for increased contact rates by concentrating hosts (Civitello et al., 2018). Here transmission is modelled as:

$$F(S,I) = \theta S \cdot ln(1 + \frac{\beta_b I}{\theta}) \tag{27}$$

and includes a heterogeneous mixing term, θ (ind^{-1} t^{-1}), with values $0 < \theta < \infty$ (Barlow, 2000). Small θ corresponds to highly aggregated infections. The model assumes that the probability of infection is higher when a susceptible individual has an infected neighbor, and so infections are clustered in space. As $\theta \to 0$, $F(S,I)$ decreases, as the natural log around I approaches 0, as the mean number of infected individuals encountered per susceptible individual is reduced. As $\theta \to \infty$, aggregation is reduced, and $F(S,I)$ again simplifies to DD transmission. This negative binomial dynamic is shown in Figure 4B.

2.3 Conclusion transmission functions

The contact rate between hosts determines the shape of transmission dynamics. There are multiple approaches for describing the spread of a pathogen through a host population, the most common distinction is between DD and FD dynamics, but there are numerous variations, reflecting differences in the biology of the pathogen and the behaviour of the host. For example, wildlife diseases are almost always modeled assuming DD dynamics, regardless of the pathogen, as their behaviour makes them more likely to approximate mass-action dynamics (White et al., 2017). In contrast, human interactions commonly depart from assumptions of DD, as people frequently form small contact networks, and might thus be better modelled using FD dynamics. However, many empirical studies have shown evidence of both FD and DD dynamics in the same system, with density-dependent dynamics more often observed at low population densities, and frequency-dependent dynamics more common at higher densities (Diekmann and Kretzschmar, 1991; Antonovics, 2017; Roberts and Heesterbeek, 2018). Unsurprisingly, more complex dynamics are also to be found in many disease systems, for example, feline retrovirus, in which infrequent encounters at low host densities make contact rates close to constant (FD), intermediate densities contact rate becomes proportional to density (DD), while at high densities there is a further increase in contact rate due to overlapping territories (Fromont et al., 1998). These and other classic case studies, such as in the example of how culling of badgers increased their intraspecific contacts and increased tuberculosis prevalence, shows the importance of contact structure in modeling wildlife diseases (Vicente et al., 2007; Pope et al., 2007; McCallum, 2016; White et al., 2017). However, as shown in Cross et al. (2004), contacts in complex networks, such as buffalo, tend to homogenize over time. Therefore at larger

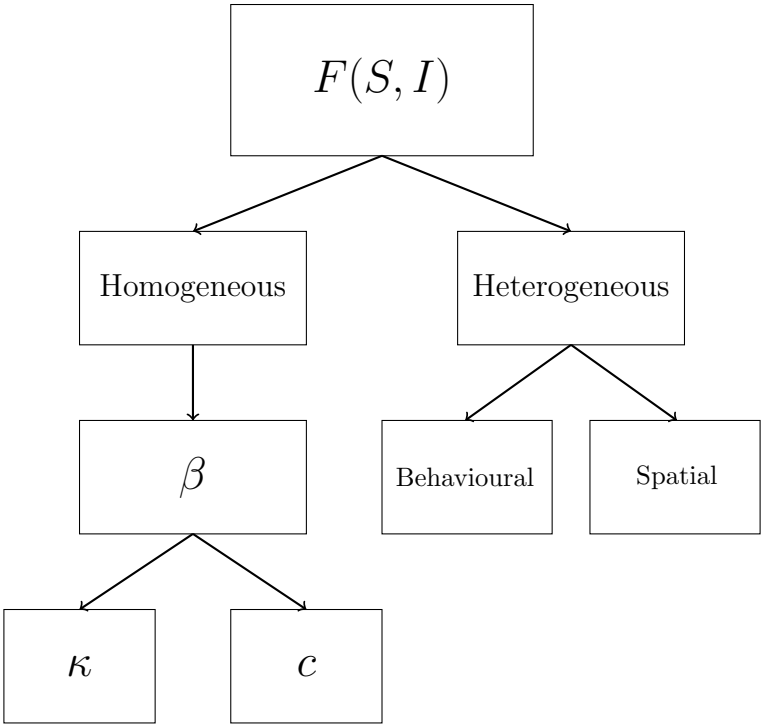

Figure 1: Graphical representation of the decomposition of the transmission function, $F(S, I)$, in traditional SIR compartmental disease models. β is the transmission parameter, κ the contact rate and c the probability of successful transmission. On the left, we decompose the transmission parameter for homogeneously distributed host populations. On the right we show heterogeneously distributed host population, of which the dynamics are more complex and therefore the transmission function is not as simply decomposed.

timescales, heterogeneity can average out, giving the modeller the opportunity to simplify models by dropping the complex heterogeneity. Modellers should therefore keep in mind both the timescales and the hosts' and pathogen's behaviour when modelling a disease accordingly.

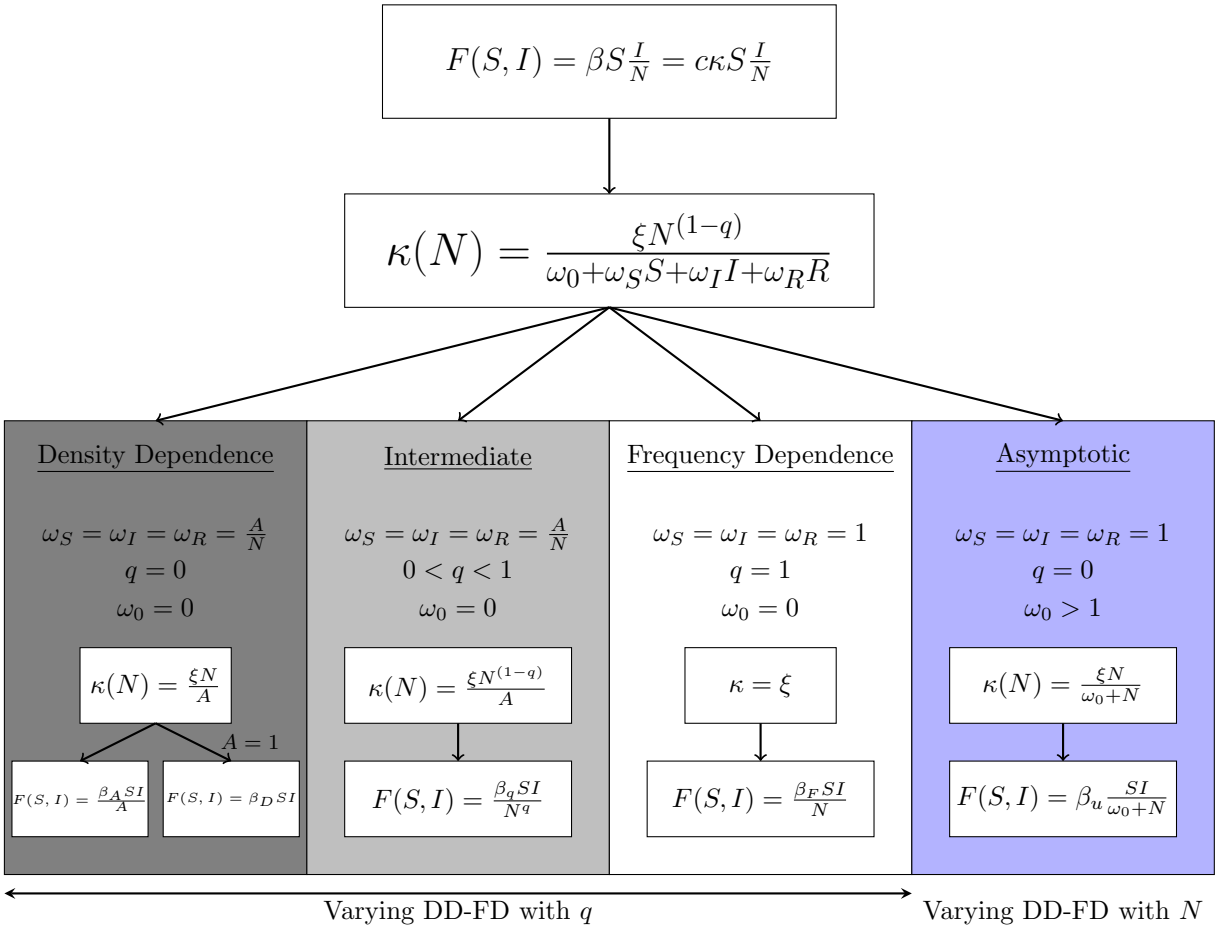

Figure 2: Flow chart showing how the contact rate function, $\kappa(N)$, determines the density- or frequency-dependent nature of the transmission function, $F(S,I)$. This in turn determines the dimensions of the transmission coefficient β_x. The first 3 panels show a spectrum from fully DD to FD, determined by parameter q. The right panel shows a transmission function that can vary from DD to FD with population size N.

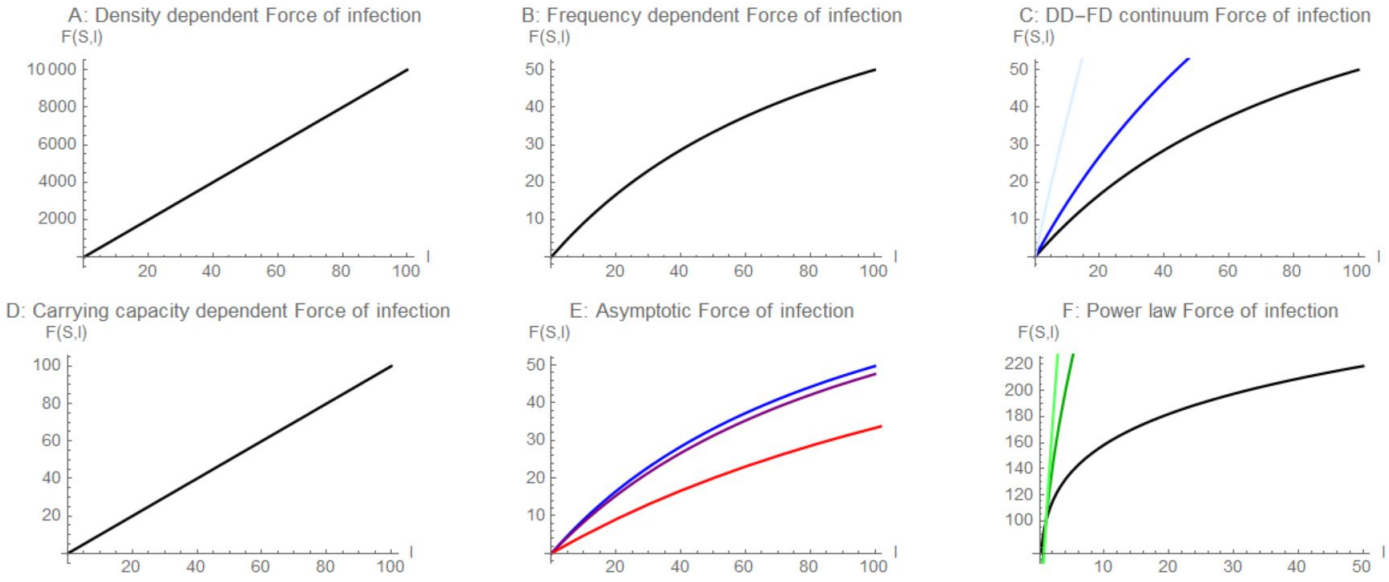

Figure 3: F(S,I) in homogeneous populations with S constant at 100 individuals. A: DD transmission. B: FD transmission C: Intermediate transmission. With light blue, $q = 0.7$, in blue $q = 0.9$ and black, $q = 1$. D:Carrying capacity with $K = 100$. E: Asymptotic transmission with in red ω_0=1, in purple ω_0=10 and in blue ω_0=100. F: Power law transmission, black $n = 0.2$, dark green $n = 0.5$, light green $n = 0.8$.

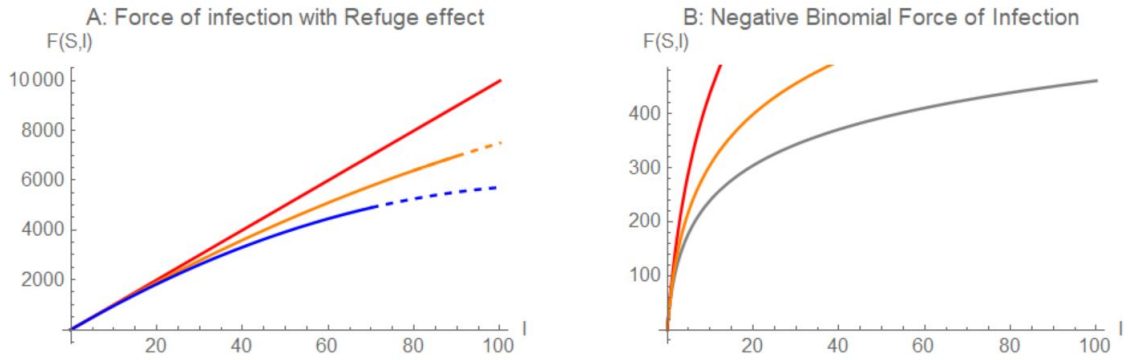

Figure 4: F(S,I) in spatially heterogeneous populations, S set at 100. A: Refuge effect.Red, $q = 1$, orange $q = 0.8$, blue $q = 0.7$. Dashed line indicates the end of the biologically relevant transmission. B: Negative binomial. Gray, $k = 1$, orange, $k = 1.5$, red, $k = 3$.

3 Transmission in multi-host systems

Pathogens that infect only a single host are rare, it is far more common for a pathogen to infect a range of host species (Woolhouse et al., 2001). This can significantly alter disease dynamics, host community structure can affect pathogen dynamics and the pathogen can affect host community structure. In multi-host systems, transmission is determined by multiple shedding hosts, each of which can contribute differently to disease prevalence depending on their inter- and intraspecific ecological interactions (Haydon et al., 2002; Streicker et al., 2013; Fenton et al., 2015). This complicates models and how we estimate transmission, as transmission may be asymmetric among hosts, with some hosts acting as pathogen reservoirs, and pathogens may express different life-history syndromes in different hosts (Haydon et al., 2002; Gandon, 2004). In this section, we describe how multi-host, single-pathogen systems are influenced by host diversity, the role of transmission, and how transmission can be defined in such systems.

3.1 Host community dynamics

The dynamics of multi-host diseases can vary greatly with the species composition of the host community. One important relationship in wildlife disease dynamics is that between disease prevalence and host community diversity. The disease-diversity relationship is a well-known and highly-disputed concept in disease ecology.

3.1.1 Host community effects on pathogen prevalence

The *dilution effect* suggests higher diversity of hosts reduces the probability of a pathogen infecting a new host, for example, directly, by reducing encounters, and therefore transmission, between hosts or indirectly, by changing total host abundance (Keesing et al., 2006, 2010) (table 3). An example of encounter reduction (also known as *frequency-dependent dilution*) is provided in FD systems according to a model by Rudolf and Antonovics (2005), where host is rescued from pathogen mediated-extinction (*apparent mutualism*) by a second host that is infected by the same disease. Here, the secondary host is assumed less competent, and therefore functions as a buffer, reducing further disease spread by replacing contacts (and therefore transmission events) with the original host, reducing frequency of contacts between competent hosts. Lyme disease is a frequently cited example of the dilution effect where diversity of vertebrates hosts decreases the risk of spillover to humans. The bacterial pathogen (*Borrelia burgdorferi*) is vectored by the black legged tick (*Ixodes scapularis*), and uses the white footed mouse (*Peromyscus leucopus*) as a primary host; however, prevalence of *B. burgdorferi* decreases when a secondary host, the eastern Chipmunk (*Tamias striatus*), increases in density (Keesing et al., 2006). Evidence for a dilution effect was also found in bovine TB in sub-Saharan African mammals, which was reduced at higher mammal density (Huang et al., 2013, 2014).

In contrast to the dilution effect, the *amplification* effect suggests that increasing the number of host species increases disease prevalence. For example, *directly*, by elevating contact rates by increasing total host density (Rudolf and Antonovics, 2005) or by the addition of a highly competent, super-spreader host (a phenomenon parallel to *the selection effect* in biodiversity science (Loreau and Hector, 2001)), or *indirectly* by changing host densities through competition Holt and Bonsall (2017) (see table 3). For instance, higher amphibian diversity is thought to have increased Chytrid disease (caused by *Batrachochytrium dendrobatidis*) prevalence in some species of frogs, as high competent, and thus amplifying species, are more abundant in species-rich habitats (Ostfeld and Keesing, 2012). Another example of the amplification effect can come from vectored diseases (often modelled as FD due to density-independent contacts) if the increased diversity offers more competent host species so that this multi-host pathogen can more easily persist in the system. (Ostfeld and Keesing, 2000). In FD, amplification only occurs depending on the *identity* of the hosts (its competence). Theory suggests that, in DD systems, amplification will always occur, as contacts increase with host density, and there is no contacts are always additive, and never substitutive (Dobson, 2004; Rudolf and Antonovics, 2005). Regardless of the competence of the added host, without contact substitution the new host will always add to the overall prevalence. Whether diversity dilutes or amplifies disease is a major conservation conundrum, and the composition of the transmission function plays an important role in determining

Mechanism:	**Dilution**	**Amplification**
FD	Direct: Encounter reduction	Indirect
DD	Indirect	Direct: Additive encounters or selection effect

Table 3: Dilution and amplification occurring through different mechanisms. Varying host competence is assumed everywhere, but in Direct Amplification in DD pathogens, where varying host competence is not required.

dilution or amplification dynamics.

The direct effect of diversity in DD transmission will always lead to disease amplification, assuming host populations do not compete and the addition of a host is *additive* to the community. Increasing the number of hosts simply increases encounter rates between S and I individuals, and will therefore always amplify the disease, regardless of the competence of the hosts. Theoretically, in DD transmission dilution through direct effects is impossible. The direct effect of diversity in FD transmission, assuming a constant number of contacts per time unit and varying competence, the addition of a less competent host will always have a diluting effect, whereas the addition of a highly competent host may have an amplifying effect. Table 3 shows how transmission type can cause dilution or amplification
Predicting the **indirect effects** of diversity is more challenging. Changes in prevalence will reflect changes in the relative abundance of hosts, and thus how communities are altered by the introduction of novel host species (Ostfeld and Keesing, 2000). Community evenness is a particularly important dimension in multi-host disease prevalence, as the relative dominance of the most competent host will largely determine disease dynamics (Ostfeld and Keesing, 2000; Sintayehu et al., 2017). For example, dilution can occur indirectly in Lyme's disease when the alternative host reduces the density of the main reservoir through competition (Ogden and Tsao, 2009).

In all cases, it is assumed that the pathogen is a generalist and that hosts differ in competence, except for direct amplification in DD systems, which assumes contacts are additive and there is no limit to contacts per time unit. However, these latter assumptions may be unrealistic for real-life systems (Ostfeld et al., 2008), and host contacts likely saturate over higher density Antonovics (2017). Thus, at a certain density, disease dynamics may switch from DD to FD dynamics, and so we might predict dilution effects to be more common at higher densities.

3.1.2 Pathogen effects on host community

While we expect a pathogen's persistence to reflect host abundance, host abundance might also be influenced by the pathogen. For example pathogen sharing among hosts can result in *apparent competition* (Holt and Pickering, 1985). A reservoir host can indirectly suppress the density of a spillover host by acting as a source of infection, and if the reservoir host has a better adapted immune system to the pathogen, it can indirectly reduce spillover host density via pathogen induced mortality. A potential example of apparent competition is spillover of the vector-transmitted Barley Yellow Dwarf virus (genus *Luteovirus*) from wild oats (*Avena fatua*), which reduced the abundance of the spillover host species (of genus *Setaria*), serving the wild oat to maintain as a dominant species (Power and Mitchell, 2004). Although disease dynamics may appear similar, and may feedback into dilution or amplification effects, apparent competition describes the effect of pathogen presence on hosts. Dilution and amplification are theories applying to pathogen prevalence in response to host diversity. Apparent competition can only happen in DD systems, or FD ones given that the pathogen is not directly diluted by the less susceptible, spillover host(s).

Apparent mutualism, on the other hand, can occur in FD systems when hosts have equal competence. In this case, the introduction of a novel host can reduce the disease prevalence in the original host community members by replacing contacts. There is no 'competition' as all hosts are equally affected by the pathogen (Holt and Bonsall, 2017), and all benefit form the addition of additional hosts. Apparent mutualism differs from dilution, as dilution explicitly describes the reduction of pathogen prevalence due to host composition, and does not focus on how host abundance is

affected by the pathogen's presence. Also, dilution requires difference in host competence, which apparent mutualism does not.
In summary, assuming FD dynamics, increasing host diversity can result in apparent mutualism when hosts are equal in their competence, and dilution can occur as a type of pathogen prevalence reduction in a focal host through increasing diversity, when hosts have unequal competence. Apparent competition can occur if hosts vary in competence and one host 'outcompetes' the others through having a superior immune system (lower pathogen related mortality). Assuming DD dynamics, increasing diversity results in amplification whether hosts have same or different competence, assuming each host species is additive to disease transmission.

3.2 Inter-species transmission

Changing host community structure changes intraspecific and interspecific contact rates, and thus disease dynamics. The interspecies transmission rate determines the effect of species richness on the outbreak potential, R_0, as host competence functions mainly though its ability to transmit a pathogen (Dobson, 2004). Transmission competence determines a host's role as amplifier or diluter, as the persistence of a disease in multiple hosts is determined by these interspecies and intraspecies transmission dynamics (section 3.1.1). For example, jackals (*Canis adustus*) need to be frequently reinoculated by Rabies virus from domestic dogs in order to support an infection in the population (Rhodes et al., 1998; Keesing et al., 2006) as their intraspecies transmission is too low to sustain endemic prevalence, making jackals more likely a diluter than an amplifier. There are many approximations for interspecies transmission, focusing on the ecology of hosts or the co-evolution of pathogens and hosts (for example by using matching-allele and gene-for-gene models, (Poullain and Nuismer, 2012)). Here, we examine a sample of those we consider most promising for addressing disease-diversity questions.

Interspecific transmission can be modeled in a WAIFW-matrix (Who Acquired Infection From Who), an n by n - matrix (here called W) showing the transmission between species in an n-species system, where intraspecies transmission can be found on the diagonal (Dobson and Foufopoulos, 2001; Diekmann and Kretzschmar, 1991):

$$W = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & & & \\ \vdots & \ddots & & \vdots \\ \beta_{n1} & & & \beta_{nn} \end{bmatrix} \tag{28}$$

In this matrix, transmission can be defined as $\beta_{i,j}$, where i is the receiving host and j the donating host (j infects i). This matrix can also be used to calculate the community R_0, which can determine the overall disease prevalence and contributions of each host (Dobson and Foufopoulos, 2001). One way to estimate the interspecific transmission rate is to take the average of the intraspecific transmission rates (Dobson, 2004).

$$\beta_{i,j} = \alpha_{i,j} \frac{\beta_{i,i} + \beta_{j,j}}{2} \tag{29}$$

where $\alpha_{i,j}$ is a scaling parameter to account for differences in transmission potential between species i and j, and determines the magnitude of diluting and amplifying behaviour of each species in the system (Dobson, 2004). This same method of averaging between species can also be applied to contact rates, κ, as defined in $\beta = \kappa * c$, if one wants to define c separately for individual host species, rather than the community average (Anguelov et al., 2014). Multi-host systems can have asymmetrical transmission between hosts. By using the average of the intraspecific transmission rates, asymmetries and other characteristics that are involved in interspecific transmission may be overlooked. For example Blancou and Aubert (1997) suggest that for a fox with rabies to infect another species, such as a dog or cat, a million times more virus particles are necessary than to infect another fox (Ostfeld et al., 2008).

3.3 Quantifying contacts

As transmission is highly dependent on the contact structure of hosts, it is important to consider this structure into the definition of transmission in models. Heterogeneity in contacts can drastically change the initial spread and final outbreak of the disease, determining the R_0 (Eames and Keeling, 2003). Craft (2015) argued that network models are required to accurately model wildlife populations. For instance, in management of wildlife diseases, defining super-spreaders and their role in the network can be critical for intervention management. As example in a simulation study of wild chimpanzees infected by a density-dependent disease, parameterized by empirical contact data, it was shown that vaccinating the most connected individuals reduces the vaccination threshold by 35% compared to random vaccination (Rushmore et al., 2014). New advances in biomonitoring methods can help in the construction of contact network models, capturing the complexity in interactions that many animals show, including territoriality, sociality and individual variation in movements (Viana et al., 2014; White et al., 2017). These networks can also be useful if we wish to describe transmission between inter-connected populations. While we do not review this complexity here, we expand on some of these exciting methods in the appendix 5.1. However, obtaining the data to construct network models is very research intensive, and simplifying assumptions of homogeneity are frequently adopted. As social networks homogenize over time Cross et al. (2004), such simplifying assumptions might not be unrealistic. Nonetheless, new methods, such as use of camera-trapping of animals at aggregation sites, can be adopted to estimate intraspecies and interspecies contact rates by calculating the average visitation rates using information on species abundances and site visits (Barasona et al., 2017). Although less accurate, such estimates can still be used to inform SIR models. Ultimately, disease transmission is determined by contact rates, and thus defining these rates accurately is a critical step if we wish to generate realistic projections of disease spread.

Box 3: Empirical approximations of intra-species transmission:

Estimating β directly: Because it is a challenge to quantify transmission rates directly, often approximations are used. For instance, viral particles activate effector T-cells, data on infection load can be approximated by measuring these T-cells from serological samples of animal hosts (Blaser et al., 2014; Almocera and Hernandez-Vargas, 2019). Individual infection status can be determined for viral pathogens by PCR, and viral pathogen load can be used to derive transmission rates (Streicker et al., 2010; Blaser et al., 2014; Almocera and Hernandez-Vargas, 2019). Genome sequencing is a promising technique for estimating transmission rates for bacterial pathogens, as was illustrated for *Mycobacterium bovis*, and can reveal spatial patterns of transmission (Biek et al., 2012). For viruses, we can adopt phylodynamic approaches to infer transmission rates from birth-death models (MacPherson et al., 2020). Basic host demography parameters, such as K, δ (host natural mortality) and r (host birth rate) are easily available. These parameters can be allometrically scaled with host body size, therefore, we can similarly scale β (De Leo and Dobson, 1996). For vector transmitted diseases, surveys of vectors associated with hosts can give an approximation of the rate of spread of the disease, and may provide insights into interspecific transmission (Lu et al., 2010).

Estimating the probability of successful transmission, c: When the contact structure of the population is known, for example through biological monitoring, it is still challenging to define the infectivity of a contact - the probability of successful transmission after contact, c (Craft, 2015; White et al., 2017). One approach used for wildlife disease dynamics is the Secondary Attack Rate (SAR), which is the ratio of the number of exposed hosts that developed the disease to the number of exposed hosts that did not (Childs et al., 2007). Here, a clear distinction must be made between the primary and secondary cases, in order to accurately calculate this rate. For sexually transmitted diseases, which follow FD dynamics, the probability of becoming infected after contact can be calculated using the binomial distribution (Childs et al., 2007). The maximum likelihood of this probability following a single contact is identical to the SAR. Estimating disease transmission remains a challenge, and with most disease models summarizing the process of transmission into a single parameter, it is important to capture this value accurately. These methods are described in further detail in appendix 5.2.1.

4 Future challenges and conclusion

Defining the shape of the transmission function between species is challenging, and seemingly small differences can have dramatic effects on predictions from multi-species models (Dobson, 2004). Given current rate of biodiversity and habitat degradation worldwide, there is an urgent need for studies on the importance of how reduced diversity and resource and spatial limitations on hosts may influence the transmission rates of their pathogens and how this might cascade to possible spillover events. Habitat fragmentation and a decrease in the carrying capacity of an ecosystem can both affect transmission (Childs et al., 2007; Lafferty and Holt, 2003), and at the interface between natural and converted landscape we can find increased interspecies transmission (Wolfe et al., 2005; Faust et al., 2018; Goldberg et al., 2008).

As most hosts and pathogens exist within multi-host systems, we need to better understand how transmission affects disease outbreaks in such systems. The next-generation matrix (Diekmann and Kretzschmar, 1991), as described in section 3.2, calculating the community R_0 and each host's relative contribution, is a promising framework for understanding the dynamics of multi-host pathogens and disease maintenance in the reservoir. However, the complexity in quantifying transmission, with asymmetries in interspecies rates and among potential hosts following a spillover event (Wolfe et al., 2007; Auld et al., 2017), can make the application of such models fraught. Additional challenges include accounting for spatial heterogeneity, although over longer timescales contact structures may appear more homogeneous (Cross et al., 2004). However, as pathogen sharing is affected by overlapping geographical range (Davies and Pedersen, 2008), investigating the effect of local versus global dynamics remains important.

It is also becoming increasingly clear that the evolutionary relationship between species plays a role in disease transmission, with strong evidence of phylogenetic signal in the likelihood of pathogen sharing among hosts (Davies and Pedersen, 2008; Farrell et al., 2019; Parker et al., 2015; Streicker et al., 2019; Olival et al., 2017). It is likely, for example, that similarity in the immune defenses of closely related species due to the evolutionary conservation of the cellular, immunological, or metabolic traits, favours virus exchange between them (Kuiken et al., 2006; Streicker et al., 2010). Similar phylogenetic signature in pest and pathogen sharing is observed in plants (Gilbert and Webb, 2007; Gilbert et al., 2012; Ssebuliba and Davies, 2021), and phylogeny is also suggested to be a strong predictor of pathogen impact, with declining severity of the effect of the disease with increasing evolutionary distance between hosts (Gilbert et al., 2015). It would be relatively straightforward to include this information in our models. For instance, following Parker et al. (2015), we can define the probability of successful transmission, c, to be dependent on the evolutionary relationship between hosts. Assuming that intra-species probability of successful transmission is $c_i = \lambda_i$, we can then simply define the probability of successful transmission of a pathogen from the recipient species, i, to the donor species j, c_{ij}, to be:

$$c_{ij} = \lambda_j \cdot \frac{1}{\psi PPD_{ij}} \tag{30}$$

where ψ is a scaling constant and PPD the Pairwise Phylogenetic Distance between the two hosts (in Million Years). While the phylogenetic diversity of hosts - the sum of the evolutionary branch lengths connecting species - (Faith, 1992) can provide a useful predictor of disease prevalence (Rolland et al., 2012; Huang et al., 2013, 2014), the phylogenetic distance separating species may be a more useful metric for scaling the probability of interspecific transmission, and in predicting novel host shifts (Poullain and Nuismer, 2012).

Spillover events are the driver of zoonotic epidemics, including those which precipitated the current COVID-19 pandemic and transmission rates provide our best predictors of future host shifts and potential emergence events (Poullain and Nuismer, 2012). Therefore, it is of utmost importance to define transmission and its components, most importantly contact rates, accurately. In this paper we decomposed transmission into separate parameters and have shown how these are involved in the outbreak process. We defined a general contact rate that can encompass all preexisting variations of transmission functions and describe how the contact rate is at the base of each of these. With this framework, and the proposed relationship between evolutionary relatedness and probability of successful transmission, we hope to provide the tools to investigate the underlying mechanisms of the effects of host diversity

on potential disease outbreaks.

References

Almocera, A. E. S. and Hernandez-Vargas, E. A. (2019). Coupling multiscale within-host dynamics and between-host transmission with recovery (SIR) dynamics. *Mathematical Biosciences*, 309(January):34–41.

Anderson, R. M. and May, R. M. (1982). Coevolution of Hosts and Parasites. *Parasitology*, 85(2):411–426.

Anderson, R. M., May, R. M., and Ng, T. W. (1992). A ge-dependent choice of sexual partners and the transm ission dynam ics of H IV in S ub -S ah aran Africa. pages 135–155.

Anguelov, R., Garba, S. M., and Usaini, S. (2014). Backward bifurcation analysis of epidemiological model with partial immunity. *Computers and Mathematics with Applications*, 68(9):931–940.

Antonovics, J. (2017). Transmission dynamics: Critical questions and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719).

Auld, S. K., Searle, C. L., and Duffy, M. A. (2017). Parasite transmission in a natural multihost-multiparasite community. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719):1–10.

Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25):6506–6511.

Barasona, J. A., Vicente, J., Díez-Delgado, I., Aznar, J., Gortázar, C., and Torres, M. J. (2017). Environmental presence of mycobacterium tuberculosis complex in aggregation points at the wildlife/livestock interface. *Transboundary and emerging diseases*, 64(4):1148–1158.

Barlow, N. D. (1991). A spatially aggregated disease/host model for bovine tb in new zealand possum populations. *Journal of applied ecology*, pages 777–793.

Barlow, N. D. (2000). Non-linear transmission and simple models for bovine tuberculosis. *Journal of Animal Ecology*, 69(4):703–713.

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

Biek, R., O'Hare, A., Wright, D., Mallon, T., McCormick, C., Orton, R. J., McDowell, S., Trewby, H., Skuce, R. A., and Kao, R. R. (2012). Whole Genome Sequencing Reveals Local Transmission Patterns of Mycobacterium bovis in Sympatric Cattle and Badger Populations. *PLoS Pathogens*, 8(11).

Bjørnstad, O. N., Finkenstädt, B. F., and Grenfell, B. T. (2002). Dynamics of measles epidemics: estimating scaling of transmission rates using a time series sir model. *Ecological monographs*, 72(2):169–184.

Blancou, J. and Aubert, M. (1997). Transmission of rabies virus: importance of the species barrier. *Bulletin de L'academie Nationale de Medecine*, 181(2):301–11.

Blaser, N., Wettstein, C., Estill, J., Vizcaya, L. S., Wandeler, G., Egger, M., and Keiser, O. (2014). Impact of viral load and the duration of primary infection on hiv transmission: systematic review and meta-analysis. *AIDS (London, England)*, 28(7):1021.

Childs, J. E., Mackenzie, J. S., and Richt, J. A. (2007). *Wildlife and emerging zoonotic diseases: the biology, circumstances and consequences of cross-species transmission*, volume 315. Springer Science & Business Media.

Civitello, D. J., Allman, B. E., Morozumi, C., and Rohr, J. R. (2018). Assessing the direct and indirect effects of food provisioning and nutrient enrichment on wildlife infectious disease dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1745).

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669).

Cressler, C. E., McLeod, D. V., Rozins, C., Van Den Hoogen, J., and Day, T. (2016). The adaptive evolution of virulence: A review of theoretical predictions and empirical tests. *Parasitology*, 143(7):915–930.

Cross, P. C., Johnson, P. L., Lloyd-Smith, J. O., and Getz, W. M. (2007). Utility of R0 as a predictor of disease invasion in structured populations. *Journal of the Royal Society Interface*, 4(13):315–324.

Cross, P. C., Lloyd-Smith, J. O., Bowers, J. A., Hay, C. T., Hofmeyr, M., and Getz, W. M. (2004). Integrating association data and disease dynamics in a social ungulate: Bovine tuberculosis in African buffalo in the Kruger National Park. *Annales Zoologici Fennici*, 41(6):879–892.

Daszak, P., Cunningham, A. A., and Hyatt, A. D. (2000). Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science*, 287(5452):443–449.

Davies, T. J. and Pedersen, A. B. (2008). Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings of the Royal Society B: Biological Sciences*, 275(1643):1695–1701.

de Jong, M. C., Diekmann, O., and Heesterbeek, H. (1995). How does transmission of infection depend on population size. *Epidemic models: their structure and relation to data*, 5(2):84–94.

De Leo, G. A. and Dobson, A. P. (1996). Allometry and simple epidemic models for microparasites. *Nature*, 379(6567):720–722.

Dieckmann, U., Metz, J. A., and Sabelis, M. W. (2005). *Adaptive dynamics of infectious diseases: in pursuit of virulence management*. Number 2. Cambridge University Press.

Diekmann, O. and Kretzschmar, M. (1991). Patterns in the effects of infectious diseases on population growth. *Journal of Mathematical Biology*, 29(6):539–570.

Dobson, A. (2004). Population Dynamics of Pathogens with Multiple Host Species. 164(november).

Dobson, A. and Foufopoulos, J. (2001). Emerging infectious pathogens of wildlife. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1411):1001–1012.

Eames, K. T. and Keeling, M. J. (2003). Contact tracing and disease control. *Proceedings of the Royal Society B: Biological Sciences*, 270(1533):2565–2571.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological conservation*, 61(1):1–10.

Farrell, M. J., Govender, D., Hajibabaei, M., Van Der Bank, M., and Davies, T. J. (2019). Bacterial diversity in the waterholes of the Kruger National Park: An eDNA metabarcoding approach. *Genome*, 62(3):229–242.

Faust, C. L., McCallum, H. I., Bloomfield, L. S., Gottdenker, N. L., Gillespie, T. R., Torney, C. J., Dobson, A. P., and Plowright, R. K. (2018). Pathogen spillover during land conversion. *Ecology Letters*, 21(4):471–483.

Fenton, A. and Pedersen, A. B. (2005). Community epidemiology framework for classifying disease threats. *Emerging Infectious Diseases*, 11(12):1815–1821.

Fenton, A., Streicker, D. G., Petchey, O. L., and Pedersen, A. B. (2015). Are all hosts created equal? Partitioning host species contributions to parasite persistence in multihost communities. *American Naturalist*, 186(5):610–622.

Ferrari, M. J., Perkins, S. E., Pomeroy, L. W., and Bjrnstad, O. N. (2011). Pathogens, social networks, and the paradox of transmission scaling. *Interdisciplinary Perspectives on Infectious Diseases*, 2011.

Fine, A. E., Bolin, C. A., Gardiner, J. C., and Kaneene, J. B. (2011). A study of the persistence of mycobacterium bovis in the environment under natural weather conditions in Michigan, USA. *Veterinary Medicine International*, 2011.

Fromont, E., Pontier, D., and Langlais, M. (1998). Dynamics of a feline retrovirus (FeLV) in host populations with variable spatial structure. *Proceedings of the Royal Society B: Biological Sciences*, 265(1401):1097–1104.

Gandon, S. (2004). Evolution of multihost parasites. *Evolution*, 58(3):455–469.

Gilbert, G. S., Briggs, H. M., and Magarey, R. (2015). The impact of plant enemies shows a phylogenetic signal. *PLoS ONE*, 10(4):1–11.

Gilbert, G. S., Magarey, R., Suiter, K., and Webb, C. O. (2012). Evolutionary tools for phytosanitary risk analysis: Phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evolutionary Applications*, 5(8):869–878.

Gilbert, G. S. and Webb, C. O. (2007). Phylogenetic signal in plant pathogen–host range. *Proceedings of the National Academy of Sciences*, 104(12):4979–4983.

Goldberg, T. L., Gillespie, T. R., Rwego, I. B., Estoff, E. L., and Chapman, C. A. (2008). Forest fragmentation as cause of bacterial transmission among nonhuman primates, humans, and livestock, uganda. *Emerging infectious diseases*, 14(9):1375.

Graham, J. P., Leibler, J. H., Price, L. B., Otte, J. M., Pfeiffer, D. U., Tiensin, T., and Silbergeld, E. K. (2008). The animal-human interface and infectious disease in industrial food animal production: Rethinking biosecurity and biocontainment. *Public Health Reports*, 123(3):282–299.

Greer, A. L., Briggs, C. J., and Collins, J. P. (2008). Testing a key assumption of host-pathogen theory: Density and disease transmission. *Oikos*, 117(11):1667–1673.

Gubbins, S., Gilligan, C. A., and Kleczkowski, A. (2000). Population dynamics of plant-parasite interactions: Thresholds for invasion. *Theoretical Population Biology*, 57(3):219–233.

Halliday, F. W. and Rohr, J. R. (2019). Measuring the shape of the biodiversity-disease relationship across systems reveals new findings and key gaps. *Nature Communications*, 10(1):1–10.

Han, B. A., O'Regan, S. M., Paul Schmidt, J., and Drake, J. M. (2020). Integrating data mining and transmission theory in the ecology of infectious diseases. *Ecology Letters*, 23(8):1178–1188.

Han, B. A., Park, A. W., Jolles, A. E., and Altizer, S. (2015). Infectious disease transmission and behavioural allometry in wild mammals. *Journal of Animal Ecology*, 84(3):637–646.

Haydon, D. T., Cleaveland, S., Taylor, L. H., and Laurenson, M. K. (2002). Identifying reservoirs of infection: A conceptual and practical challenge. *Emerging Infectious Diseases*, 8(12):1468–1473.

Hoch, T., Fourichon, C., Viet, A. F., and Seegers, H. (2008). Influence of the transmission function on a simulated pathogen spread within a population. *Epidemiology and Infection*, 136(10):1374–1382.

Hochberg, M. E. (1991). Non-linear transmission rates and the dynamics of infectious disease. *Journal of theoretical biology*, 153(3):301–321.

Holt, R. D. and Bonsall, M. B. (2017). Apparent Competition. *Annual Review of Ecology, Evolution, and Systematics*, 48:447–471.

Holt, R. D. and Pickering, J. (1985). Infectious Disease and Species Coexistence : A Model of Lotka-Volterra Form Author (s): Robert D . Holt and John Pickering Source : The American Naturalist , Vol . 126 , No . 2 (Aug ., 1985), pp . 196-211 Published by : The University of Chicago Press. *The American naturalist*, 126(2):196–211.

Huang, Z. Y., de Boer, W. F., Van Langevelde, F., Xu, C., Ben Jebara, K., Berlingieri, F., and Prins, H. H. (2013). Dilution effect in bovine tuberculosis: Risk factors for regional disease occurrence in Africa. *Proceedings of the Royal Society B: Biological Sciences*, 280(1765):1–7.

Huang, Z. Y., Xu, C., Van Langevelde, F., Prins, H. H., Ben Jebara, K., and De Boer, W. F. (2014). Dilution effect and identity effect by wildlife in the persistence and recurrence of bovine tuberculosis. *Parasitology*, 141(7):981–987.

IPBES (2020). Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services. page 108.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 266:859–867.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., and Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468(7324):647–652.

Keesing, F., Holt, R. D., and Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters*, 9(4):485–498.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.

Knell, R. J., Begon, M., and Thompson, D. J. (1996). Transmission dynamics of bacillus thuringiensis infecting plodia interpunctella: a test of the mass action assumption with an insect pathogen. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1366):75–81.

Kuiken, T., Holmes, E. C., McCauley, J., Rimmelzwaan, G. F., Williams, C. S., and Grenfell, B. T. (2006). Host species barriers to influenza virus infections. *Science*, 312(5772):394–397.

Lafferty, K. D. and Holt, R. D. (2003). How should environmental stress affect the population dynamics of disease? *Ecology Letters*, 6(7):654–664.

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., and Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359.

Loreau, M. and Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412(6842):72–76.

Lu, D. B., Wang, T. P., Rudge, J. W., Donnelly, C. A., Fang, G. R., and Webster, J. P. (2010). Contrasting reservoirs for Schistosoma japonicum between marshland and hilly regions in Anhui, China a two-year longitudinal parasitological survey. *Parasitology*, 137(1):99–110.

MacPherson, A., Louca, S., McLaughlin, A., Joy, J. B., and Pennell, M. W. (2020). A general birth-death-sampling model for epidemiology and macroevolution. *bioRxiv*.

May, R. M. and Anderson, R. M. (1984). Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences*, 72(1):83–111.

McCallum, H. (2016). Models for managing wildlife disease. *Parasitology*, 143(7):805–820.

McCallum, H., Barlow, N., and Hone, J. (2001). How should pathogen transmission be modelled? *Trends in Ecology and Evolution*, 16(6):295–300.

McCallum, H., Fenton, A., Hudson, P. J., Lee, B., Levick, B., Norman, R., Perkins, S. E., Viney, M., Wilson, A. J., and Lello, J. (2017). Breaking beta: Deconstructing the parasite transmission function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719).

McKay, B., Ebell, M., Dale, A. P., Shen, Y., and Handel, A. (2020). Virulence-mediated infectiousness and activity trade-offs and their impact on transmission potential of influenza patients: Infectiousness and Activity Trade-Offs. *Proceedings of the Royal Society B: Biological Sciences*, 287(1927).

Mollentze, N., Streicker, D. G., Murcia, P. R., Hampson, K., and Biek, R. (2020). Virulence mismatches in index hosts shape the outcomes of cross-species transmission. *Proceedings of the National Academy of Sciences of the United States of America*, 117(46):28859–28866.

Novozhilov, A. S. (2008). Heterogeneous Susceptibles-Infectives model: Mechanistic derivation of the power law transmission function. pages 1–14.

Ogden, N. H. and Tsao, J. I. (2009). Biodiversity and Lyme disease: Dilution or amplification? *Epidemics*, 1(3):196–206.

Olival, K. J., Hosseini, P. R., Zambrana-Torrelio, C., Ross, N., Bogich, T. L., and Daszak, P. (2017). Host and viral traits predict zoonotic spillover from mammals. *Nature*, 546(7660):646–650.

Ostfeld, R. S. and Keesing, F. (2000). Biodiversity series: the function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology*, 78(12):2061–2078.

Ostfeld, R. S. and Keesing, F. (2012). Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics*, 43:157–182.

Ostfeld, R. S., Keesing, F., and Eviner, V. T. (2008). *Infectious disease ecology: effects of ecosystems on disease and of disease on ecosystems*. Princeton University Press.

Park, A. W., Farrell, M. J., Schmidt, J. P., Huang, S., Dallas, T. A., Pappalardo, P., Drake, J. M., Stephens, P. R., Poulin, R., Nunn, C. L., and Davies, T. J. (2018). Characterizing the phylogenetic specialism-generalism spectrum of mammal parasites. *Proceedings of the Royal Society B: Biological Sciences*, 285(1874).

Parker, I. M., Saunders, M., Bontrager, M., Weitz, A. P., Hendricks, R., Magarey, R., Suiter, K., and Gilbert, G. S. (2015). Phylogenetic structure and host abundance drive disease pressure in communities. *Nature*, 520(7548):542–544.

Pope, L. C., Butlin, R. K., Wilson, G. J., Woodroffe, R., Erven, K., Conyers, C. M., Franklin, T., Delahay, R. J., Cheeseman, C. L., and Burke, T. (2007). Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis. *Molecular Ecology*, 16(23):4919–4929.

Poullain, V. and Nuismer, S. L. (2012). Infection genetics and the likelihood of host shifts in coevolving host-parasite interactions. *American Naturalist*, 180(5):618–628.

Power, A. G. and Mitchell, C. E. (2004). Pathogen spillover in disease epidemics. *the american naturalist*, 164(S5):S79–S89.

Rhodes, C. J., Atkinson, R. P. D., Anderson, R. M., and Macdonald, D. W. (1998). Rabies in Zimbabwe : reservoir dogs and the implications for disease control. (November 1996).

Roberts, M. G. and Heesterbeek, J. A. (2018). Quantifying the dilution effect for models in ecological epidemiology. *Journal of the Royal Society Interface*, 15(140).

Rohani, P., Green, C. J., Mantilla-Beniers, N. B., and Grenfell, B. T. (2003). Ecological interference between fatal diseases. *Nature*, 422(6934):885–888.

Rolland, J., Cadotte, M. W., Davies, J., Devictor, V., Lavergne, S., Mouquet, N., Pavoine, S., Rodrigues, A., Thuiller, W., Turcati, L., Winter, M., Zupan, L., Jabot, F., and Morlon, H. (2012). Using phylogenies in conservation: New perspectives. *Biology Letters*, 8(5):692–694.

Rudolf, V. H. and Antonovics, J. (2005). Species coexistence and pathogens with frequency-dependent transmission. *American Naturalist*, 166(1):112–118.

Rushmore, J., Caillaud, D., Hall, R. J., Stumpf, R. M., Meyers, L. A., and Altizer, S. (2014). Network-based vaccination improves prospects for disease control in wild chimpanzees. *Journal of the Royal Society Interface*, 11(97):20140349.

Sintayehu, D. W., Heitkönig, I. M., Prins, H. H., Tessema, Z. K., and De Boer, W. F. (2017). Effect of host diversity and species assemblage composition on bovine tuberculosis (bTB) risk in Ethiopian cattle. *Parasitology*, 144(6):783–792.

Sjöberg, M., Albrectsen, B., and Hjältén, J. (2000). Truncated power laws: A tool for understanding aggregation patterns in animals? *Ecology Letters*, 3(2):90–94.

Smith, M. J., Telfer, S., Kallio, E. R., Burthe, S., Cook, A. R., Lambin, X., and Begon, M. (2009). Host-pathogen time series data in wildlife support a transmission function between density and frequency dependence. *Proceedings of the National Academy of Sciences of the United States of America*, 106(19):7905–7909.

Ssebuliba, E. and Davies, T. J. (2021). Assessing the phylogenetic host breadth of millet pathogens and its implication for disease spillover. *Ecological Solutions and Evidence*, 2(1):1–11.

Streicker, D. G., Fallas González, S. L., Luconi, G., Barrientos, R. G., and Leon, B. (2019). Phylodynamics reveals extinction–recolonization dynamics underpin apparently endemic vampire bat rabies in Costa Rica. *Proceedings of the Royal Society B: Biological Sciences*, 286(1912).

Streicker, D. G., Fenton, A., and Pedersen, A. B. (2013). Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecology Letters*, 16(8):975–984.

Streicker, D. G., Turmelle, A. S., Vonhof, M. J., Kuzmin, I. V., McCracken, G. F., and Rupprecht, C. E. (2010). Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science*, 329(5992):676–679.

Tien, J. H. and Earn, D. J. (2010). Multiple transmission pathways and disease dynamics in a waterborne pathogen model. *Bulletin of Mathematical Biology*, 72(6):1506–1533.

Viana, M., Mancy, R., Biek, R., Cleaveland, S., Cross, P. C., Lloyd-Smith, J. O., and Haydon, D. T. (2014). Assembling evidence for identifying reservoirs of infection. *Trends in Ecology and Evolution*, 29(5):270–279.

Vicente, J., Delahay, R. J., Walker, N. J., and Cheeseman, C. L. (2007). Social organization and movement influence the incidence of bovine tuberculosis in an undisturbed high-density badger Meles meles population. *Journal of Animal Ecology*, 76(2):348–360.

White, L. A., Forester, J. D., and Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews*, 92(1):389–409.

Wilber, M. Q., Langwig, K. E., Kilpatrick, A. M., McCallum, H. I., and Briggs, C. J. (2016). Integral Projection Models for host–parasite systems with an application to amphibian chytrid fungus. *Methods in Ecology and Evolution*, 7(10):1182–1194.

Wolfe, N. D., Daszak, P., Kilpatrick, A. M., and Burke, D. S. (2005). Bushmeat hunting, deforestation, and prediction of zoonotic disease. *Emerging infectious diseases*, 11(12):1822.

Wolfe, N. D., Dunavan, C. P., and Diamond, J. (2007). Origins of major human infectious diseases. *Nature*, 447(7142):279–283.

Woolhouse, M. E., Taylor, L. H., and Haydon, D. T. (2001). Population biology of multihost pathogens. *Science*, 292(5519):1109–1112.

5 Appendix

5.1 Contact network model distributions

Many animals are aggregated in flocks, herds or other social networks. For directly transmitted diseases in populations with strong social and spatial structuring, we need to revisit the assumption of homogeneous and random host mixing, assumed within density-dependent transmission, as animals interact with a small network, not the entire population (Ferrari et al., 2011; Craft, 2015; White et al., 2017). Contact models are based on graph-theory, named *epidemic network models*, where the realized per capita transmission rate $\hat{\beta}$ is scaled to the mean neighbourhood size (Ferrari et al., 2011; White et al., 2017). It has been shown that the degree of connections within a population has a large effect on the disease dynamics. For example, the existence of triangular contacts (3 connected individuals) reduces both the initial spread and final disease outbreak (Keeling, 1999). The transmission across a connection is weighted against its connections, using parameter $< k >$, the mean degree number of contacts, or dispersion parameter, which can be scaled in 3 ways (Lloyd-Smith et al., 2005; Ferrari et al., 2011):

1. Independent of total population size, N
   This is similar to the assumptions of the classic frequency-dependent transmission function
2. Increasingly slower than linear $C \cdot \sqrt{N}$, C being a constant
   An intermediate between case 1 and 3.
3. Increasing linearly with population size, N
   This is analogous to the density-dependent transmission function, assuming a constant area.

In epidemic network models, the R_0 calculation is different from the original compartmental models, as it is assumed that some individuals contribute more to the spread of the disease as others, such as *superspreaders* (Craft, 2015; White et al., 2017; Lloyd-Smith et al., 2005). This is similar to some of the previously introduced multi-host models, as in both cases the population consists of individuals contributing differently to the disease spread. However, the nature of transmission described above is homogeneous, therefore not taking into account the heterogeneity of individuals, group formation and the extend of mixing within these groups. It has been shown that in heterogeneous populations the limited spread of invading pathogens creates bigger intraspecific competition than in homogeneously mixed population, resulting in a reduced success invasion (Keeling, 1999). This has a big effect on the R_0 and so should be taken into account. Fortunately, epidemic network models, that are often used to track contact in human epidemics (Childs et al., 2007), have the advantage of being able to take into account the heterogeneous nature of contacts between hosts and the topology of the contact network, by using the mean and variance of contacts. This can be done by choosing from a number of degree distributions. A couple of common examples are listed below (Ferrari et al., 2011).

5.1.1 Poisson distribution

Here, the variance to mean ratio is close to one. Due to this relatively low variance it is closest to homogeneous mixing. A Poisson distribution is based on the assumption that events happen in a certain time interval at a constant rate, irrespective of the time that has passed. An individual contacts another individual at a constant contact rate, with a constant probability of being infected, independent of the number of previously infected individuals. Increasing the number of nodes has either a slightly decreasing or more decreasing effect on $\hat{\beta}$, depending on whether $< k >$ increases linearly with population size (case 3), increases with $\sqrt{N}$ (case 2) or if it is constant with population size (case 1), respectively (Ferrari et al., 2011).

5.1.2 Power-law truncated distribution

This distribution consists of two power law functions, joint at a cutoff-group size value, at which the behavior of the group changes. This cut-off value can be an optimal group size (in mammals) or a certain minimal group size to

bring group advantage (Sjöberg et al., 2000). Here, variance to mean ratio >> 1, such that in a group of individuals where some have many contacts, whereas others almost none, the phenomenon of super-spreaders can be captured (Ferrari et al., 2011). This distribution assumes extreme heterogeneity in local contacts. $\hat{\beta}$ slightly increased (case 3), slightly decreased (case 2) or decreased more (case 1) with the number of nodes.

5.1.3 An exponential distribution

This is the intermediate case between the Poisson and truncated power-law distributions. It is another case of a memoryless distribution, like the Poisson process, as it models the time between 2 events in a Poisson process. This distribution has, just like the power-law distribution, a greater proportion of superspreaders.
$\hat{\beta}$ more or less remained the same (case 3), slightly decreased (case 2) or linearly decreased (case 1) with an increasing number of nodes.

5.2 Empirical estimations of disease transmission

As described in box 2, there are a number of ways to empirically determine the probability of successful transmission, c. A couple of interesting ones that can be used in the definition of transmission according to this review are listed below.

5.2.1 Secondary Attack Rate to estimate c

The Secondary Attack Rate, or SAR, is a common approximation of c of transmission within a population (Childs et al., 2007). It uses contact tracing, where the secondary cases are defined as the cases that arise within the time-span of incubation and infection time of a primary case.

$$\text{SAR} = \frac{\text{total secondary cases}}{\text{total susceptibles}} \tag{31}$$

This probability ratio can be used as probability of successful transmission, c.

5.2.2 Binomial model of transmission for c

Assume the probability of transmission at contact is p and the probability of escaping infection is $q = 1 - p$. Suppose the susceptible host makes n contacts with an infected host. The probability of escaping infection will be $q^n = (1-p)^n$. Therefore, we can define the probability of getting infection after contact as:

$$1 - q^n = 1 - (1 - p)^n \tag{32}$$

This is the binomial distribution. Here, p can be empirically estimated with the maximum likelihood:

$$\hat{p} = \frac{\text{number of individuals who become infected}}{\text{total number of contacts with infected individuals}} \tag{33}$$

The difference between this method and SAR, is that here the number of contacts is determined relative to the contacts with infectious hosts, whereas SAR weighs the probability of infection against the number of susceptible and exposed hosts. This function is often used for sexually transmitted diseases, so for pathogens showing FD behaviour (Childs et al., 2007).