The effects of heterogeneity in individual infectiousness on disease modeling predictions

A DISSERTATION

SUBMITTED TO THE FACULTY OF

UNIVERSITY OF MINNESOTA

BY

Lauren A. White

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Meggan Craft, James Forester

June 2018

Acknowledgements

I'd like to start by thanking my advisors and co-authors, Meggan Craft and James Forester, as well as my committee members Elizabeth Borer and Susan Jones. My third chapter would not have been possible without Montse Torremorell's expertise and empirical data on swine influenza and swine industry practices. I must also thank past and current members of the Craft Lab who have provided feedback on drafts, presentations, and ideas (including, but not limited to: Kim VanderWaal, Amy Kinsley, Jenny Blackburn, Nick Fountain-Jones, Katie Worsley-Tonks, and Marie Gilbertson).

This work would not have been possible without funding provided through the NSF Graduate Research Fellowship, NSF Doctoral Dissertation Improvement Grant, and the University of Minnesota Informatics Institute MnDRIVE Fellowship. The University of Minnesota's Department of Ecology, Evolution, and Behavior and Council of Graduate Students both provided Travel and Research Grants that made it possible for me to share this research at national and international conferences.

Finally, I must thank my husband, Rob, for his patience during this five-year journey, as well as my parents for their love and support. Thank you to my friends Catherine and Andrea for listening, and to Meredith and Noah for letting me stay with you numerous times during my visits to Minnesota. This process would have been exponentially more challenging without my Flow Yoga family, who have always helped me to keep everything into perspective.

Dedication

For Rob.

Abstract

Individual variation in infectiousness is generated by heterogeneities in the host, the pathogen, and the environment. However, many models of disease transmission, especially those designed for wildlife and livestock populations, do not typically allow for such variation in individual infectiousness. The objective of my research is to explore the effects of heterogeneity in individual infectiousness on disease modeling predictions within and across populations. My dissertation research explores three different types of heterogeneity that can alter individual infectiousness: (i) host heterogeneity resulting from individual differences in susceptibility, infectiousness, and behavioral contact rates, (ii) contact heterogeneity that arises within a population from underlying social systems and interactions; and (iii) spatial heterogeneity that arises from variation in host density as a function of resource quality and variable individual movement rates across a landscape. An improved understanding of the factors that lead to variability in individual infectiousness and the conditions that necessitate the inclusion of such variability in future disease models will be critical to address the growing global threats of zoonoses and emerging infectious diseases.

Table of Contents

List of Tables

List of Figures

Introduction

Emerging infectious diseases pose substantial risks to humans, livestock and wildlife, in the forms of human disability and death, loss of livestock productivity, and species extinctions (Daszak, Cunningham & Hyatt 2000; Jones *et al.* 2008). Since the majority of emerging infectious diseases that affect humans are zoonotic in origin (Jones *et al.* 2008), it is imperative to understand the conditions under which such pathogens spread and persist in wildlife and domestic animal populations. Although theoretical models can provide important insights into these conditions, many models of animal disease transmission do not allow for variation in individual differences in behavior and physiology. For example, an underlying assumption of traditional disease transmission models is that each individual in a population has an equal probability of contacting and infecting any other individual (Anderson & May 1979, 1991; Keeling & Eames 2005).

In natural systems, however, these assumptions do not necessarily hold true—in fact, empirical studies suggest that unequal contact rates are the rule rather than the exception (Böhm, Hutchings & White 2009; Clay *et al.* 2009; Craft *et al.* 2011). Moreover, contact rates can vary with infection-induced behavioral changes, and these changes are likely non-uniform across individuals (Lopes, Block & König 2016). Finally, there is evidence that the first individual infected in a population (the index case), and the relative composition of behavioral phenotypes (e.g. bold vs. shy continuum) can substantially affect how effectively a pathogen spreads within a population (Adelman *et al.* 2015; Keiser *et al.* 2016).

However, there is no consistent framework that outlines when individual heterogeneity in pathogen transmission is important and when it is necessary to account for those differences in sampling or interventions, even though allowing for such differences can markedly change predictions of an epidemic's duration and behavior (Keeling & Eames 2005; Meyers 2007). The assumptions of traditional models are further violated when the processes governing heterogeneity in transmission occur at multiple scales, including within hosts, within groups of hosts, between groups, and across landscapes (Restif *et al.* 2012). There is also no guarantee that the transmission process, data collection, and potential interventions are necessarily taking place at the same scale (Riley *et al.* 2015). Thus, when designing experiments to analyze factors contributing to transmission, disease ecologists face several challenges: they must determine which scale is appropriate for the system, determine if it is necessary to account for heterogeneity, and choose a sampling strategy that adequately addresses both.

Most disease models seek to predict the rate at which a pathogen will spread in a host population. This quantity is the product of the number of infected individuals (I), the number of susceptible individuals (S), and the rate (β) at which infectious individuals successfully transmit the pathogen. This transmission rate, β, can further be broken down into behavioral (β_c) and physiological (β_p) components: (i) the intraspecific behavioral rate of contact (β_c) between infected and uninfected individuals depends upon both local processes like group size, mate choice, or sickness behaviors and broad-scale processes like resource availability, host-density, and migration (Lloyd-Smith *et al.* 2005a); and (ii) the transmission efficiency (β_p) is determined by a host's innate physiological

characteristics such as immunocompetence, shedding rate, latency period, and co-infection (Lehmer *et al.* 2010; Telfer *et al.* 2010; Hawley & Altizer 2011; Lass *et al.* 2013). Notably, physiology can covary with behavior, so disease models need to incorporate possible interactions between these two components, not just heterogeneity in contact rate (Hawley *et al.* 2011).

Network models are tools that can capture individual variability in the number and rate of contacts (β_c). In these networks, *nodes* represent individuals or groups, while *edges* describe a connection or contact between nodes. Pathogens spread through the network from node to node via connecting edges. Thus, a *contact* is any interaction that could allow for transmission of an infectious agent between a pair of individuals, groups of individuals, or geographic regions. Despite their advantages, however, network models are not used extensively for domestic animal and wildlife populations, in part because of the added expense and effort required to adequately parameterize such models (Krause *et al.* 2013).

The overall objective of my dissertation research is to explore the effects of heterogeneity in behavioral and immune competence on disease modeling predictions within and across populations. Here, I have used a modeling approach to explore three different types of heterogeneity that can alter individual infectiousness: (i) host heterogeneity (host β_p and β_c); (ii) contact heterogeneity (local-scale β_c); and (iii) spatial heterogeneity (broad-scale β_c). I have modelled these types of heterogeneity in domestic and wild animal populations for specific host-pathogen systems (e.g., influenza A virus in

swine herds) and more generally for theoretical systems. These models differ in the behavioral processes that they encompass and the spatial scale at which they operate.

In Chapter 1, I conducted a review of how contact networks have been used to study macro- and microparasite transmission in wildlife. This review: (1) explains why contact heterogeneity is relevant for wildlife populations; (2) explores theoretical and applied questions that contact networks have been used to answer; (3) gives an overview of unresolved methodological issues; and (4) suggests improvements and future directions for contact network studies in wildlife.

In Chapter 2, I explored how the behavioral and physiological components of transmission (host β_c and β_p, respectively) interact to determine epidemic outcomes. I also considered how sickness-induced behavioral changes might affect epidemic outcomes. I used a dynamic network, individual-based modeling approach to allow for covariation between the behavioral and physiological components of transmission (White, Forester & Craft 2018a). My results demonstrate that: (a) individual variability in susceptibility or infectiousness, which is typically unaccounted for in disease models, can have profound effects on population-level disease dynamics; (b) when contact rate and susceptibility or infectiousness negatively covary, it takes substantially longer for epidemics to spread throughout the population; and (c) reductions in contact rate resulting from infection-induced behavioral changes can prevent the pathogen from reaching most of the population.

In Chapter 3, I investigated how heterogeneity in contact rate (local-scale β_c) and infection risk, as governed by variable movement through an industrial farm system and

production stage, could affect epidemic dynamics. Recent modelling and empirical work on influenza A virus (IAV) suggests that piglets play an important role as an endemic reservoir. The objective of this study was to test intervention strategies aimed at reducing the incidence of IAV in piglets and ideally, preventing piglets from becoming exposed in the first place. These interventions include biosecurity measures, vaccination, and management options that swine producers may employ individually or jointly to control IAV in their herds. I developed a stochastic Susceptible-Exposed-Infectious-Recovered-Vaccinated (SEIRV) model that reflects the spatial organization of a standard breeding herd and accounts for the different classes of pigs therein including gilts, sows, and piglets in various production and immune stages (White, Torremorell & Craft 2017d). The findings show that piglets are a high risk sub-group and that combined biosecurity and vaccination efforts can reduce, but are unlikely to eliminate, IAV after it has been introduced into the breeding herd.

In Chapter 4, I asked: how does individual movement behavior, governed by perceptual distance and individual selection for resource availability and conspecific density, interact with spatial heterogeneity (broad-scale β_c) via resource availability and clustering to affect epidemic dynamics? Spatial heterogeneity can sometimes lead to non-linear or counterintuitive outcomes depending on the host and pathogen system (Tracey *et al.* 2014; White, Forester & Craft 2017a). I employed a stochastic, individual-based, Susceptible Infectious Recovered (SIR) model where a resource selection function (RSF) governed individual movement choices. Our results support studies showing that, counterintuitively, increased fragmentation can promote pathogen persistence (Tracey *et*

al. 2014), but this finding was largely dependent upon movement rules controlled by perceptual range and strength of selection for conspecifics. For simulations with high conspecific density, slower recovery rates, and strong selection for conspecifics, more complex epidemic dynamics emerge, and the most fragmented landscapes are not necessarily the most conducive to outbreaks or disease persistence. Thus, our findings point to the importance of interactions between landscape structure, individual movement behavior, and pathogen transmission for determining disease dynamics.

Chapter 1. Using contact networks to explore mechanisms of parasite transmission in wildlife

White, L.A., Forester, J.D. and Craft, M.E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. Biol. Rev., 92, 389-409. doi:10.1111/brv.12236

1.1 Abstract

A hallmark assumption of traditional approaches to disease modelling is that individuals within a given population mix uniformly and at random. However, this assumption does not always hold true; contact heterogeneity or preferential associations can have a substantial impact on the duration, size, and dynamics of epidemics. Contact heterogeneity has been readily adopted in epidemiological studies of humans, but has been less studied in wildlife. While contact network studies are becoming more common for wildlife, their methodologies, fundamental assumptions, host species, and parasites vary widely. The goal of this article is to review how contact networks have been used to study macro- and microparasite transmission in wildlife. The review will: (1) explain why contact heterogeneity is relevant for wildlife populations; (2) explore theoretical and applied questions that contact networks have been used to answer; (3) give an overview of unresolved methodological issues; and (4) suggest improvements and future directions for contact network studies in wildlife.

1.2 Introduction

Why model parasite transmission in wildlife?

Wildlife diseases pose substantial challenges to species conservation, maintenance of biodiversity, ecosystem stability, livestock welfare, and public health (Daszak, Cunningham & Hyatt 2000; Lloyd-Smith *et al.* 2009; Restif *et al.* 2012), but the impacts of wildlife disease in these different areas give rise to competing priorities and ethical dilemmas when monitoring, preventing outbreaks, and deciding on interventions (McCallum & Hocking 2005). For instance, culling of wildlife is often the default management strategy for a wildlife-derived pathogen that spills over to livestock, but this strategy must be re-examined if the wildlife species in question is of conservation concern. Culling can also disrupt contact patterns in ways that are counterproductive to reducing disease prevalence (McDonald *et al.* 2008). Moreover, treating species of conservation concern can be controversial in of itself – especially when handling affects survival rates (McCallum & Hocking 2005). Limited resources, funding, and logistical challenges are also likely to constrain the number and type of interventions that can be implemented. Modelling provides an ethical and economical way to test hypotheses about which factors are most influential in the spread of the parasites and which interventions might prove most effective (Lloyd-Smith *et al.* 2009). Combined with the fact that collecting disease data can be especially difficult in wildlife systems, scientists, policy-makers, and managers should prioritize a model-informed management and data-collection approach (Restif *et al.* 2012).

Assumptions of traditional approaches to disease modelling

Disease models can further two, sometimes incompatible, objectives: (1) to deepen our understanding of the mechanisms of disease dynamics, and (2) to offer accurate or precise predictions of future epidemics or the impact of interventions (Keeling & Rohani 2008a). Levins (1966) framed this conflict more broadly, arguing that modelling the natural world will always involve irrevocable trade-offs between precision, generality, and realism – that is, it is possible to achieve two of the three qualities, but always at the expense of the third. Anderson & May (1979) popularized compartmental models in epidemiology, which arguably sacrifice realism, especially when making assumptions about how individuals come into contact with one another. In the tradition of particle physics, these compartmental or mass-action models assume that individuals mix like molecules in an ideal gas – with random mixing and no difference in contact frequency or duration between individuals (McCallum, Barlow & Hone 2001). Thus, compartmental models are general and can give precise results, but they may not realistically incorporate the fundamental contact patterns of a population if non-random mixing occurs (Meyers 2007). Here, the goal is not to undermine the utility of compartmental models in providing new ideas and inferences in epidemiology, but rather to think critically about situations where a more accurate portrayal of contact duration and frequency can improve our predictions and understanding of disease models – essentially to be able to discern when averaging across a population is no longer 'good enough'.

The traditional compartmental model is the SIR (susceptible–infectious–removed) model (Kermack & McKendrick 1927; Anderson & May 1991). Here individuals exist in any one compartment as defined by their disease status. 'Susceptible' individuals (S) become 'infected' (I) based on the transmission parameter (β) and 'removed' (R, through death or recovery with immunity) at rate γ. Transmission is most commonly modelled as either density or frequency-dependent. When contact rate increases with the density of individuals in a population, density-dependent transmission applies (Equations 1.1-1.3). Generally, animal and plant systems are modelled as density-dependent (Keeling & Rohani 2008a).

$$\frac{dS}{dt} = -\beta S \qquad \text{Equation 1.1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \qquad \text{Equation 1.2}$$

$$\frac{dR}{dt} = \gamma I \qquad \text{Equation 1.3}$$

When the number of contacts scales independently of population size (N), a system exhibits frequency-dependent transmission (Equations 1.4-1.6).

$$\frac{dS}{dt} = -\frac{\beta' SI}{N} \qquad \text{Equation 1.4}$$

$$\frac{dI}{dt} = \frac{\beta' SI}{N} - \gamma I \qquad \text{Equation 1.5}$$

$$\frac{dR}{dt} = \gamma I \qquad \text{Equation 1.6}$$

The transmission parameter (β or β') is defined as the product of the contact rate and the conditional probability of transmission given contact or transmission efficiency, but depending on the form of the transmission function will have different dimensions (Begon *et al.* 2002). For frequency-dependent transmission, the contact rate component

of the transmission parameter (β') remains constant (Begon *et al.* 2002). However, both these forms of global transmission are functionally equivalent for a population of constant size and occupying a constant area (Turner, Begon, & Bowers, 2003; Ferrari *et al.*, 2011).

It is important to note that these compartmental models describe transmission globally (i.e. within a population) for what is fundamentally a local process (i.e. between individuals) (Turner *et al.*, 2003; Ferrari *et al.*, 2011). Thus it is perhaps not surprising that observed epidemics often exhibit a range of transmission functions rather than being strictly density or frequency-dependent (Ferrari *et al.*, 2011). To account for this, other forms of transmission functions have been proposed, but density- and frequency-dependent transmission functions still predominate in the literature (Ferrari *et al.*, 2011). In some instances, density-dependent transmission has been assumed to be the same as random, homogeneous mixing, while frequency-dependent transmission has been equated correspondingly with some form of heterogeneity (Begon *et al.*, 2002; Ferrari *et al.*, 2011). In fact, Begon *et al.* (2002), argue that contact structure operates on an axis independent from that of the transmission function, such that a frequency-dependent system could have heterogeneity in local contact structure, but would appear homogeneous at a global level. Recent work suggests that observed 'intermediate' transmission may result from a transition from density-dependent transmission at low densities to frequency-dependent transmission at high densities (Davis *et al.* 2015). In much of the discussion that follows in Section 1.3, we will invoke the assumptions and limitations of density-dependent transmission, as this form is commonly used when

modelling wildlife diseases. Nevertheless, the nomenclature and choice of the transmission function are still controversial and for readers looking for greater detail on the issue we recommend McCallum, Barlow, & Hone (2001) and Begon *et al.*, (2002).

What are contact networks?

Regardless of the form of the transmission function, deriving accurate predictions of the transmission parameter, β, is challenging, especially in free-living wildlife systems (McCallum *et al.*, 2001; Caley & Ramsey, 2001). In reality, non-random association patterns that affect the contact rate component of β are common in humans, livestock, and wildlife (Mossong *et al.* 2008; Martínez-López, Perez & Sánchez-Vizcaíno 2009; Craft & Caillaud 2011a). Contact network models expand the relevance of compartmental models by incorporating these heterogeneous interactions. Contact networks represent possible transmission pathways through the population of interest. In these networks, *nodes* represent individuals or groups, while *edges* represent a connection or contact between nodes. For readers less familiar with network terminology, key network terms are italicized upon their first use and defined in Appendix A. Less commonly in models of spatial heterogeneity, nodes may represent larger geographic areas such as counties or states (Maher *et al.* 2012; Buhnerkempe *et al.* 2014; Grange *et al.* 2014). Thus, a 'contact' is any interaction that could potentially allow for transmission of an infectious agent between a pair of individuals, groups of individuals, or geographic regions. What constitutes a contact will depend on host life history, the parasite or pathogen's life cycle and its mode of transmission. For instance, transmission of Ebola virus in socially living primates may occur through aerosolized particles (Nunn *et al.* 2008; Ryan, Jones &

Dobson 2013; Rushmore *et al.* 2013), while transmission of Devil Facial Tumour Disease (DFTD) in the more solitary Tasmanian devil (*Sarcophilus harrisii*) depends on aggressive interactions between individuals (Hamede *et al.* 2009). Thus biting another Tasmanian devil may serve as an effective contact for DFTD, while simply being within a certain distance of an infected primate may serve as an effective contact for Ebola.

For contact network models describing transmission, pathogens can spread through the network from node to node *via* connecting edges. Any given node has a certain number of contacts, which is termed as a node's *degree*. The contact rates assumed by the traditional compartmental model could be considered a special case of network model since a compartmental model is equivalent to a network model where an edge exists between every single node (i.e. a fully connected network, Figure 1.1A) (Craft & Caillaud 2011a). There are a variety of metrics (e.g. *centrality*, degree, etc.) that describe an individual's position or influence in the network or that describe the properties of the network as a whole. Such metrics are important because at an individual level, they can predict the risk of infection or exposure, and at a population level they can explain observed variation in epidemic dynamics (Christley *et al.* 2005; Ames *et al.* 2011; Godfrey 2013).

Figure 1.1A demonstrates how individuals in a population with heterogeneous contact structure will have different degrees and different centralities depending on their position in the network. All nodes in Figure 1.1A have equal centrality and equal degree, while nodes with the highest normalized *betweenness centrality* are shown in red in Figure 1.1B and Figure 1.1C. In sparser networks, some nodes may be unconnected such that there may be more than one *component* in the network (Figure 1.1C).

Aims

Social network analysis in wildlife was originally used to address questions in behaviour and behavioural ecology, but social network structure has since gained recognition for its importance in governing a variety of evolutionary and ecological

processes including social evolution, co-evolution and population stability (Proulx, Promislow & Phillips 2005; Kurvers *et al.* 2014). In the field of disease ecology, contact networks address questions at the intersection of epidemiology, ecology, and animal behaviour. While such contact network studies are becoming more common in wildlife systems, their methodologies, fundamental assumptions, host species, and parasites vary widely (Table 1.1). The objective of this review is to highlight and synthesize the ways that contact networks can further our understanding of parasite transmission in wildlife, while critically analysing the way this tool has been used. This article will: (1) explain why contact heterogeneity matters for wildlife populations; (2) explore theoretical and applied questions that contact networks have been used to answer; (3) give an overview of unresolved methodological issues; and, (4) suggest improvements and future directions for contact network studies in wildlife.

This review is meant to give a comprehensive, but not exhaustive, overview of relevant literature. In Sections 1.5 and 1.6, we identify and discuss seven critical theoretical and applied questions along with four unresolved methodological questions. Table 1.1 provides a compilation of the empirical network studies cited for those questions. For each study, Table 1.1 lists the focal host, focal pathogen or disease, method of data collection, and the questions addressed by each study. The numbered topics in the last eleven columns of the table correspond to the questions highlighted in Sections 1.5 and 1.6 in order of appearance: theoretical and applied questions (Columns Q1–Q7) and unresolved methodological questions (Columns M1–M4). The grey shading indicates studies of particular relevance to each section. In some columns, there are

studies that are highlighted as pertinent, but that are not discussed explicitly in the main text under that particular heading. Even so, this remains a conservative demarcation of studies that address each respective topic, since many studies address more of the questions than are highlighted in the text or indicated in the table. Empirical articles for this review were obtained through a cross referencing of *Web of Science* and *PubMed* during September and October 2014. Search terms included: "Wildlife AND disease AND 'contact network'" and "'social network' AND disease and animals NOT livestock." Papers were also obtained by tracing references in sources already obtained through prior searches.

In this review, parasite is defined broadly in the ecological sense to encompass macro- and microparasites (Anderson & May 1979). Macroparasites are typically larger organisms (e.g. helminths, flukes, arthropods) that have free-living infectious stages outside the host, while microparasites are generally smaller (e.g. bacteria, viruses, protozoa, prions) and reproduce within the host, usually with correspondingly shorter generation times (Keeling & Rohani 2008a). In this article, microparasites will also be referred to interchangeably as pathogens. Parasites can be further classified as directly or indirectly transmitted. Directly transmitted infections result from close contact between a susceptible and infectious individual, while indirectly transmitted infections result from second-hand exposure to the infectious agent through the environment. In general, most microparasites are directly transmitted because they cannot survive for a long time outside a host, and most macroparasites are indirectly transmitted because of their free-living infectious stages (Keeling & Rohani 2008a).

1.3 Why does contact heterogeneity matter for wildlife?

Compartmental models have given rise to a ubiquitous parameter in epidemiology that underlies many disease-intervention strategies: R_0. This parameter, known as the basic reproductive number, represents the number of secondary cases arising from one infectious individual in an entirely susceptible population (Anderson & May 1991). If R_0 is greater than 1, then the pathogen has the potential to spread throughout the population, while if R_0 is less than 1, the number of infected cases should subside. Yet, contact heterogeneity or preferential associations can have a substantial impact on the duration, size, and dynamics of epidemics such that R_0 may not be a reliable predictor of disease dynamics (Keeling & Eames 2005; Meyers 2007). This concept has been readily acknowledged in epidemiological studies of humans (Ames *et al.* 2011), but has been less studied in wildlife (Craft & Caillaud 2011a). Arising from the calculation of R_0 in a density-dependent SIR model is the idea that there is a corresponding population threshold below which an epidemic cannot occur (McCallum *et al.* 2001; Lloyd-Smith *et al.* 2005a). Thus, disease control in wildlife often relies on methods like culling that are justified by assumptions that transmission rates increase with host density(Carter *et al.* 2007). However, there is often a non-linear relationship between density and parasite prevalence in wildlife, which may result from factors like sociality, territoriality, individual movement, variable reproductive rates, and multi-host reservoirs (Lloyd-Smith *et al.* 2005a; Viana *et al.* 2014).

Consider the pitfalls of assuming random-mixing and density-dependent transmission in the case of the European badger (*Meles meles*), which is a wildlife reservoir for bovine tuberculosis (bTB, *Mycobacterium bovis*). Badger culling has been implemented for many years as a means of controlling bTB in cattle, but more recent reviews and studies of this practice have found mixed evidence in support of culling (Vicente *et al.* 2007; White *et al.* 2008). In the UK, culling reduced bTB in the immediate area, but increased incidence in neighbouring areas (McDonald *et al.* 2008). In fact, large group size (i.e. high density) of badgers appears to be less of a risk factor than the perturbation of badger social systems through culling (Vicente *et al.* 2007).

While bTB is a complicated multi-host pathogen, the unintended side effects of culling arguably resulted in part from not accounting for the role of contact heterogeneity in disease transmission. Badgers live in territorial, social groups called setts. In a radio-tracking study of an undisturbed badger population, contacts between badgers of the same sett accounted for almost 90% of contacts relative to contacts between individuals of different setts (Böhm *et al.* 2008). Culling disrupts setts such that surviving badgers range into new territories, potentially increasing inter-sett contacts and the risk of bTB transmission to new locales (Carter *et al.* 2007; McDonald *et al.* 2008; Prentice *et al.* 2014). In network terminology, the undisturbed badger populations had a higher *modularity* than post-culling populations because of their stronger intra-sett associations. Communities with high modularity or greater community structure are more resistant to disease invasions (Salathé & Jones 2010). Many factors complicate bTB transmission in the UK, but eradication efforts in badgers highlight the complexities of wildlife ecology,

especially for a multi-host pathogen. It also highlights how the assumptions implicit to compartmental models may be inadequate for describing critical complexity in wildlife systems.

1.4 Overview of empirical studies

The 39 empirical studies highlighted here investigate a wide variety of hosts and parasites (Table 1.1). The use of contact network methods for describing parasite transmission in wildlife systems is still relatively novel, with the majority of studies published in 2008 or later. Roughly three-quarters of studies targeted mammalian hosts with only five studies targeting reptiles, two targeting fish, and two targeting invertebrates. Most studies targeted directly transmitted parasites; roughly a third of which dealt with bTB in different wildlife systems. Comparatively few papers addressed indirectly transmitted, vector-borne, or trophically transmitted parasites. The majority of studies relied on behavioural observations or capture–mark–recapture methods to establish contact network patterns, but new methods like proximity data loggers and genetic strain comparison of commensal microparasites are gaining traction. The definition of a 'contact' varies widely in terms of required proximity, duration of association, and permitted time-lag for shared or asynchronous space use (Table 1.1: Column M1). Replication in the same host–pathogen system is relatively infrequent. Moreover, when the studies are grouped by the types of questions that they answer, there is a marked scarcity of studies for any given question that makes it difficult to generalize

findings to other systems. A detailed discussion of those theoretical and applied questions and relevant findings follows in Sections 1.5 and 1.6.

1.5 Theoretical and applied questions that contact networks have been used to answer

Uncovering superspreaders: do wildlife populations exhibit contact heterogeneity?

The impacts of contact heterogeneity on pathogen transmission have been well demonstrated for humans and livestock (Martínez-López *et al.* 2009; Ames *et al.* 2011). However, uncovering contact heterogeneity and determining its effects on epidemic outcomes are still motivating objectives for wildlife studies (Craft & Caillaud 2011a). If a wildlife population exhibits contact heterogeneity, a corresponding question is: are there *superspreaders* in the population? Superspreaders are individuals that disproportionately contribute to the spread of pathogens by virtue of a high number of contacts – or a high degree in network terminology (Lloyd-Smith *et al.* 2005b). Identification of superspreaders is critical because they can be targeted for surveillance and control measures (Christley *et al.* 2005). Monitoring individuals or sub-populations that act as 'hubs' for the population can increase the efficiency of outbreak detection (Eubank *et al.* 2004), and targeted treatment of high-risk individuals in endangered wildlife species can enable lower coverage of vaccination than conventional methods (Haydon *et al.* 2006).

There have been multiple studies demonstrating that contact heterogeneity exists in wildlife, and that accounting for this heterogeneity while modelling the spread of pathogens is important for making predictions. For instance, deer mice (*Peromyscus maniculatus*), susceptible to Sin Nombre virus (SNV, *Bunyaviridae*: *Hantavirus*), exhibit

non-normal distributions of contacts, such that a small proportion of the individuals are responsible for a majority of contacts (Clay *et al.* 2009). This supports the concept of superspreaders or the '20/80 rule' where 20 per cent of the individuals account for 80 per cent of transmission (Clay *et al.* 2009). Similarly in brushtail possums (*Trichosurus vulpecula*), contact rate was not proportional to density suggesting frequency-dependent transmission (Ji, White & Clout 2005), and observed *degree distributions* were non-normal and best described by a negative binomial fit (Porphyre *et al.* 2008). Moreover, observed networks were *small world networks* that yielded mean estimates for R_0 roughly 1.5 times greater compared to disease simulations on randomly generated networks of comparable size (Porphyre *et al.* 2008).

However, superspreaders may not always be present, even if associations are not random. For Tasmanian devils at risk of DFTD, the degree distribution of contacts was neither random nor highly aggregated, and no particular class of individual was more highly connected overall (Hamede *et al.* 2009). Nevertheless, including these observed distributions in a contact network model resulted in a lower epidemic threshold, but a higher probability of pathogen extinction, than in an equivalent mean field model (Hamede *et al.* 2012).

What factors mediate individual variability in susceptibility and exposure?

Individual variability in contributing to the spread of pathogens is not merely limited to a heightened ability to transmit as a result of a disproportionally high number of contacts; it may also be necessary to account for super-shedders, super-susceptibles or

super-recipients, and super-movers (Craft 2015). For instance, co-infection with other parasites may make an individual more susceptible to infection or dramatically increase the rate of shedding (Cattadori, Boag & Hudson 2008; Lass *et al.* 2013). In a study that tracked the obligate blood-sucking lice (*Lemurpediculus verruculosus*) occurring on individual mouse lemurs (*Microcebus rufus*), both 'superspreader' and 'super-recipient' mouse lemurs existed concurrently in the population. Super-recipients, in this case, were individuals with the highest parasite loads but low lice turnover – that is, they did not disproportionately spread lice to conspecifics like the superspreaders did (Zohdy *et al.* 2012). Additionally, mobile individuals that are not central to a network, but that connect otherwise disparate components of a network may function as 'super-movers'. For instance, in a network study of Kenyan ungulates, zebras (*Equus burchelli*) had the highest *cut-point potential* and second highest betweenness scores (e.g. betweenness centrality) of all the studied species, suggesting that their high mobility and large home range brought them into contact with other less-wide-ranging species (VanderWaal *et al.* 2014a).

Contact networks can parse out the competing effects of variable contact rates and individual variability in susceptibility that can result from differing levels of hormones. For example, a directly transmitted helminth (*Heligmosomoides polygyrus*) had inexplicably been shown to have a roughly equal prevalence in male and female yellow-necked mice (*Apodemus flavicollis*) despite demonstrated male-biased transmission (Perkins, Ferrari & Hudson 2008). To explore the unexpected equal prevalence in both sexes, the authors used a contact network combined with a susceptible–infectious (SI)

model to test the relative roles of contact heterogeneity and sex-biased transmission. The mice exhibited dissassortative mixing where males preferentially mixed with females instead of members of the same sex – but just accounting for empirical contact patterns did not yield an equal parasite prevalence between males and females; the male-biased transmission had to be increased tenfold over the observed contact networks to yield model predictions of prevalence that matched observed values.

However, testosterone may aid male-biased infection through immunosuppressive effects and behavioural changes that increase exposure. By experimentally manipulating testosterone levels in male white-footed mice (*Peromyscus leucopus*), Grear, Perkins, & Hudson (2009) showed that males with higher levels of testosterone induced higher contact rates than untreated males, not only for themselves, but for their entire study plot. Elevated cortisol levels may also make individuals more susceptible to infection (Sapolsky 2005). In a group of Japanese macaques (*Macaca fuscata yakui*), MacIntosh *et al.* (2012) sought to parse out the effects of socially mediated exposure and individual susceptibility. The authors found that more-dominant females had higher parasite-shedding rates and higher parasite richness. Female centrality, which roughly mirrored the dominance hierarchy, also correlated with infection prevalence. However, cortisol levels did not differ among females of different rank, supporting the idea that social interaction rather than individual susceptibility was most important in predicting infection risk in this population.

Individuals may also mediate their risk of exposure by inherent temperament or behavioural type (Bell 2007). Animals can exhibit a range of consistent behaviour

roughly correlating to 'bold' and 'shy' temperaments (Natoli *et al.* 2005). For instance, in feral domestic cats (*Felis catus* L.), bold male cats are more likely to contract feline immunodeficiency virus (Natoli *et al.* 2005). Similarly, in a study of bTB risk in cattle, the most dominant animals in the herd (i.e. 'bold') were more likely to explore and interact with sedated brushtail possums, which behaved like terminally ill, tuberculosis-infected possums. Roughly 90% of bTB positive cattle were in the top 20% of the dominance hierarchy (Sauter & Morris 1995). An analogous pattern was seen in a proximity data-logger study of cattle and badgers, where cattle with the highest intra-herd contact rates also had the highest interspecific contact rates (Böhm, Hutchings & White 2009).

How do community structure and group living affect the spread of parasites?

Group living can have significant consequences for epidemic outcomes in wildlife populations. Factors like group size, variance in group size, and individual movement between groups all potentially play a role in epidemic outcomes (Cross *et al.* 2005; Rifkin, Nunn & Garamszegi 2012; Caillaud, Craft & Meyers 2013). In hierarchical populations that contain sub-grouping within larger social groups, understanding disease persistence may depend more on the transmission of the disease between sub-groups rather than characteristic properties of individual transmission events such as R_0 (Cross *et al.* 2007b). One characteristic of networks that incorporates the effects of group living is community structure; a population has community structure when group members have more connections between themselves then between members of other groups. In network terminology, a higher modularity (Q) corresponds to higher density of intra-

group connections relative to inter-group connections (Figure 1.2). Standard network measures like degree distributions may not fully capture the properties of networks with higher modularity (Salathé & Jones 2010).

According to theory, networks with less clustering or fewer sub-groups are more vulnerable to larger-sized epidemics (Keeling 1999). This is consistent for honey bees (*Apis mellifera*) where colonies with lower clustering and more *robust* networks had greater pathogen transmission potential (Naug 2008). In meerkats (*Suricata suricatta*), as group size increased, so did the clustering coefficient, suggesting that transmission of bTB might be limited by the predominance of smaller cliques within larger groups (Drewe *et al.* 2011). Similarly primates living in larger social groups have networks with higher modularity and a correspondingly lower parasite richness (Griffin & Nunn 2011).

In some populations, mobile individuals or nomads can increase the connectivity and potential transmission risk in the network. In Belding's ground squirrels (*Spermophilus beldingi*), higher *transitivity*, a measure of clustering, correlated negatively with the prevalence of *Cryptosporidium*. Moreover, highly mobile juvenile males reduced clustering in the population, which was in turn associated with higher parasite prevalence. By lowering transitivity, these juvenile males functioned as potential superspreaders (VanderWaal *et al.* 2013). By contrast, nomads moving between Serengeti lion (*Panthera leo*) prides increased the overall connectivity of the network, but did not function as critical superspreaders. As territorial animals, Serengeti lions were hypothesized to be less vulnerable to epidemics because of social segregation into prides. However, because of inherent small-world properties of the pride-to-pride network, lions

living in prides exhibited a few, key long-range contacts, which made the effects of transient nomads negligible (Craft *et al.* 2011).

Even if group dynamics are more fluid, as occurs in species exhibiting fission–fusion sociality, accounting for community structure can be important. For example, in a study of African buffalo (*Syncerus caffer*) association data, buffalo exhibited a more dynamic herd structure than previously thought; instead of maintaining highly segregated herds, the population became more uniformly mixed over longer time intervals (Cross *et al.* 2004). Association data were highly variable from month to month and year to year. Simulating disease across these networks showed that tighter clustering in one year would have made the population less permeable to disease compared to another year. Overall, even when accounting for a highly dynamic network with monthly 'rewiring' events, transmission potential was much lower than if movement between groups was truly random (Cross *et al.* 2004).

Are there feedbacks between network position and infection status?

In an evolutionary context, parasitism is theorized to be a cost of group living; generally, directly and indirectly transmitted parasites increase in prevalence, if not necessarily richness, with larger group size (Rifkin *et al.* 2012; Patterson & Ruckstuhl 2013). An intriguing question in the realm of behavioural ecology is whether being central in a network leads to higher parasite load or infection risk, and conversely, whether having a high parasite load or being infected alters behaviour such that an individual becomes more central (Godfrey 2013).

(a) How does network position affect infection status?

Several studies have investigated this question for indirectly transmitted parasites in reptiles, constructing delayed-transmission networks based on asynchronous den or crevasse use. For gidgee skinks (*Egernia stokesii*), well-connected lizards were more likely to be tick-infested and be infected with two or more protozoan parasite species (Godfrey *et al.* 2009). For the tuatara (*Sphenodon punctatus*), a solitary lizard species, territory overlap indices were used to construct directed, *weighted networks*. There was a strong relationship between *in-strength* or the sum of all edge weights directed at an individual and infection with ticks. For both males and females, connections with males were particularly good predictors of parasite load. Mite loads, by contrast, were better explained by individual-level traits like sex, body size and territory size (Godfrey *et al.* 2010). In Australian sleepy lizards (*Tiliqua rugosa*), highly connected lizards – or lizards that often used their neighbours' refuges – had higher tick loads (Leu, Kappeler & Bull 2010).

For brushtail possums experimentally infected with bTB, animals that naturally contracted bTB from experimentally infected individuals were more likely to be central to the network. Specifically they had higher *closeness* and *flow-betweenness* scores relative to individuals that did not become infected naturally (Corner, Pfeiffer & Morris 2003). By contrast, meerkats more central to the network were not *de facto* more likely to be infected with bTB. Rather infection status was behaviour dependent with grooming initiators and aggression recipients having the highest risks of infection (Drewe 2010).

(b) How does an individual's state of infection affect network position and topology?

The first two phases of transmission depend upon (1) individual behaviour that leads to contact between infectious and susceptible hosts, and (2) the transmission efficiency of the agent, which depends on the physiological characteristics of the pathogen and host. Thus, the first phase of transmission can be altered by individual changes in behaviour, which may include (1) seeking out or avoiding infectious individuals, (2) infection-induced changes in behaviour, or (3) covariation between reduced or heightened individual susceptibility and likelihood of contact with infectious individuals. For an excellent overview of the topic and review of potential mechanisms of covariation between the behavioural and physiological components of transmission, see Hawley *et al.* (2011).

While networks are commonly portrayed as a static entity across which pathogens can spread, it has been demonstrated in numerous experimental systems that individual host behaviour changes upon infection. Often, these changes make it easier for the pathogen or parasite to change hosts, especially for trophically transmitted parasites (Berdoy, Webster & Macdonald 2000; Goodman & Johnson 2011); however, behaviours exhibited during sickness like fever, lethargy, and anorexia could reduce intraspecific contacts by limiting mobility and foraging activity (Adelman, Moyers & Hawley 2014). Yet, there has been very little investigation into how infection-induced changes to individual-level behaviour might scale up to changes in contact frequency at a population level (Hawley *et al.* 2011).

Initial experimental findings suggest that accounting for infection-induced changes in behaviour may be very important. Using Trinidadian guppies (*Poecilia reticulata*) and directly transmitted trematodes (*Gyrodactylus* spp.), Croft *et al.* (2011) explored how groups of fish reacted to the introduction of either an infected or uninfected individual. They found that infected guppies associated less with the group than their uninfected counterparts. Moreover, in infected treatments, uninfected individuals were more likely to initiate shoal fission – or split the group – than in uninfected treatments, which may serve as a possible avoidance mechanism.

Even if the exact manifestations that infection may incur on contact frequency are not known, modellers can begin to explore how behavioural changes could affect epidemic dynamics. For instance, when modelling the hypothesized effects of dumb and furious rabies on raccoon (*Procyon lotor*) contact patterns, both the final outbreak size and speed of rabies spread varied in response to these simulated behavioural changes (Reynolds *et al.* 2015).

Are certain populations more vulnerable to disease epidemics?

The field of graph theory has provided numerous theoretical insights on how network structure affects epidemic behaviour (Keeling, 1999; Keeling & Eames, 2005). However, applying these findings to the realm of conservation to help protect endangered or threatened wildlife populations is relatively novel. In this context, the term 'population' may correspond to a species, sub-species, or any group of individuals within a designated geographic range. This has been done for at-risk or threatened populations

such as chimpanzees (*Pan troglodytes schweinfurthii*) and orangutans (*Pongo pygmaeus wurmbii*) (Carne *et al.* 2013), and mammal-eating killer whales (*Orcinus orca*) (Guimarães *et al.* 2007). Even individuals that may be considered relatively asocial, like Tasmanian devils or raccoons, may turn out to be more vulnerable than anticipated because the population is completely or highly well connected (Hamede *et al.* 2009; Hirsch *et al.* 2013). Similarly, populations exhibiting small world networks – consisting of mostly local, but a few long range contacts – may be particularly vulnerable to epidemics (Porphyre *et al.* 2008; Craft *et al.* 2011).

How important are heterogeneities in interspecific interactions for maintenance or spillover of multi-host pathogens?

The contraindications of badger culling illustrate the challenges of designing interventions for multi-host pathogens like bTB. Contact network studies can help elucidate intra- and interspecific contact patterns, especially when transmission pathways are unclear. In the UK, the prevailing belief was that badgers generally avoided cattle, so most management strategies to date have focused on reducing indirect transmission scenarios, i.e. exposure to bacilli from contaminated pasture or feeding troughs (Garnett, Delahay & Roper 2002; Garnett, Roper & Delahay 2003). However, in a study of interspecific contacts between badgers and cattle in the UK, contacts between badgers and cattle occurred more frequently than intraspecific contacts between badgers of different setts. Although interspecific contacts were still relatively infrequent, direct contacts occurred more often than previously suspected, and these contacts usually occurred with cattle that were very central to their intraspecific networks. This suggests

that direct interspecific interactions between badgers and cattle could be more important than inter-sett badger contacts in spreading bTB (Böhm *et al.* 2009). One of the limitations of contact network studies is the challenge associated with determining the directionality of transmission (i.e. badger to cattle *versus* cattle to badger), but there is now genetic evidence supporting interspecies transmission in this system. A whole-genome study of *M. bovis* in sympatric badger and cattle populations found evidence that suggested that *M. bovis* had likely persisted in the region for several years, despite cattle herds testing bTB negative, and that recent transmission events between the two species were probable (Biek *et al.*, 2012). Accounting for both direct and indirect transmission scenarios between badgers and cattle will be an important consideration for future management strategies (Böhm *et al.* 2009).

Contact networks can also be used to explore disease dynamics and persistence for potential wildlife reservoirs or spillover events into wildlife populations. A reservoir is a population, or group of connected populations, capable of maintaining an infection indefinitely and, further, is capable of transmitting the infection to a 'target species' for which elimination or control of the disease is the desired outcome (Haydon *et al.* 2002; Viana *et al.* 2014). The transmission of a pathogen from a reservoir to a target species constitutes a spillover event (Power & Mitchell 2004). Detailed contact network information was used to explore retroactively the dynamics of the 1994 canine distemper virus (CDV) outbreak in Serengeti lions. The question of interest was whether the Serengeti lion population would have been able to self-sustain the epidemic or whether the outbreak involved repeated spillover from sympatric hosts like hyenas (*Crocuta*

crocuta) or jackals (*Canis adustus, C. aureus,* and *C. mesomelas*). In this case, the lions were found to be a 'non-percolating' population incapable of acting as a maintenance host for CDV, thus requiring multiple spillover events from other hosts to give rise to the observed epidemic (Craft *et al.* 2009). Therefore eliminating CDV from the system will require interventions directed at other hosts rather than just lions.

Just as it is possible to identify a superspreader within a population, for multi-host pathogens, it is also possible to ask whether a particular species is functioning as a superspreader relative to the community. In a study of wild and domestic ungulate species in Kenya, association data was combined with *E. coli* genetic strain information to infer a transmission network. Based on this information, Grant's gazelle (*Gazella granti*) were found to be most central to the network, while zebra were most likely to act as bottlenecks or conduits between two otherwise unconnected sub-groups (VanderWaal *et al.* 2014a).

Are there 'trait-based' features that are predictive of superspreader status?

While many studies try to find superspreaders based on observed contact patterns, an equally important but neglected question is whether there are any 'trait-based' characteristics that are predictive of an individual being a superspreader. This is particularly important in wildlife populations where elucidating a contact network can be time consuming and costly. So far only a few studies have investigated this question, yielding mixed results. As previously mentioned in the case of Tasmanian devils, no particular class of individual was more highly connected overall, making targeted interventions against DFTD of limited utility (Hamede *et al.* 2009). Otterstatter &

Thomson (2007) explored how different social and labour roles in bumble bee (*Bombus impatiens*) colonies correlate with infection status, and found that an individual's unique position in the network was the best predictor of infection risk. Centrality was a poor predictor because infection depended specifically on the contact rate with other infected hive mates rather than just having a high contact frequency overall. In this way, their results mirror findings in human sexually transmitted disease network studies (Friedman *et al.* 1997; Wylie & Jolly 2001; Neaigus *et al.* 2001; Liljeros, Edling & Amaral 2003).

However, some studies have found that age, morphological traits, or position within the social hierarchy can indicate high-risk individuals. In honeybees, age of the individual dictates spatial organization of the colony. This means that older, foraging bees are most likely to be the first infected and thus introduce the pathogen into the colony (Naug 2008). For deer mice, body mass rather than sex or breeding condition was the best predictor of well-connected individuals (Clay *et al.* 2009). Wild chimpanzees with large families and highly ranked males were more likely to be more connected within networks (Rushmore *et al.* 2013), and targeting these classes of individuals with simulated vaccine interventions was found to be more effective than random vaccination (Rushmore *et al.* 2014). The evidence of trait-based characteristics and the corresponding, theoretical support for trait-based interventions could allow for more cost-efficient and targeted management strategies even when the entire contact structure of a population is not known. While initial findings have been mixed, this is perhaps one way that scientists and managers can generalize the results of contact network studies that otherwise originate from specific host–pathogen systems.

1.6 Unresolved methodological questions

Not all contacts are created equal: what is a 'contact'?

How researchers collect data for contact network models and the assumptions they make while doing so have implications for the resultant network and the predictions that can safely be made from it. Transmission consists of three stages: (1) the susceptible and infectious individual coming into contact, (2) the transfer of the pathogen from infectious to susceptible individual, and (3) the development of infection from that transfer. It is important to recognize that contact networks only give insight into the first step (Hamede *et al.* 2009). The type of contact needed for a possible transmission event varies with the targeted species and pathogen, and in many cases, what truly constitutes a 'contact' may still be unknown.

The variability in contact definition has been recognized as a challenge for modelling pathogen spread in humans. For example, a network representing possible transmission of a sexually transmitted disease like syphilis might contain only a subset of the edges for a network representing an airborne disease like measles (Keeling & Eames 2005). Assumptions about what constitutes a 'contact' in wildlife vary widely across systems and may be limited by the method of data collection. For example, 'contact' has been defined as: being in the same herd (Cross *et al.* 2004); being caught in the same trap on the same day (VanderWaal *et al.* 2013); two individuals found within a certain distance according to proximity data (Böhm *et al.* 2009); kernel density estimates based on trap location (Porphyre *et al.* 2008); shared but asynchronous refuge use (Godfrey *et al.* 2009); and even asynchronous territory overlap for a solitary species (Godfrey *et al.*

2010)(Table 1.1: Column M1). Different transmission modes may be captured by different measures of contact. For instance, contacts for faecal–oral parasites may be accurately captured through shared space use, while directly transmitted parasites may require more restrictive definitions of spatial and temporal proximity. Consequently, the type or length of contact may be important in determining transmission. For example: does the interaction have to be solely aggressive as is suspected for DFTD and rabies or can it be mere proximity? Network topology and density can change dramatically when different time requirements are used to limit the definition of a contact. In monthly contact networks for raccoons, increasing the amount of time required to constitute a contact decreased social connectivity measures (Hirsch *et al.* 2013). Similarly, for deer mice and Sin Nombre virus, contact frequency alone was not a predictor of infection status until combined with contact duration (Clay *et al.* 2009). Such challenges in defining a contact can even be seen in the same host for different pathogens. For instance, den sharing has been shown to be critical for bTB transmission in brushtail possums (Corner *et al.* 2003), but is of secondary importance compared to brief nocturnal interactions for transmission of commensal *E. coli* (Blyton *et al.* 2014).

The epidemiological characteristics of the parasite of interest should also be considered when defining a contact. Such factors include the length of incubation or latency and infectious periods and, in the case of indirectly transmitted parasites, the persistence of the parasite off the host. Naug & Smith (2007) experimentally manipulated the infectious period for a gut-pathogen proxy in honey bees and compared the resultant networks derived from food transfer interactions. Since only food that remains in a bee's

crop can be transferred to other bees, they created long and short 'infectious periods' by manipulating the bees' rate of digestion: higher concentrations of sucrose resulted in a slower rate of crop emptying such that the gut-pathogen proxy remained available for mouth-to-mouth transfer for longer. And *vice versa*, lower concentrations of sucrose created the shorter infectious period treatment. Not surprisingly, the network derived from the longer infectious period treatment had higher prevalence, but lower intensity of 'infection' (Naug & Smith 2007), as well as a higher mean degree and greater flow than the network for the shorter 'infectious period' (Naug 2008). Although no actual pathogen was present, this experiment hints at the importance of accounting for relevant exposure and infectious periods during data collection. Ultimately, several factors should be taken into account when defining a contact: the type and length of contact needed for a transmission event, the data collection method available for a given species, and the epidemiological characteristics of the parasite of interest.

How does the method of data collection affect the perceived network?

As suggested by the number of definitions for a 'contact', there is no standard method of data collection for association data. The contact structure of a population can be collected through capture–mark–recapture, behavioural observations, powder marking, and various forms of remote monitoring such as telemetry, proximity loggers, or cameras [see Craft & Caillaud (2011) for a summary of methods and Krause *et al.* (2013) for a review of biologging and telemetry techniques]. In the same way that different definitions of a contact will result in networks with different properties, different methods of data collection will also provide different portrayals of association data. Comparing the

perceived networks obtained from different data-collection methods and different definitions of a contact may help to determine what type of contact is important (VanderWaal *et al.* 2014b).

In a comparison of methods to quantify contact networks using radio telemetry and capture–mark–recapture data, Perkins *et al.* (2009) found that both methods of data collection produced similar contact distributions for a population of yellow-necked mice. However, while betweenness was similar, other network metrics like closeness and *connectance* were statistically different. This difference may have arisen as a result of changing rodent density, since capture–mark–recapture data were more informative at higher densities and radio-telemetry networks provided better resolution at low population densities (Perkins *et al.* 2009). Similarly, in a methods comparison study with three modes of data collection, cameras and global positioning system (GPS) units documented fewer contacts relative to proximity loggers in white-tailed deer (*Odocoileus virginianus*) (Lavelle *et al.* 2014). Other studies indicate that capture–mark–recapture methods may underestimate potential contacts. In mouse lemurs, contact networks were derived from both capture–mark–recapture of lemurs and individual tagging of an obligate lice species. The turnover of lice on lemurs was consistent with more numerous and less-spatially limited contacts than indicated by recapture data alone (Zohdy *et al.* 2012).

Are social networks effective pathways for transmission of indirectly transmitted parasites?

Although contact networks have been traditionally employed for directly transmitted parasites, an active area of interest in the realm of behavioural and disease ecology is exploring the ability of contact networks to predict indirectly transmitted parasites from asynchronous space use (Godfrey 2013). This is commonly done with 'delayed-transmission' networks, where individuals using the same space within a designated interval of time (but not necessarily at the same time) are considered to have an edge between them in the network. This edge may be weighted based on the frequency of the space-sharing or the distance between individuals, giving rise to a weighted network.

The effectiveness of using networks to predict indirectly transmitted parasite loads seems to be parasite and host dependent. For the Australian sleepy lizard and its ectoparasitic tick, transmission networks derived from asynchronous den use accurately predicted that highly connected lizards had higher tick loads (Leu *et al.* 2010). Similarly, when documenting the infection of Pygmy bluetongue lizards (*Tiliqua adelaidensis*) with ticks and nematodes, lizards infected with ticks were more connected in the network, but lizards infected with nematodes did not differ in their connectivity when compared with uninfected individuals (Fenner, Godfrey & Michael Bull 2011). In the solitary tuatara, there was a strong relationship between in-strength and infection with ticks and tick-borne protozoans, but mite loads were better explained by personal traits like sex, body size or territory size (Godfrey *et al.* 2010). Ultimately, however, the empirical results

supporting indirect pathogen/parasite transmission are largely correlative because it is very difficult to separate the contributions of spatial heterogeneity and contact frequency (Bull, Godfrey & Gordon 2012).

A new direction that some studies have taken is simultaneously to track the social network and the strains of a commensal proxy like *E. coli* or *Salmonella* species; this provides greater resolution than simply treating a parasite or pathogen as a uniform entity. In the Australian sleepy lizard, individuals sharing the same genotype of *Salmonella enterica* were more likely to be connected in the social network; by contrast, spatial proximity was not a predictor of a shared genotypes (Bull *et al.* 2012). Similarly for giraffes (*Giraffa camelopardalis*), an individual's position in the social network, but not spatial network, correlated to its position in a transmission network even for *E. coli*, which transmits *via* indirect faecal–oral routes (VanderWaal *et al.* 2014b).

While the body of work supporting the use of contact networks as transmission routes for indirectly transmitted pathogens is growing, researchers should be vigilant in defining the network in a suitable way that matches host life history and pathogen life cycle. For example, for helminths (e.g. *Capillaria tamiasstriati* and *Citellinema bifurcatum*) in chipmunks (*Tamias striatus*), transmission of faecal–orally transmitted species only correlated with contact networks that allowed for a time delay that corresponded to the parasites' off-host development stage. Social contact networks that represented synchronous space use were not representative of transmission. Moreover, transmission of trophically transmitted species (i.e. those with mobile intermediate hosts)

did not correlate with either delayed-transmission or social contact networks (Grear, Luong & Hudson 2013).

Dynamic networks and rewiring: how do temporal changes in networks affect epidemic outcomes?

Dynamic networks are networks that allow for changes in contact structure over time. These changes are called rewiring events. Blonder *et al.* (2012) warn against the consequences of treating a network as if its topology is fixed, when it is in fact changing. But for most pathogens especially in wildlife populations, it is not always clear how important it is to account for these dynamic interactions (Pellis *et al.* 2014). While dynamic network studies in wildlife are not usually employed (but see Cross *et al.*, 2004; Hamede *et al.*, 2012; Wilson *et al.*, 2014), theoretical work suggests that using static networks in lieu of dynamic networks can greatly alter epidemic outcomes (Fefferman & Ng 2007). For example, accounting for temporal changes in livestock contact networks has significant ramifications for epidemic behaviour, especially for diseases with lower reproductive numbers (i.e. $R_0 < 2$) (Chen *et al.* 2014). Another important consideration is that the pathogen itself may contribute to rewiring of the network (Eames *et al.*, 2015). Accounting for this rewiring may be more important for some host–pathogen systems than others. For instance, a highly virulent pathogen that causes high mortality will eliminate nodes in the network, and a highly mutable pathogen may re-infect recovered nodes.

In general, wildlife networks appear to be very dynamic in nature, and thus may have important implications for predicting disease dynamics (Craft & Caillaud 2011a). Observed meerkat monthly grooming and aggression networks were stable through time (Drewe *et al.* 2011), but wild chimpanzee networks were highly dynamic over time (Rushmore *et al.* 2013) as were inter- and intraspecific contact patterns for badgers and cattle (Böhm *et al.* 2009). Other species exhibit different patterns of association between breeding and non-breeding seasons including Tasmanian devils, possums, and mouse lemurs (Ji *et al.* 2005; Hamede *et al.* 2009; Zohdy *et al.* 2012; Rushmore *et al.* 2013). This means that accounting for seasonal variations in networks may be sufficient for certain species and, in some cases, knowing more about the dynamicity of networks could inform decisions for timing of interventions like oral bait vaccine administration for rabies prevention in raccoons (Reynolds *et al.* 2015).

1.7 Discussion

Using, developing, and comparing contact networks with discernment

While this review has emphasized cases where contact heterogeneity has been found to explain infection patterns in wildlife, it may be impractical and unfeasible to account for contact heterogeneity in all cases. Before undertaking a contact network study in wildlife, researchers should consider whether accounting for non-random mixing would substantially change predictions or potential interventions. Returning to Levins' modelling trade-offs, researchers must evaluate the costs of realism (Levins 1966). It is the exception, rather than the rule, that targeted interventions are realistic options in

wildlife (Vial *et al.* 2006; Haydon *et al.* 2006). For this reason, future studies should seek out trait-based characteristics that can enable network interventions even if the whole network structure is not known (Salathé & Jones 2010; Pellis *et al.* 2014). Based on work that has been done so far, it seems that directly transmitted pathogens in systems with high spatial organization or clear social hierarchy would be most likely to exhibit trait-based characteristics, i.e. social insects (Naug, 2008) or primates (Rushmore *et al.*, 2013, 2014). Uncovering populations where trait-based methods can be used effectively is essential since data collection for networks can be both costly and time intensive, and since projecting results for an entire population from a subsample is often necessary (e.g. Craft *et al.*, 2009; Reynolds *et al.*, 2015).

If researchers opt for association data, it is very important that they define contacts and collect data at intervals that are relevant to the epidemiology of the pathogen of interest (Cross *et al.* 2004; Keeling & Eames 2005; Grear *et al.* 2013). Data collection should take into account host life history and the relevant infectious and exposure periods for pathogens in order for the contact network to be meaningful. For instance, when evaluating a population's vulnerability to epidemics, combining months and years of data into a single contact network can suggest an extremely interconnected population, when in fact, the infectious period is much shorter in duration than the time between when individuals come into contact (Cross *et al.* 2005). This is especially true for very solitary species or sparse populations where contacts are infrequent.

Finally, researchers should be wary of comparing systems that do not lend themselves to comparison. As evidenced in this review, empirical studies of networks in wildlife detail

a huge range of hosts, pathogens, 'contact' definitions, and data-collection methods. Comparing results from studies on tick transmission in lizards and tuberculosis in meerkats without acknowledging the different biological and methodological processes at hand may lead to spurious conclusions.

Current dilemmas: separating correlation from causation

In many aspects of network studies, it is challenging to separate correlation from causation. For example, is a social network truly reflective of transmission patterns or would transmission be explained better by spatial heterogeneities that lead to preferential host clumping or parasite survival? For the majority of papers exploring parasite load as a function of sociality, contact occurs when different individuals share the same space even if that use is asynchronous, but empirical evidence supporting indirect parasite transmission through social networks is lacking (Bull *et al.* 2012).

The goal for understanding the dynamics of indirectly transmitted parasites or faecal–oral pathogens should be to separate the effects of shared space use and contact frequency. It is worth considering that some pathogens, which are generally thought to be directly transmitted, can persist for a long time in the environment. For instance, a large proportion of wildlife network studies involve bTB (e.g. Porphyre *et al.*, 2008; Böhm, Hutchings, & White, 2009; Drewe, 2010), which can certainly be transmitted directly, but also can persist for almost three months in soil depending on conditions (Fine *et al.* 2011). Similarly, certain prion diseases like chronic wasting disease (CWD) have been shown to remain infectious in soil for an undetermined period (Johnson *et al.* 2006), but

modelling attempts and interventions have focused on density effects and direct contacts between hosts (Almberg *et al.* 2011). Questions of environmental persistence are also important for commensal proxies like *E. coli*; not all studies have identified a relationship between social contact and strain similarity (Chiyo *et al.* 2014). Lastly, it is unclear how much turnover or variability occurs for commensal bacteria within individuals, although by some estimations this turnover may be extremely high (Blyton *et al.* 2013). Such microbial community turnover is an important consideration for contact network studies because only bacterial strains with relatively high individual turnover are likely to reflect recent contact patterns (Blyton *et al.* 2014). If turnover rates are low relative to contact rates, resolution of contact patterns may be lost, and thus, turnover rates may also be important to consider when determining sampling frequency.

Network studies that track proxy pathogens or commensal species offer the opportunity to separate out the roles of shared space and contact frequency because they glean information from both the genetics of the pathogen and the contacts of the host (Bull *et al.* 2012; VanderWaal *et al.* 2014a; b; Blyton *et al.* 2014). However, the intersection of how individual behaviour and overall network topology change in response to infection is not well studied (Croft *et al.* 2011; Hawley *et al.* 2011). With these potential limitations, it is important to recognize that contact networks identified for a commensal species may be vastly different than those for a pathogenic species (VanderWaal *et al.* 2014b).

Consequently, another priority should be to explore the reciprocal relationship of how pathogens spread on a network and conversely, how a social network changes in

response to disease. Modelling studies in humans have begun to incorporate the effects of behavioural changes in response to disease, including: adherence to vaccination programs, fear-induced contact reduction, hygiene improvement, or changes in mobility or travel (Epstein *et al.* 2008; Coelho & Codeço 2009; Funk, Salathé & Jansen 2010; Funk *et al.* 2015; Meloni *et al.* 2011). For wildlife studies, this may require dynamic network methods and experimentation that involves introducing parasitized individuals to healthy, established groups. Studies where infection status is monitored while observing contact network structure in wildlife are relatively rare (e.g. Corner, Pfeiffer, & Morris, 2003; Croft *et al.*, 2011). Disease ecologists should seek out and test systems that are amenable and ethical for observing the spread of fast-spreading, directly transmitted pathogens; social insects might provide better opportunities for this (Naug & Smith 2007; Otterstatter & Thomson 2007; Naug 2008; Konrad *et al.* 2012; Charbonneau, Blonder & Dornhaus 2013). These tractable systems would enable the continuous monitoring of an entire population, a feat that is challenging at best in larger, more mobile species. Such investigations will help to elucidate another causation–correlation dilemma: is infection a result of network position or does being infected alter host behaviour and thus a host's contacts?

A final goal should be to investigate further the relative importance of population heterogeneity (i.e. contact frequency) and intrinsic individual heterogeneity as manifested in host susceptibility or immune response. Networks can be used to elucidate population contact patterns and then test different transmission biases in order to match actual population prevalence. This has only been done in a handful of wildlife systems so far

(Perkins *et al.* 2008; Grear *et al.* 2009; MacIntosh *et al.* 2012). Collaboration between the fields of disease ecology and ecological immunology offers new opportunities for answering these questions (Hawley & Altizer 2011), and the increasing capabilities of remote biomonitoring hold great promise with smaller devices that will enable the collection of both contact data and individual physiological responses to disease (Adelman *et al.* 2014).

Future directions

When used judiciously, contact networks are a valuable extension to traditional disease modelling approaches. Contact networks can offer insights to questions of parasite transmission across a broad range of fields and applications. Cross-disciplinary efforts are already leading to novel approaches and questions from ecologists, epidemiologists, veterinarians, geneticists, and conservation scientists. Behavioural ecologists have expanded contact network use into the realm of indirectly transmitted parasites. The convergence of disease ecology with genetics (Archie, Luikart & Ezenwa 2009) and ecological immunology (Hawley & Altizer 2011) has offered some of the most promising inroads towards a better understanding of actual transmission patterns and individual susceptibility.

A novel technique that holds great promise for inferring transmission is the simultaneous tracing of parasites and contact data. This has been done selectively for individual macroparasites (Zohdy *et al.* 2012) and commensal strains of bacteria like *E. coli* and *Salmonella* spp. (Bull *et al.* 2012; Blyton *et al.* 2014; VanderWaal *et al.* 2014b).

In fact, with higher through-put metagenomic sequencing methods, it may soon be feasible to trace genetic strain information across a contact network, not just for a single pathogen of interest, but for a microbial community from any given individual (Archie & Theis 2011; Thomas, Gilbert & Meyer 2012; Kao *et al.* 2014; Zhou *et al.* 2015). One study demonstrated that adjacent chimpanzee community members exhibited different microbiota profiles, and that community members that had immigrated long ago still retained traces of their former community's microbiota profile (Degnan *et al.* 2012). At present, there is still a poor understanding of how environment and genetics interact to create an individual's microbiome (Benson *et al.* 2010), but eventually metagenomic sequencing methods could provide another tool for eliminating confounding factors like shared space, diet, or individual microbial turnover.

Technological advances also mean that real-time tracking of contact patterns will continue to become more accessible to a greater number of scientists. Biologging technologies are becoming smaller, longer lasting, further-ranging and more affordable (Krause *et al.* 2013). Scientists may be able to deploy more devices per study and potentially track smaller species that currently have device weight limitations. Additionally, more studies are beginning to parse out how data-collection method affects empirical results (e.g. Perkins *et al.*, 2009; Lavelle *et al.*, 2014). Relatively novel biomonitoring devices will offer the chance to collect simultaneous contact and physiological data to parse out questions of individual susceptibility *versus* variability in contact rate (Adelman *et al.* 2014).

With such advances, contact network studies that implement direct, experimental manipulations of natural systems should also become more feasible. Given the confounding factors inherent to any natural system, it is not surprising that wildlife network studies that use direct experimental manipulation are still rare and typically employ captive wildlife populations (Corner *et al.* 2003; Croft *et al.* 2011) – Grear *et al*. (2009) is a notable exception. In other areas of ecology, studies that have experimentally manipulated wildlife populations to study disease hint at ways in which experimental methods could be employed to deconstruct unresolved issues in contact network studies. For instance, administration of anthelmintics to rodent populations might provide an opportunity to study how social networks change in the presence or absence of parasites (Knowles *et al.*, 2013; Pedersen & Antonovics, 2013).

Ultimately the strengths of a contact network approach are also its weaknesses. Realism and precision can limit the applicability of contact data to general contexts. How can scientists extrapolate from these often limited and system-specific data sets and begin to inject realism into more general models? Approaches that are common in human studies like exponential random graph models (ERGMs) have seen little use so far in wildlife studies, but would enable researchers to begin inferring disease dynamics for a larger population from a subset of individuals (e.g. Reynolds *et al.*, 2015). Since network studies are often limited in the number of individuals that they can follow, new statistical approaches are being developed to tackle the challenges of boundary effects and sampling bias (Cross *et al.* 2012). Finally, Bayesian techniques of statistical inference may enable more generalized conclusions about the overall properties of networks or

transmission behaviour from specific data sets (Welch, Bansal & Hunter 2011; Groendyke, Welch & Hunter 2011; Lindström *et al.* 2013). Collectively, these technological and statistical improvements will make contact network studies feasible for more investigators and for new host–pathogen systems, and ideally, lead to the identification of broad-scale patterns that hold true across systems.

1.8 Conclusions

(1) The articles highlighted here address a wide variety of hosts and parasites (Table 1.1), but replication of results in the same system is infrequent. We identified 11 overarching theoretical, applied, and methodological questions that these studies collectively address; despite an increase in contact network wildlife studies in recent years, there are still only a few studies that address each transmission-related question, which leaves a lot of investigative space for future studies to explore. Most studies targeted directly transmitted parasites; a majority of that subset dealt with bTB in different wildlife systems. Comparatively few papers targeted indirectly transmitted, vector-borne, or trophically transmitted parasites. The majority of studies relied on behavioural observations or capture–mark–recapture methods to establish contact network patterns, but new methods like proximity loggers and strain sharing of commensal bacteria are becoming more popular.

(2) Several areas stand out as priorities for future contact network research: (*a*) distinguish the effects of shared space use and contact frequency, especially for indirectly transmitted parasites; (*b*) explore the reciprocal relationship of how pathogens spread on a network, and conversely, how a social network changes in response to disease; (*c*)

investigate the relative roles of contact heterogeneity and individual host variability in susceptibility or immune response on parasite transmission; and (*d*) prioritize the identification of 'trait-based' features that will make network results more applicable to other systems and allow for targeted interventions and treatment, even in populations where contact patterns are unknown.

(3) Researchers should be vigilant in defining contacts, choosing a data-collection method, and collecting data at suitable frequencies in order to best match host life history and parasite life cycle. Definitions of a contact are highly variable in the articles reviewed here (Table 1.1: Column M1). Researchers should critically examine several factors when defining a contact: the type and length of contact needed for a transmission event, the data-collection method available for a given species, and the epidemiological characteristics of the parasite of interest. More empirical and modelling work are needed to clarify how perceived contact networks differ as a result of different data-collection methods. Based on the predominance of seasonal contact patterns in many wildlife species, researchers should allow for dynamic rewiring in their network models or present strong evidence for exemption.

(4) All scientists must contend with the tension between realism and generality when modelling. Contact networks give detailed, precise information for a small sample of a particular host and pathogen during a short time period, often making generalizations and comparisons across systems challenging. Additionally, collecting empirical data for contact network studies in wildlife is time and resource intensive, so researchers must evaluate the tractability of their system to data collection. New technological advances in

the form of biomonitoring and more portable tracking devices will make contact data collection more accessible to new investigators and for new systems, while new statistical approaches including ERGMs and Bayesian inference will make it easier to extrapolate and generalize from future and existing studies. Ideally, such advances will aid in the identification of broad-scale patterns that exist across host–pathogen systems.

1.9 Acknowledgements

We would like to thank Kim VanderWaal and two anonymous reviewers for their helpful comments on this manuscript. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 00039202. J.D.F was supported by an Institute on the Environment Resident Fellowship and a UMN Grant in Aid (#22325). M.E.C. was funded by National Science Foundation (DEB-1413925), the University of Minnesota's Institute on the Environment, Agricultural Experiment Station General Agricultural Research Funds, and College of Veterinary Medicine. This material is based upon work supported by the Cooperative State Research Service, U.S. Department of Agriculture, under Project Numbers MINV 62-044, and 62-051. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.

1.10 Figures and tables

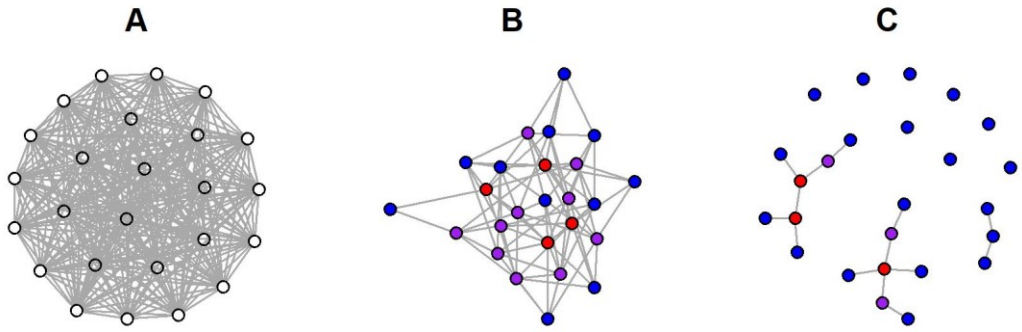


Figure 1.1. A depiction of networks with different probabilities of edge formation for 25 nodes. (A) This network is equivalent to a compartmental model where every individual has an equal probability of encountering any other individual or every node is connected to every other node. The mean degree is 24. (B) An Erdős–Rényi random graph where probability of edge formation is equal to 0.25. Colour of the nodes corresponds to normalized betweenness centrality. Red nodes are the most central, purple nodes have mid-range centrality and blue nodes are the least central. The mean degree is 6.64. (C) An Erdős–Rényi random graph where probability of edge formation is equal to 0.05. Here the network consists of multiple components and the mean degree is 1.12.

Figure 1.2. A graphical depiction of modularity for 100 individuals in a community. Colours correspond to the ten different groups within the community. Modularity (Q) is the proportion of connections within groups relative to the connections between groups within a community. Modularity is highest for completely segregated groups (A), decreases as the number of inter-group connections grow (B, C), and approaches zero for a more well-mixed community (D).

Table 1.1. Empirical studies or studies that used empirical data to study the spread of parasites on wildlife host networks. For inclusion, the host species had to be wildlife, although studies targeting captive wildlife populations were also included. The numbered topics (Columns Q1–Q7) correspond to the theoretical and applied questions highlighted in order of appearance in Section IV. The grey shading indicates studies of particular relevance to each section. In some columns, there are studies that are highlighted as pertinent, but that are not discussed explicitly in the main text under that particular heading. The letters (a) and (b) listed under Column Q4 correspond to the sub-sections for the corresponding question in the text. The numbered topics (Columns M1–M4) indicate which empirical studies pertain to the unresolved methodological questions in Section V with column M1 explicitly listing the definition of a contact for each study. Disease or parasite abbreviations: bTB, bovine tuberculosis; CDV, Canine Distemper Virus; CWD, Chronic Wasting Disease; DFTD, Devil Facial Tumour Disease; SNV, Sin Nombre Virus. Studies that did not develop a network for a specific parasite are labelled 'generic'. Transmission type abbreviations: D, direct, I; indirect; T, trophic; V, vector-borne. Data-collection method abbreviations: Behavioural obs., behavioural observations; CMR, capture–mark–recapture. GPS, global positioning system; PIT, passive integrated transponder; VHF, very high frequency.

Study	Species	Disease or pathogen	Transmission type	Data collection method	Q1: Superspreaders	Q2: Individual variability	Q3: Community structure and group living	Q4: Infection status and network position	Q5: Population vulnerability	Q6: Multi-host pathogens	Q7: Predictive trait-based features	M1: What is a "contact"?	M2: Data collection method	M3: Indirectly transmitted parasites	M4: Dynamic networks
Corner *et al.* (2003)	Brushtail possum	bTB	D	Behavioural obs.				(a)				Den sharing			
Cross *et al.* (2004)	African buffalo	bTB	D	Radio-tracking								Same herd			
Ji *et al.* (2005)	Brushtail possum	bTB	D	Proximity data loggers								< 40 cm			
Guimarães *et al.* (2007)	Killer whales	Not specified	D	Behavioural obs.								Same pod			

Otterstatter & Thomson (2007)	Bumble bees	Gut protozoan	D	Video & behavioural obs.				■			■	Physical contact ≤ 1 cm			
Böhm *et al.* (2008)	Badger	bTB	D	Radio-tracking	■		■					< 50 or < 250 m			
Naug (2008)	Honeybee	Generic gut pathogen	D	Video & behavioural obs.			■				■	Food transfer	■		
Perkins *et al.* (2008)	Yellow-necked mouse	Helminth	D	CMR		■						Same or adjacent trap within 14 days			
Porphyre *et al.* (2008)	Brushtail possum	bTB	D	CMR	■			■				Kernel density estimates			
Böhm *et al.* (2009)	Badger & cattle	bTB	D	Proximity data loggers	■	■				■		< 4 m			■
Clay *et al.* (2009)	Deer mice	SNV	D	Powder marking & PIT tags	■						■	Powder transfer; same location within 15 s			
Craft *et al.* (2009)	African lion	CDV	D	Behavioural obs.						■		Proximity			
Perkins *et al.* (2009)	Yellow-necked mouse	Generic faecal-oral parasite	I	Radio telemetry & CMR								≤ 15 m; same/adjacent trap within 5 days	■		
Godfrey *et al.* (2009)	Gidgee skinks	Protozoans and ticks	V or I	CMR			■	(a)				Asynchronous crevasse use (≤ 41 days)		■	
Grear *et al.* (2009)	White-footed mice	Generic micro- or macroparasite	D or I	CMR		■						Same or adjacent trap within 2 weeks			
Hamede *et al.* (2009)	Tasmanian devil	DFTD	D	Proximity data loggers	■				■		■	< 30 cm			■
Drewe (2010)	Meerkats	bTB	D	Behavioural obs.				(a)				Grooming or aggressive interactions			
Godfrey *et al.* (2010)	Tuatara	Ticks, mites & protozoan	V or I	Behavioural obs.				(a)				Asynchronous territory use overlap		■	
Leu *et al.* (2010)	Australian sleepy lizard	Ticks	V or I	Behavioural obs.			■	(a)				Asynchronous refuge use		■	
Craft *et al.* (2011)	African lion	CDV	D	Observation, GPS & VHF collars			■		■			Different prides ≤ 1 m & feeding events			
Croft *et al.* (2011)	Trinidadian guppy	Trematode	D	Behavioural obs.				(b)				≤ 4 body lengths			■
Drewe *et al.* (2011)	Meerkats	bTB	D	Behavioural obs.		■	■	(a)				Grooming, aggression, or foraging			■
Fenner *et al.* (2011)	Pygmy bluetongue lizards	Tick and nematode	I	Behavioural obs.			■	(a)				Burrows < 31 m apart		■	
Griffin & Nunn (2011)	Primates	Generic	D or I	Meta-analysis			■					Meta-analysis: primarily grooming			
Bull *et al.* (2012)	Australian sleepy lizard	*S. enterica* (commensal)	I	GPS & genetic strain				(a)				< 2 m		■	
Hamede *et al.* (2012)	Tasmanian devil	DFTD	D	Proximity data loggers	■				■			< 30 cm			■
MacIntosh *et al.* (2012)	Japanese macaques	Nematodes and Strongyloides	I	Behavioural obs.		■					■	Grooming			
Zohdy *et al.* (2012)	Mouse lemurs	Obligate lice	D	CMR	■	■						Overlapping trap locale & louse exchange	■		■

Carne *et al.* (2013)	Great apes & organ-utans	Generic (ex. Ebola)	D	Behavioural obs.	■		■		■			Party level (≤ 50 m)			
Grear *et al.* (2013)	Eastern chipmunks	Helminths	I and T	CMR								Same/adjacent traps; same, 1, or 2 week intervals	■	■	
Hirsch *et al.* (2013)	Raccoons	Generic (ex. Rabies, CDV)	D	Proximity data loggers					■			1–1.5 m proximity; 1, 5, 15 or 30 min/month	■		
Rushmore *et al.* (2013)	Wild chimpanzees	Generic (ex. Ebola)	D	Behavioural obs.	■						■	Party level and ≤ 5 m associations			■
VanderWaal *et al.* (2013)	Belding's ground squirrels	*Cryptosporidium*	I	CMR	■		■					Same trap on the same day			
Blyton *et al.* (2014)	Brushtail possum	*E. coli* (commensal)	I	Proximity logger & genetic strain								Proximity < 5 m & strain sharing	■	■	
Wilson *et al.* (2014)	Trinidadian guppy	Generic (ex. ectoparasites)	D	Behavioural obs.								≤ 4 body lengths			■
Lavelle *et al.* (2014)	White-tailed deer	Generic (Ex. CWD)	D	Camera, GPS & proximity logger								Photo and proximity (< 1 m); GPS < 20.4 m	■		
Rushmore *et al.* (2014)	Wild chimpanzees	Generic (Ex. Ebola)	D or I	Behavioural obs.					■		■	Proximity (< 5 m) & party level (< 50 m)			
VanderWaal *et al.* (2014a)	Giraffe	*E. coli* (commensal)	I	Behavioural obs. & genetic strain								Same group or shared *E. coli* sub-type	■	■	
VanderWaal *et al.* (2014b)	Kenyan ungulates	*E. coli* (commensal)	I	Behavioural obs. & genetic strain	■	■				■		< 50 m and shared *E. coli* sub-type	■	■	

Chapter 2. Covariation between the physiological and behavioral components of pathogen transmission: Host heterogeneity determines epidemic outcomes

White, L.A., Forester, J. D. and Craft, M. E. (In press). The role of host heterogeneity in determining epidemic outcomes: Covariation between the physiological and behavioral components of transmission. Oikos. doi: 10.1111/oik.04527

2.1 Abstract

Although heterogeneity in contact rate, physiology, and behavioral response to infection have all been empirically demonstrated in host-pathogen systems, little is known about how interactions between individual variation in behavior and physiology scale-up to affect pathogen transmission at a population level. The objective of this study is to evaluate how covariation between the behavioral and physiological components of transmission might affect epidemic outcomes in host populations. We tested the consequences of contact rate covarying with susceptibility, infectiousness, and infection status using an individual-based, dynamic network model where individuals initiate and terminate contacts with conspecifics based on their behavioral predispositions and their infection status. Our results suggest that both heterogeneity in physiology and subsequent covariation of physiology with contact rate could powerfully influence epidemic dynamics. Overall, we found that (1) individual variability in susceptibility and infectiousness can reduce the expected maximum prevalence and increase epidemic

variability; (2) when contact rate and susceptibility or infectiousness negatively covary, it takes substantially longer for epidemics to spread throughout the population, and rates of epidemic spread remained suppressed even for highly transmissible pathogens; and (3) reductions in contact rate resulting from infection-induced behavioral changes can prevent the pathogen from reaching most of the population. These effects were strongest for theoretical pathogens with lower transmissibility and for populations where the observed variation in contact rate was higher, suggesting that such heterogeneity may be most important for less infectious, more chronic diseases in wildlife. Understanding when and how variability in pathogen transmission should be modelled is a crucial next step for disease ecology.

2.2 Introduction

Direct transmission of a pathogen from one host to the next is a complex process that depends on host behavior, host physiology, and the transmission efficiency of the pathogen itself (Begon *et al.* 2002). In natural systems, it has been demonstrated that these interrelated facets of transmission can vary widely between individuals. In fact, empirical studies suggest that unequal contact rates are the rule rather than the exception (Craft & Caillaud 2011b), that contact rates can vary with infection-induced behavioral changes (Croft *et al.* 2011), and that these changes are likely non-uniform across individuals (Lopes et al. 2016). Innate and plastic heterogeneity in susceptibility to infection has been documented for several species (Dwyer, Elkinton & Buonaccorsi 1997; Beldomenico & Begon 2010; Gibson, Jokela & Lively 2016), and variability in infectiousness has also been observed, particularly when concomitant infections are

present (Cattadori, Albert & Boag 2007). Finally, there is evidence that the first individual infected in a population (i.e., the index case) and the relative composition of behavioral phenotypes (e.g., bold vs. shy continuum) can substantially alter how effectively a pathogen spreads within a population (Adelman *et al.* 2015; Keiser *et al.* 2016).

Nevertheless, it is not uncommon for disease models to overlook individual variation in behavior and physiology. This is often done for practical or necessary reasons, but has resulted in a lack of understanding of how these heterogeneities scale up to affect disease dynamics in natural populations (Beldomenico & Begon 2010; Barron *et al.* 2015). The rate at which a pathogen will spread in a host population is a function of the number of infected individuals (I), the number of susceptible individuals (S), and the rate (β) at which infectious individuals successfully transmit the pathogen. The transmission rate (β) encapsulates two separate processes that are required for a successful transmission event: (1) an appropriate contact between a susceptible and infectious individual, and (2) the actual transmission between an infected and susceptible host given contact, which depends upon the physiology of both the host and the pathogen (McCallum *et al.* 2017). See Appendix B for a schematic representing transmission and definitions of key terms. Thus, β can further be broken down into the behavioral (β_c) and physiological (β_p) components of transmission, respectively (Hawley et al., 2011). The apparent simplicity of β as a single parameter may belie non-linearities that can arise at any stage of transmission and affect a pathogen's spread through a population (McCallum *et al.* 2017).

Interactions between the behavioral and physiological components of transmission may arise under a variety of contexts for wildlife (Appendix B), and the effects of these interactions can be exacerbated by host behavior-parasite feedback loops (Ezenwa *et al.* 2016). Covariation between susceptibility and exposure to pathogens in wildlife may be mediated through both the neuroendocrine system and behavioral syndromes (Hawley *et al.* 2011). For instance, in some species, testosterone in males not only increases exposure through agonistic contacts, but can also raise males' susceptibility to parasites (Grear *et al.* 2009). There is also evidence that suites of behaviors (e.g., coping style) can mediate individual exposure risk, and that those same behavioral syndromes are often associated with distinct physiological traits as well (e.g., hypothalamic–pituitary–adrenal (HPA) axis reactivity and stress levels) (Natoli *et al.* 2005; Koolhaas 2008). Finally, covariation between infectiousness and contact rate can arise when pathogens alter host behavior to make it easier for the pathogen to spread between hosts (particularly in trophically transmitted parasites, e.g., Berdoy et al. 2000) or indirectly through sickness behaviors that reduce host activity levels such as fever, lethargy, and limited foraging (Adelman, Moyers & Hawley 2014; Welicky & Sikkel 2015). Alternatively, uninfected individuals may avoid infected conspecifics, or infected individuals may associate less with their counterparts (Croft *et al.* 2011). Modeling studies in humans have begun to incorporate the effects of behavioral changes in response to infectious disease, including adherence to vaccination programs, fear-induced contact reduction, hygiene improvement, or changes in mobility or traveling (Coelho & Codeço 2009; Funk, Salathé & Jansen 2010; Meloni *et al.* 2011), but few investigations

have been made into the effects of behavioral changes on the spread of disease in natural populations. Notably, since physiology can covary with behavior, disease models should ideally incorporate possible interactions between these two components, not just including one component or the other (Hawley *et al.* 2011).

For example, "superspreaders" are individuals that contribute a disproportionately high number of secondary cases either through an unusually high number of contacts or by being especially infectious (Lloyd-Smith *et al.* 2005b). In disease ecology, however, the focus has largely been on variability in contact rate (β_c) (VanderWaal & Ezenwa 2016; White, Forester & Craft 2017c), while individual heterogeneity in physiology, plasticity in behavior in response to infection, and possible covariation between behavioral and immune competence have been somewhat neglected (Barron *et al.* 2015). The role that physiological immunity might play in superspreading has not been fully elucidated (Hawley & Altizer 2011; VanderWaal & Ezenwa 2016), but there is evidence that some individuals are particularly vulnerable to infection (super-receivers) or particularly adept at transmitting the pathogen to others due to high infection load or shedding rates (super-shedders) (Cattadori *et al.* 2007; Zohdy *et al.* 2012).

Without modification, an underlying assumption of traditional, mean-field disease models is that every individual in a population has an equal probability of contacting and infecting any other individual (Anderson & May 1991). These compartmental models can be constructed to reflect different categories of relative risk according to factors like sex or age, which has been done for HIV (Anderson *et al.* 1986; May, Anderson & Irwin 1988). For instance, for the gypsy moth and its nuclear polyhedrosis virus, Dwyer et al.

(1997) incorporated a continuous distribution of susceptible classes and demonstrated a resulting non-linear relationship between virus density and transmission. In some cases, however, it is also important to account for variable contact rates in order to explain superspreading patterns (Lloyd-Smith *et al.* 2005b; Meyers 2007). Network models are a tool that can capture individual variability in the number and rate of contacts (β_c). With a network modeling approach, a contact (i.e., an edge) is any interaction that could allow for transmission of an infectious agent between a pair of individuals (i.e., nodes). In general, network models that account for contact heterogeneity predict less frequent, but more explosive outbreaks than their compartmental model counterparts (Lloyd-Smith *et al.* 2005b).

Many wildlife studies still employ static networks, which do not reflect real-time behavioral shifts or potentially capture changes in the network in response to disease (Masuda & Holme 2013; White *et al.* 2017c). In contrast, dynamic network models describe association patterns in real-time and allow for rewiring events in which individuals can change who they are interacting with at any given time step (Blonder *et al.* 2012). Dynamic networks can be thought of as a continuum between mass-action models, which have high mixing rates, and static network models, which have fixed and prolonged contacts (Volz & Meyers 2007; Bansal *et al.* 2010). However, the implications of using static vs. dynamic contact networks for disease model predictions are still not fully understood, and the tools for dynamic network analysis lag behind their static counterparts (Blonder *et al.* 2012; Masuda & Holme 2013). In a theoretical framework, utilizing a static network for a dynamic system was found to overestimate epidemic

predictions (Fefferman & Ng 2007; Masuda & Holme 2013). Similarly, incorporating dynamic, empirically-based interactions in livestock networks markedly changed predicted epidemic outcomes; Chen et al. (2014) incorporated temporal variability with and without changes in individuals' degree order and observed greater discrepancies in predictions for pathogens with lower values of R_0. Springer et al. (2017) found that incorporating dynamic interactions increased the theoretical transmission of cryptosporidium through wild lemur networks. However, Stehlé et al. (2011) suggested that daily aggregated networks were acceptable proxies for real-time dynamic networks for an SEIR model of conference attendees. As of now, the implications of including dynamic interactions appears to be highly system specific, and there is no clear consensus on when dynamic interactions should be incorporated into disease models (White *et al.* 2017c).

In this manuscript, we employ an individual-based, dynamic network modeling approach because dynamic networks allow us to explicitly incorporate contact heterogeneity, variability in physiology, and behavioral changes resulting from infection. Specifically, we ask: how might possible covariation in the behavioral (β_c) and the physiological (β_p) components of transmission affect epidemic dynamics? We tested scenarios where contact rate covaried with susceptibility, infectiousness, or infection status. This last scenario allowed us to investigate how infection-induced behavioral changes could potentially affect disease dynamics. For a theoretical, directly-transmitted pathogen, we evaluated how these different covariation scenarios might affect epidemic variability in the forms of: maximum prevalence reached, the time it took to reach

maximum prevalence, the realized transmission rate, and the likelihood of epidemic fade-out. By fade-out, we are referring to simulations where the pathogen never spreads beyond the initially infected individual. We conducted a random forest analysis to identify key factors that were most likely to explain these metrics. While previous contact network studies have identified the importance of contact heterogeneity within a population (Lloyd-Smith *et al.* 2005b), our results suggest that both heterogeneity in physiology and subsequent covariation of physiology with contact rate could powerfully influence epidemic dynamics.

2.3 Methods

We developed an individual-based, dynamic network model that explores how heterogeneity in individual contact behavior, susceptibility, and infectiousness can interact to affect pathogen transmission. We employed a Susceptible-Infected (SI) model to describe the spread of a pathogen through a closed population, assuming no births, deaths, or disease-related mortality (Anderson & May 1991). We used a factorial design to explore the effects of epidemiological parameters on epidemic outcomes and measured the maximum prevalence reached during 750 time steps, the number of time steps it took to reach that maximum prevalence, and the rate of epidemic spread. Simulations were conducted for a population size of 525 individuals with 100 repetitions per parameter set (Table 2.1).

Dynamic network framework

At each time step during the simulation, individuals (nodes) could form or remove contacts (edges) with conspecifics based on their intrinsic individual behavioral phenotype (i.e., contact rate, β_c). This dynamic network behavior relies on a discrete time, separable temporal exponential-family random graph model (STERGM) framework, which allows for biologically realistic variation in mean degree, duration of contacts, and disease-induced behavioral changes (Krivitsky & Handcock 2014). These models are built on an exponential random graph (ERGM) framework; ERGMs are a family of statistical models that describe random graphs (i.e., random networks) based on their underlying node attributes such as degree, betweenness, transitivity, etc. (Robins *et al.* 2007). A random graph *Y* consists of nodes, *n*, and edges, *m*, with state space: $\{Y_{ij}: i = 1 \dots, n; j = 1, \dots, n\}$. $Y_{ij} = 1$ if an edge exists between nodes *i* and *j,* and $Y_{ij} = 0$ otherwise. The basic form of an ERGM is: $P(Y = y) = \frac{\exp(\theta' g(y))}{k(y)}$, which describes the probability of observing a given network, *y*, given the space of all possible networks, *Y,* that could exist for a given set of nodes. The numerator contains both a set of model statistics $g(y)$ and coefficients corresponding to those statistics, θ. The denominator, $k(y)$, represents the sum of the numerator across all possible networks (Krivitsky & Handcock 2014). STERGMs extend into discrete time by utilizing two independent ERGMs: a formation and dissolution model. STERGMs employ the Markov assumption that the state of a network at the current time step is memoryless—so the formation and dissolution of edges is only dependent upon the current state of the network (Hanneke, Fu

& Xing 2010). We assume the simplest case for the dissolution model—that all edges have the same probability of dissolving (i.e., a Bernoulli process). For all simulations, we assumed a constant edge dissolution probability of 25 time steps (Table 2.1).

Models were constructed in R (version 3.3.2) using self-written modules in the *EpiModel* package (version 3.4.0) (Jenness et al. 2016a). The *EpiModel* package (http://www.epimodel.org/) provides a suite of pre-written and modifiable functions for simulating infectious disease dynamics, including stochastic network models that rely on temporal ERGMs from the *statnat* package. The *EpiModel* package has been used to investigate complex disease dynamics and interventions for diseases like HIV (Jenness et al. 2016b). Fully annotated sample code is provided in Appendix C, and all code and simulation data for the manuscript are available from the Dryad Digital Repository: http://dx.doi.org/... (White et al. n.d.).

Covariation: Incorporating β_c and β_p

We considered three mechanisms by which the physiological components of transmission (β_p) and contact rate (β_c) may covary: (1) susceptibility vs. contact rate; (2) infectiousness vs. contact rate; and (3) infection status vs. contact rate (Appendix B Figure B.2). For each scenario, we tested a control scenario where individuals exhibited no variation in physiology (β_p) but heterogeneity in contact rate (β_c), a null scenario where individuals exhibited heterogeneity in physiology (β_p) but no heterogeneity in contact rate (β_c), a positive covariation scenario where physiology (β_p) positively covaries with contact rate (β_c) (Appendix B Figure B.2*A*, blue line), and a negative covariation

scenario where physiology (β_p) negatively covaries with contact rate (β_c) (Appendix B Figure B.2*A*, red line).

At the start of each simulation, every individual was assigned an intrinsic contact rate (β_c) and physiological state (β_p)—either susceptibility (s) or infectiousness (κ) depending on the experiment. The behavioral component of transmission (β_c) was thus incorporated implicitly into the transmission process by determining which hosts are contacting one another at any given time step based on the dynamic network simulation. For a given set of conditions, the population was divided equally into thirds (175 individuals per sub-group) with each group assigned a higher-than-average ("high"), an average ("medium"), or a lower-than-average ("low") number of contacts (Appendix B Figure B.2*B*; Table 2.2). These behavioral phenotypes can be thought of as corresponding roughly to spectrums of individual personality (e.g., shy vs. bold) that might dictate social behavior. Empirical studies in wildlife have cited mean degrees ranging from less than one to approximately eight (Godfrey *et al.* 2009; Perkins *et al.* 2009; Hirsch *et al.* 2013). We simulated a mean degree of 4, which appears to be a reasonable approximation for social animals like macaques and prairie dogs (MacIntosh *et al.* 2012; Verdolin, Traud & Dunn 2014). For susceptibility vs. contact rate and infectiousness vs. contact rate, individuals with higher-than-average or lower-than-average contact rates had an absolute difference in mean degree of either 2 or 4. So, for example, simulations with a "low" separation of mean degree would have three separate groups with mean degrees of 2, 4, and 6 (e.g., Appendix B Figure B.2*B*), and those with a "high" separation would have three separate groups with mean degrees of 0, 4, and 8 (Table 2.1). In terms of

simulating the STERGM, the only network statistic, $g(y)$, included is mean degree and the coefficients are $\theta = [2\ 4\ 6]$ or $\theta = [0\ 4\ 8]$ for low and high variation in contact rate, respectively (Table 2.2).

We incorporate the physiological component of transmission, β_p, explicitly into the final probability of transmission given contact (i.e., the existence of an edge in the dynamic network). Depending on the experiment, β_p is represented either through susceptibility of the susceptible host, s, or infectiousness of the infected host, κ. To induce covariation, individuals were assigned physiological states (β_p) corresponding to their contact rates. For these physiological states, individuals were assigned a "low," "medium," or "high" value (0, 1, or 2, respectively) for their susceptibility (s) or infectiousness (κ)—such that the average susceptibility or infectiousness in the population would always be approximately equal to 1 (Table 2.2).

The mechanism of transmission

The possibility of transmission was evaluated at each time step if (1) two nodes shared an edge, and (2) one node was infected and one node was susceptible. The final transmission probability, $\mathbb{P}(T)$, that we used for this model is based on the intuition involved in the Reed-Frost or chain binomial models which estimate the likelihood that an individual "escapes" infection during a discrete time step (Kyvsgaard, Johansen & Carabin 2007). Instead of calculating the likelihood of an individual escaping infection from multiple infectious hosts in the population, we allow for the possibility that during a time step, multiple opportunities for transmission could occur when a susceptible and

infectious host share an edge in the dynamic network. This might correspond to discrete events like bites, coughing, sneezing, vector transfer, etc. The resulting final transmission probability is: $\mathbb{P}(T) = 1 - (1 - \tau)^{\alpha}$ where τ represents the transmission efficiency per individual interaction, and the action rate, α, represents the potential number of infectious interactions that could occur via an edge per time step. While the transmission efficiency likely represents a complex relationship between pathogen physiology and host immunocompetence, we use it here to represent the idea that, all else being equal, certain pathogens are more infectious than others on average (Appendix B). We vary transmission efficiency, τ, in our factorial design (Table 2.1) and discuss the modifications for the final transmission probability for each specific experimental scenario below.

Susceptibility vs. contact rate

For this mechanism, individuals varied in β_p via their susceptibility (*s*, likelihood of being infected given contact with an infected conspecific). A successful transmission event was dependent upon the innate susceptibility (*s*) of the susceptible contact in the susceptible-infected dyad such that the final transmission probability, $\mathbb{P}(T)$, took the form: $\mathbb{P}(T) = 1 - (1 - \min\{1, \tau \cdot s\})^{\alpha}$

Here, action rate (α) is defined as the number of possible transmission events per time step. In this scenario, we assume the action rate to be equal to one per time step for each susceptible-infectious interaction, so the final transmission probability simplifies to $\mathbb{P}(T) = \min\{1, \tau \cdot s\}$. At time step t=1, one individual was randomly selected to be the

first infected individual (i.e., the index case). If the first randomly selected individual had a susceptibility of zero (s=0), the pathogen could not propagate further.

Infectiousness vs. contact rate

For this mechanism, individuals varied in β_p via their infectiousness (κ, likelihood of successfully transmitting the pathogen given contact with an uninfected conspecific). In this model, the probability of successful transmission, $\mathbb{P}(T)$, to a susceptible individual given contact with an infectious individual was proportional to the infectiousness of the infected contact:

$$\mathbb{P}(T) = 1 - (1 - \tau)^{\alpha * \kappa}$$

In this case, infectiousness (κ) was modelled as affecting the action rate (α), which could be interpreted as the pathogen load or the amount of shedding by an infectious host per time step. At time step t=1, one individual was randomly selected to be the first infected individual (i.e., the index case). If the first randomly selected individual had an infectiousness of zero (κ = 0), the pathogen could not propagate further.

Infection status vs. contact rate: Disease-induced behavioral changes

The objective of this scenario was to test the possible effects of sickness-induced behavioral changes. For example, a very sick individual that is highly infectious might increase their contact rate (e.g., furious rabies) or decrease their contact rate because of fever, lethargy, or anorexia (Adelman *et al.* 2014; Welicky & Sikkel 2015). To consider the possibility that the magnitude of behavioral change is correlated with infectiousness (e.g., individuals with a higher pathogen load might display more extreme sickness

behaviors), we allow individuals to become either "highly infectious" or "less infectious" post-exposure with a corresponding change in contact rate (β_c) depending on the type of covariation (Table 2.3). It is worth noting that in the above scenarios (Sections 0 and 0), it is possible for a secondary correlation to result between contact rate and infection status. For example, in the positive covariation scenario for susceptibility vs. contact rate, we would expect highly susceptible individuals (who also have higher contact rates) to become infected first. This experiment differs from the previous two in that contact rate is allowed to change explicitly as a result of infection status.

To begin, we modelled a control case where no changes in contact rate (β_c) occurred post-infection and individual infectiousness was homogenous throughout the population ($\kappa = 1$ for all individuals). For the null case, there was no change in contact rate (β_c) upon infection, but individuals had heterogeneity in infectiousness (individuals were randomly assigned an infectiousness of κ =1 or 2 upon infection). For positive and negative covariation, an individual's contact rate increased or decreased upon infection respectively, and after a successful exposure, individuals had an equal likelihood of becoming highly infectious ($\kappa = 2$) or less infectious ($\kappa = 1$) (). Unlike the first two scenarios tested (0 & 0), each simulation began at t=1 with two infected individuals. For the null, positive, and negative covariation cases, these consisted of one highly infectious individual ($\kappa = 2$) and one less infectious individual ($\kappa = 1$); it was necessary to include both classes of infected individuals at the start of the simulation for the purposes of calibrating the dynamic network. All remaining susceptible individuals started with a mean degree of four. For positive covariation, less infectious individuals increased their

expected mean degree to 6, and highly infectious individuals increased their expected mean degree to 8. Likewise, for negative covariation, less infectious individuals decreased their expected mean degree to 2, and highly infectious individuals decreased their expected mean degree to 0 (Table 2.3). In the *EpiModel* package, this was achieved by using infection status itself as a network statistic via the "nodefactor" term for simulating the dynamic network (Jenness, Goodreau & Morris 2016a). This term of the model allows different sub-groups of the population to have heterogeneity in their attributes—in this case, mean degree (Jenness *et al.* 2016a). However, infected individuals were not any more likely to form edges with susceptible conspecifics than infected conspecifics, so there was no preferential mixing as a result of infection status. A necessary consequence of including infection status as a factor governing edge formation was that network density either increased (positive covariation) or decreased (negative covariation) over time.

We tested two forms of infectiousness: (1) the form described in Section 0 where infectiousness influences the action rate in the exponent of the final transmission probability $\mathbb{P}(T) = 1 - (1 - \tau)^{\alpha \cdot \kappa}$; and (2) a form where infectiousness (κ) directly modifies the probability of infection (as in Section 0), so that the final transmission probability was equal to: $\mathbb{P}(T) = 1 - (1 - \min\{1, \tau \cdot \kappa\})^{\alpha}$, which simplifies to $\mathbb{P}(T) = \min\{1, \tau \cdot \kappa\}$ when the action rate, $\alpha = 1$.

Metrics and nonlinear least square regression

We included four metrics to investigate differences in epidemic outcomes across experiments and covariation types. First, we measured the maximum prevalence reached

in 750 time steps. Because this is an SI model, the mean maximum prevalence reflects both the maximum prevalence reached by successful simulation runs and the percentage of epidemics fading-out. We also explicitly measured the time it took to reach maximum prevalence and the percentage of simulation runs fading-out for each treatment. Finally, since the contact structure of different experimental set-ups (particularly those with higher variation in contact rate) could limit the proportion of the population eligible to be infected, we measured a "realized" transmission rate (β) to estimate the rates of epidemic spread in each population. To do this we used nonlinear least square regression implemented through the *nlsLM* function in the *minpack.lm* package (version 1.2-1) in R (Elzhov 2016). We fit each individual epidemic simulation to the logistic growth equation: $I(t) = \frac{K}{1+b \cdot e^{-\beta K t}}$ where $I(t)$ is the number of infected individuals at time *t*; *K* is the carrying capacity; β is the realized transmission rate, and *b* is a scaled parameter equal to $\frac{K-I_0}{I_0}$ where I_0 is the initial population size at time zero (Derivation in Appendix D). We assigned values of I_0 appropriate to each simulation ($I_0 = 1$ for susceptibility or infectiousness vs. contact rate and $I_0 = 2$ for infection status vs. contact rate), and we allowed both *K* and β to vary to determine the best fit, although *K* was not allowed to exceed the total number of individuals in the simulation.

Random forest analysis

In simulation studies, significance testing can be less useful because an essentially unlimited sample size can result in labeling even small differences in the magnitude of outcomes as statistically significant (White *et al.* 2014). To further a descriptive approach

to the analysis of our simulation results, we have used random forest analysis—a machine learning method that can handle complex, non-linear relationships between model inputs and outputs, as well as potential collinearity between covariates (Cutler *et al.* 2007). Random forest analysis is a recursive partitioning method that combines the predictions from numerous fittings of classification trees to the same set of data (Breiman 2001; Cutler *et al.* 2007). Variable importance measures resulting from these analyses can be used to estimate the relative importance of a covariate in determining model outcomes, and unlike most univariate methods, can account for possible correlations between inputs.

To calculate variable importance, we employed a measure of permutation importance which has been demonstrated to be more robust than node impurity (Strobl *et al.* 2007; Strobl, Hothorn & Zeileis 2009). Using the *cforest* function in the *party* package in R (version 3.3.2), we simulated 10,000 trees per analysis to ensure that the order of variable importance was robust to changes in the random seed, and we calculated mean decrease in accuracy variable importance scores using the *varimp* function in the *party* package (Hothorn *et al.* 2006; Strobl *et al.* 2007, 2009). The mean decrease in accuracy describes the loss of predictive value that results from a particular variable being randomly permutated. Stated another way, higher mean decrease in accuracy scores indicate a greater importance in model prediction. For susceptibility vs. contact rate and infectiousness vs. contact rate, we included the following as covariates: transmission efficiency, separation between mean degree, type of covariation, and physiological phenotype of the index case. For infection status vs. contact rate, we included transmission efficiency, type of covariation, and form of infectiousness (i.e., in the

exponent or the product of the final transmission probability). The response variables for all three mechanisms were: maximum prevalence, time until maximum prevalence, and the realized β.

Data deposition

Data available from the Dryad Digital repository: https://doi.org/10.5061/dryad.8t201 (White, Forester & Craft 2017b).

2.4 Results

Susceptibility vs. contact rate

Allowing for variability in susceptibility of the host population (null case) reduced the maximum prevalence reached during the 750-step simulation (compared to the control case) and increased the variability of observed epidemic outcomes with at least one-quarter of epidemics fading out (Figure 2.1 & Figure 2.2). This finding was consistent across differences in mean degree and for different transmission efficiencies (Appendix E Figure E.1 and E.2). In general, for simulations with higher variation in contact rate (i.e., difference in mean degree of 4), the maximum prevalence was lower relative to corresponding simulations with smaller variations in contact rate (i.e., difference in mean degree of 2). This finding reflects the fact that one-third of the population is expected to be isolated (β_c=0) for networks constructed with higher variation in contact rate (i.e., mean degree difference = 4). Notably, positive covariation counteracted this observed difference in maximum prevalence between the control case and other covariation types, and this effect was consistent across infection probabilities (Figure 2.2A). In the case of

negative covariation, there was an observable increase in the time it took to reach maximum prevalence relative to the control, null, and positive covariation scenarios; this increase was the greatest for lower transmission efficiencies and higher variation in contact rate (Figure 2.2B). In general, the epidemics spread more quickly with higher transmission efficiency, regardless of variation in contact rate. The differences in the realized β between positive and negative covariation were largest for higher values of transmission efficiency and high contact rate variability (Figure 2.2C). Negative covariation continued to substantially suppress the realized β even at higher values of transmission efficiency.

Infectiousness vs. contact rate

Variability in infectiousness (null case) increased variability in epidemic outcomes (Figure 2.3); simulations experienced fade-out because of those individuals in the population with an infectiousness of zero ($\kappa=0$). These observations were consistent across simulated differences in mean degree and infection probabilities (Appendix E Figure E.3 & Figure E.4). As with susceptibility vs. contact rate, a larger simulated variation in contact rate within the population also decreased the maximum prevalence, even in the control case (Figure 2.4A); this was the result of a contact structure where one-third of the population was socially isolated ($\beta_c=0$). For negative covariation, there was a substantial increase in the time it took to reach maximum prevalence relative to the control, null, and positive covariation scenarios; this effect was most pronounced for lower transmission efficiencies and higher variation in contact rate (Figure 2.4B). Similar

to the results for susceptibility vs. contact rate, a faster rate of epidemic spread occurred for simulations with higher transmission efficiency regardless of variation in contact rate, and the difference in magnitude of the realized β between positive and negative covariation was largest for higher values of transmission efficiency and contact rate variability (Figure 2.4C).

Infection status vs. contact rate

For infection status vs. contact rate, we tested two ways that infectiousness might play into the final transmission probability (see Section 0), but results were consistent across these two different formulations. As with susceptibility vs. contact rate and infectiousness vs. contact rate, simply including heterogeneity in physiology increased variability in epidemic outcome (Figure 2.5; compare control vs. null cases). Reduction in contact rate upon infection (negative covariation) drastically reduced the maximum prevalence reached within 750 time steps (Figure 2.5 & Figure 2.6A), while increasing contact rate upon infection (positive covariation) had a comparatively minimal effect on increasing the maximum prevalence relative to the null and control cases (Figure 2.5 & Figure 2.6A). In general, the differences in the time it took to reach maximum prevalence for positive, negative, and null covariation were largest for lower transmission efficiencies (Figure 2.6B). This is likely because the control, null and positive covariation cases all saturated very quickly at higher transmission efficiencies (Appendix E Figure E.5 and E.6). Consistent with susceptibility vs. contact rate and infectiousness vs. contact rate, the realized transmission rate (β) was highest for higher values of transmission

efficiency, and high contact variability revealed the sharpest differences between all four types of covariation (Figure 2.6C).

Random forest results

Variable importance scores for maximum prevalence indicate that the physiological phenotype of the index case had the highest importance for susceptibility vs. contact rate and infectiousness vs. contact rate; this was followed in importance by separation in mean degree, and then type of covariation (Table 2.4). Transmission efficiency had a negligible mean decrease in accuracy for both mechanisms in predicting maximum prevalence. For infection status vs. contact rate, the type of covariation had the highest variable importance score, followed by transmission efficiency.

For time until maximum prevalence, index case also had the highest importance for susceptibility vs. contact rate and infectiousness vs. contact rate. In order of decreasing score, this was followed by transmission efficiency, type of covariation, and degree of separation. Type of covariation was most important for predicting time until maximum prevalence for infection status vs. contact rate (Table 2.4).

For the realized transmission rate (β), pathogen transmission efficiency was an informative predictor for all three experiments (Table 2.4). For both susceptibility vs. contact rate and infectiousness vs. contact rate, physiology of the index case had the highest variable importance score, but this score was of similar order of magnitude to pathogen transmission efficiency; variation in contact rate had a negligible variable importance score (two orders of magnitude lower) for both mechanisms. For infection status vs. contact rate, the transmission efficiency had the highest ranking variable

importance score, which was of similar order of magnitude to covariation type. The form of infectiousness (either in the exponent or the product of the final transmission probability) had a negligible effect in predicting all three response variables for infection status vs. contact rate.

2.5 Discussion

Accounting for contact heterogeneity has been shown to dramatically alter disease predictions (Keeling & Eames 2005); however, our results support the idea that both heterogeneity in physiology and subsequent covariation of physiology with contact rate could also powerfully influence epidemic dynamics. Overall, we found that (1) individual variability in susceptibility or infectiousness, which is typically unaccounted for in wildlife disease models, can both increase epidemic variability and the likelihood of disease fade-out; (2) when contact rate and susceptibility or infectiousness negatively covary, it takes longer for epidemics to spread throughout the population, and the rate of epidemic spread is reduced even for highly transmissible pathogens; and (3) reductions in contact rate resulting from infection-induced behavioral changes can prevent the pathogen from reaching most of the population and can dramatically limit the rate of epidemic spread, even for pathogens with high transmissibility.

Our results demonstrated that simply allowing for heterogeneity in susceptibility or infectiousness without any kind of covariation could increase variability of epidemic outcomes. An increase in the variability of epidemic outcomes (i.e., successful invasion of the population vs. fade-out) will have important implications for disease predictions, control, and interventions.

The random forest analysis highlighted the potential importance of physiological phenotype of the index case in explaining much of the observed variation in epidemic outcome for susceptibility vs. contact rate and infectiousness vs. contact rate. Much of this predictive power is likely a function of how the model structured, where roughly one-third of the population is not susceptible ($s = 0$) or not infectious ($\kappa = 0$). While such extreme physiological phenotypes might be less common in natural populations, this theoretical finding does support the results of recent empirical work where the index case and group composition of phenotype played important roles in epidemic outcomes (Adelman *et al.* 2015; Keiser *et al.* 2016). Across the three different mechanisms, negative covariation decreased maximum prevalence, increased time to reach maximum prevalence, and dampened the rate at which the disease spread through the population relative to all other types of covariation. Universally, differences between types of covariation were strongest for theoretical pathogens with lower transmission efficiency, which suggests that such heterogeneity may be most important for less infectious, more chronic diseases in wildlife such as bovine tuberculosis (Cosgrove *et al.* 2012). This finding is consistent with studies using empirically informed networks that have found dynamic interactions to be more important at lower transmissibility (Chen *et al.* 2014; Springer *et al.* 2017). Additionally, differences in the time it took to reach maximum prevalence for different types of covariation were most pronounced for simulations with higher variation in contact rate. In general, simulations with higher contact variation had higher rates of epidemic spread—with the key exception of negative covariation where

the realized transmission rate stayed roughly constant even at high values of pathogen transmission efficiency (Figure 2.2C, Figure 2.4C &Figure 2.6C).

Trends for time until maximum prevalence and the intrinsic rate of increase were consistent for susceptibility vs. contact rate and infectiousness vs. contact rate. Across parameter sets, infectiousness vs. contact rate simulations reached a higher maximum prevalence—a result of infectiousness affecting the action rate rather than transmission efficiency in the final transmission probability. The transmission efficiency positively correlated with the realized transmission rate (β) for all three experiments (Figure 2.2C, Figure 2.4C & Figure 2.6C), but overall, played a negligible role in explaining maximum prevalence, especially for susceptibility vs. contact rate and infectiousness vs. contact rate. For infection status vs. contact rate, negative covariation (i.e., decreased contact rate upon infection) dramatically reduced the maximum prevalence reached within 750 time steps relative to the other two experiments, especially for lower values of transmission efficiency. Negative covariation also increased the time it took to reach maximum prevalence for all values of transmission efficiency and decreased the rate of epidemic spread. These findings were consistent across the two different formulations of final transmission probability that were simulated. While intuitive, these results are important because reduction of activity and contact rate because of infection are well-documented (Croft *et al.* 2011; Welicky & Sikkel 2015; Lopes *et al.* 2016), but less commonly incorporated into disease models.

For simplicity of analyzing a complex model, we assumed a constant population size—no births or natural or disease-induced mortality. To limit the number of

epidemiological parameters, we also made the simplifying assumption of an SI model rather than a more complicated SIR or SEIR disease model. Another key assumption of our models was the discrete physiological (β_p) vs. social (β_c) states that were assigned to each individual. Because the formation model of STERGMs consists of a discrete set of covariates, we had to individually assign nodes to distinct behavioral phenotypes (e.g., low, medium, and high contact rates). This feature of STERGMs prevented us from testing a continuous covariation that might be more reasonable in empirical populations. Future studies could test different continuous distributions of susceptibility and infectiousness or add more discrete levels of contact rate within the population. While populations in natural settings are unlikely to replicate the exact contact structure that we employed here, it is not uncommon to for a small proportion of the population be responsible for the majority of contacts. For instance, superspreaders generally represent a much smaller proportion of the population and the resulting contact distribution is usually skewed (Clay *et al.* 2009). This is sometimes referred to as the '20/80' rule, where 20% of the individuals are responsible for 80% of the contacts (Woolhouse *et al.* 1997).

More work needs to be done to characterize the effects of static network approximations on disease modelling predictions, since our work suggests that disease-induced behavioral changes (which are not likely to be adequately captured through static network approximations) could have a substantial effect on the likelihood of successful pathogen invasion. While STERGMs are well suited to calibration with empirical data (Jenness *et al.* 2016b), wildlife host-pathogen systems with existing dynamic contact

network and individual physiological data are rare (Craft & Caillaud 2011b; White *et al.* 2017c). Another consideration for future studies is the clumping of contacts in time (known as bursts) in empirical systems. STERGM models do not necessarily capture temporal clumping because they assume an exponential probability for dissolution rate of edges (Masuda & Holme 2013). In addition, more work needs to be done to characterize differences in physiology in wild populations that result from innate genetic differences and plastic responses to infection, particularly since wild populations are often more heterogeneous and likely to experience more heterogeneous environments than those studied in labs (Dwyer *et al.* 1997). For instance, Beldomenico and Begon (2010) highlighted how natural populations may also experience additional interactions between resource availability, host density, and body condition, which can mediate host susceptibility.

Collaboration between the fields of disease ecology and ecoimmunology will likely yield more empirical study systems in which these ideas can be tested (Adelman *et al.* 2014). In particular, improvements in radiotelemetry, radio-frequency identification (RFID), and temperature sensing passive integrated transponder (PIT) tags may allow for concrete steps forward in the simultaneous collection of contact and sickness behavior (Adelman *et al.* 2014). The type of dynamic network modelling presented here could be used to explicitly investigate ratios and index cases of behavioral and physiological phenotypes in closed populations (Keiser *et al.* 2016).

Host heterogeneity in contact rate and physiology and potential covariations between these two components exist in a myriad of real life systems (Hawley *et al.* 2011;

VanderWaal & Ezenwa 2016). However, there is no consistent framework that outlines when individual heterogeneity in pathogen transmission is important and when it is necessary to account for those differences in sampling or interventions, even though allowing for such differences can markedly change predictions of an epidemic's duration and behavior (Keeling & Eames 2005; Meyers 2007). By including the heterogeneity of hosts, populations, or resources in modeling approaches, disease ecologists may develop targeted control measures that could increase the benefit-cost ratio of management strategies (Eisinger & Thulke 2008). This may occur through targeted monitoring or interventions (including vaccination, culling, treatment, etc.) on high-risk individuals, sub-populations, or spatial hot-spots that act as "hubs" for the population (Haydon *et al.* 2006). The caveat for such strategies is that the cost of identifying "super" individuals must be less than the uniform administration of an intervention (Paull *et al.* 2012). Given the time and resource-intensive nature of gathering pathogen data in wildlife populations, improved models will provide insight to the amount of research effort necessary to better capture the transmission process (Krause et al., 2013; Tompkins et al., 2011).

Understanding how and when variability in pathogen transmission should be modelled is a crucial next step for the field of disease ecology and is a critical refinement for future modeling strategies. Through an iterative approach to empirical experiments and modeling (Restif *et al.* 2012), and additional collaboration between the fields of animal behavior, ecoimmunology, and disease ecology, we can improve disease modeling predictions to account for heterogeneity in contact rate and host physiology, as well as the potential feedbacks between these critical facets of pathogen transmission.

2.6 Conclusions

These results highlight the importance of heterogeneity in physiology and the potential role that covariation between the behavioral and physiological components of pathogen transmission could play in epidemic outcomes. Simply allowing for variability in host physiology without instituting any type of covariation fostered increased epidemic variability. Random forest analysis supported the idea that much of this variation could be attributed to the physiological phenotype of the index case for susceptibility vs. contact rate and infectiousness vs. contact rate, which was not surprising, given the extreme physiological phenotypes ($s = \kappa = 0$) present in the population that contributed to fade-out events. The observed differences between different types of covariation were strongest for low transmission efficiencies and for larger variation in contact rate, with negative covariation increasing the time until maximum prevalence across mechanisms tested. This suggests that accounting for such heterogeneity may be most important for less infectious, chronic wildlife diseases and for populations that exhibit more heterogeneous contact structure. For infection status vs. contact rate, negative covariation dramatically decreased the maximum prevalence reached during the duration of the simulation, and this finding was robust to the formulation of final transmission probability. Accounting for covariation in behavior and physiology may be important for future wildlife disease models and disease modelling more broadly. More empirical and modelling work should be performed to determine the circumstances and methods for best capturing heterogeneity in pathogen transmission.

2.7 Acknowledgements

L.A.W. was funded by the National Science Foundation (GRFP-00039202 and DEB-1701069). M.E.C. was funded by National Science Foundation (DEB-1413925 and DEB-1654609) and the University of Minnesota's Office of the Vice President for Research and Academic Health Center Seed Grant. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: http://www.msi.umn.edu

Supplementary material (available online as Appendix oik-04527 at <www.oikosjournal.org/appendix/oik-04527>). Appendix 1-4.

2.8 Tables

Table 2.1. Variables and parameters used in models.

Parameter	**Levels**	**Values**
Transmission efficiency (τ)	Low, medium, high	0.025, 0.25, 0.5
Total separation between mean degree (β_c)	Low, high	2, 4, 6 (± 2); 0, 4, 8 (± 4)
Dissolution rate of edges	Constant	25 time steps
Population size	Constant	525 individuals
Total density of network/edges	Constant	Expected mean degree = 4
Duration of simulation	Constant	750 time steps
Number of simulations per parameter set	Constant	100

Table 2.2. Experimental design for Sections 2.3: Susceptibility vs. contact rate and infectiousness vs. contact rate.

Type of covariation	**Number of individuals in sub-group**	**Mean degree for "low" contact variability treatment (β_c)**	**Mean degree for "high" contact variability treatment (β_c)**	**β_p (susceptibility, s, or infectiousness, κ)**
Control	175	2	0	1
	175	4	4	1
	175	6	8	1
Null	175	4	4	$unif\{0,2\}$
	175	4	4	$unif\{0,2\}$
	175	4	4	$unif\{0,2\}$
Positive	175	2	0	0
	175	4	4	1
	175	6	8	2
Negative	175	2	0	2
	175	4	4	1
	175	6	8	0

Table 2.3. Experimental design for Section 2.3: Infection status vs. contact rate.

Type of covariation	**Mean degree pre-exposure (β_c)**	**Mean degree post-infection (β_c)**	**β_p (infectiousness, κ)- post-infection**	**Percent of individuals (post-exposure) (%)**
Control	4	4	1 (low infectious)	100
Null	4	4	1 (low infectious)	50
	4	4	2 (high infectious)	50
Positive	4	6	1 (low infectious)	50
	4	8	2 (high infectious)	50
Negative	4	2	1 (low infectious)	50
	4	0	2 (high infectious)	50

Table 2.4. Variable importance results from random forest analysis. Reported as mean decrease in accuracy scores from random forest analysis rounded to four significant figures. Higher values indicate a higher variable importance and corresponding predictive power.

Model outcome	**Variable**	**Susceptibility vs. contact rate**	**Infectiousness vs. contact rate**	**Infection status vs. contact rate**
Maximum prevalence	Variation in contact rate (β_c, separation in mean degree)	0.07515	0.1252	--
	Covariation	0.04213	0.02488	0.1060
	Transmission efficiency (τ)	-0.0001008	0.0004898	0.06941
	Physiology of the index case (β_p: *s* or κ)	0.1390	0.2591	--

	Form of infectiousness (exponent or product)	--	--	-0.0002619
Time until maximum prevalence	Variation in contact rate (β_c, separation in mean degree)	5615	7261	--
	Covariation	8652	8439	79,560
	Transmission efficiency (τ)	10,180	10,380	15,940
	Physiology of index case (β_p: *s* or κ)	12,740	14,900	--
	Form of infectiousness (exponent or product)	--	--	-7.755
Realized beta (*β*)	Variation in contact rate (β_c, separation in mean degree)	9.339e-08	1.570e-07	--
	Covariation	7.404e-07	5.623e-07	5.0133e-07
	Transmission efficiency (τ)	1.104e-06	6.402e-07	6.043e-07
	Physiology of index case (β_p: *s* or κ)	1.139e-06	6.497e-07	--
	Form of infectiousness (exponent or product)	--	--	5.436e-09

2.9 Figures

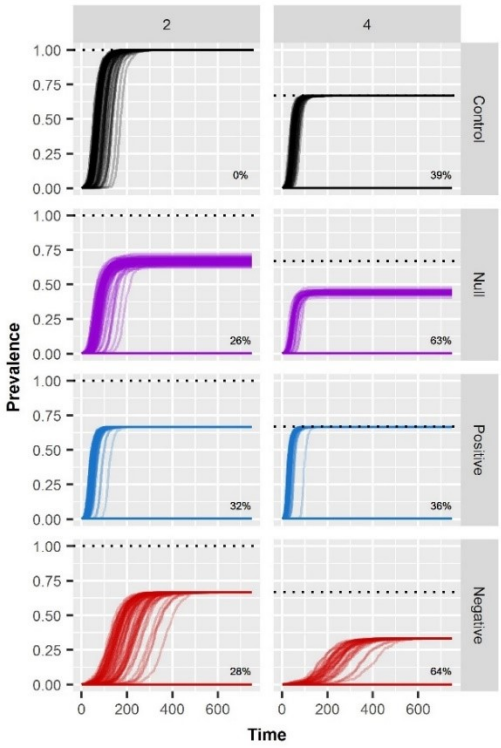
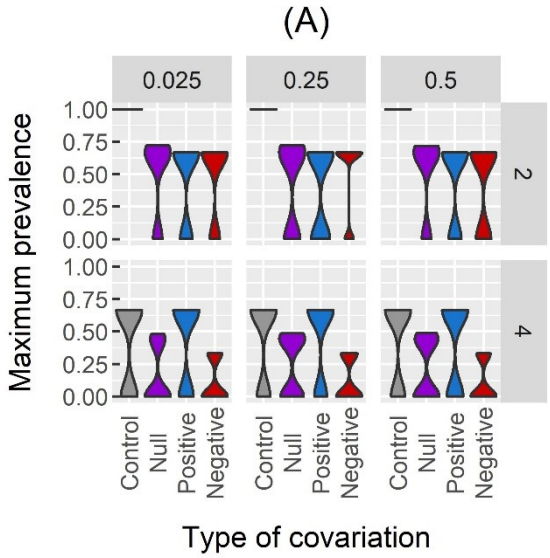
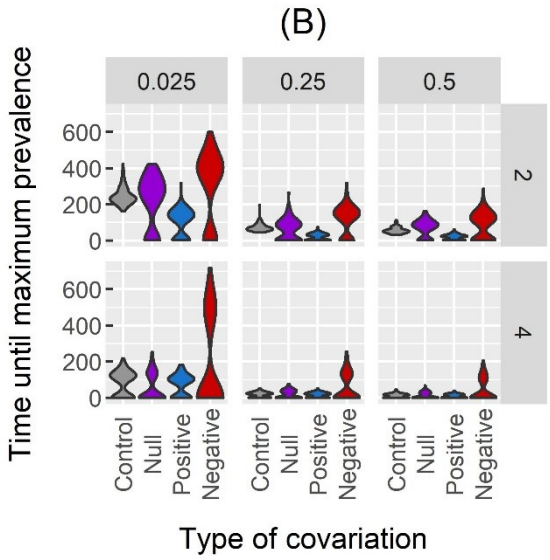


Figure 2.1. Time course of simulated epidemics for susceptibility vs. contact rate for the lowest transmission efficiency tested of τ=0.025. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in susceptibility, no covariation), null (variability in susceptibility, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent, and the numbers in the lower righthand corner of each panel describe the percentage of simulations fading-out for each treatment. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population as a β_c=0, limiting maximum prevalence to 0.66. Time courses for the corresponding medium (τ = 0.25) and high (τ = 0.5) transmission efficiencies are available in Appendix E Figures E.1 & E.2


(A)
0.025
0.25
0.5
2
4
1.00
0.75
0.50
0.25
0.00
Maximum prevalence
Control
Null
Positive
Negative
Type of covariation


(B)
0.025
0.25
0.5
2
4
600
400
200
0
Time until maximum prevalence
Control
Null
Positive
Negative
Type of covariation

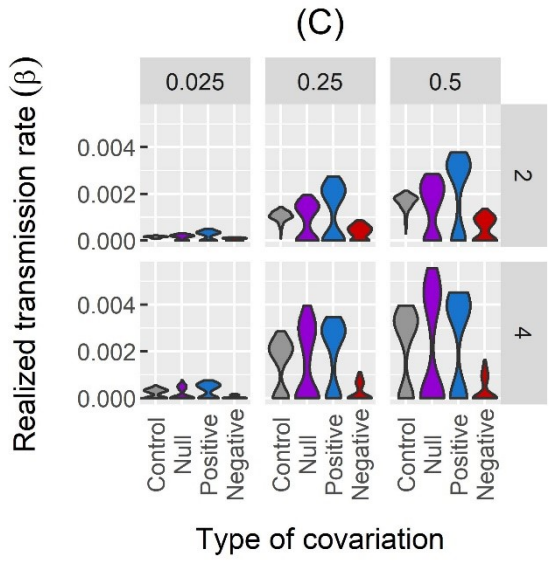

(C)
0.025
0.25
0.5
2
4
0.004
0.002
0.000
Realized transmission rate (β)
Control
Null
Positive
Negative
Type of covariation

Figure 2.2. For susceptibility vs. contact rate, violin plots depicting for all covariation types: **(A)** the maximum prevalence reached in 750 time steps; **(B)** the time it takes to reach that maximum prevalence; and **(C)** the realized transmission rate (β), which describes the rate of epidemic spread. The columns correspond to the transmission efficiency (i.e., 0.025, 0.25, and 0.5), and the rows correspond to the difference in mean degree (i.e., 2 or 4).

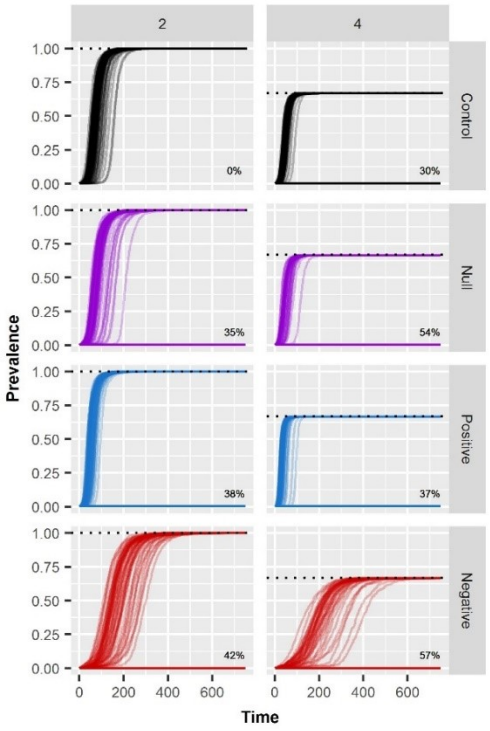


Figure 2.3. Time course of simulated epidemics for infectiousness vs. contact rate for the lowest transmission efficiency tested of τ=0.025. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in infectiousness, no covariation), null (variability in infectiousness, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent, and the numbers in the lower right-hand corner of each panel describe the percentage of simulations fading-out for each treatment. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population has a β_c=0, limiting maximum prevalence to 0.66. Time courses for the corresponding medium (τ=0.25) and high (τ=0.5) transmission efficiencies are available in Appendix E Figures E.3 & E.4.

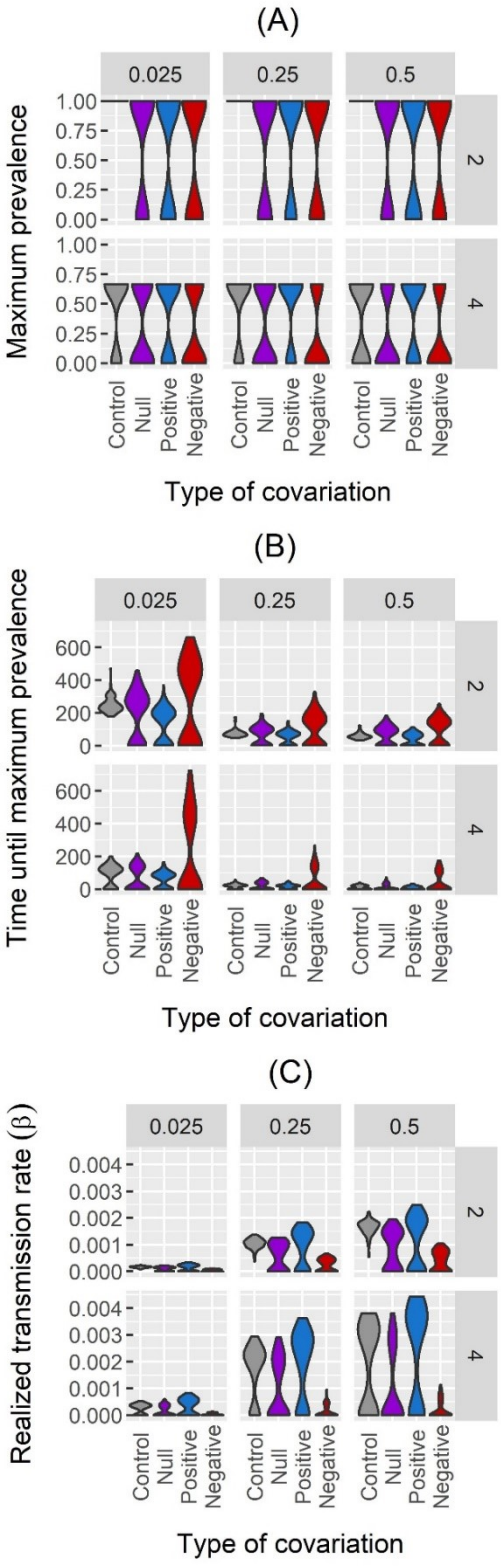

(A)
0.025
0.25
0.5
2
4
Maximum prevalence
1.00
0.75
0.50
0.25
0.00
Control
Null
Positive
Negative
Type of covariation
(B)
Time until maximum prevalence
600
400
200
0
(C)
Realized transmission rate (β)
0.004
0.003
0.002
0.001
0.000

Figure 2.4. For infectiousness vs. contact rate, violin plots depicting for all covariation types: **(A)** the maximum prevalence reached in 750 time steps; **(B)** the time it takes to reach that maximum prevalence; and **(C)** the realized transmission rate (β), which describes the rate of epidemic spread. The columns correspond to the transmission efficiency (i.e., 0.025, 0.25, and 0.5), and the rows correspond to the difference in mean degree (i.e., 2 or 4).

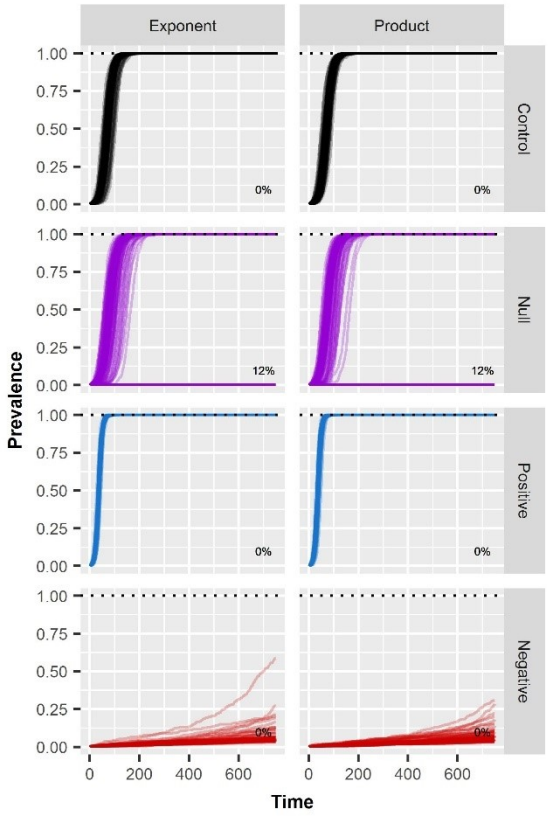


Figure 2.5. Time course of simulated epidemics for infection status vs. contact rate for the lowest transmission efficiency tested of $\tau = 0.025$. Columns correspond to how infectiousness was modelled (either in the exponent or the product of the final transmission probability), and rows correspond to the mechanism of covariation: control (all infection statuses have equal mean degree and no variability in infectiousness), null (variability in infectiousness, but no covariation with contact rate), positive covariation, and negative covariation. Individual trials are shown as semi-transparent, and the numbers in the lower righthand corner of each panel describe the percentage of simulations fading-out for each treatment. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. Time courses for the corresponding medium ($\tau = 0.25$) and high ($\tau = 0.5$) transmission efficiencies are available in Appendix E Figures E.5 & E.6.

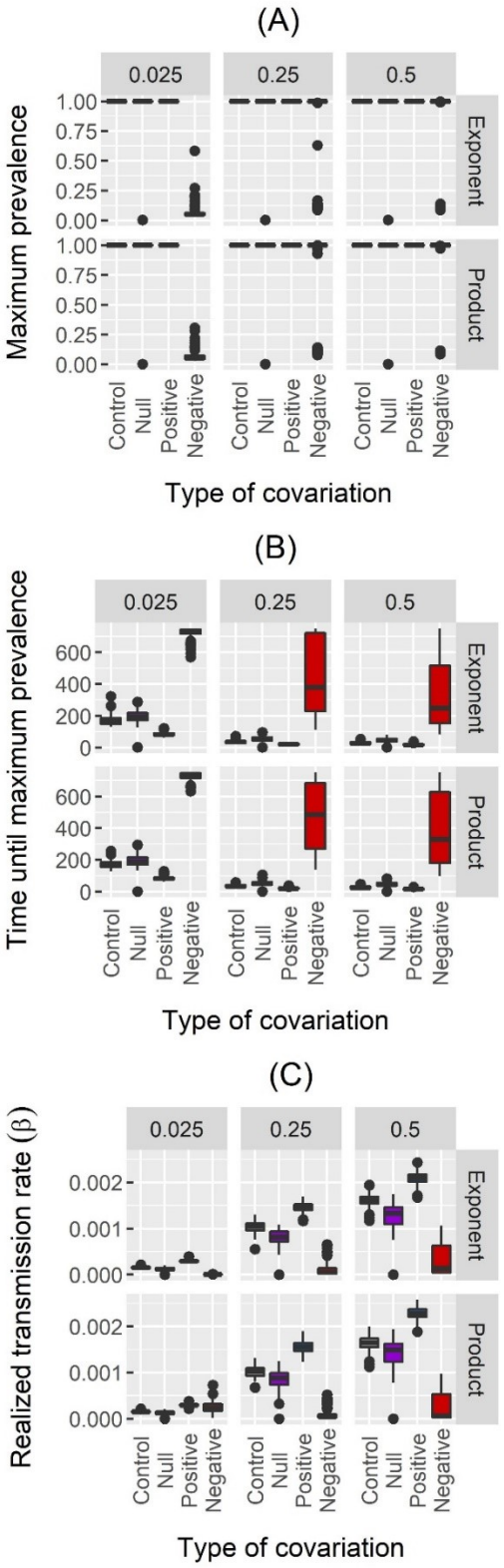

(A)
0.025
0.25
0.5
Exponent
Product
Maximum prevalence
1.00
0.75
0.50
0.25
0.00
Control
Null
Positive
Negative
Type of covariation
(B)
Time until maximum prevalence
600
400
200
0
(C)
Realized transmission rate (β)
0.002
0.001
0.000

Figure 2.6. For infection status vs. contact rate, box and whisker plots depicting for all covariation types: **(A)** the maximum prevalence reached in 750 time steps; **(B)** the time it takes to reach that maximum prevalence; and **(C)** the realized transmission rate (β), which describes the rate of epidemic spread. The columns correspond to the transmission efficiency (i.e., 0.025, 0.25, and 0.5) and the rows correspond to way that individual infectiousness affected the final transmission probability (i.e., in the exponent or the product). Note: in this case, we elected to display results with a box and whisker plot rather than a violin plot because the violin plots poorly portrayed some of the distinct point values and dichotomous epidemic outcomes.

Chapter 3. Influenza A virus in swine breeding herds: Combination of vaccination and biosecurity practices can reduce likelihood of endemic piglet reservoir

White. L.A., Torremorell, M. and Craft, M.E. (2017). Influenza A virus in swine breeding herds: Combination of vaccination and biosecurity practices can reduce likelihood of endemic piglet reservoir. Prev. Vet. Med., 138, 55-69. doi: 10.1016/j.prevetmed.2016.12.013

3.1 Abstract

Recent modelling and empirical work on influenza A virus (IAV) suggests that piglets play an important role as an endemic reservoir. The objective of this study is to test intervention strategies aimed at reducing the incidence of IAV in piglets and ideally, preventing piglets from becoming exposed in the first place. These interventions include biosecurity measures, vaccination, and management options that swine producers may employ individually or jointly to control IAV in their herds. We have developed a stochastic Susceptible-Exposed-Infectious-Recovered-Vaccinated (SEIRV) model that reflects the spatial organization of a standard breeding herd and accounts for the different production classes of pigs therein. Notably, this model allows for loss of immunity for vaccinated and recovered animals, and for vaccinated animals to have different latency and infectious periods from unvaccinated animals as suggested by the literature. The interventions tested include: (1) varied timing of gilt introductions to the breeding herd, (2) gilt separation (no indirect transmission to or from the gilt development unit), (3) gilt

vaccination upon arrival to the farm, (4) early weaning, and (5) vaccination strategies of sows with different timing (mass and pre-farrow) and efficacy (homologous vs. heterologous). We conducted a Latin Hypercube Sampling and Partial Rank Correlation Coefficient (LHS-PRCC) analysis combined with a random forest analysis to assess the relative importance of each epidemiological parameter in determining epidemic outcomes. In concert, mass vaccination, early weaning of piglets (removal 0-7 days after birth), gilt separation, gilt vaccination, and longer periods between introductions of gilts (6 months) were the most effective at reducing prevalence. Endemic prevalence overall was reduced by 51% relative to the null case; endemic prevalence in piglets was reduced by 74%; and IAV was eliminated completely from the herd in 23% of all simulations. Importantly, elimination of IAV was most likely to occur within the first few days of an epidemic. The latency period, infectious period, duration of immunity, and transmission rate for piglets with maternal immunity had the highest correlation with three separate measures of IAV prevalence; therefore, these are parameters that warrant increased attention for obtaining empirical estimates. Our findings support other studies suggesting that piglets play a key role in maintaining IAV in breeding herds. We recommend biosecurity measures in combination with targeted homologous vaccination or vaccines that provide wider cross-protective immunity to prevent incursions of virus to the farm and subsequent establishment of an infected piglet reservoir.

3.2 Introduction

Influenza A virus (IAV) is a globally endemic infection that causes significant morbidity in swine, poses a substantial public health risk to humans, and inflicts a

considerable financial burden on the swine industry (Torremorell *et al.* 2012; Vincent *et al.* 2014). Swine are an important host for IAV because they are susceptible to strains from multiple species, and thus can foster the reassortment of avian, swine and human strains (Torremorell *et al.* 2012). Interspecies transmission events of IAV are well documented, including transmission events from swine to people and people to swine (Myers *et al.* 2006; Myers, Olsen & Gray 2007; Nelson *et al.* 2014; Nelson & Vincent 2015; Choi *et al.* 2015). The rapid rate of change in IAV strains due to both genetic drift and genetic shift ensure that novel strains are likely to arise (Torremorell *et al.* 2012). The 2009 H1N1 pandemic illustrates the potential zoonotic risk that is present for humans from such reassortment (Girard *et al.* 2010). While reduced growth rates and morbidity of infected animals are the primary financial concern for the swine industry, public concern during outbreaks can also have substantial consequences (Sandbulte *et al.* 2015). For instance, swine industry revenue losses in excess of $1 billion dollars have been estimated in the aftermath of the 2009 H1N1 pandemic (Pappaioanou & Gramer 2010).

In swine, direct contact with infected pigs is thought to be the primary transmission route for IAV, but transmission may also occur through aerosols and fomites (Torremorell *et al.* 2012; Allerson, Cardona & Torremorell 2013a). IAV infections are common at the herd level and are difficult to eliminate on a farm (Torremorell *et al.* 2012). Prolonged infection at the population level has public health implications because of an increased risk of exposure for swine workers (Myers *et al.* 2006, 2007; Allerson *et al.* 2013b). Persistent infection is likely dependent on many farm-

specific factors including housing structure, the frequency of intra- and inter-farm movement rates of specific subpopulations, and animal density (Allerson *et al.* 2013b). One hypothesis for the prolonged persistence of IAV in swine herds is the recurring availability of susceptible individuals (Brown 2000; Allerson *et al.* 2013b; Pitzer *et al.* 2016). This hypothesis is supported by several documented cases of IAV positive neonatal piglets in the absence of IAV positive sows (Larsen *et al.* 2010; Allerson *et al.* 2013b). Notably, it has been demonstrated that piglets can still become infected during the lactation phase (Corzo *et al.* 2012). Gilt introductions into breeding herds may also play an important role in IAV persistence on farms (Diaz *et al.* 2015).

Producers commonly employ vaccination and biosecurity measures to reduce IAV incidence in their herds (Torremorell *et al.* 2012). A 2006 survey conducted by the National Animal Health Monitoring System (NAHMS) found that 29% of the farms with 100 or more animals were vaccinated for either H1N1 or H3N2 strains (NAHMS, 2008). Currently, all commercial IAV vaccines in the U.S. are inactivated (Sandbulte *et al.* 2015). Such vaccines are only effective for identical or antigenically similar strains of IAV (Sandbulte *et al.* 2015). Custom-manufactured autogenous vaccines, which are derived from a strain circulating within a specific herd, are also in widespread use (Sandbulte *et al.* 2015). Thus, vaccines may also be classified as homologous or heterologous, depending on whether they activate an effective immune response for a specific strain of IAV. Homologous vaccines should match the circulating strain exactly and induce complete immunity, while heterologous vaccines do not and may induce only partial immunity. Farmers may routinely vaccinate pregnant sows (pre-farrow

vaccination) in order to impart maternal immunity to piglets or may vaccinate the entire herd at regular intervals (mass vaccination). A case study demonstrated that strain-specific, mass vaccination can decrease IAV shedding at the herd level (Corzo *et al.* 2012). Herd closure and partial depopulation have also been successfully used to eliminate IAV from a swine farm (Torremorell *et al.* 2009).

While vaccination has been shown to reduce transmission rates in experimental settings, it is unlikely to prevent transmission completely, and the implications for herd level transmission are less clear (Romagosa *et al.* 2011; Torremorell *et al.* 2012; Allerson *et al.* 2013c). Results of vaccination efforts in large production systems have been mixed (Beaudoin *et al.* 2012; Sandbulte *et al.* 2015). This is likely because of high mutation rates, high strain diversity, and potential interference from maternally derived immunity (Kitikoon *et al.* 2006; Thacker & Janke 2008).

Currently, the dynamics of IAV circulating in swine herds remains poorly understood (Dorjee *et al.* 2013; Vincent *et al.* 2014). Mathematical models are important tools for understanding disease transmission and dynamics because they allow us to make predictions about the spread of pathogens and to explore "what if" scenarios, including the effects of potential interventions (Lloyd-Smith *et al.* 2009). To date, very little mathematical modelling has been done on IAV in swine (Coburn, Wagner & Blower 2009; Dorjee *et al.* 2013); to our knowledge, only two models have been conducted (Reynolds, Torremorell & Craft 2014; Pitzer *et al.* 2016). Both of these studies support the idea that piglets play a pivotal role in maintaining IAV endemicity. Reynolds et al. (2014) employed a deterministic model to assess homologous and heterologous vaccine

strategies in both a breeding and a wean to finish farm. Their results suggest that vaccination alone does not effectively remove influenza from breeding herds; homologous vaccination eliminated transmission in sows and gilts, but not in piglets. Pitzer et al. (2016) combined mathematical models and phylogenetic analysis to posit that a critical herd size of at least 3,000 individuals can maintain IAV and to suggest that high piglet turnover play a key role in maintaining IAV endemically.

The goal of this study is to test a variety of intervention strategies that swine producers may feasibly employ to reduce the incidence of IAV in piglets and prevent piglets from becoming exposed in the first place. Here we develop a stochastic, metapopulation model to emulate the intra-herd dynamics of a typical Midwestern swine breeding farm. The motivation for developing a stochastic model was to examine the cases in which epidemiological parameter variability and experimental interventions might lead to IAV elimination from a breeding herd. This model and subsequent analyses differ from previous modelling efforts in several key ways: (1) a stochastic approach allows for the analysis of both parameters and interventions leading to extinction events, (2) vaccinated and recovered animals can lose their immunity, (3) vaccinated animals have separate latency and recovery rates as suggested by the literature, and (4) an extensive global sensitivity analysis was conducted to assess the importance of parameters in contributing to disease outcomes.

We began by developing a null model where no interventions were implemented. In the absence of interventions, we tested how farm size and the frequency and source of IAV introduction might affect IAV dynamics (Figure 3.1). Using expected mean or mode

values from the literature and the resulting null model, we first explored transmission dynamics in a breeding herd without any interventions, then tested fourteen different types of management interventions independently, and finally tested a combination of interventions to see if any could successfully reduce IAV incidence. In total, we tested fifteen experimental scenarios. In addition, we tested the robustness of these results to variation in epidemiological parameters and the relative importance of each parameter by conducting a global sensitivity analysis and a random forest analysis.

3.3 Materials and methods

Model development and parameterization

We have developed a stochastic, Susceptible-Exposed-Infectious-Recovered-Vaccinated (SEIRV) simulation of IAV dynamics in a swine breeding herd. Animals are categorized in classes according to both their IAV immune status and their production stage (Table 3.1). This metapopulation model is density-dependent, reflects the spatial organization of a standard swine breeding herd, and thus, accounts for the various production stages including gilts, sows, and piglets (Figure 3.2). The spatial subpopulations of the model correspond to a gilt development unit (GDU), a breeding/gestation unit, and a farrowing unit as found on a typical farrow-to-wean facility. Gilts are introduced into the GDU where they mature for 10 weeks before moving to the breeding/gestation unit. In the breeding/gestation unit, sows are bred and gestate for 16 weeks. Then, sows move to the farrowing unit where they give birth during the first week and remain until their piglets are weaned after an additional three weeks.

Sows have an equal probability of giving birth on any day during the first week of introduction to the farrowing unit, which means that piglets may be a minimum of 21 days and a maximum of 28 days old upon weaning. Weaned sows return to the breeding/gestation area where they are either culled (50%) or rebred. Movement of animals between stages and sub-populations occurs on a weekly basis.

Although there is a trade-off with increased model complexity, we have separated out the different production classes because this allows for the administration of interventions to specific groups and for different production classes to have separate epidemiological parameters (Table 3.1). In particular, we separate piglets based on their age (one to four weeks old) because piglets have been found to be IAV positive even in the absence of infection of sows or gilts (Larsen *et al.* 2010; Allerson *et al.* 2013b). Also, including weekly classes for piglets allows for the testing of different early weaning times. Animals are assumed to mix homogenously within their respective production units.

This continuous Markov chain model allows for stochastic births, non-disease related deaths, direct and indirect infection, and loss of both vaccine-induced and naturally-acquired immunity via Gillespie's Direct Method (Gillespie 1977; Keeling & Rohani 2008b). The rates for each event, j, are calculated based only on the current properties of the system at each time step. At each time step, we evaluated the transition rates (λ_j) for 161 separate events (Equations in Appendix F Table F.1). The time between events is represented by an exponential distribution with an expected (mean) value equal to the inverse of the sum of the transition rates:

$E(x) = 1/\sum_{j=1}^{161} \lambda_j dt$, and the probability that the next event will be of type *j* is equal to $\lambda_j dt/\sum_{j=1}^{161} \lambda_j dt$ (Legrand *et al.* 2004). The epidemiological and demographic parameters that determine these rates are shown in Table 3.3.

Parameterization

Epidemiological parameters are derived from recent experimental work that characterizes the rate of IAV transmission (β) between pigs for piglets, adults, and vaccinated animals (Romagosa *et al.* 2011; Allerson *et al.* 2013a; c). Transmission values are converted from frequency-dependent (β') to density-dependent (β) values by dividing β' by the number of individuals in each respective experimental treatment group for Allerson et al. 2013a, 2013c, and Romagosa et al. 2011 (Begon *et al.* 2002; Reynolds *et al.* 2014). We assume density-dependent transmission because population size fluctuates in the model due to births and culling and because IAV is a directly transmitted virus (Keeling and Rohani, 2008). Since pigs in a breeding herd are in a constrained area, contacts are likely to increase overall with density, and thus the force of infection is likely to increase correspondingly.

Expected values (Table 3.3) were the mean values for normally-distributed parameters, but could also be mode values for asymmetrical triangle distributions or simply an informed point estimate within a uniform distribution. For instance, the expected values for indirect transmission were extrapolated from experimental data, assuming that $\beta_{ind} \approx \beta_d/178$ (Allerson et al., 2013a; Reynolds et al., 2014); during Latin Hypercube Sampling (LHS), the range for indirect transmission was assumed to be

from 0 to the lower limit for the corresponding direct contact rate. In this way, indirect transmission is always less than direct transmission for each simulation, but there is no forced correlation between the two parameters. The ranges of transmission parameters explored during LHS for the homologous vaccine treatments still allowed for minimal potential direct and indirect transmission events and did not offer "perfect" immunity (Romagosa *et al.* 2011; Allerson *et al.* 2013c). Empirical work suggests that there may be differences in the *latency* (the period of time between first exposure and infectiousness: $1/\sigma$) and infectious periods (duration of infectiousness: $1/\gamma$) between vaccinated and unvaccinated groups (Romagosa *et al.* 2011). To test the importance of these potential differences, we included a separate exposed and infectious (VE and VI) classes for vaccinated animals.

Strain type

While several studies in IAV vaccine efficacy have been performed, few have characterized inter-animal transmission rates between pigs (Van Reeth *et al.* 2001; Lange *et al.* 2009). Based on the available empirical work, here we model a single, generic H1N1 strain. The empirical studies that inform the parameters for this model utilized specific challenge strains (e.g., delta cluster H1N1 [A/Sw/ MN/07002083/07] and β cluster H1N1 [A/Sw/IA/00239/04]) and vaccine formulations (Romagosa *et al.* 2011; Allerson *et al.* 2013a; c). Given the substantial strain diversity of IAV, it would be reasonable to expect some variation in these parameters for different challenge strain-vaccine pairings. By testing the range defined by the reported 95% CIs in the literature

for the LHS-PRCC sensitivity analysis, we have attempted to account for some of this conceivable variation.

Testing assumptions and initial conditions

To compare the effects of different initial conditions and the effectiveness of each intervention, we computed four metrics for each scenario: the maximum prevalence of infected individuals observed at the infection peak, the mean endemic prevalence (the average prevalence for all animals following the initial infection peak from t= 6-12 months), the mean endemic prevalence in piglets (the average prevalence for piglets from t= 6-12 months), and the probability of stochastic IAV extinction (the percentage of trials where no IAV was present on the farm by the end of the simulation's duration).

Farm size

Over the past two decades, the U.S. swine industry has consolidated dramatically, skewing heavily to fewer production sites with larger holdings. By 2009, 86% of the U.S. swine inventory existed in farms with 2,000 or more total animals (McBride & Key 2013). However, smaller sites with less than 2,000 animals are more likely to have a breeding herd. 57.6 percent of breeding herds have fewer than 250 sows and gilts, 6.8 percent have 250-499, and 35.6 percent have more than 500 sows and gilts (USDA 2015). Using a null model without any additional interventions, we tested a range of farm sizes varying from approximately 500 to 5000 total animals (Figure 3.1). We observed the effect of farm size on maximum prevalence, endemic prevalence, endemic prevalence in piglets, and the likelihood of stochastic extinction (Appendix F Table F.2).

Source and frequency of IAV introduction

We evaluated the effects of both source and frequency of IAV introduction on disease dynamics within the breeding herd. For source of infection, we tested introductions of a single infected gilt to the GDU and a single sow to the breeding/gestation unit or the farrowing unit respectively. For frequency of IAV introduction, we compared scenarios where a single infected gilt was introduced to the GDU at t=0 and where a single infected gilt was introduced weekly into the GDU (Figure 3.1). For each case, we computed peak prevalence, endemic prevalence, endemic prevalence in piglets, and the likelihood of stochastic extinction (Appendix F Table F.3).

Null model

We introduced one new IAV strain into an immunologically naïve farm (unless animals were vaccinated as specified by an intervention) via a single infected gilt. Each simulation began with a total of 3500 pigs on the farm: 2000 piglets, 330 gilts and 1170 sows, and continued for a period of 365 days (1 year). The population size for the null model fluctuated between approximately roughly 3,000 and 4,000 individual pigs. Our null model assumes that no vaccinations are given to any class, that 30 gilts are introduced to the GDU on a weekly basis, that indirect transmission can occur between each sub-population of the farm, and that piglets spend 21-28 days with their mothers in the farrowing unit after birth (Table 3.2, Experiment #0). We have modeled indirect transmission as only occurring between different sub-populations (e.g., a gilt in the GDU could only experience indirect infection from animals in either the farrowing or

breeding/gestation unit). There is evidence that such indirect transmission may occur even if a farm is observing high standard biosecurity practices (Allerson *et al.* 2013a). Once recovered, piglets did not lose immunity. We did not explicitly model strain mutation, but because simulations last for up to a year, we assumed that recovered sows and gilts had an exponentially waning natural immunity that lasts an average of 8-16 weeks (Table 3.3).

Interventions

Beginning with this null model, we tested five intervention types that include: (1) varied timing of gilt introductions to the breeding herd, (2) gilt separation (no indirect transmission to or from the GDU), (3) gilt vaccination upon arrival to the farm, (4) early weaning, and (5) vaccination strategies with different timing (mass and pre-farrow) and efficacy (homologous vs. heterologous). The types of interventions used in each experiment are described in Table 3.2.

Timing of gilt introductions:

From a 2012 USDA survey, about one-third of large farms introduce gilts at least every two weeks (USDA 2015). We tested intervals of one week, one month, three months, and 6 months between introductions (Table 3.2, Experiment #1, 2 & 3 respectively). In interventions where the timing of gilt introductions was altered, the number of gilts introduced at each interval was equal to 30 multiplied by the number of weeks that had elapsed since the prior introduction; this kept the number of introduced gilts consistent across experiments.

Gilt separation

When gilt separation (Table 3.2, #4) was implemented, no indirect transmission between the GDU and other sub-populations of the farm was possible. However, movement of infectious individuals was still possible via the weekly movement of gilts from the GDU to the breeding/gestation area.

Gilt vaccination

For the gilt vaccination intervention, all gilts, with the exception of the single infected "seeder" gilt, were of vaccinated status upon arrival to the farm. We modeled two types of vaccine efficacy: homologous and heterologous (Table 3.2, #5, 6).

Early weaning

In the null case, piglets remain with their mothers for 3-4 weeks (21-28 days). We tested early weaning of 0-7 days, 7-14, and 14-21 days after birth (Table 3.2, #7, 8 & 9 respectively). It is important to note here that early weaning causes a commensurate reduction in the overall population size of the farm from the removal of piglets. For instance, the total population oscillates between 1,500-2,500 animals, 2,000-3,000 animals, and 2,500-3,500 animals for early weaning after 0-7 days, 7-14 days, and 14-21 days respectively.

Vaccine efficacy and strategy

As with gilt vaccination, we tested both homologous and heterologous vaccines in combination with two types of vaccination strategy: mass vaccination and pre-farrow vaccination. For mass vaccination interventions (Table 3.2, #10, 11 & 12), we assumed

that piglets began with vaccine-induced maternal immunity and that all mature animals (sows and gilts) had vaccine-induced immunity at t=0. This would be the most likely profile for an IAV naïve farm with an ongoing mass vaccination program. We tested two time frames for revaccination of sows and gilts: every two months (#10 & 12) and every six months (#11). Mass vaccination every two months represents an ideal scenario, since the immunity for commonly used vaccines on the market wanes after 8-10 weeks (Pfizer Animal Health 2011; Merck Animal Health 2013). However, vaccinating the herd every 4-6 months is a more realistic timeframe for most producers given time and resource constraints. For pre-farrow interventions (Table 3.2, #12 & 13), sows were vaccinated when they reached the pre-farrow sow compartment (S_3) three weeks prior to farrowing, and give birth to piglets with vaccine induced maternal immunity (S_{15}). The strength of this immunity for the piglet depends on whether the sow received a heterologous or homologous vaccine.

Combination strategy

After performing simulations where only one intervention type was altered at a time (Experiments #0-14), we selected the final "optimal" combination of interventions (Experiment #15) based on the interventions that decreased at least one of these four metrics relative to the null case (Experiment #0).

Uncertainty and sensitivity analyses

Two sources of uncertainty may contribute to variation in a stochastic model. The first, aleatory uncertainty, results from the inherent stochasticity in the model (e.g., births,

deaths, and transmission events). The second type of uncertainty, epistemic uncertainty, results from variation in the model parameters themselves. In a stochastic model, it is necessary to analyze the sources of uncertainty in series because they cannot necessarily be teased apart simultaneously (Marino *et al.* 2008). Therefore first we assessed the aleatory uncertainty inherent to the model by running each experiment 100 times using the expected parameter values Table 3.3 and calculating the coefficient of variation (Table 3.4) (Marino *et al.* 2008).

To understand the extent of epistemic uncertainty in the model, we conducted a multivariate sensitivity analysis using Latin Hypercube Sampling (LHS) and partial rank correlation coefficient (PRCC) analysis. LHS-PRCC is a sampling based global sensitivity analysis that has been used for disease models (Seaholm, Ackerman & Wu 1988; Blower & Dowlatabadi 1994; Legrand *et al.* 2008; Wu *et al.* 2013). For LHS, the minimum required sample size (N) is $N \geq K+1$ or $N \geq 4/3*K$ where K is the number of parameters included in the LHS (Blower & Dowlatabadi 1994). However, replicating the same LHS parameter sets rather than just increasing the number of independent samples can improve the accuracy of PRCC because it decreases aleatory uncertainty and reduces computational intensity (Marino *et al.* 2008). Marino et al. (2008) conducted a sensitivity analysis for a 12 parameter stochastic model and found 3 replicates of 100-division LHS to be equivalent to 1,000 random samples in the ability of PRCC to discern the effects of significant parameters. Here we performed LHS for 15 parameters with 100 divisions per parameter and then replicated that same parameter set three times for each intervention

for a total of 300 simulations per intervention and 4,800 simulations in total (100 x 3 x 15).

Distributions for all 15 parameters for LHS sampling were normal, triangular, or uniform as informed by the literature (Table 3.3). Uniform distributions were utilized when less information was available, particularly for indirect transmission rates and for loss of immunity (Rushton *et al.* 2000). We conducted PRCC on the averaged values of the three replicated sets for each intervention and for three evaluation criteria (maximum prevalence, endemic prevalence, and endemic prevalence in piglets) (Tenhumberg *et al.* 2004). Specifically, we used Pearsons's partial correlation coefficients to detect monotonic relationships between parameters and outputs after accounting for the effects of all other parameters (Marino *et al.* 2008). All LHS simulations and PRCC analyses were conducted in MATLAB (R2015b).

For the analysis of both aleatory and epistemic uncertainty, we have chosen to focus on describing and discussing differences in the magnitude in the outcomes and PRCC results. Arguably, significance testing can lose its utility for simulation studies because an essentially unlimited sample size can contribute to even small differences in the magnitude of outcomes being designated as statistically significant (White *et al.* 2014).

To aid in this descriptive approach, we have included a random forest analysis to examine the relative importance of both interventions and parameters leading to IAV extirpation. Random forest analysis is a flexible classification tool that can handle non-linear relationships between a large number of predictor variables. As a random-based

ensemble method, it is considered an improvement on traditional classification or regression trees because it compensates for overfitting by sampling a large number of trees instead of just one (Breiman 2001; Liaw & Wiener 2002a; Cutler *et al.* 2007). This method has been used to analyze and predict complex epidemiological data for other disease models (Herrick, Huettmann & Lindgren 2013; Kane *et al.* 2014). We used variable importance plots to evaluate the effect of both interventions and parameters on IAV endemic prevalence, maximum prevalence, and likelihood of extinction. The mean decrease in accuracy describes the loss of predictive value that results from a particular variable being randomly permutated. We used a measure of permutation importance rather than node impurity to explore variable importance because these measures have been shown to be more robust to the numbers and types of input variables (Strobl *et al.* 2007, 2009). These analyses were conducted in R (version 3.2.2) using the *cforest* function in the *party* package. For each random forest analysis, we simulated 10,000 trees, which was of sufficient ensemble size such that the order of variable importance did not vary with the random seed (Strobl *et al.* 2009).

3.4 Results

Farm size

Persistence of IAV in the breeding herd was robust to farm size. All farm sizes exhibited a similar maximum and endemic prevalence (Appendix F Table F.2). After farm size increases beyond 200 sows and gilts or a total inventory of 500 animals, we observed no difference in the percentage of stochastic extinctions.

Frequency and source of IAV introduction

The dynamics of IAV infection for the null model were robust to changes in source of infection or frequency of IAV introduction (Appendix F Table F.3). There were no observable differences in maximum prevalence, endemic prevalence, or likelihood of stochastic extinction across the tested scenarios.

Aleatory uncertainty

Null scenario (Experiment #0)

For the null case, there is a single epidemic peak followed by the maintenance of a steady level of endemicity. However, the observed pattern of infection differs for each of the production classes (Figure 3.3). The initial spike in infection is due to IAV spreading from the GDU to other sub-populations of the farm (particularly sows) and occurs within the first week following IAV introduction to the farm (mean time of peak prevalence= 4.39 ± 0.36 days). After this initial peak, the endemic level of infection, ranging from ~400-600 animals, results from the continual infection of newly born piglets. Sows and gilts are less likely to be infected during the endemic phase.

Gilt introduction (Experiments #1-3)

Changing the timing of gilt introductions only marginally reduced the three prevalence measures (Table 3.4) and did not affect the probability of IAV extinction. The most noticeable effects were for increasing the period between gilt introductions from one week to six months (Experiment #3) which decreased maximum prevalence by 0.9

percent, endemic prevalence by 1.9 percent and endemic prevalence in piglets by 3.6 percent. It is worth noting however that a slightly lower mean endemic prevalence came at the cost of large spikes in gilt prevalence when gilt introductions occurred (Figure 3.3).

Gilt separation (Experiment #4)

Gilt separation reduced the maximum prevalence by 4.6 percent, but had negligible effects on endemic prevalence, endemic prevalence in piglets, and the likelihood of IAV extinction.

Gilt vaccination (Experiments #5 & 6)

Homologous gilt vaccination (Experiment #5) reduced maximum prevalence by 9.8 percent and endemic prevalence by 23.8 percent, but had comparatively little effect on endemic prevalence in piglets with a reduction of only 2.0 percent observed. The effects of heterologous gilt vaccination (Experiment 6) were less dramatic, similarly reducing maximum prevalence by 9.2, but only reducing endemic prevalence and endemic prevalence in piglets by 4.6 and 1.0 percent respectively.

Early weaning (Experiments #7-9)

Early weaning of piglets had little effect on maximum prevalence. Early weaning of piglets after 0-7 days (Experiment #7) reduced endemic prevalence and endemic prevalence in piglets by 26 and 61 percent respectively. Counterintuitively, early weaning at later stages (7-14 days and 14-21 days) increased prevalence relative to the null. This is because while the number of infected individuals in each class did not increase, the

relative numbers of animals on the farm decreased from the loss of piglets. After the initial infection peak, the total number of infected individuals per time step was approximately 250 individuals for 0-7day treatment, 500 individuals for 7-14 day treatment, and 600 individuals for 14-21 day treatment. In contrast, in the null model, total population size oscillates between 3,000-4,000 individuals with approximately 500 infected individuals per time step after the initial infection peak.

Vaccine strategy and efficacy (Experiments #10-14)

Homologous mass vaccination every two months (Experiment #10) reduced the maximum prevalence by 51.3 percent, the endemic prevalence by 24.3 and the endemic prevalence in piglets by 24.0 percent. Most importantly, this treatment also increased the likelihood of IAV extinction to 24 percent, although this extinction was most likely to occur within the first few days of IAV introduction. Some of this effect was still observed even when homologous mass vaccination was executed every six months (Experiment #11) with reductions of 47.2, 18.2 and 18.1 percent to maximum, endemic and endemic prevalence in piglets respectively. The effects on endemic prevalence were less noticeable for heterologous mass vaccination (Experiment #12), although the peak prevalence was still reduced by 37.7 percent. Both homologous and pre-farrow vaccination (Experiments #12 &13) had a negligible effect on endemic prevalence overall and endemic prevalence in piglets, but reduced peak prevalence by 13.4 and 12.5 percent respectively.

Combination strategy (Experiment #15)

The combination strategy (#15) included homologous mass vaccination every two months, early weaning of piglets (removal 0-7 days after birth), gilt separation, homologous gilt vaccination, and a longer period between introductions of gilts (6 months). Relative to the null, this combination of interventions was the most effective at reducing prevalence (Table 3.4, Figure 3.5, Figure 3.6). This strategy reduced the maximum prevalence by 65.4 percent, the endemic prevalence by 51.1 percent, and the endemic prevalence in piglets by 74.0 percent. Additionally, in 23 percent of simulated cases, IAV was eliminated from the breeding herd, but as with homologous mass vaccination (#10 & 11) this extinction occurred within the first few days after IAV introduction.

Epistemic Uncertainty (Global Sensitivity Analysis)

The amount of variation observed when using expected values across treatments is small relative to the epistemic uncertainty (Figure 3.4). Despite using informed (i.e. normal and triangle) parameter distributions for most parameters, the amount of epistemic uncertainty contributed substantially to the observed variation (Figure 3.5);

PRCC analysis

Based on the PRCC results (Figure 3.5), the recovery rate (γ) was negatively correlated with endemic prevalence across all experiments. This suggests that as the infectious period decreases endemic prevalence will decrease. For experiments employing early weaning, the reciprocal of latency period (σ), the loss of immunity rate

for recovered sows and gilts (ω_r) and the direct transmission rate for piglets with maternal immunity (β_d^{pm}) were positively correlated with endemic prevalence of IAV in the herd. Specifically, with shorter latency periods (σ^{-1}) for exposed animals and duration of immunity (ω_r^{-1}) for recovered animals, the endemic prevalence during the simulations increased.

For the endemic prevalence in piglets (Appendix F Table F.4), similar trends were observed for the recovery rate (γ), the reciprocal of the latency period (σ), and the direct transmission rate for piglets with maternal immunity (β_d^{pm}). However, duration of immunity (ω_r^{-1}) for recovered animals was not detected as having a substantial effect on endemic prevalence in piglets.

Finally, for the maximum prevalence (Appendix F Table F.5), the recovery rate (γ) had a strong negative correlation, and the reciprocal of the latency period (σ) had a strong positive correlation with maximum prevalence across all experiments. Again, this suggests that as the infectious periods decrease and latency periods increase, the maximum observed prevalence will decrease. Across metrics and experiments, both recovery rate and latency period were most likely to influence epidemic outcomes.

Random forest analysis

For the endemic prevalence of infected pigs, the top four variables with the highest importance scores were recovery rate (γ), early weaning interventions, the loss of immunity rate for recovered sows and gilts (ω_r), and the inverse of the latency period (σ) (Figure 3.6A). These results were consistent with the correlation results of the PRCC

analysis (Table 3.5). For the maximum prevalence of infected pigs, the inverse of the latency period (σ), vaccination strategy (mass vs. pre-farrow), the recovery rate (γ), and vaccine efficacy (heterologous vs. homologous) were the four top ranking factors in variable importance (Figure 3.6B). Again, the inverse of the latency period (σ) and the recovery rate were identified as having strong effects on maximum prevalence in the PRCC analysis (Appendix F Figure F.5). For the probability of stochastic extinction (Figure 3.6C), the direct transmission rate for piglets with maternal immunity (β_d^{pm}), the indirect transmission rate for piglets (β_{ind}^{p}), the indirect transmission rate for vaccinated animals ($\beta_{ind\ vax}$), and the inverse of the latency period (σ) were variables with the highest prediction power.

3.5 Discussion

Our findings support other modelling and empirical studies that suggest that younger animals play a pivotal role in maintaining IAV in swine herds (Beaudoin *et al.* 2012; Allerson *et al.* 2013b; Reynolds *et al.* 2014; Diaz *et al.* 2015; Pitzer *et al.* 2016). The two other modelling studies conducted on IAV have found similar patterns of persistence in piglets (Reynolds *et al.* 2014; Pitzer *et al.* 2016). However, our model predicts IAV persistence even in very small herd sizes as low as 500 total animals in contrast to the critical herd size estimate of 3,000 animals predicted by Pitzer et al. (2016).

In many ways, the model provides a "worst case scenario" because it allows for reinfection and imperfect immunity both from vaccination and maternal immunity. Under

these circumstances, many common management and biosecurity strategies that were tested did not produce a significant decrease in IAV prevalence. Based on the behavior of the model, once IAV had spread from the GDU to establish residency in other parts of the farm, stochastic extinctions became very unlikely; stochastic extinctions were most likely to occur within the first few days after IAV introduction. For instance, the effects of increased spacing between gilt introduction and gilt separation were minimal. It seems that the continual animal movement between production units makes preventing indirect transmission less effective. The indirect transmission rate for sows and gilts (β_{ind}) did not correlate strongly with prevalence values, which is logical since stochastic extinctions are unlikely in unvaccinated populations. However, the indirect transmission rate for vaccinated animals ($\beta_{ind\ vax}$) and the indirect transmission rate for piglets (β_{ind}^{p}) appear to contribute to the likelihood of stochastic extinction according to the random forest variable importance metric (Figure 3.6C). This suggests that the likelihood of spreading IAV to other areas of the farm within the first few days of introduction and particularly likelihood of introduction to the vulnerable piglet population could be especially important for facilitating extinction events. PRCC and random forest analysis results suggest that this relationship between indirect transmission and stochastic extinction would not be as important once IAV has become endemic.

Homologous vaccination was the best overall intervention in our model as it decreased all three prevalence metrics and increased the likelihood of stochastic extinction to nearly a fourth. Encouragingly, these effects were observed for both the two- and six-month administration frequencies. However, while homologous vaccination

caused substantial reductions in IAV prevalence it did not completely eliminate sow and gilt transmission as predicted in a previous deterministic model (Reynolds *et al.* 2014). These results were somewhat surprising since the vaccination interventions (#10-14) differed from other interventions in that a large portion of adult animals had vaccine derived immunity at t=0. Given these conditions, extinction would not be unexpected, but was in fact relatively rare. This finding highlights the potential importance of preventive measures and the difficulty of eliminating IAV once it becomes endemic. While early weaning was a particularly effective strategy for reducing IAV prevalence in piglets, it was only effective in the more drastic case of removing piglets within a week after birth. The benefits of this intervention must be offset by the potential production costs to the farmer, and is unlikely to be sustainable for longer durations as simulated here. It may be that a "dry spell", i.e., a removal of piglets and cessation of births would be necessary to provide the best chances for complete IAV elimination which may be accomplished in batch farrowing production systems.

Although American swine producers commonly rely on some form of pre-farrow vaccination (National Animal Health Monitoring System 2008), we found that both the homologous and heterologous pre-farrow vaccination only had a slight effect on maximum endemic prevalence and a negligible effect on the other three metrics. Similarly, heterologous mass vaccination had a slight effect in reducing maximum prevalence, but a negligible effect in the realm of endemicity or extinction. Although heterologous vaccination did not have a substantial effect on the evaluation criteria measured here, vaccines are still valuable from a clinical perspective. Under experimental

conditions, vaccines can decrease lung lesions and improve clinical signs (Van Reeth *et al.* 2001; Lange *et al.* 2009). So from a production standpoint, heterologous vaccines may still play a valuable role in herd health, but not necessarily by mitigating transmission at the herd level. One question that warrants further investigation is whether decreasing endemic prevalence in the breeding herd is expected to modify the incidence in subsequent production phases; in part that will depend on IAV prevalence levels in piglets at weaning. However, it is unclear whether transmission can be prevented in low prevalence situations given that the waning of maternal antibodies will result in an increase of susceptibility in piglets or whether transmission can be prevented only through complete extinction of IAV prior to weaning.

The relative inefficacy of heterologous vaccine interventions in reducing IAV prevalence highlight both the importance of matching the vaccine to any circulating strains and the potential value of biosecurity and herd management interventions beyond vaccination. Matching circulating strains will likely require autogenous vaccines in many cases and only commercial multivalent vaccines when appropriate. Vaccine-associated enhanced respiratory disease (VAERD) is also a potential concern. VAERD can result when maternal or vaccine-induced antibodies poorly match a strain that later infects a pig, and causes heightened clinical signs or complications (Sandbulte *et al.* 2015). In clinical studies, pigs challenged with a monovalent heterologous vaccine prior to infection had worse clinical signs and higher levels of pro-inflammatory cytokines, although VAERD has not been observed definitively under field conditions (Gauger *et al.* 2011; Sandbulte *et al.* 2015). Genetically engineered modified live vaccines (MLV) and

vectored vaccines could provide better options for targeted protection and improved cross-immunity in the future (Vincent *et al.* 2008, 2012; Gauger *et al.* 2011; Sandbulte *et al.* 2015).

Some limitations of our study and avenues for future directions are as follows. First, while we explored the effects of farm size on the null model, we did not test the effects of farm size for each of the interventions. It is important to note that stochastic effects and possible pathogen extinction events would be more likely for smaller herds – arguably this stochasticity in smaller populations could make interventions more effective. We did not account for any potential seasonality in transmission or introductions, but based on recent empirical work this may be worth future consideration (Beaudoin *et al.* 2012; Diaz *et al.* 2015). Empirical work suggests that there may be differences in the *latency* $(1/\sigma)$ and infectious periods $(1/\gamma)$ between piglets with and without maternal immunity (Allerson *et al.* 2013c). We did not incorporate these differences in an effort to constrain the numbers of parameters included in the model, but this could be important to include for future models, especially since the *latency* and infectious periods have been highlighted as key parameters by the sensitivity analysis. Finally, apart from allowing for loss of susceptibility, we did not explicitly account for mutation in the circulating IAV strain, introductions of different strains, or partial immunity existing from exposure to prior strains. Our parameters were based on a limited number of transmission studies for specific H1N1 challenge strains and vaccine formulations (Romagosa *et al.* 2011; Allerson *et al.* 2013a; c).

Working together, scientists, veterinarians, and producers can engage in an iterative process with modelers to conduct model-guided fieldwork that can both inform and test new hypotheses (Restif *et al.* 2012). The comparatively large amount of epistemic uncertainty generated in the model supports the idea that slight differences in epidemiological parameters in this system can have dramatic consequences. More experimental work is needed to characterize transmission rates for different IAV subtypes and to better understand the effects of cross-reactivity and cross-immunity between strains. The existing literature supports the idea of significant variation between production classes and vaccinated and non-vaccinated animals (Romagosa *et al.* 2011*, and b*oth the LHS-PRCC and random forest analyses consistently suggested that the latency and infectious periods play a key role in determining epidemic outcomes. To our knowledge, our study is the only one to have tested a range of latency and infectious periods and to allow for distinct latency and infectious periods for vaccinated individuals. These potential differences will be important to include in future modelling work and to consider in future empirical tests of different vaccine/strain pairings.

3.6 Conclusions

Several of the vaccination and management strategies that were tested here did not significantly reduce IAV prevalence in the simulated herd. Homologous mass vaccination and early weaning were the most efficacious interventions. The combination of frequent homologous mass vaccination, early weaning, gilt separation, gilt vaccination and longer periods between gilt introductions reduced endemic prevalence overall by 51% relative to the null scenario, endemic prevalence in piglets by 74%, and eliminated

IAV completely in 23% of all simulations. Our findings support recent empirical and modelling work which suggests that piglets play a pivotal role in IAV persistence. Based on these results, biosecurity measures designed to prevent IAV from reaching piglets combined with strain-targeted (homologous) vaccines or vaccines inducing a broader protective immune response are likely the best option for producers to control IAV in their herds.

3.7 Acknowledgements

The authors would like to thank Amy Kinsley for her comments on the manuscript. This material is based upon work supported by the National Science Foundation under Grant Nos. GRFP-00039202 and DEB-1413925, University of Minnesota's Office of the Vice President for Research and an Academic Health Center Seed Grant, and the Cooperative State Research Service, USDA, under Project No. MINV-62-044. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: http://www.msi.umn.edu

3.8 Figures

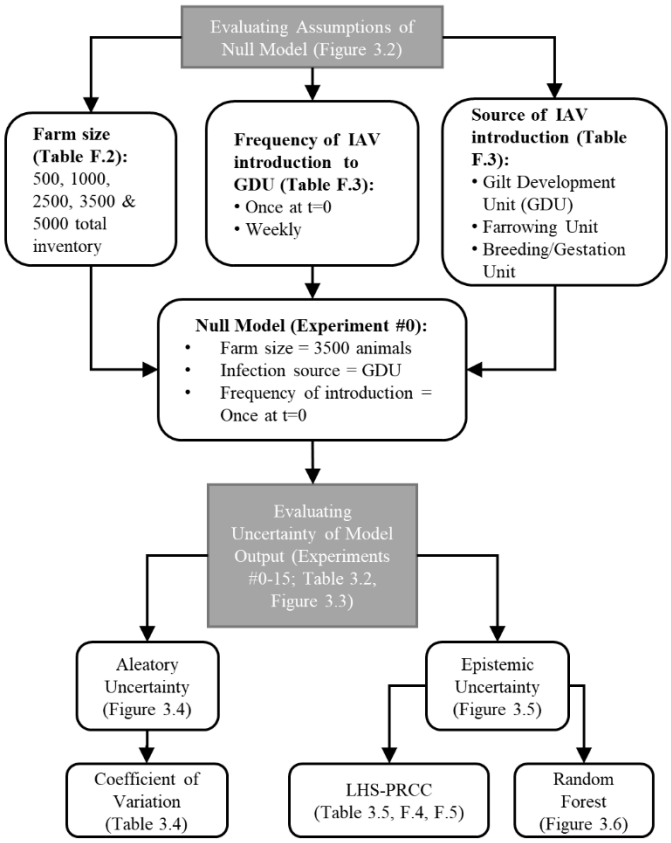


Figure 3.1. A visual depiction of the work flow and analyses included in the paper. We began by evaluating assumptions used for the null model including farm size and frequency and source of IAV introduction. When the null model proved robust to these assumptions, we explored aleatory and epistemic uncertainty in the model output. Corresponding table and figure numbers for each analysis are included.

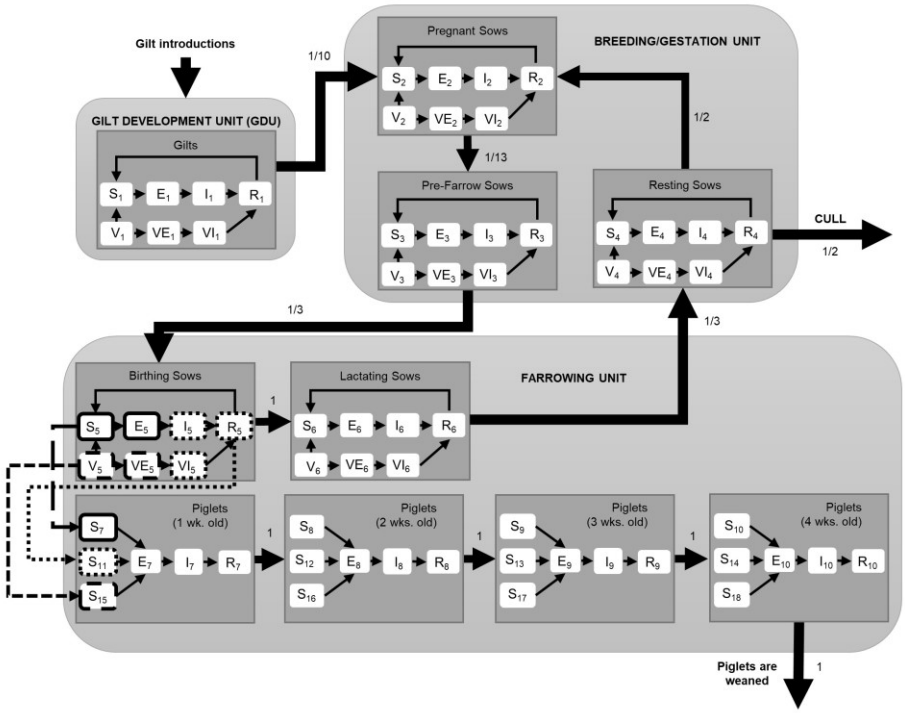


Figure 3.2. Model structure accounts for the different classes of pigs (gilts, sows, and piglets) in various production and immune stages: S-Susceptible, E-Exposed, I-Infectious, R-Recovered, V-Vaccinated, VE-Vaccinated and Exposed, VI- Vaccinated and Infected. Subscripts correspond to the classes listed in Table 3.1. Arrows leading from compartment to compartment represent the weekly probability that animals will move from one stage to the next—thus, these fractions also represent the reciprocal of the average number of weeks spent in that stage. For example, gilts spend 10 weeks in the GDU, so every week 1/10th of the gilt population moves to the breeding/gestation unit. The long-dashed, dotted, and dashed lines leading from the "Birthing Sows" to the "Piglets (1 wk. old)" compartment represent the different types of maternal immunity that sows impart to their piglets, respectively: Susceptible (S_5) and exposed sows (E_5) give birth to piglets with no maternal immunity for IAV (S_7); Infected (I_5), vaccinated-infected (VI_5), and recovered (R_5) sows give birth to piglets with maternal immunity (S_{11}); and vaccinated (V_5) and vaccinated-exposed (VE_5) sows give birth to piglets with vaccine-induced maternal immunity (S_{15}).

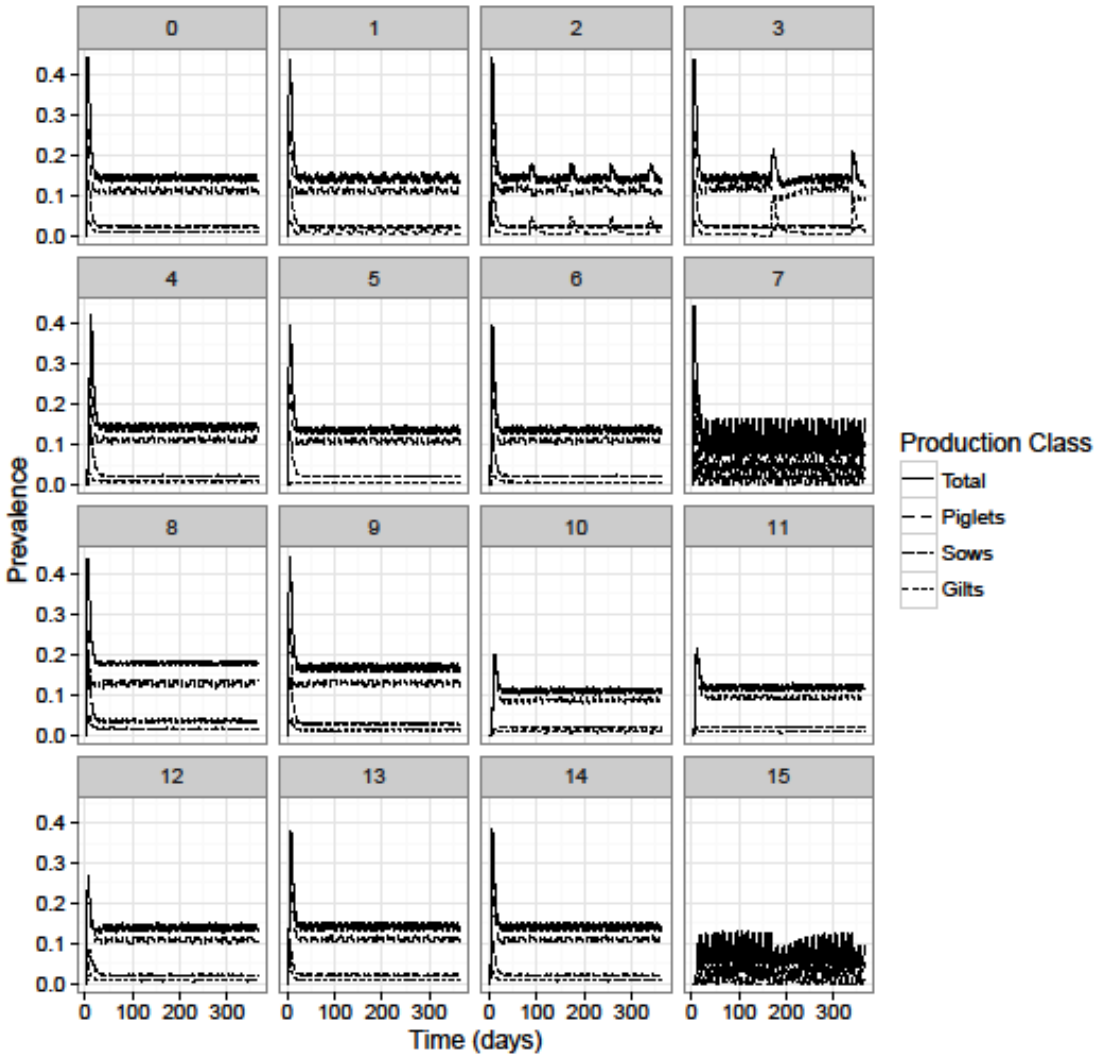


Figure 3.3. Predicted prevalence of IAV in a breeding herd for Experiments #0-15 using expected parameter values. Experiment numbers correspond to experiments described in Table 3.2. Mean prevalence of infected piglets, sows, and gilts relative to the total prevalence of infected pigs in the herd are shown—piglets experience the brunt of infection after the initial outbreak. The order of the epidemiological curves reflects the order of the legend labels in decreasing prevalence.

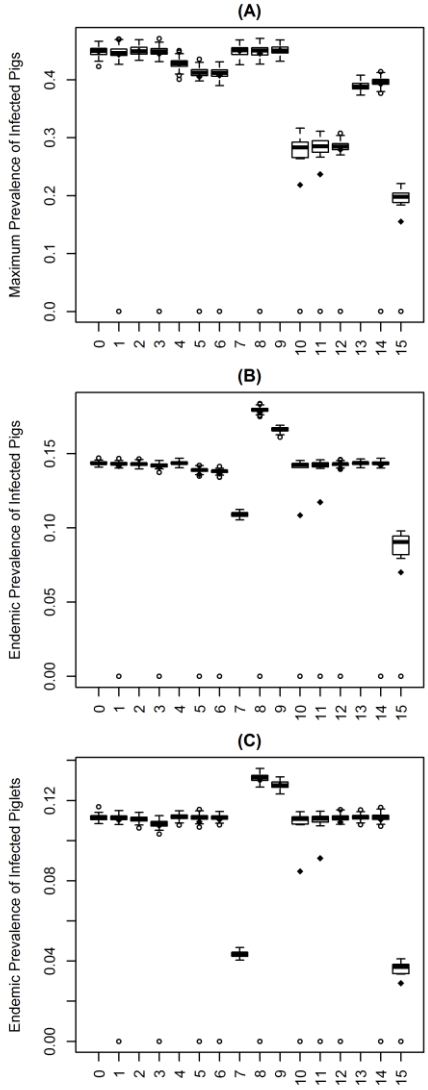


Figure 3.4. Results of the uncertainty analysis using expected values from Table 3.3 for three evaluation criteria: (A) Maximum prevalence of infected pigs per treatment; (B) Endemic prevalence of infected pigs per treatment; (C) Endemic prevalence of piglets per treatment. Box and whisker plots display for each experiment: the median, the first (25%) and third quartiles (75%), and the minimum and maximum values observed. The whiskers extend to the most extreme data point which is less than or equal to 1.5 times the interquartile (Q1-Q3) range from the box. Open circles are data points that exceed that interval. The mean for each experiment is shown by a solid diamond.

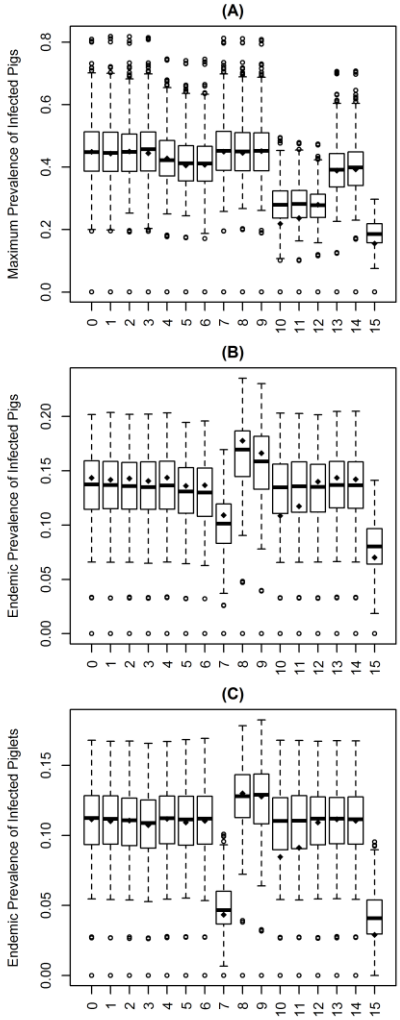


Figure 3.5. Results of the sensitivity analysis from LHS sampling of parameter distributions from Table 3.3 for three evaluation criteria: (A) Maximum prevalence of infected pigs per treatment; (B) Endemic prevalence of infected pigs per treatment; (C) Endemic prevalence of piglets per treatment. Box and whisker plots display for each experiment: the median, the first (25%) and third quartiles (75%), and the minimum and maximum values observed. The whiskers extend to the most extreme data point which is less than or equal to 1.5 times the interquartile (Q1-Q3) range from the box. Open circles are data points that exceed that interval. The mean for each experiment is shown by a solid diamond.

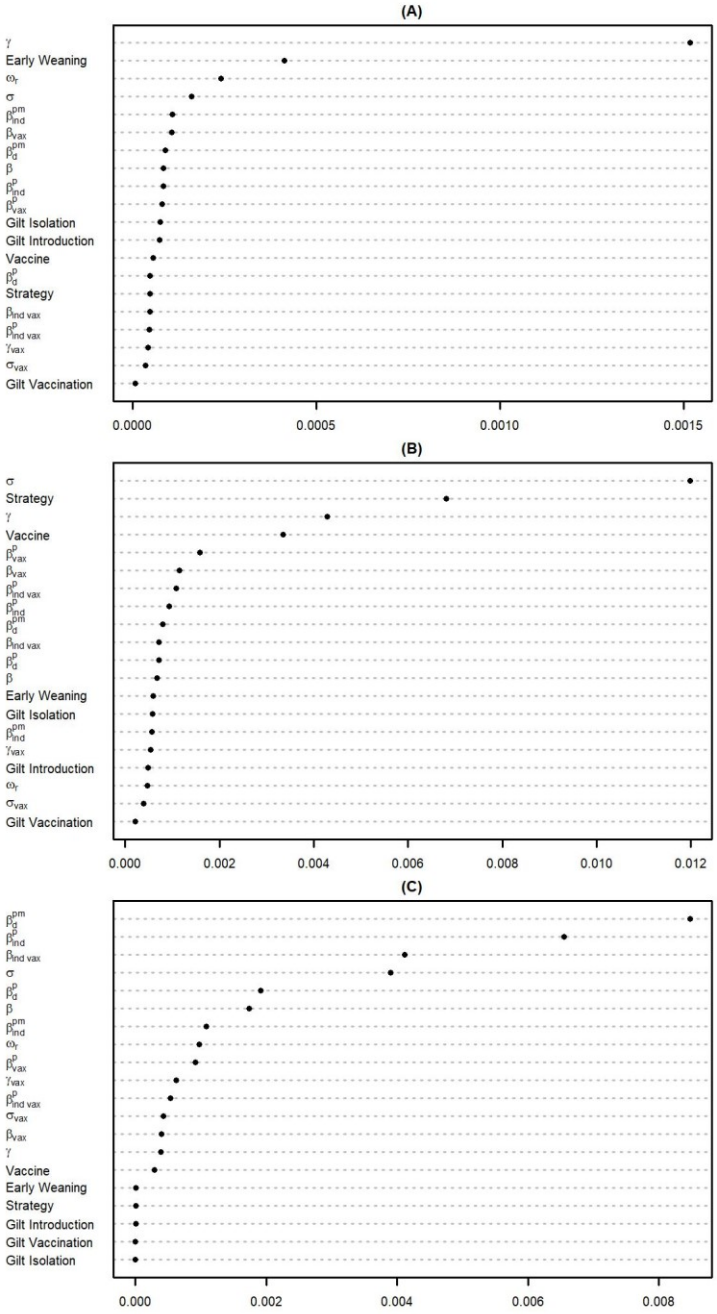


Figure 3.6. Variable importance plots from random forest analysis for three outputs: (A) Endemic prevalence of infected pigs; (B) Maximum prevalence of infected pigs; (C) Likelihood of stochastic extinction. Both (A) and (B) are regression analyses while (C) is a classification analysis. All three report variable importance results in terms of mean decrease in accuracy (values ranging from 0-1). A higher

value indicates greater variable importance. Each analysis was conducted with an ensemble of 10,000 trees. Y-axis abbreviations: Strategy = strategy of vaccination (possible values are: null, mass vaccination, or pre-farrow vaccination). Vaccine= vaccine efficacy (possible values: n/a, homologous, heterologous). Parameter symbols are defined in Table 3.3.

.

3.9 Tables

Table 3.1. Classes of animals categorized by their influenza immune status (columns: Susceptible, Exposed, Infectious, Recovered, or Vaccinated) and production stage (rows: i =1, 2, 3…). These classes correspond to the labels in Figure 3.2 and the equations in Appendix F Table F.1.

Susceptible (S_i)		**Exposed (E_i)**		**Infectious (I_i)**		**Recovered (R_i)**		**Vaccinated (V_i)**		**Vaccinated-Exposed (VE)**		**Vaccinated-Infected (VI)**	
S_1	Gilts	E_1	Gilts	I_1	Gilts	R_1	Gilts	V_1	Gilts	VE_1	Gilts	VI_1	Gilts
S_2	Pregnant sows	E_2	Pregnant sows	I_2	Pregnant sows	R_2	Pregnant sows	V_2	Pregnant sows	VE_2	Pregnant sows	VI_2	Pregnant sows
S_3	Pre-farrow sows	E_3	Pre-farrow sows	I_3	Pre-farrow sows	R_3	Pre-farrow sows	V_3	Pre-farrow sows	VE_3	Pre-farrow sows	VI_3	Pre-farrow sows
S_4	Resting sows	E_4	Resting sows	I_4	Resting sows	R_4	Resting sows	V_4	Resting sows	VE_4	Resting sows	VI_4	Resting sows
S_5	Birthing sows	E_5	Birthing sows	I_5	Birthing sows	R_5	Birthing sows	V_5	Birthing sows	VE_5	Birthing sows	VI_5	Birthing sows
S_6	Lactating sows	E_6	Lactating sows	I_6	Lactating sows	R_6	Lactating sows	V_6	Lactating sows	VE_6	Lactating sows	VI_6	Lactating sows
S_7- S_{10}	Piglets (no immunity)- wks. 1-4	E_7- E_{10}	Piglets- wks. 1-4	I_7-I_{10}	Piglets- wks. 1-4	R_7-R_{10}	Piglets- wks. 1-4						
S_{11}-S_{14}	Piglets (maternal immunity)- wks. 1-4												
S_{15}- S_{18}	Piglets (Vx-induced maternal immunity)- wks. 1-4												

Table 3.2. A description of which interventions are used for each experiment. Specifically, each column represents an experiment number, and each row represents an intervention type. Note that for each intervention type (e.g., early weaning), there are several alternatives (0-7 days, 7-14 days, etc.). Each experiment employs only one alternative from each type of intervention. A shaded "X" indicates that the particular experiment utilizes the designated intervention.

Intervention		Experiment #															
Type	**Description**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Timing of Gilt Introduction	**1 week (Null)**	X				X	X	X	X	X	X	X	X	X	X	X	
	1 month		X														
	3 months			X													
	6 months				X												X
Gilt Isolation	**No (Null)**	X	X	X	X		X	X	X	X	X	X	X	X	X	X	
	Yes					X											X
Gilt Vaccination	**No (Null)**	X	X	X	X	X			X	X	X	X	X	X	X	X	X
	Homologous							X									
	Heterologous						X										
Early weaning	**0-7 days**								X								X
	7-14 days									X							
	14-21 days										X						
	21-28 days (Null)	X	X	X	X	X	X	X				X	X	X	X	X	

Vaccination	**No vaccination**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**						
	Mass Homologous-every 2 mo.											**X**					**X**
	Mass Homologous-every 6 mo.												**X**				
	Mass Heterologous-every 2 mo.													**X**			
	Pre-Farrow Homologous														**X**		
	Pre-Farrow Heterologous															**X**	

Table 3.3. Input parameters for the model. These epidemiological and demographical parameters govern the transition rates of stochastic events in the model (described in Appendix F Table F.1). Expected values were mean values in the traditional sense for normally-distributed parameters, but could also be mode values for asymmetrical triangle distributions or simply a point estimate uniform distributions. Parameters with ranges in parentheses were varied for the LHS-PRCC sensitivity analysis.

Parameter	Description	Expected value (Range)	Distribution	Source
β_d	Direct transmission rate for unvaccinated sows & gilts	0.285 (0.091-0.9) animal^{-1}day^{-1}	Triangle	(Romagosa *et al.* 2011)
β_{ind}	Indirect transmission rate for sows and gilts	0.016 (0-0.091) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
$\beta_{d\ HE}$	Direct transmission rate for sows & gilts vaccinated with heterologous vaccine	0.0275 (0.001-0.115) animal^{-1}day^{-1}	Triangle	(Romagosa *et al.* 2011)
$\beta_{ind\ HE}$	Indirect transmission rate for sows & gilts vaccinated with heterologous vaccine	0.00017 (0-0.001) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
$\beta_{d\ HO}$	Direct transmission rate for sows and gilts vaccinated with homologous vaccine	0 (0-0.052) animal^{-1}day^{-1}	Triangle	(Romagosa *et al.* 2011)
$\beta_{ind\ HO}$	Indirect transmission rate for sows and gilts vaccinated with homologous vaccine	0 (0- 0.00029) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
β_d^p	Direct transmission rate for piglets	0.218 (0.147-0.310) animal^{-1}day^{-1}	Normal	Transmission rate for non-vaccinated treatment (Allerson *et al.* 2013c)
β_d^{pm}	Direct transmission rate for piglets with maternal immunity	0.014 (0.001-0.061) animal^{-1}day^{-1}	Triangle	(Allerson *et al.* 2013c)
β_{ind}^p	Indirect transmission rate for piglets	0.001 (0-0.147) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
β_{ind}^{pm}	Indirect transmission rate for piglets with maternal immunity	0.00008 (0- 0.001) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
$\beta_{d\ HE}^p$	Direct transmission rate for piglets whose mothers received a heterologous vaccine	0.174 (0.118-0.246) animal^{-1}day^{-1}	Normal	(Allerson *et al.* 2013c)

$\beta_{ind\ HE}^{p}$	Indirect transmission rate for piglets whose mothers received a heterologous vaccine	$9.78 \cdot 10^{-4}$ (0-0.118) animal^{-1} day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
$\beta_{d\ HO}^{p}$	Direct transmission rate for piglets whose mothers received a homologous vaccine	0.014 (0.001-0.061) animal^{-1}day^{-1}	Triangle	(Allerson *et al.* 2013c)
$\beta_{ind\ HO}^{p}$	Indirect transmission rate for piglets whose mothers received a homologous vaccine[2]	7.86 x 10^{-5} (0-0.001) animal^{-1}day^{-1}	Uniform	(Allerson et al., 2013a; Reynolds et al., 2014)
$1/\sigma$	Latency period	2.83 (1.69-3.97) days	Normal	Days between exposure and excretion for contact pigs in non-vaccinated group (Romagosa *et al.* 2011)
$1/\sigma_{HE}$	Latency period (reciprocal of exposure rate)	6.87 (2.7-11.04) days	Normal	Days between exposure and excretion for contact pigs in HE-vaccinated group (Romagosa *et al.* 2011)
σ_{HO}	Latency period (reciprocal of exposure rate)	6.87 (2.7-11.04) days	Uniform	Days between exposure and excretion for contact pigs in HO-vaccinated group (Romagosa *et al.* 2011); Note: no infectious HO pigs were observed in this experiment, so we tested a uniform distribution based on the HE-vaccinated group.
$1/\gamma$	Infectious period (Reciprocal of recovery rate)	4.5 (3.43-5.57) days	Normal	Infectious period from contact pigs in non-vaccinated group (Romagosa *et al.* 2011)
$1/\gamma_{HE}$	Infectious period (Reciprocal of recovery rate)	3.50 (1.66-5.34) days	Normal	Infectious period from contact pigs in HE-vaccinated group (Romagosa *et al.* 2011)
$1/\gamma_{HO}$	Infectious period (Reciprocal of recovery rate)	0 (1.66-5.34) days	Uniform	Infectious period from contact pigs in HO-vaccinated group (Romagosa *et al.* 2011); Note: no infectious HO pigs were observed in this experiment, so we tested a uniform distribution based on the HE-vaccinated group.
b	Probability of giving birth on a given day for birthing sows	1/7 day^{-1}	Fixed	Birthing sows have a period of one week to give birth, but they may give birth on any day of that week.
p	Number of piglets born alive per litter	12.29 ± 0.71 piglets	Normal- not varied in sensitivity analysis	(PigCHAMP 2014)
μ	Natural death rate for sows and gilts- not due to culling	0.0004 day^{-1}	Fixed	(Reynolds *et al.* 2014)
μ^{P}	Natural death rate for piglets	0.005 day^{-1}	Fixed	(Reynolds *et al.* 2014)

$1/\omega_r$	Duration of immunity rate for recovered sows & gilts (Inverse of loss of immunity rate)	56 (56-112) days	Uniform	Expert opinion M. Torremorell- immunity can be expected to last for 8-16 weeks for the same strain
$1/\omega_v$	Duration of immunity for vaccinated sows & gilts (Inverse of loss of immunity rate)	56 days	Fixed	Immunity can be expected to last at least 8-10 weeks for common vaccines (Pfizer Animal Health 2011; Merck Animal Health 2013)

Table 3.4. Comparison of model output for expected values of the model. For each experiment (#0-15), one hundred simulations were conducted using the expected values in Table 3.3. Since only the stochastic elements of the model are varying rather than the parameters, this characterizes the amount of aleatory uncertainty inherent to the model. Peak prevalence, endemic prevalence, endemic prevalence in piglets, probability of IAV extinction, and time until stochastic extinction are reported.

Experiment	Maximum Prevalence			Endemic Prevalence (Total Inventory)			Endemic Prevalence (Piglets)			IAV Elimination	
	Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Percent Extinction	Time to Extinction
0	0.4497	0.4490 ± 0.0078	1.7419	0.143	0.144 ±0.001	0.806	0.112	0.112 ± 0.001	1.321	0%	n/a
1	0.4464	0.4435 ± 0.0454	10.2292	0.143	0.142 ± 0.014	10.142	0.111	0.110 ± 0.011	10.191	1%	1 day
2	0.4494	0.4502 ± 0.0083	1.8380	0.143	0.143 ± 0.001	0.931	0.111	0.111 ± 0.001	1.264	0%	n/a
3	0.4491	0.4447 ±0.0454	10.1997	0.142	0.141 ± 0.014	10.148	0.108	0.107 ± 0.011	10.218	1%	1 day
4	0.4282	0.4282± 0.0082	1.9215	0.144	0.144 ± 0.001	0.862	0.112	0.112 ± 0.001	1.244	0%	n/a
5	0.4123	0.4051± 0.0583	14.4012	0.112	0.109 ± 0.016	14.421	0.112	0.109 ± 0.016	14.421	2%	1 day
6	0.4119	0.4076 ± 0.0417	10.2192	0.138	0.137 ± 0.014	10.142	0.111	0.110 ± 0.011	10.167	1%	1 day
7	0.4508	0.4495 ± 0.0086	1.8328	0.109	0.109 ± 0.002	1.557	0.043	0.043 ± 0.001	3.328	0%	n/a
8	0.4495	0.4489 ± 0.0082	1.8918	0.179	0.178 ± 0.018	10.151	0.131	0.130 ± 0.013	10.216	1%	1 day
9	0.4505	0.4511 ± 0.0075	1.6558	0.166	0.166 ± 0.002	0.936	0.128	0.128 ± 0.002	1.374	0%	n/a
10	0.2834	0.2185 ± 0.1230	56.2934	0.143	0.109 ± 0.061	56.485	0.111	0.085 ± 0.048	56.495	24%	1.25 days
11	0.2854	0.2370 ± 0.1113	46.9461	0.143	0.117 ± 0.055	47.098	0.111	0.091 ± 0.043	47.109	18%	1.17 days
12	0.2847	0.2797 ± 0.0406	14.5189	0.143	0.140 ± 0.020	14.384	0.112	0.109 ± 0.016	14.413	2%	1 day
13	0.3881	0.3887 ± 0.0069	1.7771	0.144	0.144 ± 0.001	0.859	0.111	0.112 ± 0.001	1.197	0%	n/a
14	0.3967	0.3926 ± 0.0401	10.2158	0.143	0.142 ± 0.014	10.137	0.112	0.110 ± 0.011	10.205	1%	1 day
15	0.1982	0.1554 ± 0.0850	54.7137	0.090	0.070 ± 0.039	55.282	0.037	0.029 ± 0.016	55.173	23%	1.26 days

Table 3.5. Partial rank correlation coefficient (PRCC) between all 15 parameters (rows) and the endemic prevalence of IAV for each experiment (columns). PRCC for the other two prevalence criteria are included in Appendix F (Table F.2 & Table F.3). Values closest to 1 or -1 indicate the strongest correlations, while values close to zero indicate little to no correlation. Bolded values have a magnitude ≥ 0.30. Shaded areas represent parameters that were not used/relevant for a particular experiment (e.g., the direct transmission rate for vaccinated animals was not employed in the null case [Experiment 0]).

Parameter	**#0**	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**	**#13**	**#14**	**#15**
β	0.14	0.14	0.14	0.13	0.14	0.13	0.15	0.19	0.18	0.15	0.16	0.15	0.17	0.12	0.14	0.14
β_{ind}	0.04	0.04	0.03	0.02	0.01	0.01	0.00	0.08	0.07	0.05	0.06	0.18	0.04	0.04	0.03	0.08
β_d^p	-0.10	-0.11	-0.11	-0.11	-0.10	-0.04	-0.10	-0.18	-0.13	-0.11	-0.21	-0.15	-0.10	-0.09	-0.11	0.02
β_{ind}^p	-0.08	-0.08	-0.07	-0.08	-0.06	-0.07	-0.07	-0.11	-0.12	-0.09	0.02	-0.07	-0.10	-0.06	-0.08	0.04
β_d^{pm}	-0.14	-0.13	-0.14	-0.14	-0.14	-0.14	-0.14	**0.31**	-0.07	-0.13	-0.06	-0.06	-0.11	-0.15	-0.14	**0.38**
β_{ind}^{pm}	-0.05	-0.06	-0.05	-0.05	-0.04	-0.08	-0.07	-0.03	-0.01	-0.04	-0.10	-0.05	-0.01	-0.08	-0.05	-0.10
σ	0.13	0.13	0.13	0.14	0.14	0.13	0.14	**0.65**	**0.40**	0.20	0.17	0.15	0.15	0.13	0.13	**0.42**
γ	**-0.80**	**-0.80**	**-0.80**	**-0.80**	**-0.79**	**-0.79**	**-0.80**	**-0.71**	**-0.78**	**-0.80**	**-0.71**	**-0.76**	**-0.80**	**-0.80**	**-0.80**	**-0.58**
β_{vax}						-0.15	-0.16				0.02	-0.08	-0.18	-0.13	-0.16	0.10
β_{vax}^p						-0.19	-0.20				-0.22	-0.16	-0.23	-0.19	-0.23	-0.08
$\beta_{ind\ vax}$						0.07	0.09				0.09	0.11	0.11	0.07	0.08	0.11
$\beta_{ind\ vax}^p$						0.12	-0.16				0.12	0.18	-0.12	0.10	-0.16	0.13
ω_r	0.13	0.13	0.13	0.13	0.10	0.12	0.09	**0.48**	0.23	0.17	0.25	0.17	0.11	0.14	0.13	**0.36**
σ_{vax}						-0.15	-0.07				-0.18	-0.17	-0.09	-0.14	-0.08	-0.17
γ_{vax}						0.15	0.08				0.18	0.18	0.08	0.14	0.06	0.16

Chapter 4. Disease outbreak thresholds emerge from interactions between movement behavior, landscape structure, and epidemiology

Lauren A. White, James D. Forester & Meggan E. Craft

Keywords: spatial heterogeneity, landscape fragmentation, disease model, individual-based model, resource selection function, perceptual range

4.1 Abstract

Disease models have provided conflicting evidence as to whether spatial heterogeneity promotes or impedes pathogen persistence. Moreover, there has been limited theoretical investigation into how animal movement behavior interacts with the spatial organization of resources (e.g., clustered, random, uniform) across a landscape to affect infectious disease dynamics. Importantly, spatial heterogeneity of resources can sometimes lead to non-linear or counterintuitive outcomes depending on the host and pathogen system. There is a clear need to develop a general theoretical framework that could be used to create testable predictions for specific host-pathogen systems. Here, we develop an individual-based model integrated with movement ecology approaches to investigate how host movement behaviors interact with landscape heterogeneity (in the form of various levels of resource abundance and clustering) to affect pathogen dynamics. For most of the parameter space, our results support the counterintuitive idea that fragmentation promotes pathogen persistence, but this finding was largely dependent upon perceptual range of the

host, conspecific density, and recovery rate. For simulations with high conspecific density, slower recovery rates, and larger perceptual ranges, more complex disease dynamics emerged, and the most fragmented landscapes were not necessarily the most conducive to outbreaks or pathogen persistence. These results point to the importance of interactions between landscape structure, individual movement behavior, and pathogen transmission for predicting and understanding disease dynamics.

4.2 Significance Statement

Understanding how emerging infectious and zoonotic diseases spread through space and time is critical for predicting outbreaks and designing interventions; disease models are important tools for realizing these goals. Currently, humans are altering the environment in unprecedented ways through urbanization, habitat fragmentation, and climate change. However, the consequences of increasingly heterogeneous landscapes on pathogen transmission and persistence remain unclear. By synthesizing mathematical modeling and movement ecology approaches, we examined how wildlife movement patterns interact with broad-scale, landscape structure to affect population-level disease dynamics. We found that habitat fragmentation could counterintuitively promote disease outbreaks, but that for higher wildlife densities and longer infectious periods, small differences in how hosts navigated their environments could dramatically alter observed disease dynamics.

4.3 Introduction

Spatial heterogeneity—differences occurring across a geographic landscape—may arise from intrinsic differences between locations (e.g., resource abundance, quality, connectivity) or can emerge from stochastic or dynamic processes within populations (e.g., demographics, conspecific density) (Keeling & Rohani 2008a). The majority of disease models that incorporate spatial heterogeneity have focused primarily on a few well-studied wildlife systems (e.g., rabies and bovine tuberculosis [bTB]) or have been conducted in a purely theoretical context (White, Forester & Craft 2018b). Theoretical and simulation studies have provided evidence both for and against spatial heterogeneity promoting pathogen persistence (Hagenaars, Donnelly & Ferguson 2004; Rees *et al.* 2013; Tracey *et al.* 2014), and the relative importance of local and long-distance processes is often unknown, except for in some well-studied diseases like rabies (White, Forester & Craft 2018b). Importantly, spatial heterogeneity may lead to non-linear or counterintuitive outcomes depending on the host and pathogen system (Rees *et al.* 2013; Tracey *et al.* 2014).

Foraging, migration, and dispersal play an important role in creating spatial heterogeneity (Lloyd-Smith 2010). Host movement and dispersal patterns can vary considerably (Brown & Crone 2016), and infection with parasites can further alter those patterns (Welicky & Sikkel 2015). Both perceptual range (how far an individual can perceive habitat, in order to be able to make movement choices) and movement capacity (the ability and efficiency with which an individual can move) can affect the realized connectivity of habitat patches in heterogeneous landscapes (Lima & Zollner 1996).

While some models have explored the sensitivity of disease dynamics to dispersal and migration rates (White *et al.* 2018b), few studies compare the effects of different movement rules over a spatially-explicit landscape on pathogen transmission [but see: (Lane-deGraaf *et al.* 2013; Tracey *et al.* 2014)]. Moreover, disease models with mechanistic representations of animal movement remain rare (Fofana & Hurford 2017). In the realm of movement ecology, resource selection functions (RSFs) models have fostered a better understanding of how organisms navigate their surroundings (Dougherty *et al.* 2018; White *et al.* 2018b). While several studies have utilized RSFs to estimate inter-species contact risk, RSFs are not commonly used to infer transmission events (White *et al.* 2018b). Moreover, cues such as conspecific density can also be important drivers of individual movement decisions, but are rarely utilized when modeling habitat selection (Campomizzi *et al.* 2008). In addition, it is difficult to validate movement model predictions in the context of pathogen transmission because overlapping movement and pathogen transmission datasets are still uncommon (Dougherty *et al.* 2018). Many existing spatial disease models that explicitly incorporate individual movement rely on random or correlated random walks (Tracey *et al.* 2014; Fofana & Hurford 2017). While adequate for some species at specific temporal scales, these approaches do not necessarily capture an individual's response to its immediate surroundings or its memory that might favor revisitation or avoidance of previous sites (Smouse *et al.* 2010; Oliveira-Santos *et al.* 2016).

Overall, we lack a mechanistic understanding of how host movement and habitat preferences across heterogeneous landscapes affect pathogen dynamics. Here we asked:

how does individual movement behavior, governed by perceptual range and individual selection for resource availability and conspecific density, interact with spatial heterogeneity (via resource availability and clustering) to affect infectious disease dynamics? Specifically, we examined outbreak behavior through two questions: (*i*) which epidemiological, movement, or landscape factors led to a successful outbreak (defined as spreading beyond the initially infected individual)?; and, (*ii*) given at least one secondary case, which factors best predicted maximum prevalence and duration? We developed an individual-based, susceptible-infected-recovered (SIR) model for a theoretical, host-pathogen system where a resource selection function governed host movement choices and pathogens were directly transmitted, assuming a density-dependent transmission function. We varied both landscape structure and movement parameters, and we quantified disease dynamics by the maximum prevalence and the duration of the outbreak. In addition, we compared our spatially-explicit model output to a comparable SIR model that assumes homogeneous mixing. We propose a general, theoretical framework to generate testable predictions for specific host-pathogen systems existing on complex landscapes.

4.4 Results

Which factors determine whether an outbreak is successful? The outbreak data are heavily skewed—most initial pathogen introductions never transmitted successfully to a second individual. Based on random forest analysis, recovery rate (γ) had the largest effect on whether there were any secondary cases beyond the initially

infected individual (Figure 4.1*A*). While conspecific density (*d*) was the next most influential parameter, it had less than one-fourth of the mean decrease in accuracy of recovery rate, suggesting that pathogen infectious period $(1/\gamma)$ played an outsized role in determining whether an outbreak resulted in secondary cases (Figure 4.1*A,* Appendix G Figure G.1*A*).

Given secondary cases, which parameters most influence maximum prevalence and duration? If an outbreak spread successfully beyond the initially infected individual, the interaction of recovery rate, perceptual range, and conspecific density had strong effects on both maximum prevalence and duration for the simulations (Figure 4.2). This finding was supported by the random forest analysis which identified recovery rate, perceptual range, and conspecific density as the top three parameters in predictive value for both maximum prevalence and duration (Figure 4.1*B* and *C,* Appendix G Figure G.1*B* and *C*). Overall, simulations with faster recovery rates ($\gamma = 0.4$), lower conspecific densities (d=0.25), and smaller perceptual ranges (r=1, 2) had fewer successful outbreaks, and outbreaks that were successful reached fewer individuals and were shorter in duration (Figure 4.2 and Appendix G Figure G.2).

Beyond these top three influential parameters, the order of variable importance as identified by random forest analysis differed for maximum prevalence and duration. For maximum prevalence, selection for resources (β_1) and proportion of available habitat (p) outweighed strength of selection for conspecifics (β_2 & β_3) and degree of patchiness (H) (Figure 1*B* and Appendix G Figure G.1*B*). In contrast, for mean duration, the strength of

selection for conspecifics (β_2 & β_3) outweighed strength of selection for resources (β_1) and landscape structure (p & H) (Figure 4.1*C* and Appendix G Figure G.1*C*).

Effects of landscape structure and individual movement behavior. When holding the influential parameters of recovery rate, perceptual range, and conspecific density constant, interactions between landscape structure and individual movement behavior emerged. For simulations with lower conspecific density (d = 0.25) and faster recovery rates ($\gamma = 0.4$), epidemics were rarely successful (Appendix G Figure G.3 *E* & *F*). For simulations with lower conspecific density (d = 0.25) and slower recovery rates ($\gamma = 0.1$ & 0.2), more fragmented landscapes ($H \leq 0.5$) with lower resource availability ($p \leq 0.5$) exhibited larger outbreak size and lasted longer for more RSF combinations (Appendix G Figure G.3 *A*-*D*); this same pattern was observed for systems with both higher conspecific density (d = 0.50) and faster recovery rates ($\gamma = 0.2, 0.4$) (Appendix G Figure G.4 *C*-*F*). In general, positive selection for resources ($\beta_1 = 3, 6$) led to higher outbreak peaks and longer lasting outbreaks compared to random selection for habitat ($\beta_1 = 0$) (Figure 4.3, Appendix G Figure G.5). However, for simulations with higher conspecific densities (d = 0.5), slower recovery rates ($\gamma = 0.1$) and higher perceptual ranges (r = 3), we observed more complex dynamics (Figure 4.3, and Appendix G Figure G.4 *A* and *B,* Figure G.6) where RSFs interacted with landscape structure to determine epidemic dynamics (Figure 4.4).

For the parameter space exhibiting a higher proportion of successful outbreaks across RSFs (r=3, γ= 0.1, d=0.5), maximum prevalence was higher across more RSFs for more

clustered habitat (H=0.1) with lower resource availability (p=0.25) (Figure 4.3*A*). However, this did not necessarily correlate with duration. While outbreaks lasted longer on average for more RSFs in fragmented landscapes (H=0.1 and p=0.25), outbreaks lasted longer at intermediate clustering (H=0.5) for certain RSFs (Figure 4.3*B,* Figure 4.4*B*). Overall, when hosts exhibited strong positive selection for both conspecifics ($\beta_1 = 3, 6$ & $\beta_2 = 1,2$) and resources, disease outbreaks had the largest variation in duration for patchy landscapes (H=0.1) with an intermediate proportion of available habitat (p=0.5) (Figure 4.3*B*). However, regardless of landscape structure, stronger selection for the presence of other conspecifics (e.g., $\beta_2 = 2$) reliably increased the observed duration of outbreaks (Figure 4.3*B*). This threshold behavior for selection for conspecifics—where selection for conspecifics supersedes landscape structure—was not observed for landscapes with lower conspecific densities and faster recovery rates (Appendix G Figure G.3 and Figure G.5).

How do these results compare to models that assume homogenous mixing?

Compared to the spatially-explicit movement model described here, the equivalent homogenous mixing model with density-dependent transmission consistently overestimated the maximum prevalence reached for both conspecific densities and all three simulated recovery rates (Appendix G Figure G.7*A*). In particular, for the higher perceptual range (r=3), slowest recovery rate ($\gamma = 0.1$), and higher conspecific density treatment (d= 0.5), the homogeneous mixing model did not capture the skewed nature of observed epidemics' duration (Appendix G Figure G.7*B*).

4.5 Discussion

We were expecting landscape structure to have a substantial impact on simulated disease dynamics. While this was true to some extent, two of the top three covariates in variable importance as determined by random forest analysis (i.e., recovery rate and conspecific density) had nothing to do with landscape structure per se. However, perceptual range played a key role in determining maximum prevalence and duration of outbreaks (Figure 4.1*B* and *C*), and perceptual range functionally defines the landscape that an individual host perceives. Other modeling studies have verified the potential importance of perceptual range in determining landscape connectivity (Pe'er & Kramer-Schadt 2008), and our results emphasize the concomitant implications for pathogen spread and persistence. This suggests that the incorporation of plastic perceptual ranges may also be important for future disease models (Olden *et al.* 2004).

When holding these top three influential parameters (recovery rate, conspecific density, and perceptual range) constant, fragmentation promoted pathogen outbreaks and persistence for most of the explored parameter space, particularly for simulations with combinations of (*i*) lower conspecific densities and slower recovery rates (Appendix G Figure G.3*A*-*D*) or (*ii*) higher conspecific densities and faster recovery rates (Appendix G Figure G.4*C*-*F*). However, this pattern was highly dependent upon hosts being able to perceive more of their habitat to be able to make movement decisions (e.g., a larger perceptual range) (Appendix G Figure G.3 and Figure G.4). Positive selection for resources was also necessary to elicit differences in pathogen spread in response to landscape structure, particularly at faster recover rates (Figure 3 and Appendix G Figure

G.5). In an applied setting, these results highlight the potential role of resource hotspots and resource provisioning in altering not only animal movement patterns, but subsequent pathogen transmission (Becker & Hall 2016). Resource hotspots can occur naturally, e.g., carcasses acting as landscape hotspots for transmission of rabies in jackals (Borchering *et al.* 2017), or artificially, through human supplementation, e.g., *Mycoplasma gallisepticum* transmission at bird feeders (Dhondt *et al.* 2007) or brucellosis transmission from supplemental feeding of elk in Yellowstone National Park (Cross *et al.* 2007a). While such selection could apply to foraging choices, it could also apply to selection for burrows or dens. For example, in desert tortoises (*Gopherus agassizii*), contacts primarily occurring in underground burrows are thought to drive transmission of *Mycoplasma agassizii* (Aiello *et al.* 2016).

At higher perceptual range (r=3), slower recovery rates ($\gamma = 0.1$), and higher conspecific densities (d=0.5), we captured more nuanced and complex behavior that resulted from interactions between landscape structure, movement behavior, and recovery rate (Figure 4.3 and Figure 4.4). Notably, a comparable homogenous-mixing, density-dependent SIR model did not capture the skewed distribution of epidemic duration found for simulations in this parameter space (Appendix G Figure G.7*B*). Outbreaks reaching the most individuals generally occurred in more fragmented landscapes ($H = 0.1, p = 0.1$) (Figure 4.3*A*), but outbreaks in patchy/medium proportion habitat landscapes ($H = 0.1, p = 0.5$) lasted longer for some RSFs and exhibited more variation in observed duration (Figure 4.3*B*, Figure 4.4*B*, Appendix G Figure G.6*F*). Also, threshold behavior was exhibited for selection for conspecifics in this parameter space; very strong selection

for conspecifics ($\beta_2 = 2$) promoted longer lasting outbreaks with higher maximum prevalence regardless of landscape structure (Figure 4.3). This was interesting because relatively small differences in RSF values resulted in substantially different disease dynamics (Appendix G Figure G.8). For such regimes, we suggest that it may be important to model individual responses to landscape structure in order to better capture the dynamics of a given disease.

The model presented here best describes direct transmission of a single infectious agent (or limited indirect transmission, as defined by aerosolized transmission or limited fomite persistence relative to movement timesteps) within a single host species experiencing density-dependent transmission. Thus, these results are applicable to host-pathogen systems that have previously been favored in a spatial modelling context including rabies and bTB (Dougherty *et al.* 2018; White *et al.* 2018b). For example, a recent model of raccoon rabies found that inadequate levels of vaccination in continuous, poor-quality habitat could prove counterproductive, leading to outbreaks (Rees *et al.* 2013). Our findings are also relevant to emerging pathogens ranging from Ebola or respiratory viruses among primate species to bat-to-bat transmission of Hendra virus (Plowright *et al.* 2011; Rushmore *et al.* 2013). For example, a recent spatially-structured model for Hendra virus in fruit bats found that habitat loss led to congregation in urban roosting sites and reduced migration, which could aid in disease persistence and in spillover to humans (Plowright *et al.* 2011). Since conspecific density played a key role in determining the relationship between outbreak success and fragmentation, this work might be particularly relevant to wildlife populations where host densities vary widely

through time [e.g., Moepia virus or Hantavirus in rodents where direct transmission via agonistic interactions are known to be important (Clay *et al.* 2009; Goyens *et al.* 2013)].

We recognize that many host-pathogen systems experience more complex transmission cycles than represented by this model; we would expect dynamics to differ with the incorporation of demographic processes (births and deaths), disease related mortality from a more lethal pathogen (e.g., rabies), a substantial incubation period, or chronic infection (e.g., bTB). While this work does not explicitly address the potential effects of pathogen co-infection (Knowles *et al.* 2013), multi-host pathogens [e.g., canine distemper virus (Craft *et al.* 2008)], indirect transmission or environmental persistence [e.g., chronic wasting disease (Almberg *et al.* 2011)], or vector foraging behavior [e.g., Lyme disease (Li *et al.* 2016)], the results of this research could easily be extended to more biologically complex systems.

Future studies could build in additional landscape complexity or focus on more realistic mechanisms governing movement choices. For instance, we modeled pathogen transmission across theoretical, binary landscapes. Realistically, movement choices are influenced by a variety of landscape covariates, not just presence or absence of resources (Forester, Im & Rathouz 2009). Similarly, we considered each cell in the landscape to be equally permeable—other than relative distance from the host, there was no cost to traversing bad relative to good quality habitat. In addition, we did not model depletion of resources explicitly, but rather through reaction of individuals to other conspecifics (β_2 & β_3). This helped us limit the number of assumptions built into the model, but certainly simplifies the proximate mechanisms of foraging behavior (Rands *et al.* 2004).

Finally, for a given simulation, all individuals responded the same way to the presence of resources and conspecifics. We note that for many systems there will be individual variability in foraging and social behavior, and these differences may merit further consideration for some host-pathogen systems (Klein *et al.* 2017). For instance, a recent study of raccoons and contact risk for rabies determined that a subset of individuals were responsible for the majority of risky contacts despite all individuals following the same movement rules (Tardy *et al.* 2018). Finally, we ignored the fact that animal movement behavior can change with infection (Welicky & Sikkel 2015).

A recent systematic review investigating the relationship between pathogen transmission and anthropogenic land-use change found that a majority of studies linked anthropogenic change with higher transmission risk (Gottdenker *et al.* 2014). Overall, our work has important implications for how pathogens spread across fragmented and human-influenced landscapes and supports other modelling studies that have suggested that fragmentation can have non-linear effects on pathogen persistence in specific host-pathogens systems (Bonnell *et al.* 2010; Rees *et al.* 2013; Tracey *et al.* 2014) and that spatial hotspots of transmission can emerge from limited high quality resource sites (Bonnell *et al.* 2010; Benavides *et al.* 2012; Nunn, Thrall & Kappeler 2014). This work provides a theoretical, mechanistic framework that can be expanded to specific host-pathogen systems to provide testable hypotheses about the influence of landscape structure and movement behavior on disease dynamics, thus providing a critical bridge between the disciplines of movement and disease ecology (Fofana & Hurford 2017; Dougherty *et al.* 2018; White *et al.* 2018b). We hope this model inspires additional

consideration of how landscape structure may influence disease dynamics and foster investigation of these questions in specific host-pathogen systems and applied management settings.

4.6 Methods

We developed a stochastic, individual-based, Susceptible Infectious Recovered (SIR) model for a non-lethal pathogen in a closed population (i.e., no births, deaths, immigration, or emigration). Individuals could move across a spatially explicit, discrete lattice landscape where resource presence or absence was variable in space (Appendix G Figure G.9). Using the mid-point displacement algorithm, we generated theoretical, binary, neutral landscapes that varied in the proportion of available habitat ($\boldsymbol{p}$) and the degree of habitat clumping (Hurst exponent, $\boldsymbol{H}$) (Turner, Gardner & O'Neill 2001). We assumed a torus shape (i.e., wrapped boundaries) for these landscapes to avoid edge effects (With 1997).

A resource selection function (RSF) and a two-dimensional movement kernel modulated individual movement choices across these theoretical landscapes. For a habitat of $\boldsymbol{k = 1, \ldots, m}$ discrete grid cells, the probability of an individual moving from current location, $\boldsymbol{a}$, to new location, $\boldsymbol{b}$, over a fixed temporal time step is: $\mathbb{P}(\boldsymbol{a\ to\ b}) = \boldsymbol{\phi(a,b)w_b / \sum_{k=1}^{m}[\phi(a,c_k)w_k]}$ where $\boldsymbol{\phi}(\cdot)$ is a two-dimensional movement kernel in the absence of habitat selection, $\boldsymbol{c_k}$ represents the center point of each grid cell, and $\boldsymbol{w_b}$ and $\boldsymbol{w_k}$ are RSFs governing an individual's movement preferences for cells $\boldsymbol{b}$ and $\boldsymbol{k}$.

Generally, an RSF for any given cell $\boldsymbol{j}$, takes the form: $\boldsymbol{w_j = exp(\beta_1 \cdot R_j + \beta_2 \cdot N_j + \beta_3 \cdot N_j^2)}$. The parameters, $\boldsymbol{R_j}$ and $\boldsymbol{N_j}$, correspond to the resource quality and number of conspecifics respectively in cell $\boldsymbol{j}$. The coefficients $\boldsymbol{\beta_1}, \boldsymbol{\beta_2}$, and $\boldsymbol{\beta_3}$ govern the strength of selection that resource quality and the number of conspecifics play in habitat choice (Forester *et al.* 2009). In particular, because conspecific density changes through time across a static resource landscape, the inclusion of the quadratic term $\boldsymbol{\beta_3}$ allows us to test three biologically feasible scenarios in response to the presence of conspecifics: (*i*) individuals avoid conspecifics; (*ii*) individuals are attracted to conspecifics (signaling good quality habitat) until there are trade-offs with resource depletion; (*iii*) individuals make movement choices irrespective of conspecifics (for more on how we calibrated values for the RSF, please see Appendix G). We began by assuming a Moore neighborhood (8 neighboring cells), but subsequently tested the effects of larger, continuous perception and relocation kernels (Table 4.1). We assumed the simplest case where the movement kernel is inversely proportional to radial distance from the center point of the current grid cell and acts in the absence of resource availability: $\boldsymbol{\phi(r) = \frac{1}{2\pi r^2}}$, where $r = \sqrt{\left(x_a - x_{c_k}\right)^2 + \left(y_a - y_{c_k}\right)^2}$

The equation gives an inverse distance weight (i.e., $\frac{1}{r}$) that is multiplied by the circumference at that distance to account for a uniform circular distribution (i.e., $\frac{1}{2\pi r}$). We conducted 100 replications per parameter set, and we used a factorial design to explore the relative effects of recovery rate, conspecific density, landscape structure, resource selection functions, and perceptual range on disease dynamics (Table 4.1 and

Appendix G Figure G.9). We calibrated the transmission probability (β) in order to be able to test reasonable values of the basic reproductive rate, R_0 (for an SIR model, $R_0 \sim \beta/\gamma$). Under that framework, we tested vales of $R_0 \sim 0.5, 1, \text{and } 2$ in our simulations (Table 4.1, Appendix G Text).

Each simulation began with a single infectious individual and continued until there were no remaining infectious individuals on the landscape. At each time step, every individual in the simulation had the opportunity to evaluate their surrounding environment and move to a cell within their perceptual range. Once all individuals had been given the opportunity to relocate, the possibility of transmission was evaluated; we assumed that transmission could occur only between individuals of the same cell during the same time step. The probability of a susceptible individual becoming infected was represented by: $\mathbb{P}(T) = 1 - (1 - \beta)^I$ where β is the transmission rate and I is the number of infectious individuals in the same cell. Following potential transmission events, infected individuals could also recover with a probability, γ, at each time step. This corresponded to an average infectious period of the pathogen of 1/ γ (Table 4.1).

We compared the spatially-explicit model results with a simpler stochastic SIR model that assumes density-dependent transmission and homogeneous mixing. This was simulated with a Reed Frost model where cumulative probability of transmission during at time step, τ, is equal to $\mathbb{P}_\tau = 1 - (1 - \beta)^{I_\tau}$ where β equals the per contact transmission risk and I_τ equals the number of infectious individuals at time step, τ. The number of infected individuals at the next time step is then given by: $I_\tau = Binomial\ (S_\tau, \mathbb{P}_\tau)$. Unlike for the spatially-explicit simulations, this probability was

evaluated for the entire population, not just a single cell in the landscape. The results from this simple stochastic simulation were verified with the output from comparable deterministic ordinary differential equations (Appendix G Text).

Finally, we used random forest analysis—a machine learning method—to tease apart the relative contributions of parameters to outbreak outcomes. As a recursive partitioning method, random forest analysis fits a single predictive model by synthesizing the predictions from numerous classification or regression trees (Breiman 2001; Cutler *et al.* 2007). The random forest approach has several advantages for ecological data; most notably, this approach can handle complicated, non-linear, and potentially collinear relationships between predictor variables (Breiman 2001; Cutler *et al.* 2007). This approach also avoids some of the pitfalls of using a frequentist approach to analyze simulation results, since sample size in simulation studies is arbitrary and can result in significant p-values regardless of effect size (White *et al.* 2014). Other disease model studies have used this approach to better understand complex data with multiple predictors (Herrick *et al.* 2013; Kane *et al.* 2014).

Variable importance measures from random forest analysis describe the relative role that a covariate plays in deciding model outcomes (Cutler *et al.* 2007). We used the *randomForest* package in R (Liaw & Wiener 2002b) to calculate variable importance scores so we could understand what factors affected three separate response variables: (*i*) outbreak success (did the pathogen spread beyond the initially infected individual?), and given successful transmission (*ii*) maximum prevalence, and (*iii*) duration of the outbreak. For covariates, we included conspecific density, transmission rate, recovery

rate, landscape structure (p & H), individual movement preferences (as governed by the RSF: $\beta_1,\ \beta_2,\ \beta_3$), and perceptual range (Table 4.1). We report variable importance scores in terms of mean decrease in accuracy, which is equivalent to percent increase in mean square error (MSE) for regression random forest analyses (Liaw & Wiener 2002b). Mean decrease in accuracy corresponds to the loss of predictive value for the model when a parameter is permutated randomly, rather than using its given value (Cutler *et al.* 2007). We report raw variable importance measures that have not been scaled by the standard error, as these values may be less biased for correlated predictors (Strobl *et al.* 2009). We also corroborated variable importance results by conducting a secondary analysis using the *cforest* function from the *party* package in R (Strobl *et al.* 2009). This approach has been shown to have a more robust estimate of variable importance (Strobl *et al.* 2009); however, this comes with a computational cost. After reducing our analysis to 1000 trees (so that computational time was tractable in *cforest*), we found that our main conclusions were still supported with the only changes being the order of lower-ranked variables with very similar importance values (Appendix G Figure G.1). All simulations and analyses were conducted in R (version 3.3.2). Code and simulation results are available at: https://github.com/whit1951/landscape-sim.

Acknowledgements

L.A.W. was funded by the National Science Foundation (GRFP-00039202 and DEB-1701069) and University of Minnesota Informatics Institute. M.E.C. was funded by National Science Foundation (DEB-1413925 and DEB-1654609), the University of

Minnesota's Office of the Vice President for Research and Academic Health Center Seed Grant. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: http://www.msi.umn.edu

4.7 Figure Legends

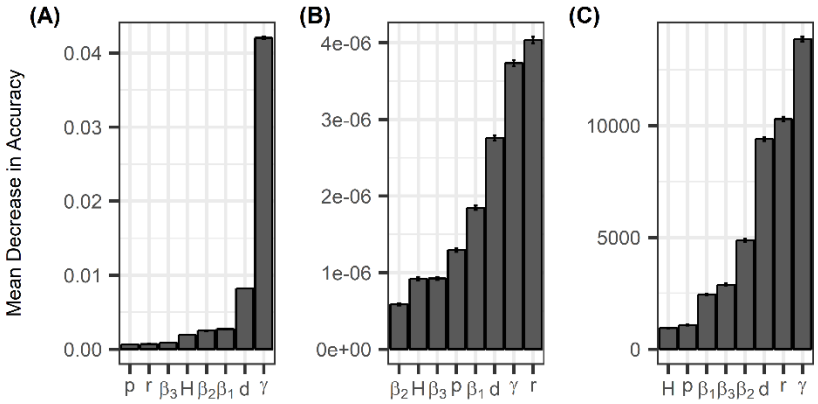


Figure 4.1. Random forest regression analysis results describing variable importance for (A) outbreak success (did the pathogen spread beyond the initially infected individual?), (B) maximum prevalence given outbreak success, and (C) outbreak duration given outbreak success. Parameter descriptions are provided in Table 4.1. Bar charts display an un-scaled mean decrease in accuracy for parameters; higher values for mean decrease in accuracy correspond to parameters with higher predictive ability. Error bars reflect standard deviation of mean decrease in accuracy. The *randomForest* package in R (with 10,000 trees) was used for this analysis. While included in the random forest analysis, transmission rate (β) received a variable importance score of zero for all three metrics (as would be expected since it did not vary in the factorial design) and is not depicted in Figure 4.1.

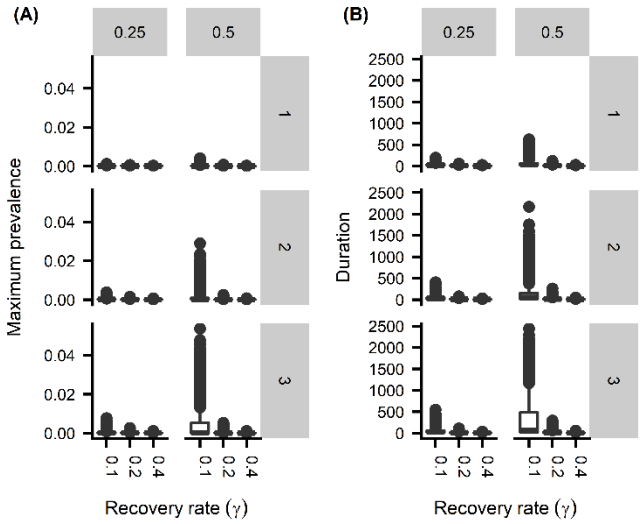


Figure 4.2. Effects of recovery rate (x-axis) on (A) maximum prevalence and (B) duration for successful outbreaks (number of secondary cases ≥ 1) across different conspecific densities (columns) and perceptual ranges (rows). These plots are combined for all resource selection functions (RSFs) and for all landscape structures. To observe these outcomes for individual resource selection functions, see Appendix G Figure G.1.

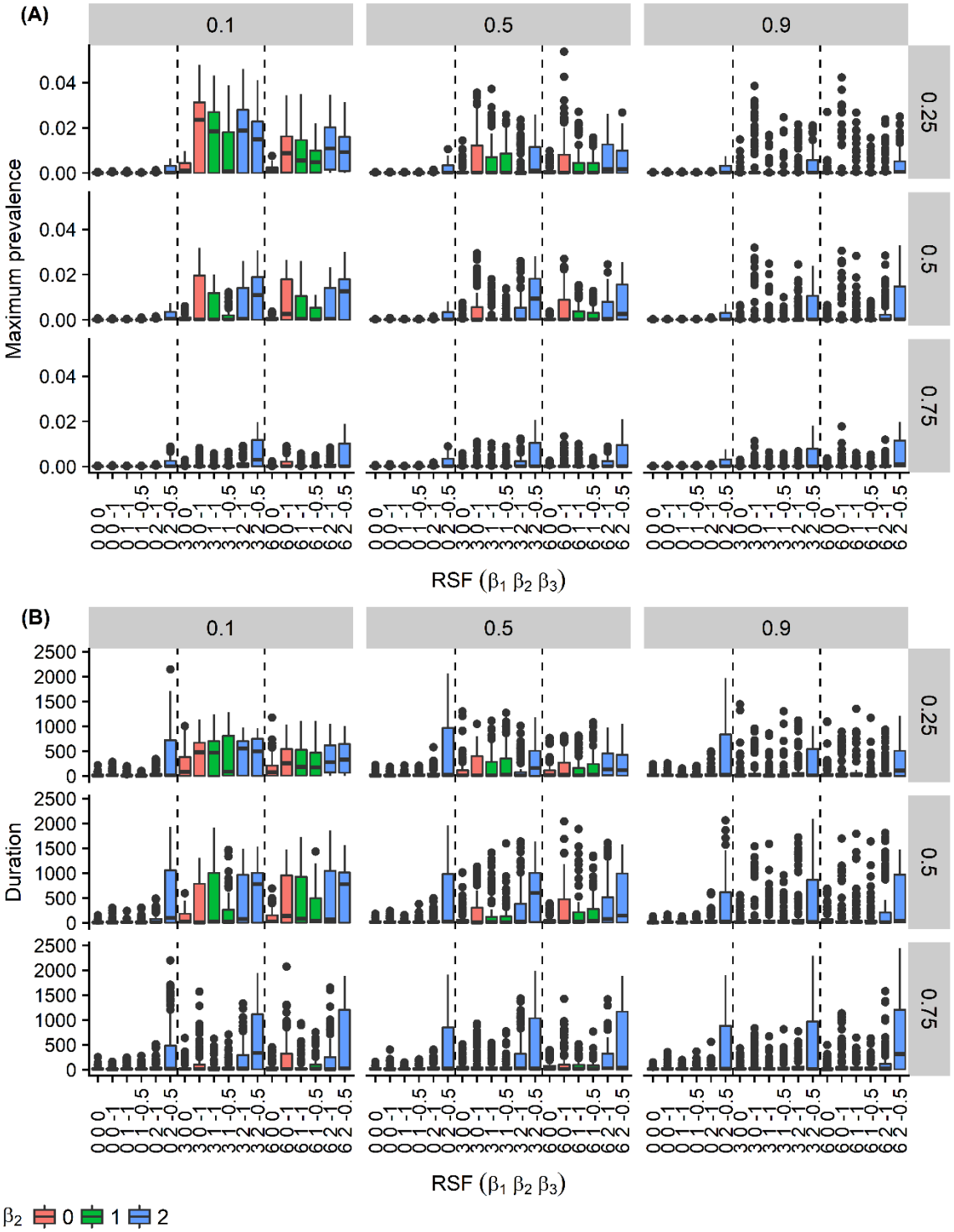


Figure 4.3. Box plots of (*A*) maximum prevalence and (*B*) duration for a subset of the simulations where γ = 0.1, d = 0.5, r = 3. Columns correspond to Hurst exponent (H; lower values correspond to higher

clustering) and rows correspond to proportion available habitat (p). Dashed lines represent separate regimes of random, medium or strong selection for resources (β_1).

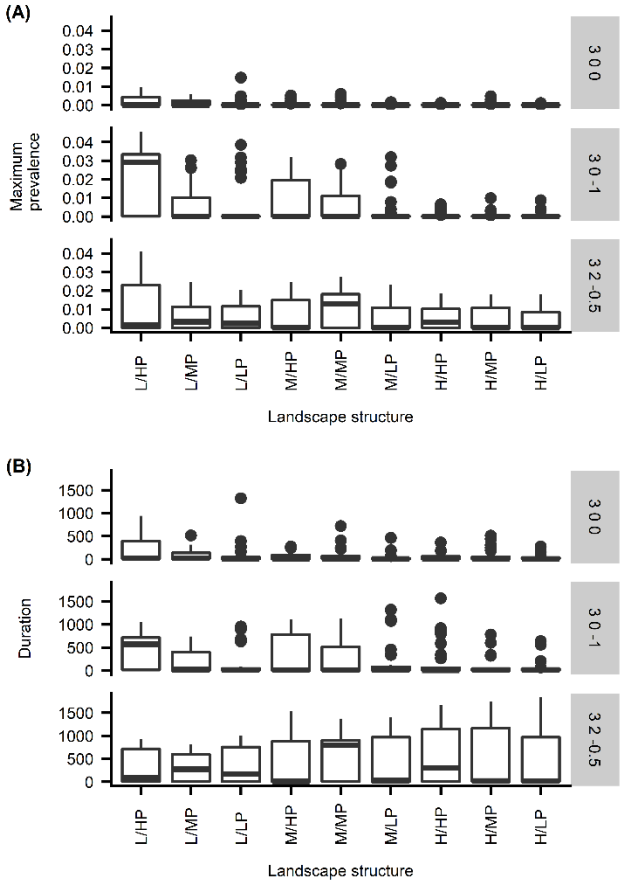


Figure 4.4. Box plots of (A) maximum prevalence and (B) duration for three RSFs when recovery rate (γ)= 0.1, conspecific density (d)=0.5, and perceptual range (r)=3. The RSFs (rows: [β_1 β_2 β_3]) correspond to three biological scenarios: (i) positive selection for resources (β_1=3, β_2=0, β_3=0), (ii) positive selection for resources with conspecific avoidance (β_1=3, β_2=0, β_3=-1) and (iii) positive selection for resources with conspecific attraction (β_1=3, β_2=2, β_3=-0.5) (Appendix G Figure G.7). Landscape structure abbreviations take the form of proportion available habitat(p)/patchiness (Hurst exponent, H): "L" = low, "M" =medium,

"H" = high proportion available habitat. "LP" =low patchiness, "MP" =medium patchiness, and "HP" =high patchiness.

4.8 Tables

Table 4.1. Factorial design of 2,916 parameter combinations encompassing epidemiology, movement behavior, and landscape structure (Appendix G, Figure G.8). We conducted 100 simulations per parameter set. For more information on how the RSF parameters (β_1, β_2, β_3) were calibrated and how these values correspond to movement choices, see the SI Appendix.

Parameter	**Levels**	**Values**
Conspecific density (d) (individuals/unit area of simulated landscape)	Low, medium	0.25, 0.5
Transmission rate (β)	Constant	0.2
Recovery rate (γ); conversely, infectious period ($1/\gamma$)	Slow, medium, fast; long, medium, short infectious periods	0.1, 0.2, 0.4 time^{-1}; 10, 5, 2.5 time steps
Proportion of available habitat (p)	Low, medium, high	0.25, 0.50, 0.75
Clustering of habitat (H)	Low, medium, high	0.1, 0.5, 0.9
Strength of selection for resources (β_1)	None (random), low, high	0, 3, 6
Strength of selection for conspecifics (β_2, β_3)	None (random), avoidance, attraction	(0, 0), (0, -1), (1, -1), (1, -0.5), (2, -1), (2, -0.5)
Perceptual range (r)	Low, medium, high	1, 2, 3

Conclusion

The four chapters of this dissertation sought to understand the consequences of heterogeneity in individual infectiousness on disease dynamics at different scales ranging from the individual host to the landscape. Collectively, these results demonstrate that heterogeneity can have important ramifications for population-level disease outcomes regardless of the scale at which it occurs.

At the level of the individual (Chapter 2), simply allowing for individual variability in susceptibility or infectiousness altered epidemic outcomes; however, negative covariation between physiology and contact rate had the strongest effect for all mechanisms we tested, and was most profound for infection-induced behavioral changes. This finding suggests that sickness-induced behavioral changes may be an overlooked mechanism in disease modeling of wildlife populations.

At the population level (Chapter 3), explicitly modelling the different swine production classes and the physiological variability across sub-populations provided important clues to how influenza A virus (IAV) can become endemic in a swine breeding herd. Notably, piglets played an outsized role in maintaining infection after introduction of IAV, and early weaning of piglets was one of the most effective control strategies, although eradication remained difficult once IAV had been established on the farm.

At a landscape level (Chapter 4), more fragmented landscapes generally promoted longer lasting outbreaks with higher epidemic peaks. However, we observed threshold behavior where more complex disease dynamics emerged for landscapes with

higher conspecific densities, pathogens with longer infectious periods, and host that could perceive and move through greater proportions of the landscape. Under these conditions, small differences in how hosts navigated their environment (through strength of selection for resources and conspecifics) dramatically altered the landscape structure that was most conducive to pathogen persistence and outbreak size.

The mathematical models in this dissertation (except for Chapter 3) represent direct transmission of a single pathogen species within a single host species. Thus, while this work does not explicitly address the potential effects of pathogen co-infection (Cattadori, Boag & Hudson 2008; Cattadori *et al.* 2014; Knowles *et al.* 2013), multi-host pathogens (Craft *et al.* 2008; Lloyd-Smith *et al.* 2009; Almberg, Cross & Smith 2010; VanderWaal *et al.* 2014a), indirect transmission (Leu, Kappeler & Bull 2010; Almberg *et al.* 2011), or vector foraging behavior (Webb *et al.* 2006; Leu *et al.* 2010), the results of this research could easily be extended to more biologically complex systems. For instance, changes in host behavior or susceptibility as a result of co-infection (Cattadori *et al.* 2008, 2014; Knowles *et al.* 2013) could be modeled as differences in host β_p and β_c. Likewise, the relative intra-species vs. inter-species contact rates of a multi-host pathogen (Böhm *et al.* 2008, 2009) could be represented via a multi-host network.

Finally, while this dissertation addresses many possible causes of variation in individual infectiousness that operate at different scales, it does not integrate across these scales; however, ultimately such integration will be necessary to truly understand how pathogens invade and persist in different populations (Lloyd-Smith *et al.* 2009; Tompkins *et al.* 2011). This research will aid in the design of future empirical studies of pathogen

transmission in wildlife and livestock by helping investigators answer three questions: (i) Which scale captures the transmission mechanism of the system? (ii) Is it necessary to account for heterogeneity? And (iii) which sampling strategy adequately addresses both factors? Understanding how (and when) variability in individual infectiousness should be modelled is a crucial next step for the field of disease ecology; moreover, such insight is a critical refinement for modeling strategies that address the growing global threats of zoonoses and emerging infectious diseases.

References

Adelman, J.S., Moyers, S.C., Farine, D.R. & Hawley, D.M. (2015) Feeder use predicts both acquisition and transmission of a contagious pathogen in a North American songbird. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20151429.

Adelman, J.S., Moyers, S.C. & Hawley, D.M. (2014) Using remote biomonitoring to understand heterogeneity in immune-responses and disease-dynamics in small, free-living animals. *Integrative and Comparative Biology*, **54**, 377–386.

Aiello, C.M., Nussear, K.E., Esque, T.C., Emblidge, P.G., Sah, P., Bansal, S. & Hudson, P.J. (2016) Host contact and shedding patterns clarify variation in pathogen exposure and transmission in threatened tortoise Gopherus agassizii: Implications for disease modelling and management (ed I Montgomery). *Journal of Animal Ecology*, **85**, 829–842.

Allerson, M.W., Cardona, C.J. & Torremorell, M. (2013a) Indirect transmission of influenza A virus between pig populations under two different biosecurity settings. *PLoS ONE*, **8**, 2–10.

Allerson, M.W., Davies, P.R., Gramer, M.R. & Torremorell, M. (2013b) Infection dynamics of pandemic 2009 H1N1 influenza virus in a two-site swine herd. *Transboundary and Emerging Diseases*, **61**, 490–499.

Allerson, M.W., Deen, J., Detmer, S.E., Gramer, M.R., Joo, H.S., Romagosa, A. & Torremorell, M. (2013c) The impact of maternally derived immunity on influenza A virus transmission in neonatal pig populations. *Vaccine*, **31**, 500–505.

Almberg, E.S., Cross, P.C., Johnson, C.J., Heisey, D.M. & Richards, B.J. (2011) Modeling routes of chronic wasting disease transmission: environmental prion persistence promotes deer population decline and extinction. *PloS ONE*, **6**, e19896.

Almberg, E.S., Cross, P.C. & Smith, D.W. (2010) Persistence of canine distemper virus in the Greater Yellowstone ecosystem's carnivore community. *Ecological Applications*, **20**, 2058–2074.

Ames, G.M., George, D.B., Hampson, C.P., Kanarek, A.R., McBee, C.D., Lockwood, D.R., … Webb, C.T. (2011) Using network properties to predict disease dynamics on human contact networks. *Proceedings of the Royal Society B*, **278**, 3544–3550.

Anderson, R. & May, R. (1979) Population biology of infectious diseases: Part I. *Nature*, **260**, 361–366.

Anderson, R. & May, R. (1991) *Infectious Diseases of Humans*. Oxford University Press, Oxford.

Anderson, R.M., Medley, G.F., May, R.M. & Johnson, A.M. (1986) A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA Journal of Mathematics Applied in Medicine and Biology*, **3**, 229–263.

Antonovics, J. (2017) Transmission dynamics: critical questions and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160087.

Archie, E.A., Luikart, G. & Ezenwa, V.O. (2009) Infecting epidemiology with genetics: A new frontier in disease ecology. *Trends in Ecology & Evolution*, **24**, 21–30.

Archie, E.A. & Theis, K.R. (2011) Animal behaviour meets microbial ecology. *Animal Behaviour*, **82**, 425–436.

Bansal, S., Read, J., Pourbohloul, B. & Meyers, L.A. (2010) The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, **4**, 478–489.

Barron, D., Gervasi, S., Pruitt, J. & Martin, L. (2015) Behavioral competence: how host behaviors can interact to influence parasite transmission risk. *Current Opinion in Behavioral Sciences*, **6**, 35–40.

Beaudoin, A., Johnson, S., Davies, P., Bender, J. & Gramer, M. (2012) Characterization of influenza a outbreaks in Minnesota swine herds and measures taken to reduce the risk of zoonotic transmission. *Zoonoses and public health*, **59**, 96–106.

Becker, D.J. & Hall, R.J. (2016) Heterogeneity in patch quality buffers metapopulations from pathogen impacts. *Theoretical Ecology*, **9**, 197–205.

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

Beldomenico, P.M. & Begon, M. (2010) Disease spread, susceptibility and infection intensity: vicious circles? *Trends in Ecology and Evolution*, **25**, 21–27.

Beldomenico, P.M., Telfer, S., Lukomski, L., Gebert, S., Bennett, M. & Begon, M. (2009) Host condition and individual risk of cowpox virus infection in natural animal populations: cause or effect? *Epidemiology and infection*, **137**, 1295–1301.

Bell, A.M. (2007) Future directions in behavioural syndromes research. *Proceedings of*

the Royal Society B, **274**, 755–61.

Benavides, J., Walsh, P.D., Meyers, L.A., Raymond, M. & Caillaud, D. (2012) Transmission of infectious diseases en route to habitat hotspots. *PLoS ONE*, **7**, e31290.

Benson, A.K., Kelly, S.A., Legge, R., Ma, F., Low, S.J., Kim, J., … Pomp, D. (2010) Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 18933–8.

Berdoy, M., Webster, J.P. & Macdonald, D.W. (2000) Fatal attraction in rats infected with *Toxoplasma gondii*. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1591–1594.

Biek, R., O'Hare, A., Wright, D., Mallon, T., McCormick, C., Orton, R.J., … Kao, R.R. (2012) Whole genome sequencing reveals local transmission patterns of Mycobacterium bovis in sympatric cattle and badger populations. *PLoS Pathogens*, **8**, e1003008.

Blonder, B., Wey, T.W., Dornhaus, A., James, R. & Sih, A. (2012) Temporal dynamics and network analysis. *Methods in Ecology and Evolution*, **3**, 958–972.

Blower, S.M. & Dowlatabadi, H. (1994) Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *International Statistical Review*, **62**, 229–243.

Blyton, M.D.J., Banks, S.C., Peakall, R. & Gordon, D.M. (2013) High temporal variability in commensal Escherichia coli strain communities of a herbivorous

marsupial. *Environmental Microbiology*, **15**, 2162–72.

Blyton, M.D.J., Banks, S.C., Peakall, R., Lindenmayer, D.B. & Gordon, D.M. (2014) Not all types of host contacts are equal when it comes to E. coli transmission. *Ecology Letters*, **17**, 970–8.

Böhm, M., Hutchings, M.R. & White, P.C.L. (2009) Contact networks in a wildlife-livestock host community: Identifying high-risk individuals in the transmission of bovine TB among badgers and cattle. *PloS ONE*, **4**, e5016.

Böhm, M., Palphramand, K., Newton-Cross, G., Hutchings, M.R. & White, P.C.L. (2008) Dynamic interactions among badgers: implications for sociality and disease transmission. *Journal of Animal Ecology*, **77**, 735–745.

Bonnell, T.R., Sengupta, R.R., Chapman, C.A. & Goldberg, T.L. (2010) An agent-based model of red colobus resources and disease dynamics implicates key resource sites as hot spots of disease transmission. *Ecological Modelling*, **221**, 2491–2500.

Borchering, R.K., Bellan, S.E., Flynn, J.M., Pulliam, J.R.C. & McKinley, S.A. (2017) Resource-driven encounters among consumers and implications for the spread of infectious disease. *Journal of the Royal Society, Interface*, **14**, 20170555.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Brown, I.H. (2000) The epidemiology and evolution of influenza viruses in pigs. *Veterinary Microbiology*, **74**, 29–46.

Brown, L.M. & Crone, E.E. (2016) Individual variation changes dispersal distance and area requirements of a checkerspot butterfly. *Ecology*, **97**, 106–115.

Buhnerkempe, M.G., Tildesley, M.J., Lindström, T., Grear, D.A., Portacci, K., Miller,

R.S., … Webb, C.T. (2014) The impact of movements and animal density on continental scale cattle disease outbreaks in the United States. *PLoS ONE*, **9**, e91724.

Bull, C.M., Godfrey, S.S. & Gordon, D.M. (2012) Social networks and the spread of Salmonella in a sleepy lizard population. *Molecular Ecology*, **21**, 4386–92.

Caillaud, D., Craft, M. & Meyers, L. (2013) Epidemiological effects of group size variation in social species. *Journal of the Royal Society Interface*, **10**, 20130206.

Campomizzi, A.J., Butcher, J.A., Farrell, S.L., Snelgrove, A.G., Collier, B.A., Gutzwiller, K.J., … Wilkins, R.N. (2008) Conspecific attraction is a missing component in wildlife habitat modeling. *Journal of Wildlife Management*, **72**, 331–336.

Carne, C., Semple, S., Morrogh-Bernard, H., Zuberbühler, K. & Lehmann, J. (2013) Predicting the vulnerability of great apes to disease: the role of superspreaders and their potential vaccination. *PLoS ONE*, **8**, e84642.

Carter, S.P., Delahay, R.J., Smith, G.C., Macdonald, D.W., Riordan, P., Etherington, T.R., … Cheeseman, C.L. (2007) Culling-induced social perturbation in Eurasian badgers Meles meles and the management of TB in cattle: an analysis of a critical problem in applied ecology. *Proceedings of the Royal Society B*, **274**, 2769–77.

Cattadori, I.M., Albert, R. & Boag, B. (2007) Variation in host susceptibility and infectiousness generated by co-infection: the myxoma–*Trichostrongylus retortaeformis* case in wild rabbits. *Journal of the Royal Society, Interface / the Royal Society*, **4**, 831–40.

Cattadori, I.M., Boag, B. & Hudson, P.J. (2008) Parasite co-infection and interaction as drivers of host heterogeneity. *International Journal for Parasitology*, **38**, 371–380.

Cattadori, I.M., Wagner, B.R., Wodzinski, L.A., Pathak, A.K., Poole, A. & Boag, B. (2014) Infections do not predict shedding in co-infections with two helminths from a natural system. *Ecology*, **95**, 1684–1692.

Charbonneau, D., Blonder, B. & Dornhaus, A. (2013) Social Insects: A Model System for Network Dynamics. *Temporal Networks*, Understanding Complex Systems (eds P. Holme & J. Saramäki), pp. 217–244. Springer Berlin Heidelberg, Berlin, Heidelberg.

Chen, S., White, B.J., Sanderson, M.W., Amrine, D.E., Ilany, A. & Lanzas, C. (2014) Highly dynamic animal contact network and implications on disease transmission. *Scientific Reports*, **4**, 4472.

Chiyo, P.I., Grieneisen, L.E., Wittemyer, G., Moss, C.J., Lee, P.C., Douglas-Hamilton, I. & Archie, E.A. (2014) The influence of social structure, habitat, and host traits on the transmission of Escherichia coli in wild elephants. (ed AL Roca). *PloS ONE*, **9**, e93408.

Choi, M.J., Torremorell, M., Bender, J.B., Smith, K., Boxrud, D., Ertl, J.R., … Lynfield, R. (2015) Live animal markets in Minnesota: A potential source for emergence of novel influenza A viruses and interspecies transmission. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, **61**, 1355–62.

Christley, R.M., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R. &

Turner, J. (2005) Infection in social networks: using network analysis to identify high-risk individuals. *American Journal of Epidemiology*, **162**, 1024–1031.

Clay, C.A., Lehmer, E.M., Previtali, A., St. Jeor, S. & Dearing, M.D. (2009) Contact heterogeneity in deer mice: implications for Sin Nombre virus transmission. *Proceedings of the Royal Society B*, **276**, 1305–1312.

Coburn, B.J., Wagner, B.G. & Blower, S. (2009) Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC Medicine*, **7**, 30.

Coelho, F.C. & Codeço, C.T. (2009) Dynamic modeling of vaccinating behavior as a function of individual beliefs. *PLoS Computational Biology*, **5**, e1000425.

Corner, L.A.L., Pfeiffer, D.U. & Morris, R.S. (2003) Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (Trichosurus vulpecula). *Preventive Veterinary Medicine*, **59**, 147–167.

Corzo, C.A., Gramer, M., Kuhn, M., Mohr, M. & Morrison, R. (2012) Observations regarding influenza A virus shedding in a swine breeding farm after mass vaccination. *Journal of Swine Health and Production*, **20**, 283–289.

Cosgrove, M.K., Campa, H., Ramsey, D.S.L., Schmitt, S.M. & O'Brien, D.J. (2012) Modeling vaccination and targeted removal of white-tailed deer in Michigan for bovine tuberculosis control. *Wildlife Society Bulletin*, **36**, 676–684.

Craft, M.E. (2015) Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140107–20140107.

Craft, M.E. & Caillaud, D. (2011a) Network models: an underutilized tool in wildlife

epidemiology? *Interdisciplinary Perspectives on Infectious Diseases*, **2011**, 676949.

Craft, M.E. & Caillaud, D. (2011b) Network models: an underutilized tool in wildlife epidemiology? *Interdisciplinary Perspectives on Infectious Disease*, **2011**, 676949.

Craft, M., Hawthorne, P., Packer, C. & Dobson, A. (2008) Dynamics of a multihost pathogen in a carnivore community. *Journal of Animal Ecology*, **77**, 1257–1264.

Craft, M.E., Volz, E., Packer, C. & Meyers, L.A. (2009) Distinguishing epidemic waves from disease spillover in a wildlife population. *Proceedings of the Royal Society B*, **276**, 1777–85.

Craft, M.E., Volz, E., Packer, C. & Meyers, L.A. (2011) Disease transmission in territorial populations: The small-world network of Serengeti lions. *Journal of The Royal Society Interface*, **8**, 776–786.

Croft, D.P., Edenbrow, M., Darden, S.K., Ramnarine, I.W., van Oosterhout, C. & Cable, J. (2011) Effect of gyrodactylid ectoparasites on host behaviour and social network structure in guppies *Poecilia reticulata*. *Behavioral Ecology and Sociobiology*, **65**, 2219–2227.

Cross, P.C., Creech, T.G., Ebinger, M.R., Heisey, D.M., Irvine, K.M. & Creel, S. (2012) Wildlife contact analysis: emerging methods, questions, and challenges. *Behavioral Ecology and Sociobiology*, **66**, 1437–1447.

Cross, P.C., Edwards, W.H., Scurlock, B.M., Maichak, E.J. & Rogerson, J.D. (2007a) Effects of management and climate on elk brucellosis in the Greater Yellowstone Ecosystem. *Ecological applications : a publication of the Ecological Society of America*, **17**, 957–964.

Cross, P.C., Johnson, P.L.F., Lloyd-Smith, J.O. & Getz, W.M. (2007b) Utility of R_0 as a predictor of disease invasion in structured populations. *Journal of the Royal Society Interface*, **4**, 315–324.

Cross, P.C., Lloyd-Smith, J.O., Bowers, J.A., Hay, C.T., Hofmeyr, M. & 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**, 879–892.

Cross, P.C., Lloyd-Smith, J.O., Johnson, P.L.F. & Getz, W.M. (2005) Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. *Ecology Letters*, **8**, 587–595.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007) Random forests for classification in ecology. *Ecology*, **88**, 2783–2792.

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

Davis, S., Abbasi, B., Shah, S., Telfer, S. & Begon, M. (2015) Spatial analyses of wildlife contact networks. *Journal of the Royal Society Interface*, **12**, 20141004.

Degnan, P., Pusey, A., Lonsdorf, E., Goodall, J., Wroblewski, E., Wilson, M., … Ochman, H. (2012) Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 12024–13039.

Dhondt, A.A., Dhondt, K. V., Hawley, D.M. & Jennelle, C.S. (2007) Experimental evidence for transmission of *Mycoplasma gallisepticum* in house finches by fomites.

Avian Pathology, **36**, 205–208.

Diaz, A., Perez, A., Sreevatsan, S., Davies, P., Culhane, M. & Torremorell, M. (2015) Association between influenza A virus infection and pigs subpopulations in endemically infected breeding herds. *Plos One*, **10**, e0129213.

Dorjee, S., Poljak, Z., Revie, C.W., Bridgland, J., McNab, B., Leger, E. & Sanchez, J. (2013) A review of simulation modelling approaches used for the spread of zoonotic influenza viruses in animal and human populations. *Zoonoses and Public Health*, **60**, 383–411.

Dougherty, E.R., Seidel, D.P., Carlson, C.J., Spiegel, O. & Wayne, M. (2018) Going through the motions: incorporating movement analyses into disease research. *Ecology Letters*.

Drewe, J.A. (2010) Who infects whom? Social networks and tuberculosis transmission in wild meerkats. *Proceedings of the Royal Society B*, **277**, 633–42.

Drewe, J.A., Eames, K.T.D., Madden, J.R. & Pearce, G.P. (2011) Integrating contact network structure into tuberculosis epidemiology in meerkats in South Africa: implications for control. *Preventive Veterinary Medicine*, **101**, 113–120.

Dwyer, G., Elkinton, J.S. & Buonaccorsi, J.P. (1997) Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. *The American Naturalist*, **150**, 685–707.

Eames, K., Bansal, S., Frost, S. & Riley, S. (2015) Six challenges in measuring contact networks for use in modelling. *Epidemics*, **10**, 72–77.

Eisinger, D. & Thulke, H. (2008) Spatial pattern formation facilitates eradication of

infectious diseases. *Journal of Applied Ecology*, **45**, 415–423.

Elzhov, T.V. et al. (2016) minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds. R package version 1.2-1.

Epstein, J.M., Parker, J., Cummings, D. & Hammond, R.A. (2008) Coupled contagion dynamics of fear and disease: mathematical and computational explorations. *PloS ONE*, **3**, e3955.

Eubank, S., Guclu, H., Kumar, V.S.A., Marathe, M. V, Srinivasan, A., Toroczkai, Z. & Wang, N. (2004) Modelling disease outbreaks in realistic urban social networks. *Nature*, **429**, 180–184.

Ezenwa, V.O., Archie, E.A., Craft, M.E., Hawley, D.M., Martin, L.B., Moore, J. & White, L. (2016) Host behaviour – parasite feedback: an essential link between animal behaviour and disease ecology. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 20153078.

Fefferman, N. & Ng, K. (2007) How disease models in static networks can fail to approximate disease in dynamic networks. *Physical Review E*, **76**, 31919.

Fenner, A.L., Godfrey, S.S. & Michael Bull, C. (2011) Using social networks to deduce whether residents or dispersers spread parasites in a lizard population. *The Journal of Animal Ecology*, **80**, 835–43.

Ferrari, M.J., Perkins, S.E., Pomeroy, L.W. & Bjørnstad, O.N. (2011) Pathogens, social networks, and the paradox of transmission scaling. *Interdisciplinary Perspectives on Infectious Diseases*, **2011**, 267049.

Fine, A.E., Bolin, C.A., Gardiner, J.C. & 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**, 765430.

Fofana, A.M. & Hurford, A. (2017) Mechanistic movement models to understand epidemic spread. *Philosophical Transactions of the Royal Society B*, **372**, 20160086.

Forester, J.D., Im, H.K. & Rathouz, P.J. (2009) Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, **90**, 3554–3565.

Friedman, S.R., Neaigus, A., Jose, B., Curtis, R., Goldstein, M., Ildefonso, G., … Des Jarlais, D.C. (1997) Sociometric risk networks and risk for HIV infection. *American Journal of Public Health*, **87**, 1289–96.

Funk, S., Bansal, S., Bauch, C.T., Eames, K.T.D., Edmunds, W.J., Galvani, A.P. & Klepac, P. (2015) Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics*, **10**, 21–25.

Funk, S., Salathé, M. & Jansen, V.A.A. (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, **7**, 1247–1256.

Garnett, B.T., Delahay, R.J. & Roper, T.J. (2002) Use of cattle farm resources by badgers (Meles meles) and risk of bovine tuberculosis (Mycobacterium bovis) transmission to cattle. *Proceedings of the Royal Society B*, **269**, 1487–91.

Garnett, B., Roper, T. & Delahay, R. (2003) Use of cattle troughs by badgers (Meles meles). *Applied Animal Behaviour Science*, **80**, 1–8.

Gauger, P.C., Vincent, A.L., Loving, C.L., Lager, K.M., Janke, B.H., Kehrli, M.E. & Roth, J.A. (2011) Enhanced pneumonia and disease in pigs vaccinated with an inactivated human-like (δ-cluster) H1N2 vaccine and challenged with pandemic 2009 H1N1 influenza virus. *Vaccine*, **29**, 2712–9.

Gibson, A.K., Jokela, J. & Lively, C.M. (2016) Fine-scale spatial covariation between infection prevalence and susceptibility in a natural population. *The American Naturalist*, **188**, 1–14.

Gillespie, D.T. (1977) Exact stochastic simulation of couple chemical reactions. *The Journal of Physical Chemistry*, **81**, 2340–2361.

Girard, M.P., Tam, J.S., Assossou, O.M. & Kieny, M.P. (2010) The 2009 A (H1N1) influenza virus pandemic: A review. *Vaccine*, **28**, 4895–902.

Godfrey, S.S. (2013) Networks and the ecology of parasite transmission: A framework for wildlife parasitology. *International Journal for Parasitology: Parasites and Wildlife*, **2**, 235–245.

Godfrey, S.S., Bull, C.M., James, R. & Murray, K. (2009) Network structure and parasite transmission in a group living lizard, the gidgee skink, *Egernia stokesii*. *Behavioral Ecology and Sociobiology*, **63**, 1045–1056.

Godfrey, S.S., Moore, J.A., Nelson, N.J. & Bull, C.M. (2010) Social network structure and parasite infection patterns in a territorial reptile, the tuatara (*Sphenodon punctatus*). *International Journal for Parasitology*, **40**, 1575–1585.

Goodman, B.A. & Johnson, P.T.J. (2011) Disease and the extended phenotype: parasites control host performance and survival through induced changes in body plan. *PloS*

ONE, **6**, e20193.

Gottdenker, N.L., Streicker, D.G., Faust, C.L. & Carroll, C.R. (2014) Anthropogenic land use change and infectious diseases: A review of the evidence. *EcoHealth*, **11**, 619–632.

Goyens, J., Reijniers, J., Borremans, B. & Leirs, H. (2013) Density thresholds for Mopeia virus invasion and persistence in its host *Mastomys natalensis*. *Journal of Theoretical Biology*, **317**, 55–61.

Grange, Z.L., Van Andel, M., French, N.P. & Gartrell, B.D. (2014) Network analysis of translocated Takahe populations to identify disease surveillance targets. *Conservation Biology*, **28**, 518–528.

Grear, D.A., Luong, L. & Hudson, P. (2013) Network transmission inference: host behavior and parasite life cycle make social networks meaningful in disease ecology. *Ecological Applications*, **23**, 1906–1914.

Grear, D.A., Perkins, S.E. & Hudson, P.J. (2009) Does elevated testosterone result in increased exposure and transmission of parasites? *Ecology Letters*, **12**, 528–537.

Griffin, R.H. & Nunn, C.L. (2011) Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology*, **26**, 779–800.

Groendyke, C., Welch, D. & Hunter, D.R. (2011) Bayesian inference for contact networks given epidemic data. *Scandinavian Journal of Statistics*, **38**, 600–616.

Guimarães, P., de Menezes, M., Baird, R., Lusseau, D. & dos Reis, S. (2007) Vulnerability of a killer whale social network to disease outbreaks. *Physical Review E*, **76**, 42901.

Hagenaars, T.J., Donnelly, C.A. & Ferguson, N.M. (2004) Spatial heterogeneity and the persistence of infectious diseases. *Journal of Theoretical Biology*, **229**, 349–359.

Hamede, R.K., Bashford, J., Jones, M. & McCallum, H. (2012) Simulating devil facial tumour disease outbreaks across empirically derived contact networks. *Journal of Applied Ecology*, **49**, 447–456.

Hamede, R.K., Bashford, J., McCallum, H. & Jones, M. (2009) Contact networks in a wild Tasmanian devil (Sarcophilus harrisii) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters*, **12**, 1147–57.

Hanneke, S., Fu, W. & Xing, E.P. (2010) Discrete temporal models of social networks. *Electronic Journal of Statistics*, **4**, 585–605.

Hawley, D.M. & Altizer, S.M. (2011) Disease ecology meets ecological immunology: understanding the links between organismal immunity and infection dynamics in natural populations. *Functional Ecology*, **25**, 48–60.

Hawley, D.M., Etienne, R.S., Ezenwa, V.O. & Jolles, A.E. (2011) Does animal behavior underlie covariation between hosts' exposure to infectious agents and susceptibility to infection? Implications for disease dynamics. *Integrative and Comparative Biology*, **51**, 528–539.

Haydon, D., Cleaveland, S., Taylor, L.H. & Laurenson, M.K. (2002) Identifying reservoirs of infection: a conceptual and practical challenge. *Emerging Infectious Diseases*, **8**, 1468–73.

Haydon, D.T., Randall, D.A., Matthews, L., Knobel, D.L., Tallents, L.A., Gravenor,

M.B., … Laurenson, M.K. (2006) Low-coverage vaccination strategies for the conservation of endangered species. *Nature*, **443**, 692–695.

Herrick, K.A., Huettmann, F. & Lindgren, M.A. (2013) A global model of avian influenza prediction in wild birds: The importance of northern regions. *Veterinary Research*, **44**, 42.

Hirsch, B.T., Prange, S., Hauver, S. a & Gehrt, S.D. (2013) Raccoon social networks and the potential for disease transmission. *PloS ONE*, **8**, e75830.

Hothorn, T., Buhlmann, P., Dudoit, S., Molinaro, A. & Van der Laan, M. (2006) Survival ensembles. *Biostatistics*, **7**, 355–373.

Jenness, S.M., Goodreau, S.M. & Morris, M. (2016a) EpiModel: mathematical modeling of infectious disease. URL https://cran.r-project.org/package=EpiModel [accessed 30 October 2016]

Jenness, S.M., Goodreau, S.M., Rosenberg, E., Beylerian, E.N., Hoover, K.W., Smith, D.K. & Sullivan, P. (2016b) Impact of the Centers for Disease Control's HIV preexposure prophylaxis guidelines for men who have sex with men in the United States. *Journal of Infectious Diseases*, **214**, 1800–1807.

Ji, W., White, P.C.L. & Clout, M.N. (2005) Contact rates between possums revealed by proximity data loggers. *Journal of Applied Ecology*, **42**, 595–604.

Johnson, C.J., Phillips, K.E., Schramm, P.T., McKenzie, D. & Aiken, J.M. (2006) Prions adhere to soil minerals and remain infectious. *PloS Pathogens*, **2**, e32.

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L. & Daszak, P. (2008) Global trends in emerging infectious diseases. *Nature*, **451**, 990–993.

Kane, M.J., Price, N., Scotch, M. & Rabinowitz, P. (2014) Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC bioinformatics*, **15**, 276.

Kao, R.R., Haydon, D.T., Lycett, S.J. & Murcia, P.R. (2014) Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends in Microbiology*, **22**, 282–91.

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

Keeling, M.J. & Eames, K.T.D. (2005) Networks and epidemic models. *Journal of the Royal Society Interface*, **2**, 295–307.

Keeling, M. & Rohani, P. (2008a) *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton, New Jersey.

Keeling, M. & Rohani, P. (2008b) *Modeling of Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton, New Jersey.

Keiser, C.N., Howell, K.A., Pinter-Wollman, N. & Pruitt, J.N. (2016) Personality composition alters the transmission of cuticular bacteria in social groups. *Biology Letters*, **12**, 20160297.

Kermack, W. & McKendrick, A. (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, **115**, 700–721.

Kitikoon, P., Nilubol, D., Erickson, B.J., Janke, B.H., Hoover, T.C., Sornsen, S.A. & Thacker, E.L. (2006) The immune response and maternal antibody interference to a heterologous H1N1 swine influenza virus infection following vaccination.

Veterinary immunology and immunopathology, **112**, 117–28.

Klein, S., Pasquaretta, C., Barron, A.B., Devaud, J.-M. & Lihoreau, M. (2017) Inter-individual variability in the foraging behaviour of traplining bumblebees. *Scientific Reports*, **7**, 4561.

Knowles, S.C.L., Fenton, A., Petchey, O.L., Jones, T.R., Barber, R. & Pedersen, A.B. (2013) Stability of within-host-parasite communities in a wild mammal system. *Proceedings of the Royal Society B*, **280**, 20130598.

Konrad, M., Vyleta, M.L., Theis, F.J., Stock, M., Tragust, S., Klatt, M., … Cremer, S. (2012) Social transfer of pathogenic fungus promotes active immunisation in ant colonies. *PLoS Biology*, **10**, e1001300.

Koolhaas, J.M. (2008) Coping style and immunity in animals: making sense of individual variation. *Brain, Behavior, and Immunity*, **22**, 662–667.

Krause, J., Krause, S., Arlinghaus, R., Psorakis, I., Roberts, S. & Rutz, C. (2013) Reality mining of animal social systems. *Trends in Ecology & Evolution*, **28**, 541–51.

Krivitsky, P.N. & Handcock, M.S. (2014) A separable model for dynamic networks. *J R Stat Soc Series B Stat Methodol*, **76**, 29–46.

Kurvers, R.H.J.M., Krause, J., Croft, D.P., Wilson, A.D.M. & Wolf, M. (2014) The evolutionary and ecological consequences of animal social networks: Emerging issues. *Trends in Ecology & Evolution*, **29**, 326–35.

Kyvsgaard, N.C., Johansen, M.V. & Carabin, H. (2007) Simulating transmission and control of *Taenia solium* infections using a Reed-Frost stochastic model. *International Journal for Parasitology*, **37**, 547–558.

Lane-deGraaf, K.E., Kennedy, R.C., Niaz Arifin, S.M., Madey, G.R., Fuentes, A. & Hollocher, H. (2013) A test of agent-based models as a tool for predicting patterns of pathogen transmission in complex landscapes. *BMC Ecology*, **13**, 35.

Lange, E., Kalthoff, D., Blohm, U., Teifke, J.P., Breithaupt, A., Maresch, C., … Vahlenkamp, T.W. (2009) Pathogenesis and transmission of the novel swine-origin influenza virus A/H1N1 after experimental infection of pigs. *The Journal of general virology*, **90**, 2119–23.

Larsen, L.E., Nielsen, C.K., Aakerblom, S., Hjulsager, C.K., Nielsen, J.P., Stege, H., … Lau, L. (2010) Dynamics of swine influenza infections in the farrowing unit of a Danish sow herd. *21st International Pig Veterinary Society Congress*, p. Vancouver, Canada.

Lass, S., Hudson, P., Thakar, J., Saric, J., Harvill, E., Albert, R. & Perkins, S.E. (2013) Generating super-shedders: co-infection increases bacterial load and egg production of a gastrointestinal helminth. *Journal of the Royal Society Interface*, **10**, 20120588.

Lavelle, M.J., Fischer, J.W., Phillips, G.E., Hildreth, A.M., Campbell, T.A., Hewitt, D.G., … Vercauteren, K.C. (2014) Assessing risk of disease transmission: direct implications for an indirect science. *BioScience*, **64**, 524–530.

Legrand, J., Sanchez, A., Le Pont, F., Camacho, L. & Larouze, B. (2008) Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. *PloS one*, **3**, e2100.

Legrand, J., Viboud, C., Boelle, P.Y., Valleron, A.J. & Flahault, A. (2004) Modelling responses to a smallpox epidemic taking into account uncertainty. *Epidemiology and*

infection, **132**, 19–25.

Lehmer, E.M., Jones, J.D., Bego, M.G., Varner, J.M., St. Jeor, S., Clay, C. a & Dearing, M.D. (2010) Long-term patterns of immune investment by wild deer mice infected with Sin Nombre virus. *Physiological and Biochemical Zoology*, **83**, 847–57.

Leu, S.T., Kappeler, P.M. & Bull, C.M. (2010) Refuge sharing network predicts ectoparasite load in a lizard. *Behavioral Ecology and Sociobiology*, **64**, 1495–1503.

Levins, R. (1966) The strategy of model building in population biology. *American Scientist*, **54**, 421–431.

Li, S., Gilbert, L., Harrison, P.A. & Rounsevell, M.D.A. (2016) Modelling the seasonality of Lyme disease risk and the potential impacts of a warming climate within the heterogeneous landscapes of Scotland. *Journal of the Royal Society Interface*, **13**, 20160140.

Liaw, A. & Wiener, M. (2002a) Classification and regression by randomForest. *R News*, **2/3**, 18–22.

Liaw, A. & Wiener, M. (2002b) Classification and regression by randomForest. *R News*, **2**, 18–22.

Liljeros, F., Edling, C.R. & Amaral, L.A.N. (2003) Sexual networks: implications for the transmission of sexually transmitted infections. *Microbes and Infection*, **5**, 189–196.

Lima, S.L. & Zollner, P.A. (1996) Towards a behavioral ecology of ecological landscape. *TREE*, **11**, 131–135.

Lindström, T., Grear, D. a, Buhnerkempe, M., Webb, C.T., Miller, R.S., Portacci, K. & Wennergren, U. (2013) A bayesian approach for modeling cattle movements in the

United States: scaling up a partially observed network. *PLoS ONE*, **8**, e53432.

Lloyd-Smith, J.O. (2010) Modeling density dependence in heterogeneous landscapes: dispersal as a case study. *Journal of Theoretical Biology*, **265**, 160–166.

Lloyd-Smith, J.O., Cross, P.C., Briggs, C.J., Daugherty, M., Getz, W.M., Latto, J., … Swei, A. (2005a) Should we expect population thresholds for wildlife disease? *Trends in Ecology & Evolution*, **20**, 511–519.

Lloyd-Smith, J.O., George, D., Pepin, K.M., Pitzer, V.E., Pulliam, J.R.C., Dobson, A.P., … Grenfell, B.T. (2009) Epidemic dynamics at the human-animal interface. *Science*, **326**, 1362–1367.

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

Lopes, P.C., Block, P. & König, B. (2016) Infection-induced behavioural changes reduce connectivity and the potential for disease spread in wild mice contact networks. *Scientific Reports*, **6**, 31790.

MacIntosh, A.J.J., Jacobs, A., Garcia, C., Shimizu, K., Mouri, K., Huffman, M.A. & Hernandez, A.D. (2012) Monkeys in the middle: parasite transmission through the social network of a wild primate. *PloS ONE*, **7**, e51144.

Maher, S.P., Kramer, A.M., Pulliam, J.T., Zokan, M.A., Bowden, S.E., Barton, H.D., … Drake, J.M. (2012) Spread of white-nose syndrome on a network regulated by geography and climate. *Nature Communications*, **3**, 1306.

Marino, S., Hogue, I.B., Ray, C.J. & Kirschner, D.E. (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of*

Theoretical Biology, **254**, 178–196.

Martínez-López, B., Perez, A.M. & Sánchez-Vizcaíno, J.M. (2009) Social network analysis. Review of general concepts and use in preventive veterinary medicine. *Transboundary and Emerging Diseases*, **56**, 109–20.

Masuda, N. & Holme, P. (2013) Predicting and controlling infectious disease epidemics using temporal networks. *F1000prime reports*, **5**, 6.

May, R.M., Anderson, R.M. & Irwin, M.E. (1988) The transmission dynamics of human immunodeficiency virus (HIV). *Philosophical Transactions of the Royal Society Lond. B*, **321**, 565–607.

McBride, W.D. & Key, N. (2013) U.S. Hog Production From 1992 to 2009: Technology, Restructuring, and Productivity Growth.

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

McCallum, H., Fenton, A., Hudson, P.J., Lee, B., Levick, B., Norman, R., … Mccallum, H. (2017) Breaking beta: deconstructing the parasite transmission function. *Philosophical Transactions B*, **372**, 20160084.

McCallum, H. & Hocking, B.A. (2005) Reflecting on ethical and legal issues in wildlife disease. *Bioethics*, **19**, 336–347.

McDonald, R.A., Delahay, R.J., Carter, S.P., Smith, G.C. & Cheeseman, C.L. (2008) Perturbing implications of wildlife ecology for disease control. *Trends in Ecology & Evolution*, **23**, 53–6.

Meloni, S., Perra, N., Arenas, A., Gómez, S., Moreno, Y. & Vespignani, A. (2011)

Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific Reports*, **1**, 62.

Merck Animal Health. (2013) MaxiVax Excell® 5.0: Swine Influenza Vaccine H1N1 and H3N2. URL http://www.merck-animal-health-usa.com/binaries/MaxiVac_Excell_5-0_Product_Bulletin_tcm96-161541.pdf [accessed 15 January 2015]

Meyers, L.A. (2007) Contact network epidemiology: bond percolation applied to infectious disease prediction and control. *American Mathematical Society*, **44**, 63–86.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., … Edmunds, W.J. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, **5**, e74.

Myers, K.P., Olsen, C.W. & Gray, G.C. (2007) Cases of swine influenza in humans: a review of the literature. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, **44**, 1084–8.

Myers, K.P., Olsen, C.W., Setterquist, S.F., Capuano, A.W., Donham, K.J., Thacker, E.L., … Gray, G.C. (2006) Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, **42**, 14–20.

National Animal Health Monitoring System. (2008) *Vaccination against Mycoplasma Pneumonia, Swine Influenza, and PRRS in Breeding Females, 2000 and 2006.* Fort Collins, Colorado.

Natoli, E., Say, L., Cafazzo, S., Bonanni, R., Schmid, M. & Pontier, D. (2005) Bold

attitude makes male urban feral domestic cats more vulnerable to Feline Immunodeficiency Virus. *Neuroscience and Biobehavioral Reviews*, **29**, 151–157.

Naug, D. (2008) Structure of the social network and its influence on transmission dynamics in a honeybee colony. *Behavioral Ecology and Sociobiology*, **62**, 1719–1725.

Naug, D. & Smith, B. (2007) Experimentally induced change in infectious period affects transmission dynamics in a social group. *Proceedings of the Royal Society B*, **274**, 61–65.

Neaigus, A., Friedman, S.R., Kottiri, B.J. & Des Jarlais, D.C. (2001) HIV risk networks and HIV transmission among injecting drug users. *Evaluation and Program Planning*, **24**, 221–226.

Nelson, M.I. & Vincent, A.L. (2015) Reverse zoonosis of influenza to swine: new perspectives on the human-animal interface. *Trends in microbiology*, **23**, 142–53.

Nelson, M.I., Wentworth, D.E., Culhane, M.R., Vincent, A.L., Viboud, C., LaPointe, M.P., … Detmer, S.E. (2014) Introductions and evolution of human-origin seasonal influenza a viruses in multinational swine populations. *Journal of virology*, **88**, 10110–9.

Nunn, C.L., Thrall, P.H. & Kappeler, P.M. (2014) Shared resources and disease dynamics in spatially structured populations. *Ecological Modelling*, **272**, 198–207.

Nunn, C.L., Thrall, P.H., Stewart, K. & Harcourt, A.H. (2008) Emerging infectious diseases and animal social systems. *Evolutionary Ecology*, **22**, 519–543.

Olden, J.D., Schooley, R.L., Monroe, J.B. & Poff, N.L. (2004) Context-dependent

perceptual ranges and their relevance to animal movements in landscapes. *Journal of Animal Ecology*, **73**, 1190–1194.

Oliveira-Santos, L.G.R., Forester, J.D., Piovezan, U., Tomas, W.M. & Fernandez, F.A.S. (2016) Incorporating animal spatial memory in step selection functions (ed J Fryxell). *Journal of Animal Ecology*, **85**, 516–524.

Otterstatter, M.C. & Thomson, J.D. (2007) Contact networks and transmission of an intestinal pathogen in bumble bee (Bombus impatiens) colonies. *Oecologia*, **154**, 411–21.

Pappaioanou, M. & Gramer, M. (2010) Lessons from pandemic H1N1 2009 to improve prevention, detection, and response to influenza pandemics from a One Health perspective. *ILAR journal / National Research Council, Institute of Laboratory Animal Resources*, **51**, 268–280.

Patterson, J.E.H. & Ruckstuhl, K.E. (2013) Parasite infection and host group size: a meta-analytical review. *Parasitology*, **140**, 803–13.

Paull, S.H., Song, S., McClure, K.M., Sackett, L.C., Kilpatrick, A.M. & Johnson, P.T.J. (2012) From superspreaders to disease hotspots: linking transmission across hosts and space. *Frontiers in Ecology and the Environment*, **10**, 75–82.

Pe'er, G. & Kramer-Schadt, S. (2008) Incorporating the perceptual range of animals into connectivity models. *Ecological Modelling*, **213**, 73–85.

Pedersen, A.B. & Antonovics, J. (2013) Anthelmintic treatment alters the parasite community in a wild mouse host. *Biology letters*, **9**, 20130205.

Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V. & Trapman, P. (2014)

Eight challenges for network epidemic models. *Epidemics*, 3–7.

Perkins, S.E., Cagnacci, F., Stradiotto, A., Arnoldi, D. & Hudson, P.J. (2009) Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology*, **78**, 1015–1022.

Perkins, S.E., Ferrari, M.F. & Hudson, P.J. (2008) The effects of social structure and sex-biased transmission on macroparasite infection. *Parasitology*, **135**, 1561–9.

Pfizer Animal Health. (2011) FluSure XP®. URL https://www.zoetisus.com/products/pages/flusure/documents/flusurexp_h1n2_fxp09022_prodsheet.pdf [accessed 15 January 2016]

PigCHAMP. (2014) 2014 Benchmarking Summaries. URL http://www.pigchamp.com/Portals/0/Documents/Benchmarking Summaries/USA 2014.pdf [accessed 15 January 2016]

Pitzer, V.E., Aguas, R., Riley, S., Loeffen, W.L.A., Wood, J.L.N., Grenfell, B.T. & Pitzer, V.E. (2016) High turnover drives prolonged persistence of influenza in managed pig herds. *Journal of the Royal Society, Interface / the Royal Society*, **13**, 519–522.

Plowright, R.K., Foley, P., Field, H.E., Dobson, A.P., Foley, J.E., Eby, P. & Daszak, P. (2011) Urban habituation, ecological connectivity and epidemic dampening: The emergence of Hendra virus from flying foxes (*Pteropus* spp.). *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3703–12.

Porphyre, T., Stevenson, M., Jackson, R. & McKenzie, J. (2008) Influence of contact heterogeneity on TB reproduction ratio R0 in a free-living brushtail possum

Trichosurus vulpecula population. *Veterinary Research*, **39**.

Power, A. & Mitchell, C. (2004) Pathogen spillover in disease epidemics. *The American Naturalist*, **164**, S79–S89.

Prentice, J.C., Marion, G., White, P.C.L., Davidson, R.S. & Hutchings, M.R. (2014) Demographic processes drive increases in wildlife disease following population reduction. *PloS ONE*, **9**, e86563.

Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005) Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, **20**, 345–53.

Rands, S.A., Pettifor, R.A., Rowcliffe, J.M. & Cowlishaw, G. (2004) State-dependent foraging rules for social animals in selfish herds. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2613–2620.

Rees, E.E., Pond, B.A., Tinline, R.R. & Bélanger, D. (2013) Modelling the effect of landscape heterogeneity on the efficacy of vaccination for wildlife infectious disease control. *Journal of Applied Ecology*, **50**, 881–891.

Van Reeth, K., Labarque, G., De Clercq, S. & Pensaert, M. (2001) Efficacy of vaccination of pigs with different H1N1 swine influenza viruses using a recent challenge strain and different parameters of protection. *Vaccine*, **19**, 4479–4486.

Restif, O., Hayman, D.T.S., Pulliam, J.R.C., Plowright, R.K., George, D.B., Luis, A.D., … Webb, C.T. (2012) Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecology Letters*, **15**, 1083–1094.

Reynolds, J.J.H., Hirsch, B.T., Gehrt, S.D. & Craft, M.E. (2015) Raccoon contact

networks predict seasonal susceptibility to rabies outbreaks and limitations of vaccination (ed S Altizer). *Journal of Animal Ecology*, **84**, 1720–1731.

Reynolds, J.J.H., Torremorell, M. & Craft, M.E. (2014) Mathematical modeling of influenza A virus dynamics within swine farms and the effects of vaccination. *PLoS ONE*, **9**, e106177.

Rifkin, J.L., Nunn, C.L. & Garamszegi, L.Z. (2012) Do animals living in larger groups experience greater parasitism? A meta-analysis. *The American Naturalist*, **180**, 70–82.

Riley, S., Eames, K., Isham, V., Mollison, D. & Trapman, P. (2015) Five challenges for spatial epidemic models. *Epidemics*, **10**, 68–71.

Robins, G., Pattison, P., Kalish, Y. & Lusher, D. (2007) An introduction to exponential random graph (p*) models for social networks. *Social Networks*, **29**, 173–191.

Romagosa, A., Allerson, M.W., Gramer, M., Joo, H.S., Deen, J., Detmer, S. & Torremorell, M. (2011) Vaccination of influenza a virus decreases transmission rates in pigs. *Veterinary Research*, **42**, 120.

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

Rushmore, J., Caillaud, D., Matamba, L., Stumpf, R.M., Borgatti, S.P. & Altizer, S. (2013) Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. *Journal of Animal Ecology*, **82**, 976–986.

Rushton, S.P., Barreto, G.W., Cormack, R.M., Macdonald, D.W. & Fuller, R. (2000)

Modelling the effects of mink and habitat fragmentation on the water vole. *Journal of Applied Ecology*, **37**, 475–490.

Ryan, S.J., Jones, J.H. & Dobson, A.P. (2013) Interactions between social structure, demography, and transmission determine disease persistence in primates. *PLoS ONE*, **8**, e76863.

Salathé, M. & Jones, J. (2010) Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology*, **6**, e1000736.

Sandbulte, M., Spickler, A., Zaabel, P. & Roth, J. (2015) Optimal use of vaccines for control of influenza A virus in swine. *Vaccines*, **3**, 22–73.

Sapolsky, R. (2005) The influence of social hierarchy on primate health. *Science*, **308**, 648–652.

Sauter, C.M. & Morris, R.S. (1995) Dominance hierarchies in cattle and red deer (Cervus elaphus): their possible relationship to the transmission of bovine tuberculosis. *New Zealand Veterinary Journal*, **43**, 301–5.

Seaholm, S.K., Ackerman, E. & Wu, S.C. (1988) Latin hypercube sampling and the sensitivity analysis of a Monte Carlo epidemic model. *International journal of bio-medical computing*, **23**, 97–112.

Smouse, P.E., Focardi, S., Moorcroft, P.R., Kie, J.G., Forester, J.D. & Morales, J.M. (2010) Stochastic modelling of animal movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2201–2211.

Springer, A., Kappeler, P.M. & Nunn, C.L. (2017) Dynamic vs. static social networks in models of parasite transmission: predicting *Cryptosporidium* spread in wild lemurs.

Journal of Animal Ecology, **86**, 419–433.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Colizza, V., Isella, L., … Vanhems, P. (2011) Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine*, **9**, 87.

Strobl, C., Boulesteix, A.-L., Zeileis, A. & Hothorn, T. (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, **8**, 25.

Strobl, C., Hothorn, T. & Zeileis, A. (2009) Party on! A new, conditional variable-importance measure for random forests available in the party package. *The R Journal*, **1**, 14–17.

Tardy, O., Massé, A., Pelletier, F. & Fortin, D. (2018) Interplay between contact risk, conspecific density, and landscape connectivity: An individual-based modeling framework. *Ecological Modelling*, **373**, 25–38.

Telfer, S., Lambin, X., Birtles, R., Beldomenico, P., Burthe, S., Paterson, S. & Begon, M. (2010) Species interactions in a parasite community drive infection risk in a wildlife population. *Science*, **330**, 243–246.

Tenhumberg, B., Tyre, A.J., Pople, A.R. & Possignham, H.P. (2004) Do harvest refuges buffer kangaroos against evolutionary responses to selective harvesting ? *Ecology*, **85**, 2003–2017.

Thacker, E. & Janke, B. (2008) Swine influenza virus: zoonotic potential and vaccination strategies for the control of avian and swine influenzas. *The Journal of Infectious Diseases*, **197**, S19–S24.

Thomas, T., Gilbert, J. & Meyer, F. (2012) Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation*, **2**, 1–12.

Tompkins, D.M., Dunn, A.M., Smith, M.J. & Telfer, S. (2011) Wildlife diseases: from individuals to ecosystems. *Journal of Animal Ecology*, **80**, 19–38.

Torremorell, M., Allerson, M.W., Corzo, C., Diaz, A. & Gramer, M. (2012) Transmission of influenza A virus in pigs. *Transboundary and Emerging Diseases*, **59**, 68–84.

Torremorell, M., Juarez, A., Chavez, E., Yescas, J., Doporto, J.M. & Gramer, M. (2009) Procedures to eliminate H3N2 swine influenza virus from a pig herd. *The Veterinary record*, **165**, 74–7.

Tracey, J.A., Bevins, S.N., Vandewoude, S. & Crooks, K.R. (2014) An agent-based movement model to assess the impact of landscape fragmentation on disease transmission. *Ecosphere*, **5**, 1–24.

Turner, J., Begon, M. & Bowers, R.G. (2003) Modelling pathogen transmission: The interrelationship between local and global approaches. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 105–112.

Turner, M.G., Gardner, R.H. & O'Neill, R. (2001) *Landscape Ecology in Theory and Practice: Pattern and Processs*, 2nd ed. Springer-Verlag, New York.

USDA. (2015) *Swine 2012 Part I: Baseline Reference of Swine Health and Management in the United States, 2012*. Fort Collins, CO.

VanderWaal, K.L., Atwill, E.R., Hooper, S., Buckle, K. & McCowan, B. (2013) Network structure and prevalence of Cryptosporidium in Belding's ground squirrels. *Behavioral Ecology Sociobiology*, **2013**, 1951–1959.

VanderWaal, K.L., Atwill, E.R., Isbell, L.A. & McCowan, B. (2014a) Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation*, **169**, 136–146.

VanderWaal, K.L., Atwill, E.R., Isbell, L.A. & McCowan, B. (2014b) Linking social and pathogen transmission networks using microbial genetics in giraffe (Giraffa camelopardalis). *Journal of Animal Ecology*, **83**, 406–414.

VanderWaal, K.L. & Ezenwa, V.O. (2016) Heterogeneity in pathogen transmission: Mechanisms and methodology. *Functional Ecology*, **30**, 1606–1622.

Vedhara, K., Gill, S., Eldesouky, L., Campbell, B.K., Arevalo, J.M.G., Ma, J. & Cole, S.W. (2015) Personality and gene expression: Do individual differences exist in the leukocyte transcriptome? *Psychoneuroendocrinology*, **52**, 72–82.

Verdolin, J.L., Traud, A.L. & Dunn, R.R. (2014) Key players and hierarchical organization of prairie dog social networks. *Ecological Complexity*, **19**, 140–147.

Vial, F., Cleaveland, S., Rasmussen, G. & Haydon, D.T. (2006) Development of vaccination strategies for the management of rabies in African wild dogs. *Biological Conservation*, **131**, 180–192.

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

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

348–60.

Vincent, A., Awada, L., Brown, I., Chen, H., Claes, F., Dauphin, G., … Ciacci-Zanella, J. (2014) Review of influenza A Virus in swine worldwide: A call for increased surveillance and research. *Zoonoses and Public Health*, **61**, 4–17.

Vincent, A.L., Ma, W., Lager, K.M., Janke, B.H. & Richt, J. a. (2008) Swine Influenza Viruses: A North American Perspective. *Advances in Virus Research*, pp. 127–154.

Vincent, A.L., Ma, W., Lager, K.M., Richt, J.A., Janke, B.H., Sandbulte, M.R., … García-Sastre, A. (2012) Live attenuated influenza vaccine provides superior protection from heterologous infection in pigs with maternal antibodies without inducing vaccine-associated enhanced respiratory disease. *Journal of virology*, **86**, 10597–605.

Volz, E. & Meyers, L.A. (2007) Susceptible-infected-recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 2925–2934.

Webb, C.T., Brooks, C.P., Gage, K.L. & Antolin, M.F. (2006) Classic flea-borne transmission does not drive plague epizootics in prairie dogs. *Proceedings of the National Academy of Sciences*, **103**, 6236–6241.

Welch, D., Bansal, S. & Hunter, D.R. (2011) Statistical inference to advance network models in epidemiology. *Epidemics*, **3**, 38–45.

Welicky, R.L. & Sikkel, P.C. (2015) Decreased movement related to parasite infection in a diel migratory coral reef fish. *Behavioral Ecology and Sociobiology*, **69**, 1437–1446.

White, P.C.L., Böhm, M., Marion, G. & Hutchings, M.R. (2008) Control of bovine tuberculosis in British livestock: there is no "silver bullet". *Trends in Microbiology*, **16**, 420–7.

White, L.A., Forester, J.D. & Craft, M.E. (2017a) REVIEW: Dynamic, spatial models of parasite transmission in wildlife: their structure, applications, and remaining challenges. *Journal of Animal Ecology*.

White, L.A., Forester, J.D. & Craft, M.E. (2017b) Data from: Covariation between the physiological and behavioral components of pathogen transmission: Host heterogeneity determines epidemic outcomes. Dryad Digital Repository. URL https://doi.org/10.5061/dryad.8t201

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

White, L.A., Forester, J.D. & Craft, M.E. (2018a) Covariation between the physiological and behavioral components of pathogen transmission: host heterogeneity determines epidemic outcomes. *Oikos*, **127**, 538–552.

White, L.A., Forester, J.D. & Craft, M.E. (2018b) Dynamic, spatial models of parasite transmission in wildlife: Their structure, applications and remaining challenges. *Journal of Animal Ecology*, **87**, 559–580.

White, J.W., Rassweiler, A., Samhouri, J.F., Stier, A.C. & White, C. (2014) Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos*, **123**, 385–388.

White, L.A., Torremorell, M. & Craft, M.E. (2017d) Influenza A virus in swine breeding

herds: Combination of vaccination and biosecurity practices can reduce likelihood of endemic piglet reservoir. *Preventive Veterinary Medicine*, **138**, 55–69.

Wilson, A.D.M., Krause, S., James, R., Croft, D.P., Ramnarine, I.W., Borner, K.K., … Krause, J. (2014) Dynamic social networks in guppies (Poecilia reticulata). *Behavioral Ecology and Sociobiology*, **68**, 915–925.

With, K.A. (1997) The application of neutral landscape models in conservation biology. *Conservation Biology*, **11**, 1069–1080.

Woolhouse, M.E., Dye, C., Etard, J.F., Smith, T., Charlwood, J.D., Garnett, G.P., … Anderson, R.M. (1997) Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 338–342.

Wu, J., Dhingra, R., Gambhir, M. & Remais, J. V. (2013) Sensitivity analysis of infectious disease models: methods, advances and their application. *Journal of the Royal Society, Interface / the Royal Society*, **10**, 20121018.

Wylie, J. & Jolly, A. (2001) Patterns of chlamydia and gonorrhea infection in sexual networks in Manitoba, Canada. *Sexually Transmitted Diseases*, **28**, 14–24.

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G. & Alvarez-Cohen, L. (2015) High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. *mBio*, **6**, e02288-14.

Zohdy, S., Kemp, A.D., Durden, L.A., Wright, P.C. & Jernvall, J. (2012) Mapping the social network: tracking lice in a wild primate (*Microcebus rufus*) population to infer social contacts and vector potential. *BMC Ecology*, **12**, 4.

Appendix A. Glossary of network terms for Chapter 1

- **Centrality –** various metrics that indicate how important or influential an individual is to a network. Depending on the particular measurement, it may emphasize a high number of contacts (degree), being connected to others of high degree, or functioning as a control point or bridge between different parts of the network (betweenness).
- **Betweenness –** a node with high betweenness connects other nodes *via* the shortest possible path length serving as a key 'intermediary'.
    - **Flow-betweenness –** evaluates not just how often an individual is on the shortest geodesic path length between two other individuals, but also incorporates other indirect routes of 'flow' between two individuals that could circumvent the target individual (Corner *et al.*, 2003).
    - **Closeness –** the inverse of the sum of path lengths to all other nodes in the network; a node with high closeness has comparatively short path lengths to all other nodes.
- **Component –** a group of nodes in a network that are connected and thus offer a continuous transmission pathway; if all nodes are connected in a network, then the entire network is a single component.

- **Connectance –** the ratio of realized edges in a network relative to the total possible number of edges that could exist between nodes (May, 1974; Perkins *et al.*, 2009; Poisot & Gravel, 2014).
- **Cut-point potential –** difference between expected betweenness based on degree and actual betweenness; individuals with high cut-point potential connect disparate parts of the network (VanderWaal *et al.*, 2014*b*).
- **Degree –** the number of contacts of a given node, or similarly, the number of edges of a given node.
- **Degree distribution –** The distribution of contact numbers in the population; non-random associations will be reflected by a skewed degree distribution.
- **Edge –** a connection between two individuals or nodes in a contact network; if the association matrix, A_{ij}=1, than individuals *i* and *j* have an edge between them.
- **In-degree or in-strength –** the number of edges coming into a node in a directed network. In-degree corresponds to an unweighted network and in-strength to a weighted network.
- **Modularity (Q) –** a measure of the amount of community structure; the ratio of intra-group connections to inter-group connections. There are many methods of calculating modularity, but it typically ranges in value from 0 to 1 with higher values corresponding to greater community structure (Salathé & Jones, 2010).

- **Node –** a potential contact in the network; a node can take on many different hierarchical levels in disease modelling: an individual, a group, a population in a meta-population, etc.; most of the studies in this review treat individuals as nodes.
- **Robustness –** the resilience of a network to the random removal of individuals (nodes) or connections between them (edges).
- **Small world network –** a network with mostly local, but a few long-range contacts that often make the network more susceptible to epidemics.
- **Superspreaders –** individuals with an unexpectedly high number of contacts that contribute disproportionately to the spread of pathogens (Lloyd-Smith *et al.*, 2005*b*).
- **Transitivity (global clustering coefficient) –** the proportion of triangles in a network to linear triplets in a network; asks how often contacts of a given node are also contacts of each other (VanderWaal *et al.*, 2013).
- **Weighted network –** a network that incorporates interaction strengths between individuals. This can be ideal for exploring the effects of interaction frequency and duration in transmission (MacIntosh *et al.*, 2012).

Appendix B. Explanation of model terms and conceptual experimental design for Chapter 2

Transmission is the complex process of the successful establishment of a pathogen in a new host, which depends on the pathogen, host, and environment (Antonovics 2017). For a directly transmitted pathogen, this requires (1) contact between an infected host and a susceptible host, and then (2) depends upon the infectiousness and susceptibility of the infected and susceptible host, respectively, as well as the transmissibility of the pathogen itself (Fig B1). The **transmission rate (β)** is a term used in mathematical models that describes how quickly (per unit time) susceptible individuals in a population become infected (Begon *et al.* 2002). It encompasses both behavioral (β_c) and physiological (β_p) aspects of transmission (Hawley *et al.* 2011), each of which may have non-linear components (McCallum *et al.* 2017).

Contact rate (β_c) is the frequency with which hosts interact with each other. In the specific context of pathogen transmission, a contact represents a possible transmission event between an infected and a susceptible host pair. Hosts within a population may be considered superspreaders if they have higher than average contact rates (Lloyd-Smith *et al.* 2005b). The contact rate depends upon both local processes like group size, mate choice, or sickness behaviors as well as broad-scale processes like resource availability, host density, and migration (Lloyd-Smith *et al.* 2005a). Contact rate may also be altered by pathogen-induced changes in behavior or sickness behaviors (Ezenwa *et al.* 2016); we consider this scenario in our third experiment: infection status vs. contact rate. In all

models in this manuscript, contact rate is governed by the underlying dynamic network model (Fig B.2; Section 2.1).

The **physiological component of transmission (β_p)**, which we incorporate into the model as probability of transmission given an eligible contact between hosts, can be affected by a variety of factors including a host's innate physiological characteristics such as immunocompetence, shedding rate, latency period, and co-infection (Lehmer *et al.* 2010; Telfer *et al.* 2010; Hawley & Altizer 2011; Lass *et al.* 2013). We incorporate these possible factors into the model via host susceptibility, host infectiousness, and pathogen transmission efficiency (Fig B.2). **Susceptibility (s)** describes a host's innate immune and physiological response to pathogen exposure that helps determines whether a transmission event is successful. Host susceptibility may be affected by body condition (Beldomenico *et al.* 2009) and coinfection (Cattadori *et al.* 2008). Some studies suggest that a hosts' immunocompetence and sociality are linked, e.g., more extraverted individuals may be at higher risk of exposure, but may also have more active immune systems (Natoli *et al.* 2005; Vedhara *et al.* 2015). For the purposes of this paper, susceptibility only modifies the final probability of transmission, not the contact rate (Section 2.4). Thus, we consider susceptibility to be a host-driven trait. **Infectiousness (κ)** describes how effectively an infected host transmits pathogens to susceptible hosts. Depending on the system, this might be behavioral via sneezing, coughing, or biting, but it could also correspond to the pathogen load or shedding rate. For the purposes of the model, we consider infectiousness to be an intrinsic, host-determined trait (Section 2.5). Apart from the final experimental scenario of infection status vs. contact rate (Section

2.6), we do not consider any feedbacks of infectiousness on host behavior. Finally, we use **transmission efficiency (τ)** to describe transmissibility of the pathogen itself. This reflects the idea that, all else being equal and host physiology non-withstanding, some pathogens are more infectious than others.

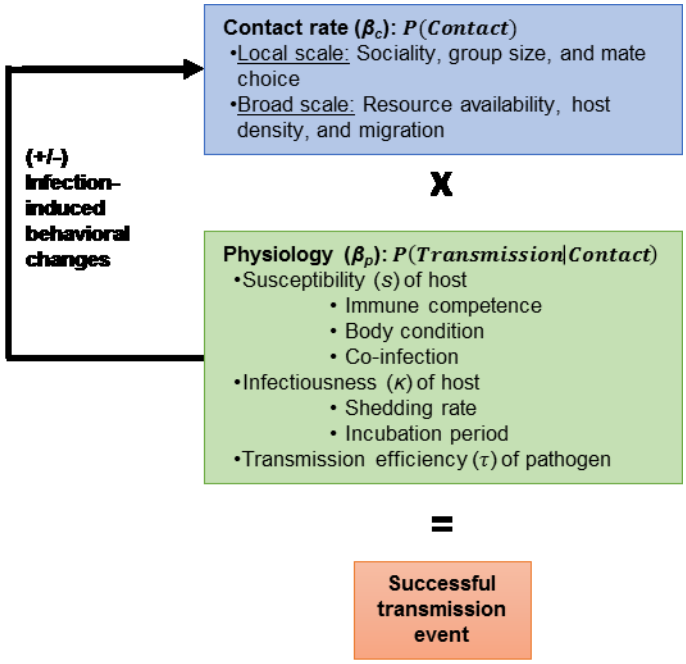


Figure B.1. A successful transmission event first requires contact between a susceptible and infected host (blue rectangle). Given contact, a variety of factors affecting host physiology and pathogen transmission

efficiency may affect the likelihood of transmission (green rectangle). Infection can result in infection-induced behavioral changes that induce a positive or negative feedback on contact rate (black arrow).

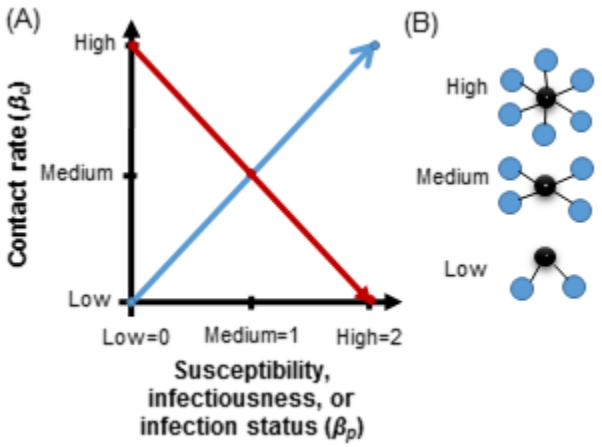


Figure B.2. A conceptual representation of the experimental design for this study. **(A)** Possible forms of covariation between the behavioral and physiological components of transmission, which are represented as contact rate (β_c) and susceptibility, infectiousness, and infection status (β_p). Positive covariation is shown in blue; negative covariation is shown in red. **(B)** The behavioral component of transmission (β_c) is incorporated in the dynamic network model via contact rate. This panel shows how contact rate is reflected as mean degree (average number of contacts per time step) in the dynamic network. From top to bottom, the focal node (black) has either higher than average, average, or a lower than average number of contacts (blue)—corresponding to the y-axis of panel A.

Appendix C. Methods supplement for Chapter 2

Experiment 1: Susceptibility vs. contact rate

Begin by loading the libraries and setting the random seed

```
require(EpiModel)

require(parallel)

set.seed(4321)
```

Parameter specifications

```
type.description<-"SvsCR" #Susceptibility vs. contact rate

degree.diff<-2 #degree difference: 2="low"; 4="high"

infection.p<-0.025 #infection probability: high=0.5, med=0.25, low=0.025

file_name<-"SvsCR2low"
```

Initialize the networks

We create an undirected network with 525 individuals or nodes. We also set up other specifications for the simulation: how long each simulation should run, the number of repetitions per parameter set, and whether we'd like to run the simulations in parallel.

```
num<-525; # number of individuals in simulation (ideally divisible by 3)

nw <- network.initialize(n = num, directed = FALSE) #initialize network

duration<-100 #duration of simulations

nsim<-10 #number of disease simulations to run per set of conditions

ncores<- detectCores() #number of cores (either on PC or to optimize on supercomputer)
```

Set up the vertex attributes for the network

Here we establish a gregarious attribute for our nodes which corresponds to β_c. The population is divided evenly into "low", "medium", and "high" contact rate groups of 175 individuals each.

```
greg<-c("low","med","high") #add "gregariousness" attribute to nodes in network
#greg<-sample(greg, num, replace=TRUE)
greg<-rep(greg, times=num/length(greg))
nw <- set.vertex.attribute(nw, "gregarious", greg)
#check distribution of gregarious phenotype in population
sum(get.vertex.attribute(nw, "gregarious")=="low")

## [1] 175

sum(get.vertex.attribute(nw, "gregarious")=="med")

## [1] 175

sum(get.vertex.attribute(nw, "gregarious")=="high")

## [1] 175
```

Model fit

We set up our formation and dissolution equations for the STERGM in this section. The target stats are done in terms of edge numbers, so if you want to think in terms of mean degree you must convert your target stats accordingly, as we do below. Our dissolution formula is very simple and only has a mean edge duration rate. You can read more about ERGM terms here.

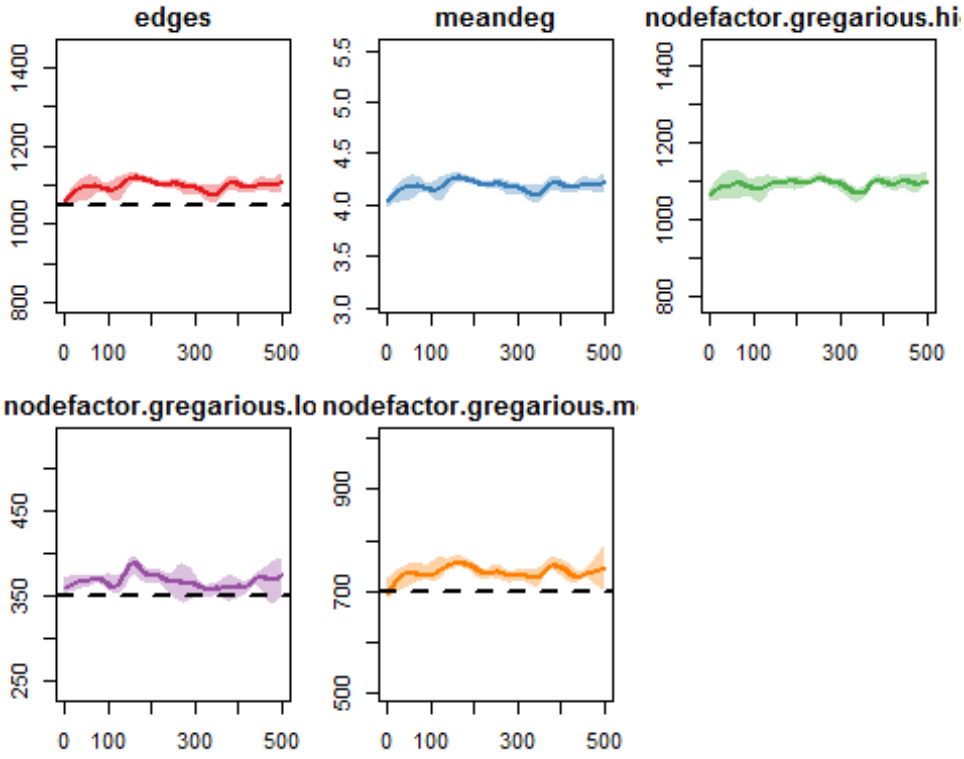
```
formation <- ~edges + nodefactor("gregarious")
mean.degree<-4
target.stats <- c((num/2)*mean.degree, (num/3)*(mean.degree-degree.diff), (num/3)*mean.
degree) #mean number of edges (number of nodes/2*degree), number of edges for gregariou
s=low individuals (number of nodes/3*target degree), number of edges for gregarious= me
d
coef.diss <- dissolution_coefs(dissolution = ~offset(edges), duration = 25)
```

Model diagnostics

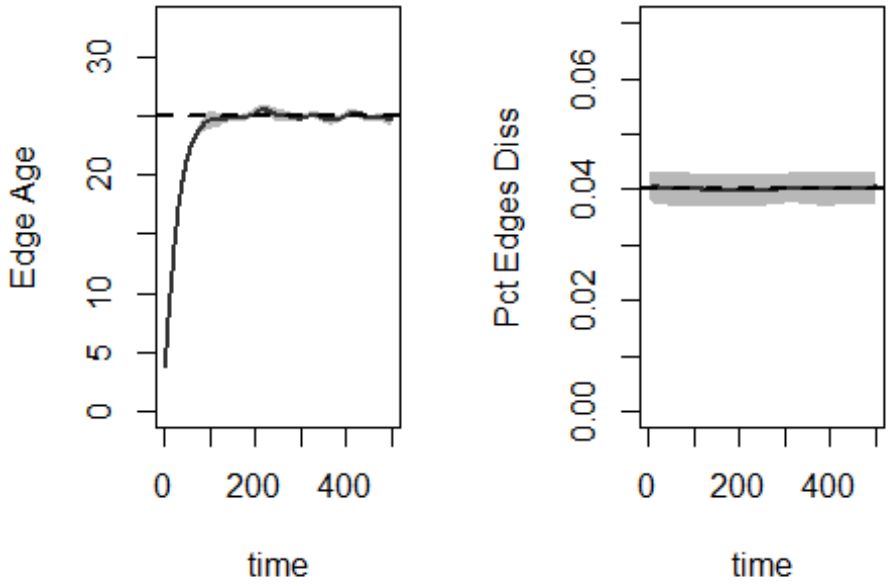
Before we simulate disease on the networks, we use the `netest` function to fit a temporal ERGM using our specified formation and dissolution parameters. The `netdx` function runs diagnostics on the fitted networks to see how well the simulated networks match the target specifications. We can view diagnostic results in table format and graphically.

```
est1 <- netest(nw, formation, target.stats, coef.diss, edapprox = TRUE)
dx <- netdx(est1, nsims = 5, nsteps = 500,
            nwstats.formula = ~edges + meandeg+ nodefactor("gregarious", base = 0))
##
## Network Diagnostics
## -----------------------
## - Simulating 5 networks
##   |*****|
## - Calculating formation statistics
## - Calculating duration statistics
##   |*****|
## - Calculating dissolution statistics
```

```
##   |*****|
##

dx

## EpiModel Network Diagnostics
## =======================
## Diagnostic Method: Dynamic
## Simulations: 5
## Time Steps per Sim: 500
##
## Formation Diagnostics
## ----------------------
##                            Target Sim Mean Pct Diff Sim SD
## edges                        1050 1096.762    0.045 34.802
## meandeg                        NA    4.178       NA  0.133
## nodefactor.gregarious.high     NA 1090.554       NA 39.269
## nodefactor.gregarious.low     350  366.605    0.047 22.649
## nodefactor.gregarious.med     700  736.365    0.052 31.200
##
## Dissolution Diagnostics
## ----------------------
##                 Target Sim Mean Pct Diff Sim SD
## Edge Duration    25.00   23.736   -0.051 23.129
## Pct Edges Diss    0.04    0.040    0.001  0.006

plot(dx, plots.joined=FALSE) #Remember for node factor plots, we are looking at number
of edges in the network
```




```
par(mfrow = c(1, 2))

plot(dx, type = "duration")

plot(dx, type = "dissolution")
```



Induce covariation

In order to induce covariation with the "gregarious" node factor, we write some custom modules that feed into EpiModel's `netsim` function. These functions modify the `dat` object (which contains all the network information) at time step `at`.

```
#Positive correlation between susceptibility and contact rate

suscept_positive <- function(dat, at) {

  ## Attributes

  if (at == 2) {

     n <- sum(dat$attr$active == 1)


    for (j in 1:n){

      if(dat$attr$gregarious[j]=="med"){
```

```
        dat$attr$susceptibility[j]<-1
      }
      if(dat$attr$gregarious[j]=="low"){
        dat$attr$susceptibility[j]<-0
      }
      if(dat$attr$gregarious[j]=="high"){
        dat$attr$susceptibility[j]<-2
      }
    }
    dat$nw <- set.vertex.attribute(dat$nw, "susceptibility", dat$attr$susceptibility)
    index.case<-which(dat$attr$status=="i")
  }

  ## Summary statistics
  if (at == 2) {
    dat$epi$susceptibility <- mean(dat$attr$susceptibility, na.rm = TRUE)
    dat$epi$index.case<-index.case
    dat$epi$suscept.index.case<-dat$attr$susceptibility[index.case]
  }
  return(dat)
}

#Negative covariation between susceptibility and contact rate
suscept_negative <- function(dat, at) {
  ## Attributes
  if (at == 2) {
     n <- sum(dat$attr$active == 1)
```

```
    for (j in 1:n){
      if(dat$attr$gregarious[j]=="med"){
        dat$attr$susceptibility[j]<-1
      }
      if(dat$attr$gregarious[j]=="low"){
        dat$attr$susceptibility[j]<-2
      }
      if(dat$attr$gregarious[j]=="high"){
        dat$attr$susceptibility[j]<-0
      }
    }
     dat$nw <- set.vertex.attribute(dat$nw, "susceptibility", dat$attr$susceptibility)
     index.case<-which(dat$attr$status=="i")
  }

  ## Summary statistics
  if (at == 2) {
    dat$epi$susceptibility <- mean(dat$attr$susceptibility, na.rm = TRUE)
    dat$epi$index.case<-index.case
    dat$epi$suscept.index.case<-dat$attr$susceptibility[index.case]
  }
  return(dat)
}

#Null covariation between susceptibility and contact rate
suscept_null <- function(dat, at) {
  ## Attributes
  if (at == 2) {
```

```
    n <- sum(dat$attr$active == 1)


    suscept<-c(0, 1, 2)
    dat$attr$susceptibility<-sample(suscept, n, replace=TRUE) #Susceptibility randomly
assigned


    dat$nw <- set.vertex.attribute(dat$nw, "susceptibility", dat$attr$susceptibility)
    index.case<-which(dat$attr$status=="i")
  }


  ## Summary statistics
  if (at == 2) {
    dat$epi$susceptibility <- mean(dat$attr$susceptibility, na.rm = TRUE)
    dat$epi$index.case<-index.case
    dat$epi$suscept.index.case<-dat$attr$susceptibility[index.case]
  }


  return(dat)
}
```

Refine transmission function

We also update the standard transmission function from the `EpiModel` package so that it incorporates individual β_p via susceptibility. In particular, we update the final transmisison probability, $P(t)$.

```
infection_susceptible<-function (dat, at)
{
  active <- dat$attr$active
```

```
  status <- dat$attr$status
  susceptibility<-dat$attr$susceptibility
  modes <- dat$param$modes
  mode <- idmode(dat$nw) #id numbers for a bipartite network
  inf.prob <- dat$param$inf.prob
  inf.prob.m2 <- dat$param$inf.prob.m2
  act.rate <- dat$param$act.rate
  nw <- dat$nw
  tea.status <- dat$control$tea.status
  idsSus <- which(active == 1 & status == "s")
  idsInf <- which(active == 1 & status == "i" & susceptibility>0) ##changed here: if in
itially initially exposed individual has a susceptibility of "0"-- will not propagate i
nfection
  nActive <- sum(active == 1)
  nElig <- length(idsInf)
  nInf <- nInfM2 <- totInf <- 0
  if (nElig > 0 && nElig < nActive) {
    del <- discord_edgelist(dat, idsInf, idsSus, at) #returns data frame with set of ed
ges
    #in which the status of two partners is one is susceptible and one is infected
    if (!(is.null(del))) { #if one or more such edges exist...
      inds <- which(get.vertex.pid(nw) %in% del$sus)
      # browser()
      # del$suscept<- susceptibility[inds]
      del$suscept<-susceptibility[del$sus]
#       any(get.vertex.pid(nw) == del$sus)
#       del$suscept <- dat$attr$susceptibility[get.vertex.pid(nw, del$sus)]
      del$infDur <- at - dat$attr$infTime[del$inf] #how long has each node been infecte
```

```
d?
      del$infDur[del$infDur == 0] <- 1
      linf.prob <- length(inf.prob)
      if (is.null(inf.prob.m2)) { #is the network bipartite?
        del$transProb <- ifelse(del$infDur <= linf.prob, #if the length of the infectio
us probability vector is less than or equal to infection
                                inf.prob[del$infDur], inf.prob[linf.prob]) #then...if n
ot, else....
      }
      else {
        del$transProb <- ifelse(del$sus <= nw %n% "bipartite",
                                ifelse(del$infDur <= linf.prob, inf.prob[del$infDur],
                                       inf.prob[linf.prob]), ifelse(del$infDur <=
                                                                      linf.prob, inf.pr
ob.m2[del$infDur], inf.prob.m2[linf.prob]))
      }
      #inter.eff- efficacy of an intervention
      #inter.start- time at which intervention starts
      #if there is an intervention and the current is after the intervention start time
, then...
      if (!is.null(dat$param$inter.eff) && at >= dat$param$inter.start) {
        del$transProb <- del$transProb * (1 - dat$param$inter.eff)
      }
      lact.rate <- length(act.rate) #act.rate-average number of transmissilbe acts per
partnership per unit time
      del$actRate <- ifelse(del$infDur <= lact.rate, act.rate[del$infDur],
                            act.rate[lact.rate])
      #del$transProb<- apply(del[,c('transProb', 'susceptibility')], 1, function(x) { (
```

```
x[1]) *(x[2])} )
      del$finalProb <- 1 - (1 - del$transProb*del$suscept)^del$actRate
      del$finalProb[which(del$finalProb>1)]<-1
#       browser()
      transmit <- rbinom(nrow(del), 1, del$finalProb)
      del <- del[which(transmit == 1), ]
      idsNewInf <- unique(del$sus)
      nInf <- sum(mode[idsNewInf] == 1)
      nInfM2 <- sum(mode[idsNewInf] == 2)
      totInf <- nInf + nInfM2
      if (totInf > 0) {
        if (tea.status == TRUE) {
          nw <- activate.vertex.attribute(nw, prefix = "testatus",
                                          value = "i", onset = at, terminus = Inf,
                                          v = idsNewInf)
        }
        dat$attr$status[idsNewInf] <- "i"
        dat$attr$infTime[idsNewInf] <- at
        form <- get_nwparam(dat)$formation
        fterms <- get_formula_terms(form)
        if ("status" %in% fterms) {
          nw <- set.vertex.attribute(nw, "status", dat$attr$status)
        }
      }
      if (any(names(nw$gal) %in% "vertex.pid")) {
        del$sus <- get.vertex.pid(nw, del$sus)
        del$inf <- get.vertex.pid(nw, del$inf)
      }
```

```
    }
  }
  if (totInf > 0) {
    del <- del[!duplicated(del$sus), ]
    if (at == 2) {
      dat$stats$transmat <- del
    }
    else {
      dat$stats$transmat <- rbind(dat$stats$transmat, del)
    }
  }
  if (at == 2) {
    dat$epi$si.flow <- c(0, nInf)
    if (modes == 2) {
      dat$epi$si.flow.m2 <- c(0, nInfM2)
    }
  }
  else {
    dat$epi$si.flow[at] <- nInf
    if (modes == 2) {
      dat$epi$si.flow.m2[at] <- nInfM2
    }
  }
  dat$nw <- nw
  return(dat)
}
```

Set up for running epidemic simulations

These components of the simulations are universal, so we set them up first here.

```
param <- param.net(inf.prob = infection.p, act.rate = 1) #Set infection probability and
action rate
init<-init.net(i.num=1) #Start with one infected individual
```

Positive covariation

The control.net function is where the custom modules can be implemented. We have added a `susceptibility.FUN` which induces that type of covariation that we want. We also indicate that we want to use the customized transmission function from above rather than the default and that we would like to look at at summary statistics for each 'gregarious' class. If desired, the output can be saved as a .Rdata file.

```
control1 <- control.net(type = "SI", nsteps = duration, nsims= nsim, ncores=ncores, epi
.by = "gregarious", infection.FUN=infection_susceptible, susceptibility.FUN=suscept_pos
itive)
sim1 <- netsim(est1, param, init, control1)
#save(sim1, file=paste(file_name,"positive.RData", sep="_"))
```

Negative covariation

```
control2 <- control.net(type = "SI", nsteps = duration, nsims = nsim, ncores=ncores, ep
i.by = "gregarious", infection.FUN=infection_susceptible, susceptibility.FUN=suscept_ne
gative)
sim2 <- netsim(est1, param, init, control2)
#save(sim2, file=paste(file_name,"negative.RData", sep="_"))
```

Null

Same contact structure, variability in susceptibility, and no covariation.

```
control3 <- control.net(type = "SI", nsteps = duration, nsims = nsim, ncores=ncores, ep
i.by = "gregarious", infection.FUN=infection_susceptible, susceptibility.FUN=suscept_nu
ll)
sim3 <- netsim(est1, param, init, control3)
#save(sim3, file=paste(file_name,"null.RData", sep="_"))
```

Control

Same contact structure, but no difference in susceptibility [i.e., susceptibility, *s*, always equals 1] and no covariation.

```
control4 <- control.net(type = "SI", nsteps = duration, nsims = nsim, ncores=ncores, ep
i.by = "gregarious")
sim4 <- netsim(est1, param, init, control4)
#save(sim4, file=paste(file_name,"control.RData", sep="_"))
```

Pull summary data from simulations

Here are a few functions to pull summary data from the simulated `netsim` ojects.

```
colMax<-function(data) sapply(data, max, na.rm=TRUE) #What is the maximum value reached
?
timeMax<-function(data) sapply(data, which.max) #When is max value reached?
CI95<-function(data) apply(data,1, quantile, probs=c(2.5, 97.5)/100) #What is 95% quant
ile CI?
suscept2<-function(simu, nsim){
  suscept_2nd<-matrix(NA, nrow = nsim, ncol = 1)
```

```
  for (i in 1:nsim){
    #store<-simu$network[[i]] #access stored dynamicNetwork object
    trans<-get_transmat(simu, sim=i) #access transmission matrix
    head(trans)
    if(length(trans)==0){
      suscept_2nd[i,1]<-NA
    }
    #susceptibility of first cases (after index case)
    else{
      trans[1,1] #time of first transmission event
      rowz<-which(trans[,1]==trans[1,1])
      suscept_2nd[i,1]<-mean(trans[rowz,4]) #mean of susceptibility of secondary cases
    }
  }
  suscept_2nd<-as.matrix(suscept_2nd)
  return(suscept_2nd)
}
```

We use those functions to pull summary statistics from each of our four scenarios.

```
positive<-sim1$epi$i.num
index.case.positive<-sim1$epi$suscept.index.case
suscept_2nd.positive<-suscept2(sim1, nsim)
write.csv(positive,file=paste(file_name,"positive.csv", sep="_"))
mean_positive<-rowMeans(positive)
maximum_positive<-colMax(positive)
peak_time_positive<-timeMax(positive)
quants_positive<-t(as.matrix(CI95(positive)))
```

```
negative<-sim2$epi$i.num

index.case.negative<-sim2$epi$suscept.index.case

suscept_2nd.negative<-suscept2(sim2, nsim)

write.csv(negative,file=paste(file_name,"negative.csv", sep="_"))

mean_negative<-rowMeans(negative)

maximum_negative<-colMax(negative)

peak_time_negative<-timeMax(negative)

quants_negative<-t(as.matrix(CI95(negative)))


null<-sim3$epi$i.num

index.case.null<-sim3$epi$suscept.index.case

suscept_2nd.null<-suscept2(sim3, nsim)

write.csv(null,file=paste(file_name,"null.csv", sep="_"))

mean_null<-rowMeans(null)

maximum_null<-colMax(null)

peak_time_null<-timeMax(null)

quants_null<-t(as.matrix(CI95(null)))


control<-sim4$epi$i.num

index.case.control<-rep(1,nsim)

suscept_2nd.control<-rep(1,nsim)

write.csv(control,file=paste(file_name,"control.csv", sep="_"))

mean_control<-rowMeans(control)

maximum_control<-colMax(control)

peak_time_control<-timeMax(control)

quants_control<-t(as.matrix(CI95(control)))
```

Create summary data frames

Since the simulation objects themselves can be quite large, especially for longer runs or more repeats, these summary dataframes can be saved as .csv files for easy access later.

```
timecourse.df<-data.frame(mean_positive=mean_positive, postive_CI=quants_positive, mean
_negative=mean_negative, negative_CI=quants_negative, mean_null=mean_null, null_CI=quan
ts_null, mean_control=mean_control, control_CI=quants_control)
summary.df<-data.frame(maximum_positive, peak_time_positive, maximum_negative, peak_tim
e_negative, maximum_null, peak_time_null, maximum_control, peak_time_control)
write.csv(timecourse.df, file=paste(file_name,"timecourse.csv", sep="_"))
write.csv(summary.df, file=paste(file_name,"summary.csv", sep="_"))
```

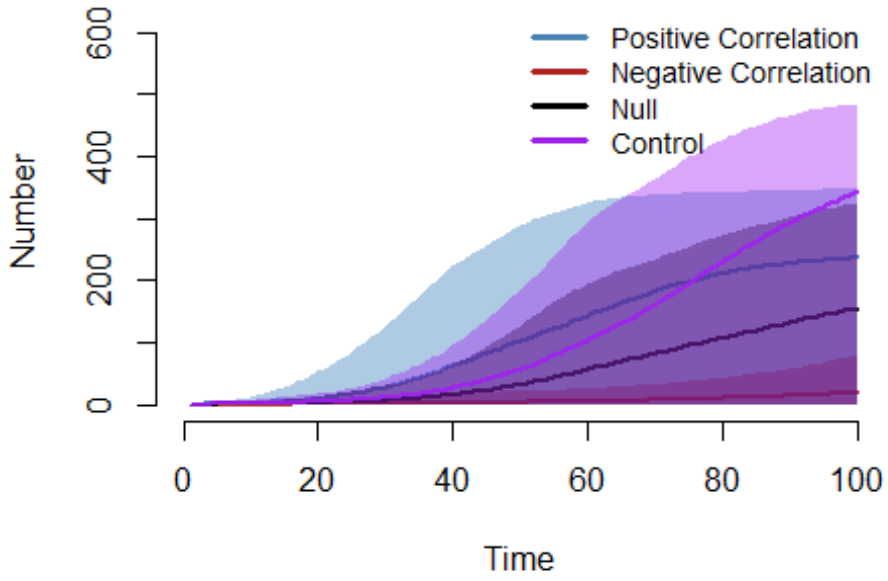
We also set up data frames in a long data format for random forest analysis and figure generation.

```
covariation<-c("positive", "negative", "null", "control")
covariation<-rep(covariation, each=nsim)
type<-rep(type.description, times=4*nsim)
infection.prob<-rep(infection.p, times=4*nsim)
maximum<-c(maximum_positive, maximum_negative, maximum_null, maximum_control)
peak_time<-c(peak_time_positive, peak_time_negative, peak_time_null, peak_time_control)
index.case<-as.numeric(c(index.case.positive, index.case.negative, index.case.null, ind
ex.case.control))
suscept_2nd<-as.numeric(c(suscept_2nd.positive, suscept_2nd.negative, suscept_2nd.null,
suscept_2nd.control))
long.df<-data.frame(type, infection.prob, covariation, maximum, peak_time, index.case,
suscept_2nd)
write.csv(long.df, file=paste(file_name,"long.csv", sep="_"))
```

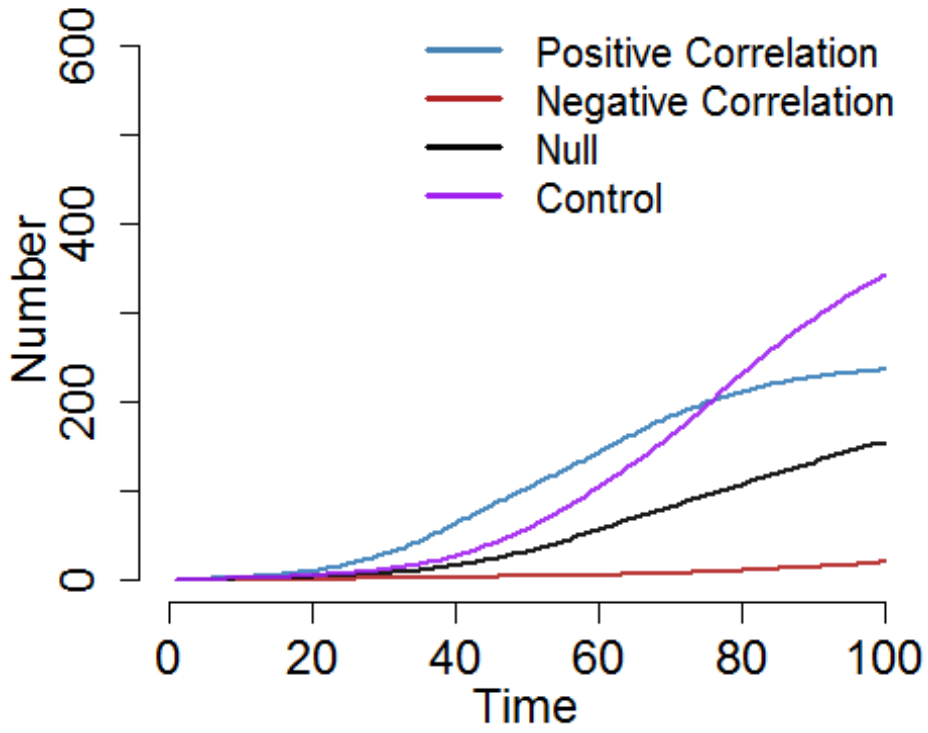
Plot data

Last, but not least, we can do a quick visual check and plot our simulation data.

```
par(mfrow = c(1,1))
plot(sim1, ylim=c(0,num+100), y = "i.num", sim.lines = FALSE, qnts = 1)
plot(sim2, y = "i.num", sim.lines = FALSE, qnts = 1,
     mean.col = "firebrick", qnts.col = "firebrick", add = TRUE)
plot(sim3, y = "i.num", sim.lines = FALSE, qnts = 1,
     mean.col = "black", qnts.col = "black", add = TRUE)
plot(sim4, y = "i.num", sim.lines = FALSE, qnts = 1,
     mean.col = "purple", qnts.col = "purple", add = TRUE)
legend("topright", c("Positive Correlation", "Negative Correlation", "Null", "Control")
, lty = 1, lwd = 3,
       col = c("steelblue", "firebrick", "black", "purple"), cex = 0.9, bty = "n")
```




```
par(mfrow = c(1,1), mar = c(3.5,4,1,1),mgp = c(2.25,1,0), cex.axis=1.5,cex.lab=1.5)
par(lwd=2)
plot(sim1, ylim=c(0,num+100), y = "i.num", sim.lines = FALSE, qnts=FALSE) #, ylim=c(0,0
.6))
plot(sim2, y = "i.num", sim.lines = FALSE, qnts=FALSE, mean.col = "firebrick", add = TR
UE)
plot(sim3, y = "i.num", sim.lines = FALSE, qnts=FALSE, mean.col = "black", add = TRUE)
plot(sim4, y = "i.num", sim.lines = FALSE, qnts=FALSE,mean.col = "purple", add = TRUE)
legend("topright", c("Positive Correlation", "Negative Correlation", "Null", "Control")
, lty = 1, lwd = 3, col = c("steelblue", "firebrick", "black", "purple"), cex = 1.25, b
ty = "n")
```



Session Information

```
## Record version information used in this analysis
print(sessionInfo(), locale = TRUE)

## R version 3.3.2 (2016-10-31)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 14393)
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
```

```
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] parallel  stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
## [1] EpiModel_1.2.7       tergm_3.4.0          ergm_3.6.0
## [4] statnet.common_3.3.0 networkDynamic_0.9.0 network_1.13.0
## [7] deSolve_1.13
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.6        knitr_1.14         magrittr_1.5
##  [4] MASS_7.3-45        doParallel_1.0.10  ape_3.5
##  [7] lattice_0.20-34    foreach_1.4.3      stringr_1.1.0
## [10] tools_3.3.2        grid_3.3.2         lpSolve_5.6.13
## [13] nlme_3.1-128       coda_0.18-1        iterators_1.0.8
## [16] htmltools_0.3.5    lazyeval_0.2.0     yaml_2.1.13
## [19] digest_0.6.10      Matrix_1.2-7.1     RColorBrewer_1.1-2
## [22] formatR_1.4        codetools_0.2-15   trust_0.1-7
## [25] robustbase_0.92-6  evaluate_0.9       rmarkdown_1.0
## [28] stringi_1.1.1      compiler_3.3.2     DEoptimR_1.0-6
```

Appendix D. Derivation of logistic growth equation for Chapter 2

We sought to measure the rate at which the pathogen was spreading through the theoretical population, or functionally, a realized transmission rate or "realized β". We start with an underlying SI model, but with the addition of some fraction of the susceptible pool, *r*, that cannot be infected due to social isolation or lack of susceptibility (as a result of experimental design with variable β_c and/or β_p):

$$\frac{dS}{dt} = -\beta(1-r)SI$$

$$\frac{dI}{dt} = \beta(1-r)SI$$

We further assume that the carrying capacity, *K* (in this context, the maximum number of individuals that can be infected), can vary and reflects the sum of the number of eligible susceptible infected individuals and infected individuals:

$$K = (1-r)S + I \rightarrow S = \frac{K-I}{1-r}$$

We than solve for $\frac{dI}{dt}$ in terms of I:

$$\frac{dI}{dt} = \beta I(1-r)\frac{K-I}{1-r} = \beta I(K-I)$$

We can put this in the recognizable, logistic growth form by multiplying the right side by K/K:

$$\frac{dI}{dt} = \beta IK\frac{(K-I)}{K} = \beta KI\left(1-\frac{I}{K}\right)$$

After integrating with partial fractions and using $I(t=0) = I_0$ as the initial condition, we get:

$$I(t) = \frac{K}{1+\frac{(K-I_o)}{I_0}e^{-\beta Kt}}$$

Appendix E. Additional time course simulations for Chapter 2

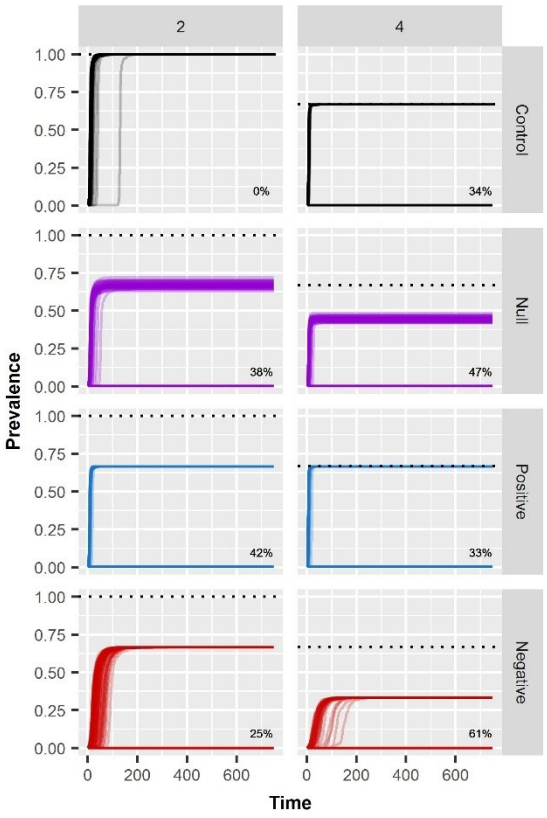


Figure E.1. Time course of simulated epidemics for susceptibility vs. contact rate for an infection probability of $\tau = 0.25$. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in susceptibility, no covariation), null (variability in susceptibility, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower righthand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population has a β_c=0 limiting maximum prevalence to 0.66.

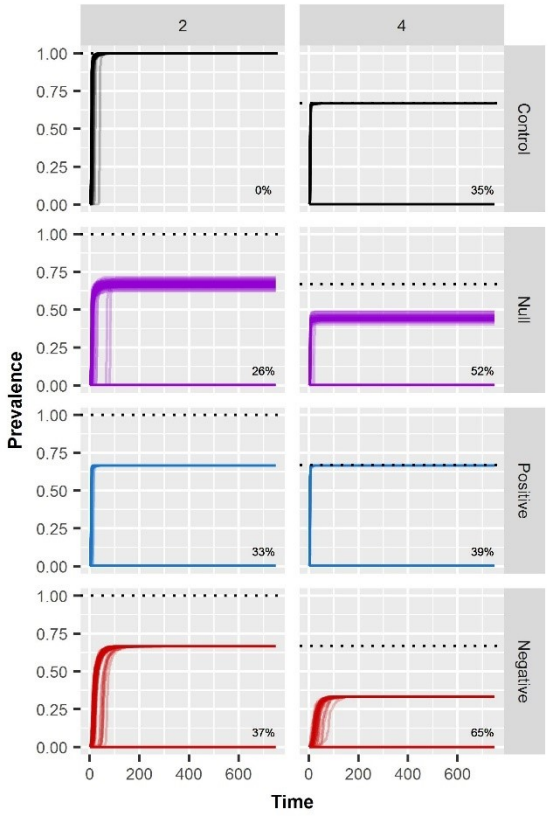


Figure E.2. Time course of simulated epidemics for susceptibility vs. contact rate for an infection probability of $\tau = 0.5$. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in susceptibility, no covariation), null (variability in susceptibility, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower righthand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population has a $\beta_c = 0$ limiting maximum prevalence to 0.66.

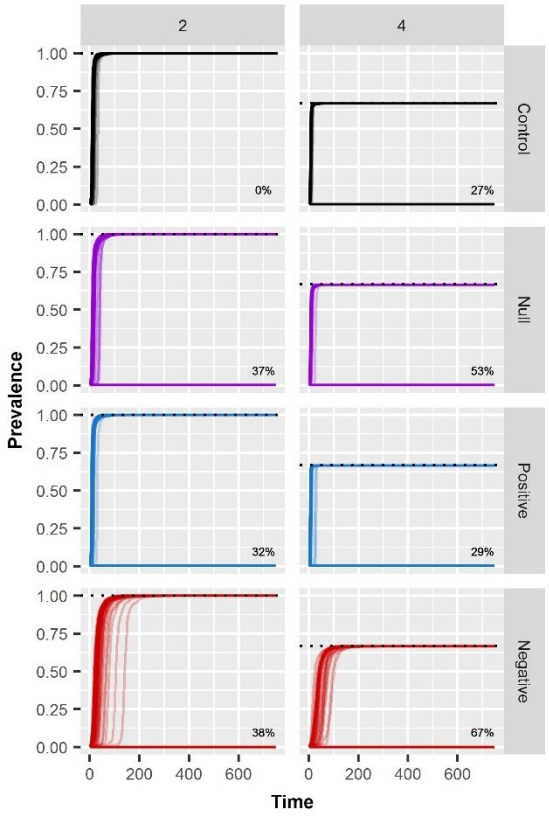


Figure E.3. Time course of simulated epidemics for infectiousness vs. contact rate for the medium infection probability tested of $\tau = 0.25$. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in infectiousness, no covariation), null (variability in infectiousness, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower righthand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population has a β_c=0 limiting maximum prevalence to 0.66.

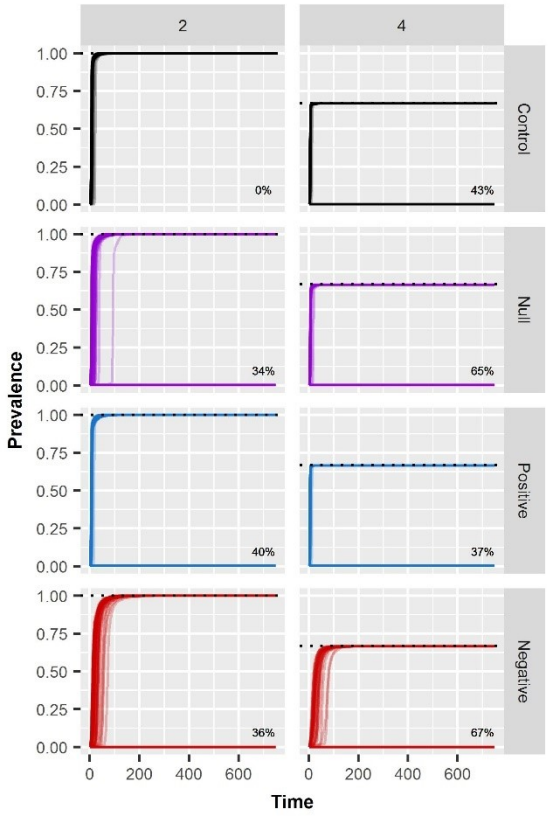


Figure E.4. Time course of simulated epidemics for infectiousness vs. contact rate for the high infection probability tested of τ= 0.5. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in infectiousness, no covariation), null (variability in infectiousness, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower righthand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure. For higher variations in contact rate, one-third of the population has a β_c=0 limiting maximum prevalence to 0.66.

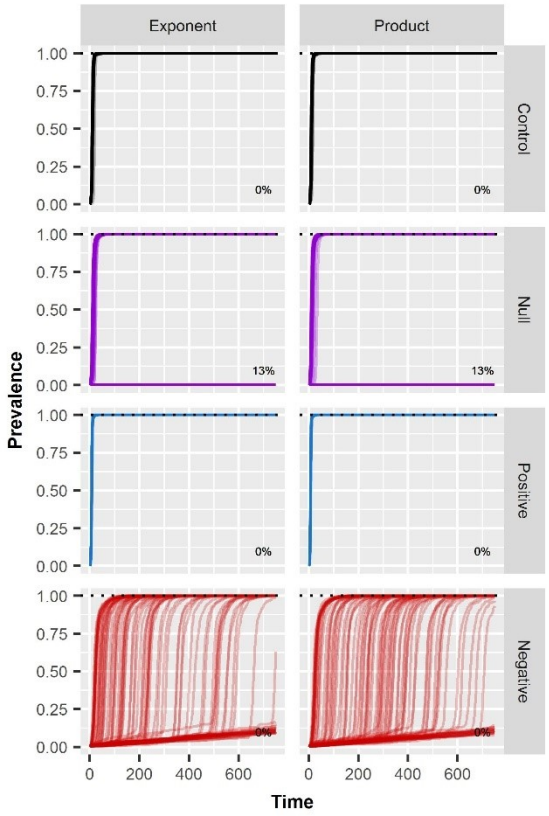


Figure E.5. Time course of simulated epidemics for infection status vs. contact rate for the medium infection probability tested of τ=0.25. Columns correspond to how infectiousness was modelled (either in the exponent or the product of the final transmission probability), and rows correspond to the mechanism of covariation: control (all infection statuses have equal mean degree), null (variability in susceptibility, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower righthand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure.

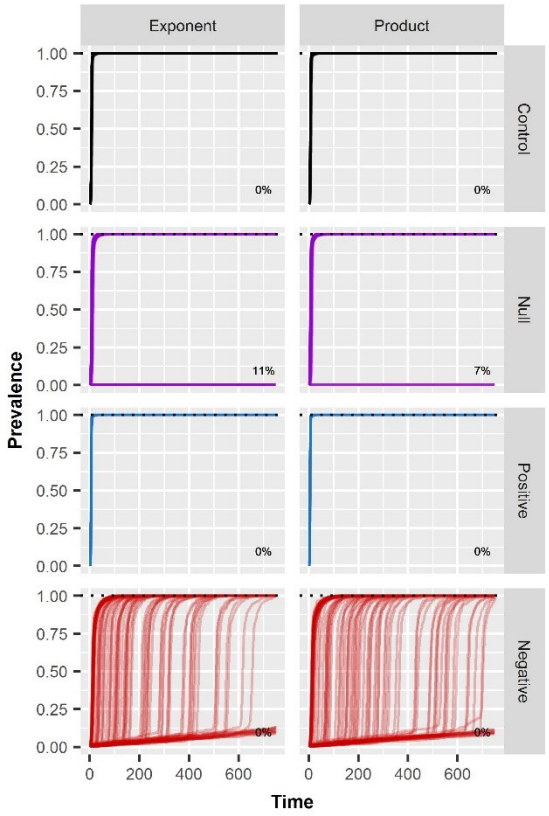


Figure E.6. Time course of simulated epidemics for infection status vs. contact rate for the high infection probability tested of $\tau = 0.5$. Columns correspond to how infectiousness was modelled (either in the exponent or the product of the final transmission probability), and rows correspond to the mechanism of covariation: control (all infection statuses have equal mean degree), null (variability in susceptibility, no covariation), positive covariation, and negative covariation. Individual trials are shown as semi-transparent lines and the colors—black, purple, blue, and red—correspond to control, null, positive, and negative covariation cases respectively. Percentages in the lower right hand corner of each panel describe the percentage of epidemics fading out. The dashed lines in each panel correspond to the expected maximum prevalence based on contact structure.

Appendix F. Supplementary tables for Chapter 3

Table F.1. The events (j) and the transition rates (λ_j) that govern their implementation in the model utilizing Gillespie's Direct algorithm. Classes of animals are listed in Table 3.2, and their spatial relationships are displayed in Figure 3.1. Epidemiological and demographic parameters are described in Table 3.2.

Description	j	Rate (λ_j)	Event/Transition
Direct transmission: Sows and gilts	1	$\beta_d \cdot S_1 \cdot (I_1 + VI_1)$	$S_1 \rightarrow S_1\text{-}1$; $E_1 \rightarrow E_1\text{+}1$
	2	$\beta_d \cdot S_2(I_2 + I_3 + I_4 + VI_2 + VI_3 + VI_4)$	$S_2 \rightarrow S_2\text{-}1$; $E_2 \rightarrow E_2\text{+}1$
	3	$\beta_d \cdot S_3(I_2 + I_3 + I_4 + VI_2 + VI_3 + VI_4)$	$S_3 \rightarrow S_3\text{-}1$; $E_3 \rightarrow E_3\text{+}1$
	4	$\beta_d \cdot S_4(I_2 + I_3 + I_4 + VI_2 + VI_3 + VI_4)$	$S_4 \rightarrow S_4\text{-}1$; $E_4 \rightarrow E_4\text{+}1$
	5	$\beta_d \cdot S_5(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_5 \rightarrow S_5\text{-}1$; $E_5 \rightarrow E_5\text{+}1$
	6	$\beta_d \cdot S_6(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_6 \rightarrow S_6\text{-}1$; $E_6 \rightarrow E_6\text{+}1$
Direct transmission: Piglets (no maternal immunity)	7	$\beta_d^p \cdot S_7(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_7 \rightarrow S_7\text{-}1$; $E_7 \rightarrow E_7\text{+}1$
	8	$\beta_d^p \cdot S_8(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_8 \rightarrow S_8\text{-}1$; $E_8 \rightarrow E_8\text{+}1$
	9	$\beta_d^p \cdot S_9(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_9 \rightarrow S_9\text{-}1$; $E_9 \rightarrow E_9\text{+}1$
	10	$\beta_d^p \cdot S_{10}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{10} \rightarrow S_{10}\text{-}1$; $E_{10} \rightarrow E_{10}\text{+}1$
Direct transmission: Piglets (with maternal immunity)	11	$\beta_d^{pm} \cdot S_{11}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{11} \rightarrow S_{11}\text{-}1$; $E_7 \rightarrow E_7\text{+}1$
	12	$\beta_d^{pm} \cdot S_{12}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{12} \rightarrow S_{12}\text{-}1$; $E_8 \rightarrow E_8\text{+}1$
	13	$\beta_d^{pm} \cdot S_{13}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{13} \rightarrow S_{13}\text{-}1$; $E_9 \rightarrow E_9\text{+}1$
	14	$\beta_d^{pm} \cdot S_{14}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{14} \rightarrow S_{14}\text{-}1$; $E_{10} \rightarrow E_{10}\text{+}1$
Direct transmission: Piglets (vaccine-induced immunity)	15	$\beta_{vax}^p \cdot S_{15}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{15} \rightarrow S_{15}\text{-}1$; $E_7 \rightarrow E_7\text{+}1$
	16	$\beta_{vax}^p \cdot S_{16}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{16} \rightarrow S_{16}\text{-}1$; $E_8 \rightarrow E_8\text{+}17$
	17	$\beta_{vax}^p \cdot S_{17}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{17} \rightarrow S_{17}\text{-}1$; $E_9 \rightarrow E_9\text{+}1$
	18	$\beta_{vax}^p \cdot S_{18}(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_{18} \rightarrow S_{18}\text{-}1$; $E_{10} \rightarrow E_{10}\text{+}1$

Indirect transmission	19	$\beta_{ind} \cdot S_1(\sum_{i=2}^{10} I_i + \sum_{i=2}^{6} VI_i)$	$S_1 \rightarrow S_1\text{-}1$; $E_1 \rightarrow E_1+1$
	20	$\beta_{ind} \cdot S_2(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	$S_2 \rightarrow S_2\text{-}1$; $E_2 \rightarrow E_2+1$
	21	$\beta_{ind} \cdot S_3(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	$S_3 \rightarrow S_3\text{-}1$; $E_3 \rightarrow E_3+1$
	22	$\beta_{ind} \cdot S_4(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	$S_4 \rightarrow S_4\text{-}1$; $E_4 \rightarrow E_4+1$
	23	$\beta_{ind} \cdot S_5(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_5 \rightarrow S_5\text{-}1$; $E_5 \rightarrow E_5+1$
	24	$\beta_{ind} \cdot S_6(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_6 \rightarrow S_6\text{-}1$; $E_6 \rightarrow E_6+1$
	25	$\beta_{ind}^{p} \cdot S_7(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_7 \rightarrow S_7\text{-}1$; $E_7 \rightarrow E_7+1$
	26	$\beta_{ind}^{p} \cdot S_8(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_8 \rightarrow S_8\text{-}1$; $E_8 \rightarrow E_8+1$
	27	$\beta_{ind}^{p} \cdot S_9(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_9 \rightarrow S_9\text{-}1$; $E_9 \rightarrow E_9+1$
	28	$\beta_{ind}^{p} \cdot S_{10}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{10} \rightarrow S_{10}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
	29	$\beta_{ind}^{pm} \cdot S_{11}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{11} \rightarrow S_{11}\text{-}1$; $E_7 \rightarrow E_7+1$
	30	$\beta_{ind}^{pm} \cdot S_{12}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{12} \rightarrow S_{12}\text{-}1$; $E_8 \rightarrow E_8+1$
	31	$\beta_{ind}^{pm} \cdot S_{13}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{13} \rightarrow S_{13}\text{-}1$; $E_9 \rightarrow E_9+1$
	32	$\beta_{ind}^{pm} \cdot S_{14}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{14} \rightarrow S_{14}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
	33	$\beta_{ind\,vax}^{p} \cdot S_{15}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{15} \rightarrow S_{15}\text{-}1$; $E_7 \rightarrow E_7+1$
	34	$\beta_{ind\,vax}^{p} \cdot S_{16}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{16} \rightarrow S_{16}\text{-}1$; $E_8 \rightarrow E_8+17$
	35	$\beta_{ind\,vax}^{p} \cdot S_{17}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{17} \rightarrow S_{17}\text{-}1$; $E_9 \rightarrow E_9+1$
	36	$\beta_{ind\,vax}^{p} \cdot S_{18}(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	$S_{18} \rightarrow S_{18}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
Indirect transmission with gilt isolation (Utilized for Experiments #4 & 15)	19b	0	$S_1 \rightarrow S_1\text{-}1$; $E_1 \rightarrow E_1+1$
	20b	$\beta_{ind} \cdot S_2(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_2 \rightarrow S_2\text{-}1$; $E_2 \rightarrow E_2+1$
	21b	$\beta_{ind} \cdot S_3(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_3 \rightarrow S_3\text{-}1$; $E_3 \rightarrow E_3+1$
	22b	$\beta_{ind} \cdot S_4(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	$S_4 \rightarrow S_4\text{-}1$; $E_4 \rightarrow E_4+1$
	23b	$\beta_{ind} \cdot S_5(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_5 \rightarrow S_5\text{-}1$; $E_5 \rightarrow E_5+1$
	24b	$\beta_{ind} \cdot S_6(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_6 \rightarrow S_6\text{-}1$; $E_6 \rightarrow E_6+1$
	25b	$\beta_{ind}^{p} \cdot S_7(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_7 \rightarrow S_7\text{-}1$; $E_7 \rightarrow E_7+1$
	26b	$\beta_{ind}^{p} \cdot S_8(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_8 \rightarrow S_8\text{-}1$; $E_8 \rightarrow E_8+1$

	27b	$\beta_{ind}^{p} \cdot S_9(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_9 \rightarrow S_9\text{-}1$; $E_9 \rightarrow E_9+1$
	28b	$\beta_{ind}^{p} \cdot S_{10}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{10} \rightarrow S_{10}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
	29b	$\beta_{ind}^{pm} \cdot S_{11}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{11} \rightarrow S_{11}\text{-}1$; $E_7 \rightarrow E_7+1$
	30b	$\beta_{ind}^{pm} \cdot S_{12}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{12} \rightarrow S_{12}\text{-}1$; $E_8 \rightarrow E_8+1$
	31b	$\beta_{ind}^{pm} \cdot S_{13}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{13} \rightarrow S_{13}\text{-}1$; $E_9 \rightarrow E_9+1$
	32b	$\beta_{ind}^{pm} \cdot S_{14}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{14} \rightarrow S_{14}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
	33b	$\beta_{ind\ vax}^{p} \cdot S_{15}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{15} \rightarrow S_{15}\text{-}1$; $E_7 \rightarrow E_7+1$
	34b	$\beta_{ind\ vax}^{p} \cdot S_{16}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{16} \rightarrow S_{16}\text{-}1$; $E_8 \rightarrow E_8+17$
	35b	$\beta_{ind\ vax}^{p} \cdot S_{17}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{17} \rightarrow S_{17}\text{-}1$; $E_9 \rightarrow E_9+1$
	36b	$\beta_{ind\ vax}^{p} \cdot S_{18}(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	$S_{18} \rightarrow S_{18}\text{-}1$; $E_{10} \rightarrow E_{10}+1$
Exposed to Infected	37	$\sigma \cdot E_1$	$E_1 \rightarrow E_1\text{-}1$; $I_1 \rightarrow I_1+1$
	38	$\sigma \cdot E_2$	$E_2 \rightarrow E_2\text{-}1$; $I_2 \rightarrow I_2+1$
	39	$\sigma \cdot E_3$	$E_3 \rightarrow E_3\text{-}1$; $I_3 \rightarrow I_3+1$
	40	$\sigma \cdot E_4$	$E_4 \rightarrow E_4\text{-}1$; $I_4 \rightarrow I_4+1$
	41	$\sigma \cdot E_5$	$E_5 \rightarrow E_5\text{-}1$; $I_5 \rightarrow I_5+1$
	42	$\sigma \cdot E_6$	$E_6 \rightarrow E_6\text{-}1$; $I_6 \rightarrow I_6+1$
	43	$\sigma \cdot E_7$	$E_7 \rightarrow E_7\text{-}1$; $I_7 \rightarrow I_7+1$
	44	$\sigma \cdot E_8$	$E_8 \rightarrow E_8\text{-}1$; $I_8 \rightarrow I_8+1$
	45	$\sigma \cdot E_9$	$E_9 \rightarrow E_9\text{-}1$; $I_9 \rightarrow I_9+1$
	46	$\sigma \cdot E_{10}$	$E_{10} \rightarrow E_{10}\text{-}1$; $I_{10} \rightarrow I_{10}+1$
Infected to Recovered	47	$\gamma \cdot I_1$	$I_1 \rightarrow I_1\text{-}1$; $R_1 \rightarrow R_1+1$
	48	$\gamma \cdot I_2$	$I_2 \rightarrow I_2\text{-}1$; $R_2 \rightarrow R_2+1$
	49	$\gamma \cdot I_3$	$I_3 \rightarrow I_3\text{-}1$; $R_3 \rightarrow R_3+1$
	50	$\gamma \cdot I_4$	$I_4 \rightarrow I_4\text{-}1$; $R_4 \rightarrow R_4+1$
	51	$\gamma \cdot I_5$	$I_5 \rightarrow I_5\text{-}1$; $R_5 \rightarrow R_5+1$
	52	$\gamma \cdot I_6$	$I_6 \rightarrow I_6\text{-}1$; $R_6 \rightarrow R_6+1$
	53	$\gamma \cdot I_7$	$I_7 \rightarrow I_7\text{-}1$; $R_7 \rightarrow R_7+1$

	54	$\gamma \cdot I_8$	$I_8 \rightarrow I_8$-1; $R_8 \rightarrow R_8$+1
	55	$\gamma \cdot I_9$	$I_9 \rightarrow I_9$-1; $R_9 \rightarrow R_9$+1
	56	$\gamma \cdot I_{10}$	$I_{10} \rightarrow I_{10}$-1; $R_{10} \rightarrow R_{10}$+1
Non-disease related mortality: Susceptible sows and gilts	57	$\mu \cdot S_1$	$S_1 \rightarrow S_1$-1
	58	$\mu \cdot S_2$	$S_2 \rightarrow S_2$-1
	59	$\mu \cdot S_3$	$S_3 \rightarrow S_3$-1
	60	$\mu \cdot S_4$	$S_4 \rightarrow S_4$-1
	61	$\mu \cdot S_5$	$S_5 \rightarrow S_5$-1
	62	$\mu \cdot S_6$	$S_6 \rightarrow S_6$-1
Non-disease related mortality: Susceptible piglets	63	$\mu^p \cdot S_7$	$S_7 \rightarrow S_7$-1
	64	$\mu^p \cdot S_8$	$S_8 \rightarrow S_8$-1
	65	$\mu^p \cdot S_9$	$S_9 \rightarrow S_9$-1
	66	$\mu^p \cdot S_{10}$	$S_{10} \rightarrow S_{10}$-1
	67	$\mu^p \cdot S_{11}$	$S_{11} \rightarrow S_{11}$-1
	68	$\mu^p \cdot S_{12}$	$S_{12} \rightarrow S_{12}$-1
	69	$\mu^p \cdot S_{13}$	$S_{13} \rightarrow S_{13}$-1
	70	$\mu^p \cdot S_{14}$	$S_{14} \rightarrow S_{14}$-1
	71	$\mu^p \cdot S_{15}$	$S_{15} \rightarrow S_{15}$-1
	72	$\mu^p \cdot S_{16}$	$S_{16} \rightarrow S_{16}$-1
	73	$\mu^p \cdot S_{17}$	$S_{17} \rightarrow S_{17}$-1
	74	$\mu^p \cdot S_{18}$	$S_{18} \rightarrow S_{18}$-1
Non-disease related mortality: Exposed animals	75	$\mu \cdot E_1$	$E_1 \rightarrow E_1$-1
	76	$\mu \cdot E_2$	$E_2 \rightarrow E_2$-1
	77	$\mu \cdot E_3$	$E_3 \rightarrow E_3$-1
	78	$\mu \cdot E_4$	$E_4 \rightarrow E_4$-1
	79	$\mu \cdot E_5$	$E_5 \rightarrow E_5$-1
	80	$\mu \cdot E_6$	$E_6 \rightarrow E_6$-1
	81	$\mu^p \cdot E_7$	$E_7 \rightarrow E_7$-1

	82	$\mu^p \cdot E_8$	$E_8 \rightarrow E_8$-1
	83	$\mu^p \cdot E_9$	$E_9 \rightarrow E_9$-1
	84	$\mu^p \cdot E_{10}$	$E_{10} \rightarrow E_{10}$-1
Non-disease related mortality: Infected animals	85	$\mu \cdot I_1$	$I_1 \rightarrow I_1$-1
	86	$\mu \cdot I_2$	$I_2 \rightarrow I_2$-1
	87	$\mu \cdot I_3$	$I_3 \rightarrow I_3$-1
	88	$\mu \cdot I_4$	$I_4 \rightarrow I_4$-1
	89	$\mu \cdot I_5$	$I_5 \rightarrow I_5$-1
	90	$\mu \cdot I_6$	$I_6 \rightarrow I_6$-1
	91	$\mu^p \cdot I_7$	$I_7 \rightarrow I_7$-1
	92	$\mu^p \cdot I_8$	$I_8 \rightarrow I_8$-1
	93	$\mu^p \cdot I_9$	$I_9 \rightarrow I_9$-1
	94	$\mu^p \cdot I_{10}$	$I_{10} \rightarrow I_{10}$-1
Non-disease related mortality: Recovered animals	95	$\mu \cdot R_1$	$R_1 \rightarrow R_1$-1
	96	$\mu \cdot R_2$	$R_2 \rightarrow R_2$-1
	97	$\mu \cdot R_3$	$R_3 \rightarrow R_3$-1
	98	$\mu \cdot R_4$	$R_4 \rightarrow R_4$-1
	99	$\mu \cdot R_5$	$R_5 \rightarrow R_5$-1
	100	$\mu \cdot R_6$	$R_6 \rightarrow R_6$-1
	101	$\mu^p \cdot R_7$	$R_7 \rightarrow R_7$-1
	102	$\mu^p \cdot R_8$	$R_8 \rightarrow R_8$-1
	103	$\mu^p \cdot R_9$	$R_9 \rightarrow R_9$-1
	104	$\mu^p \cdot R_{10}$	$R_{10} \rightarrow R_{10}$-1
Non-disease related mortality: Vaccinated animals	105	$\mu \cdot V_1$	$V_1 \rightarrow V_1$-1
	106	$\mu \cdot V_2$	$V_2 \rightarrow V_2$-1
	107	$\mu \cdot V_3$	$V_3 \rightarrow V_3$-1
	108	$\mu \cdot V_4$	$V_4 \rightarrow V_4$-1
	109	$\mu \cdot V_5$	$V_5 \rightarrow V_5$-1

	110	$\mu \cdot V_6$	V_6→V_6-1
Vaccinated animals revert to being susceptible	111	$\omega_v \cdot V_1$	V_1→V_1-1; S_1→ S_1+1
	112	$\omega_v \cdot V_2$	V_2→V_2-1; S_2→ S_2+1
	113	$\omega_v \cdot V_3$	V_3→V_3-1; S_3→ S_3+1
	114	$\omega_v \cdot V_4$	V_4→V_4-1; S_4→ S_4+1
	115	$\omega_v \cdot V_5$	V_5→V_5-1; S_5→ S_5+1
	116	$\omega_v \cdot V_6$	V_6→V_6-1; S_6→ S_6+1
Vaccinated animals become infected directly	117	$\beta_{vax} \cdot V_1 \cdot (I_1 + VI_1)$	V_1→V_1-1; VE_1→ VE_1+1
	118	$\beta_{vax} \cdot V_2(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	V_2→V_2-1; VE_2→ VE_2+1
	119	$\beta_{vax} \cdot V_3(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	V_3→V_3-1; VE_3→ VE_3+1
	120	$\beta_{vax} \cdot V_4(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	V_4→V_4-1; VE_4→ VE_4+1
	121	$\beta_{vax} \cdot V_5(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	V_5→V_5-1; VE_5→ VE_5+1
	122	$\beta_{vax} \cdot V_6(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	V_6→V_6-1; VE_6→ VE_6+1
Vaccinated animals become exposed indirectly	123	$\beta_{vax} \cdot V_1(\sum_{i=2}^{10} I_i + \sum_{i=2}^{6} VI_i)$	V_1→V_1-1; VE_1→ VE_1+1
	124	$\beta_{vax} \cdot V_2(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	V_2→V_2-1; VE_2→ VE_2+1
	125	$\beta_{vax} \cdot V_3(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	V_3→V_3-1; VE_3→ VE_3+1
	126	$\beta_{vax} \cdot V_4(I_1 + \sum_{i=5}^{10} I_i + VI_1 + VI_5 + VI_6)$	V_4→V_4-1; VE_4→ VE_4+1
	127	$\beta_{vax} \cdot V_5(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	V_5→V_5-1; VE_5→ VE_5+1
	128	$\beta_{vax} \cdot V_6(\sum_{i=1}^{4} I_i + \sum_{i=1}^{4} VI_i)$	V_6→V_6-1; VE_6→ VE_6+1
Vaccinated animals become exposed indirectly with gilt isolation (Experiments #4 & 15)	123b	0	V_1→V_1-1; VE_1→ VE_1+1
	124b	$\beta_{vax} \cdot V_2(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	V_2→V_2-1; VE_2→ VE_2+1
	125b	$\beta_{vax} \cdot V_3(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	V_3→V_3-1; VE_3→ VE_3+1
	126b	$\beta_{vax} \cdot V_4(\sum_{i=5}^{10} I_i + VI_5 + VI_6)$	V_4→V_4-1; VE_4→ VE_4+1
	127b	$\beta_{vax} \cdot V_5(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	V_5→V_5-1; VE_5→ VE_5+1
	128b	$\beta_{vax} \cdot V_6(\sum_{i=2}^{4} I_i + \sum_{i=2}^{4} VI_i)$	V_6→V_6-1; VE_6→ VE_6+1
Sows give birth	129	$b(S_5 + E_5)$	S_7→S_7+ $N(\mu = 12.29, \sigma^2 = 0.50)$

	130	$b(I_5 + VI_5 + R_5)$	$S_{11} \rightarrow S_{11} + N(\mu = 12.29, \sigma^2 = 0.50)$
	131	$b(V_5 + VE_5)$	$S_{15} \rightarrow S_{15} + N(\mu = 12.29, \sigma^2 = 0.50)$
Recovered animals lose immunity	132	$\omega_r \cdot R_1$	$R_1 \rightarrow R_1\text{-}1;\ S_1 \rightarrow S_1\text{+}1$
	133	$\omega_r \cdot R_2$	$R_2 \rightarrow R_2\text{-}1;\ S_2 \rightarrow S_2\text{+}1$
	134	$\omega_r \cdot R_3$	$R_3 \rightarrow R_3\text{-}1;\ S_3 \rightarrow S_3\text{+}1$
	135	$\omega_r \cdot R_4$	$R_4 \rightarrow R_4\text{-}1;\ S_4 \rightarrow S_4\text{+}1$
	136	$\omega_r \cdot R_5$	$R_5 \rightarrow R_5\text{-}1;\ S_5 \rightarrow S_5\text{+}1$
	137	$\omega_r \cdot R_6$	$R_6 \rightarrow R_6\text{-}1;\ S_6 \rightarrow S_6\text{+}1$
Exposed vaccinated (VE) animals become infected	138	$\sigma_{vax} \cdot VE_1$	$VE_1 \rightarrow VE_1\text{-}1;\ VI_1 \rightarrow VI_1\text{+}1$
	139	$\sigma_{vax} \cdot VE_2$	$VE_2 \rightarrow VE_2\text{-}1;\ VI_2 \rightarrow VI_2\text{+}1$
	140	$\sigma_{vax} \cdot VE_3$	$VE_3 \rightarrow VE_3\text{-}1;\ VI_3 \rightarrow VI_3\text{+}1$
	141	$\sigma_{vax} \cdot VE_4$	$VE_4 \rightarrow VE_4\text{-}1;\ VI_4 \rightarrow VI_4\text{+}1$
	142	$\sigma_{vax} \cdot VE_5$	$VE_5 \rightarrow VE_5\text{-}1;\ VI_5 \rightarrow VI_5\text{+}1$
	143	$\sigma_{vax} \cdot VE_6$	$VE_6 \rightarrow VE_6\text{-}1;\ VI_6 \rightarrow VI_6\text{+}1$
Infected vaccinated (VI) animals recover	144	$\gamma_{vax} \cdot VI_1$	$VI_1 \rightarrow VI_1\text{-}1;\ R_1 \rightarrow R_1\text{+}1$
	145	$\gamma_{vax} \cdot VI_2$	$VI_2 \rightarrow VI_2\text{-}1;\ R_2 \rightarrow R_2\text{+}1$
	146	$\gamma_{vax} \cdot VI_3$	$VI_3 \rightarrow VI_3\text{-}1;\ R_3 \rightarrow R_3\text{+}1$
	147	$\gamma_{vax} \cdot VI_4$	$VI_4 \rightarrow VI_4\text{-}1;\ R_4 \rightarrow R_4\text{+}1$
	148	$\gamma_{vax} \cdot VI_5$	$VI_5 \rightarrow VI_5\text{-}1;\ R_5 \rightarrow R_5\text{+}1$
	149	$\gamma_{vax} \cdot VI_6$	$VI_6 \rightarrow VI_6\text{-}1;\ R_6 \rightarrow R_6\text{+}1$
Non-disease related mortality for exposed vaccinated (VE) gilts and sows	150	$\mu \cdot VE_1$	$VE_1 \rightarrow VE_1\text{-}1$
	151	$\mu \cdot VE_2$	$VE_2 \rightarrow VE_2\text{-}1$
	152	$\mu \cdot VE_3$	$VE_3 \rightarrow VE_3\text{-}1$
	153	$\mu \cdot VE_4$	$VE_4 \rightarrow VE_4\text{-}1$
	154	$\mu \cdot VE_5$	$VE_5 \rightarrow VE_5\text{-}1$
	155	$\mu \cdot VE_6$	$VE_6 \rightarrow VE_6\text{-}1$

Non-disease related mortality for infected vaccinated (VI) gilts and sows	156	$\mu \cdot VI_1$	$VI_1 \rightarrow VI_1$-1
	157	$\mu \cdot VI_2$	$VI_2 \rightarrow VI_2$-1
	158	$\mu \cdot VI_3$	$VI_3 \rightarrow VI_3$-1
	159	$\mu \cdot VI_4$	$VI_4 \rightarrow VI_4$-1
	160	$\mu \cdot VI_5$	$VI_5 \rightarrow VI_5$-1
	161	$\mu \cdot VI_6$	$VI_6 \rightarrow VI_6$-1

Table F.2. The effects of farm size on peak prevalence, endemic prevalence, endemic prevalence in piglets, and probability of IAV extinction. Time to extinction describes the average number of days that infection persisted on the farm before stochastic extinction occurred. One hundred trials were conducted for each total inventory size. For each farm size, the proportions of animals in each production class were calibrated to maintain the population size at equilibrium.

Number of sows and gilts	Total inventory	Maximum prevalence			Endemic Prevalence (Total Inventory)			Endemic Prevalence (Piglets)			Extinction	
		Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Percent Extinction	Time to Extinction
200	500 (~350-650)	0.451	0.440 ± 0.080	18.109	0.142	0.138 ± 0.025	17.818	0.109	0.105 ± 0.019	18.020	3%	1.33 days
400	1000 (~800-1200)	0.451	0.452 ± 0.016	3.449	0.143	0.143 ± 0.002	1.613	0.111	0.110 ± 0.003	2.274	0%	n/a
1000	2500 (~2000-3000)	0.450	0.450 ± 0.008	1.877	0.144	0.143 ± 0.001	1.003	0.111	0.111 ± 0.001	1.329	0%	n/a
1500	3500 (~3000-4000)	0.4497	0.4490 ± 0.0078	1.7419	0.143	0.144 ±0.001	0.806	0.112	0.112± 0.001	1.321	0%	n/a
2000	5000 (~4500-5500)	0.446	0.442 ± 0.045	10.167	0.143	0.142 ± 0.014	10.123	0.111	0.110 ± 0.011	10.159	1%	1 day

Table F.3. The effects of source and frequency of IAV introduction on peak prevalence, endemic prevalence, endemic prevalence in piglets, and probability of IAV extinction. Time to extinction describes the average number of days that infection persisted on the farm before stochastic extinction occurred. One hundred trials were conducted for each case. These trials were conducted for a farm size of 1500 sows and gilts—consistent with the null model used for subsequent analyses.

Location and frequency of IAV introduction	Maximum			Endemic			Endemic Piglets			Extinction	
	Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Median	Mean (±SD)	CV	Percent Extinction	Time to Extinction
GDU- Single introduction	0.4497	0.4490 ± 0.0078	1.7419	0.143	0.144 ±0.001	0.806	0.112	0.112± 0.001	1.321	0%	n/a
GDU-Weekly gilt introduction	0.447	0.447 ± 0.009	1.893	0.144	0.144 ± 0.001	0.901	0.112	0.112± 0.002	1.398	0%	n/a
Breeding area (single introduction)	0.447	0.446 ± 0.008	1.704	0.144	0.144 ± 0.001	0.877	0.112	0.112 ± 0.002	1.496	0%	n/a
Farrowing area (single introduction)	0.446	0.447 ± 0.008	1.744	0.144	0.144 ± 0.001	0.780	0.112	0.111 ± 0.001	1.215	0%	n/a

Table F.4. Partial rank correlation coefficients (PRCC) between all 15 parameters (rows) and the endemic prevalence of infected piglets for each experiment (columns). Values closest to 1 or -1 indicate the strongest correlations, while values close to zero indicate little correlation. Bolded values have a magnitude ≥ 0.30. Shaded areas represent parameters that were not used/relevant for a particular experiment (e.g., the direct transmission rate for vaccinated animals was not employed in the null case [Experiment 0]).

Parameter	#0	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
β	0.13	0.14	0.14	0.13	0.13	0.12	0.13	0.18	0.17	0.15	0.16	0.14	0.15	0.12	0.13	0.14
β_{ind}	0.04	0.03	0.03	0.03	0.01	0.01	0.03	0.06	0.08	0.05	0.06	0.18	0.05	0.04	0.04	0.08
β_d^p	-0.10	-0.10	-0.10	-0.11	-0.10	-0.04	-0.10	-0.15	-0.12	-0.10	-0.21	-0.15	-0.10	-0.09	-0.11	0.02
β_{ind}^p	-0.08	-0.08	-0.08	-0.08	-0.07	-0.07	-0.08	-0.13	-0.14	-0.10	0.03	-0.07	-0.11	-0.06	-0.08	0.00
β_d^{pm}	-0.14	-0.12	-0.13	-0.14	-0.14	-0.14	-0.13	**0.52**	-0.04	-0.13	-0.06	-0.06	-0.11	-0.15	-0.14	**0.54**
β_{ind}^{pm}	-0.05	-0.06	-0.05	-0.06	-0.05	-0.08	-0.06	0.00	0.00	-0.04	-0.10	-0.05	-0.01	-0.08	-0.05	-0.05
σ	0.13	0.13	0.13	0.14	0.14	0.13	0.13	**0.75**	**0.45**	0.21	0.16	0.15	0.15	0.13	0.13	**0.60**
γ	**-0.80**	**-0.80**	**-0.80**	**-0.80**	**-0.79**	**-0.79**	**-0.81**	**-0.50**	**-0.77**	**-0.80**	**-0.71**	**-0.76**	**-0.81**	**-0.80**	**-0.80**	**-0.42**
β_{vax}						-0.13	-0.15				0.02	-0.08	-0.17	-0.13	-0.16	0.10
β_{vax}^p						-0.18	-0.21				-0.21	-0.16	-0.24	-0.19	-0.22	-0.12
$\beta_{ind\ vax}$						0.06	0.08				0.08	0.11	0.10	0.07	0.08	0.10
$\beta_{ind\ vax}^p$						0.11	-0.14				0.12	0.18	-0.10	0.11	-0.16	0.05
ω_r	-0.05	-0.05	-0.06	-0.06	-0.07	-0.06	-0.05	0.18	0.01	-0.02	0.11	0.03	-0.04	-0.03	-0.04	0.11
σ_{vax}						-0.15	-0.09				-0.19	-0.17	-0.10	-0.14	-0.08	-0.06
γ_{vax}						0.14	0.07				0.18	0.18	0.08	0.14	0.06	0.06

Table F.5. Partial rank correlation coefficients (PRCC) between all 15 parameters (rows) and the maximum prevalence of infected animals for each experiment (columns). Values closest to 1 or -1 indicate the strongest correlations, while values close to zero indicate little correlation. Bolded values have a magnitude ≥ 0.30. Shaded areas represent parameters that were not used/relevant for a particular experiment (e.g., the direct transmission rate for vaccinated animals was not employed in the null case [Experiment #0]).

Parameter	#0	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
β	0.18	0.19	0.19	0.17	0.18	0.18	0.16	0.18	0.18	0.17	0.16	0.19	0.17	0.16	0.17	0.12
β_{ind}	0.06	0.07	0.07	0.07	0.05	0.05	0.05	0.10	0.06	0.06	0.11	0.19	0.09	0.08	0.06	0.10
β_d^p	-0.09	-0.09	-0.07	-0.08	-0.09	-0.01	-0.10	-0.15	-0.09	-0.09	-0.16	-0.12	-0.10	-0.07	-0.08	0.05
β_{ind}^p	-0.17	-0.19	-0.19	-0.19	-0.16	-0.21	-0.20	-0.17	-0.18	-0.19	-0.07	-0.16	-0.20	-0.18	-0.19	0.02
β_d^{pm}	-0.03	-0.04	-0.03	-0.03	-0.03	-0.04	-0.04	-0.09	-0.02	-0.04	0.06	0.02	-0.03	-0.05	-0.02	0.18
β_{ind}^{pm}	0.01	0.02	0.00	0.01	0.04	0.00	0.03	-0.01	0.01	0.02	-0.01	0.04	0.07	0.01	-0.01	-0.11
σ	**0.80**	**0.80**	**0.80**	**0.80**	**0.78**	**0.79**	**0.81**	**0.80**	**0.80**	**0.81**	**0.72**	**0.71**	**0.79**	**0.80**	**0.80**	**0.39**
γ	**-0.64**	**-0.65**	**-0.65**	**-0.65**	**-0.66**	**-0.63**	**-0.69**	**-0.66**	**-0.65**	**-0.65**	**-0.51**	**-0.63**	**-0.69**	**-0.67**	**-0.66**	**-0.53**
β_{vax}						-0.12	-0.08				0.03	-0.10	-0.11	-0.10	-0.09	0.12
β_{vax}^p						-0.22	-0.22				-0.23	-0.14	-0.26	-0.18	-0.23	-0.08
$\beta_{ind\ vax}$						0.09	0.12				0.10	0.13	0.11	0.11	0.10	0.13
$\beta_{ind\ vax}^p$						-0.07	-0.21				-0.08	0.00	-0.18	-0.09	-0.20	0.12
ω_r	0.01	0.02	0.02	0.00	0.00	-0.02	0.02	0.00	0.03	0.02	0.17	0.08	0.03	0.02	0.01	0.13
σ_{vax}						0.06	-0.13				0.02	0.03	-0.14	0.07	-0.12	-0.15
γ_{vax}						-0.06	-0.08				-0.03	-0.02	-0.04	-0.07	-0.06	0.15

Appendix G. Supplementary tables for Chapter 4

Additional figures

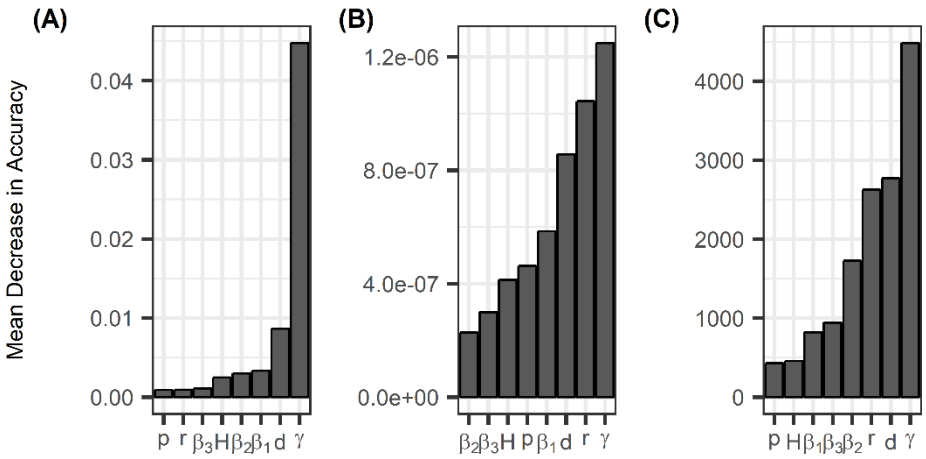


Figure G.1. Random forest regression analysis results describing variable importance for (*A*) outbreak success (did the pathogen spread beyond the initially infected individual?), (*B*) maximum prevalence given outbreak success, and (*C*) outbreak duration given outbreak success. Parameter descriptions are provided in Table 4.1. Bar charts display mean decrease in accuracy for parameters. Higher values for mean decrease in accuracy correspond to parameters with higher predictive ability. The *cforest* function from the party package in R (with 1,000 trees) was used for this analysis, and generally upheld findings from the *randomForest* package (Figure 4.1). The order of variable importance for outbreak success is the same (compare to Figure 4.1A). For both maximum prevalence and outbreak duration, both analyses identified recovery rate, perceptual distance, and conspecific density as the top three parameters.

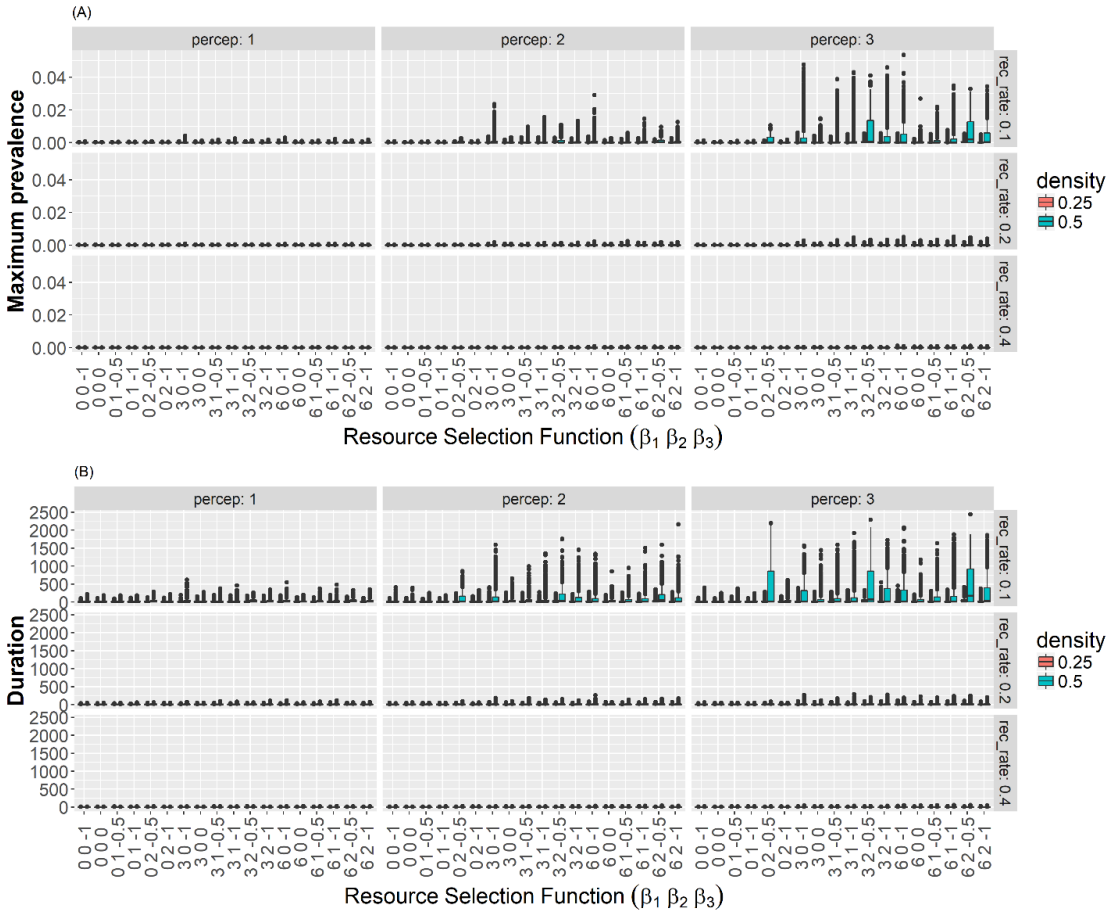


Figure G.2. Box plots of (A) maximum prevalence and (B) duration across perceptual ranges (columns), recovery rates (rows), and resource selection functions (x-axis) for low (d=0.25) and high (d=0.50) conspecific density treatments.

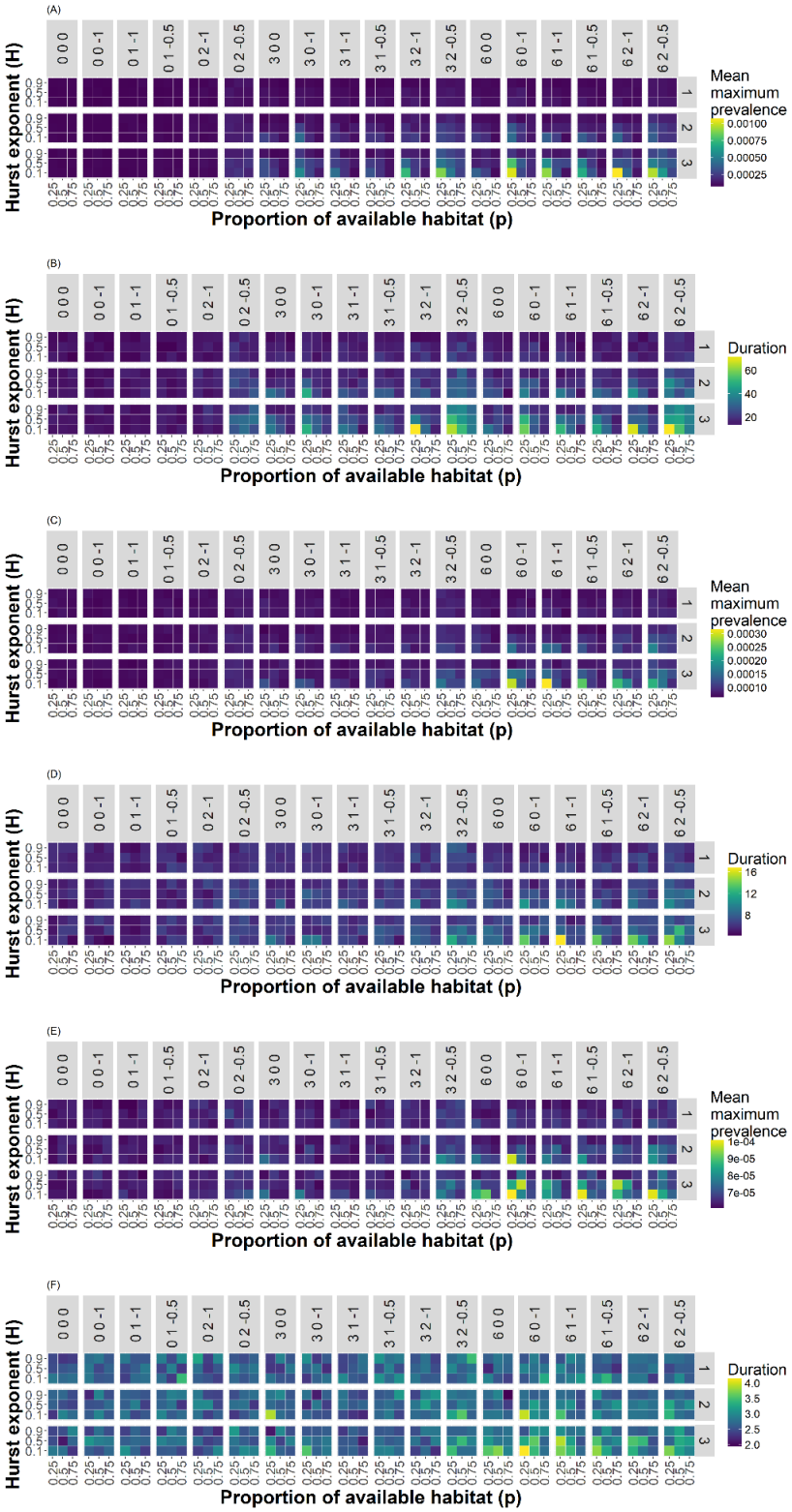

(A)
(B)
(C)
(D)
(E)
(F)
Hurst exponent (H)
Proportion of available habitat (p)
Mean maximum prevalence
Duration

Figure G.3. Heatmaps for a subset of simulations where conspecific density, d= 0.25: (A) mean maximum prevalence and (B) mean duration when recovery rate, γ=0.1; (C) mean maximum prevalence and (D) mean duration when recovery rate, γ=0.2; and (E) mean maximum prevalence and (F) mean duration when recovery rate, γ=0.4. The x-axis corresponds to the proportion of available habitat (p) tested in each landscape and the y-axis corresponds to the Hurst exponent (H) or patchiness of the simulated landscapes. Columns correspond to RSF combinations and rows correspond to perceptual range. Note that color scales are different between maps.

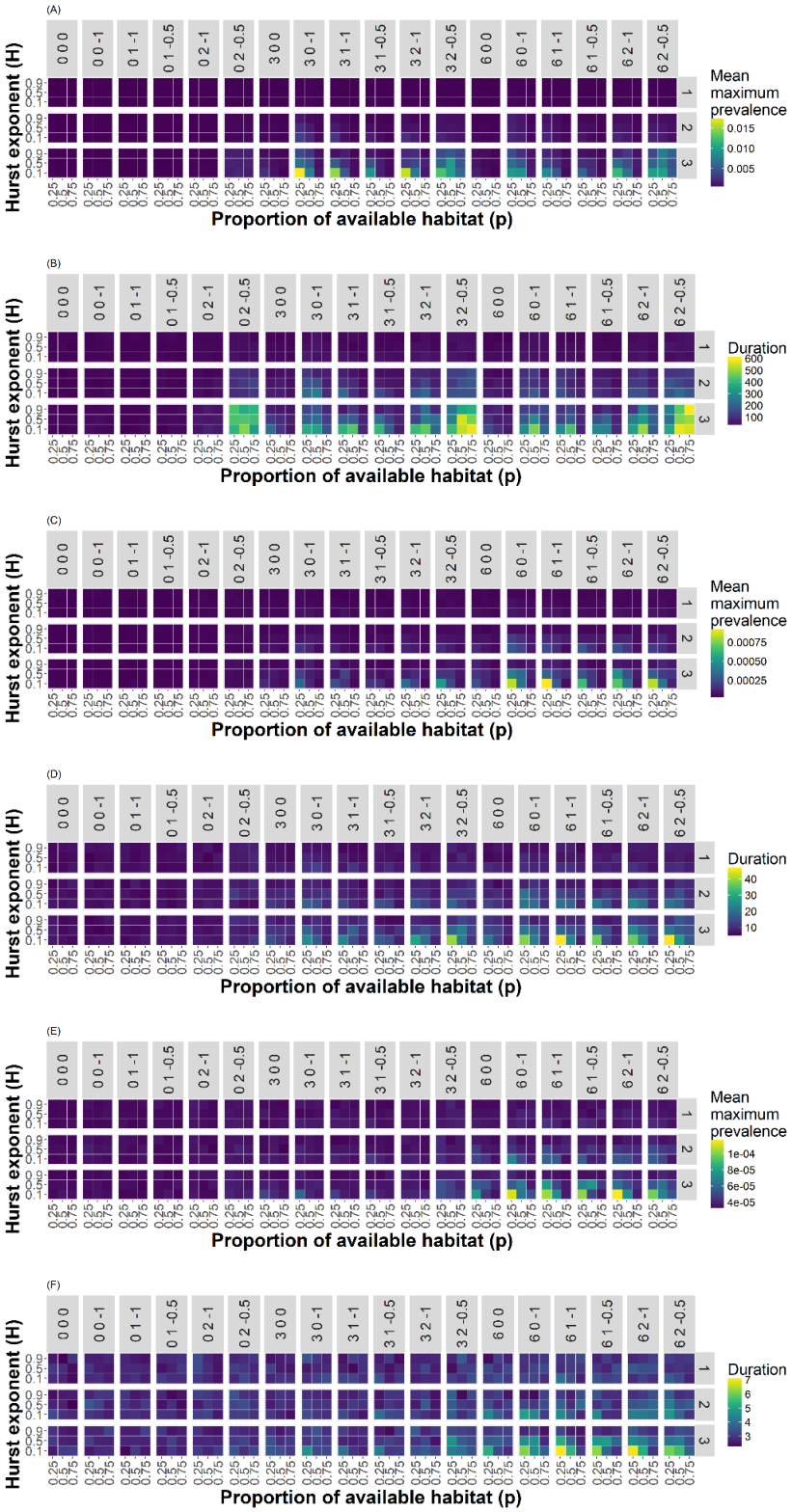

(A)
Hurst exponent (H)
Proportion of available habitat (p)
Mean maximum prevalence
0.015
0.010
0.005
(B)
Duration
600
500
400
300
200
100
(C)
Mean maximum prevalence
0.00075
0.00050
0.00025
(D)
Duration
40
30
20
10
(E)
Mean maximum prevalence
1e-04
8e-05
6e-05
4e-05
(F)
Duration
7
6
5
4
3

Figure G.4. Heatmaps for a subset of simulations where conspecific density, d= 0.5: (A) mean maximum prevalence and (B) mean duration when recovery rate, γ = 0.1; (C) mean maximum prevalence and (D) mean duration when recovery rate, γ = 0.2; and (E) mean maximum prevalence and (F) mean duration when recovery rate, γ = 0.4. The x-axis corresponds to the proportion of available habitat (p) tested in each landscape and the y-axis corresponds to the Hurst exponent (H) or patchiness of the simulated landscapes. Columns correspond to RSF combinations and rows correspond to perceptual range. Note that color scales are different between maps.

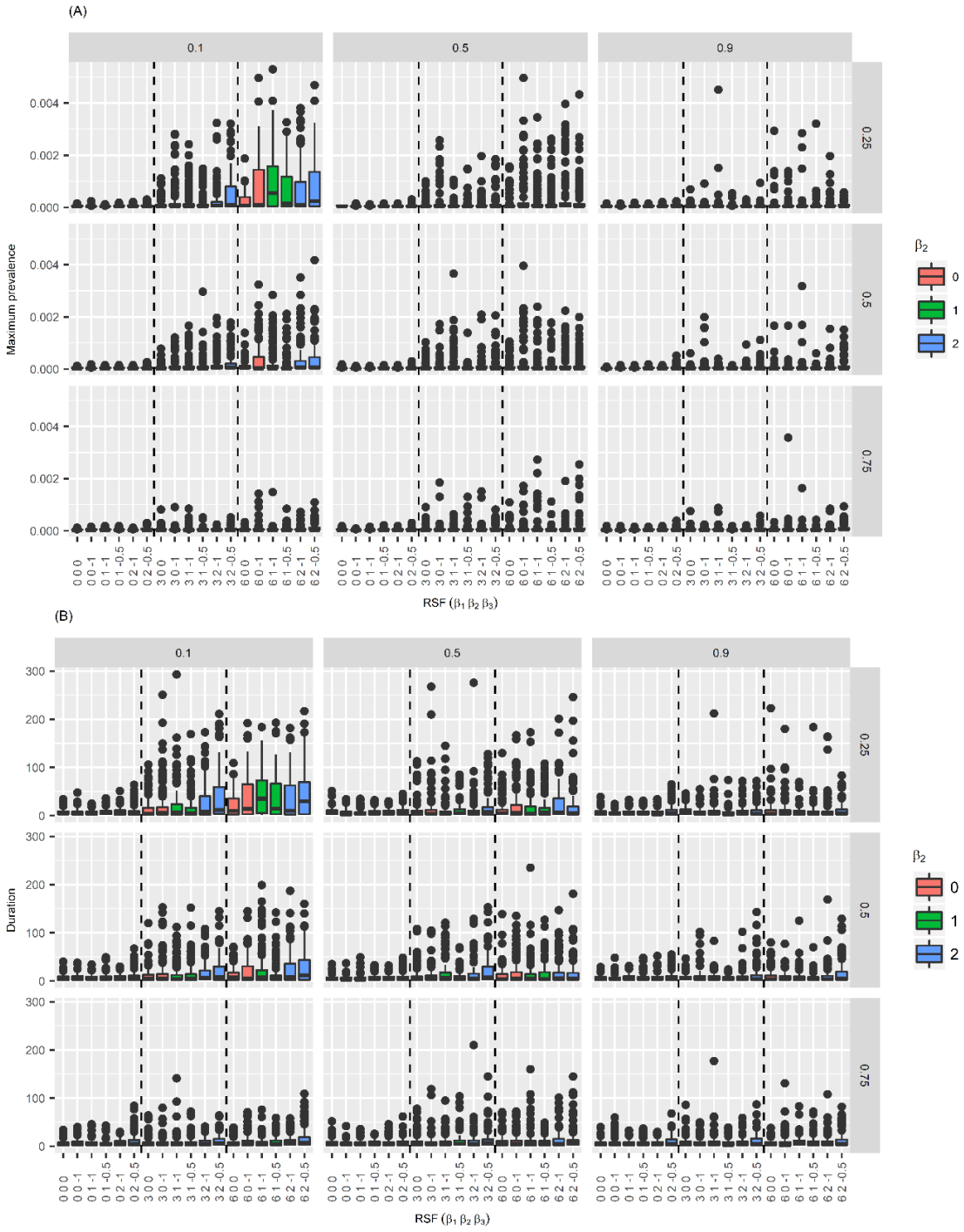


Figure G.5. Box plots of (*A*) maximum prevalence and (*B*) duration for a subset of the simulations where recovery rate (γ) =0.2, conspecific density (d) = 0.5, and perceptual range (r) = 3. Columns correspond to Hurst exponent (H; lower values correspond to higher clustering) and rows correspond to proportion

available habitat (p). Dashed lines represent separate regimes (from left to right) of random, medium or strong selection for resources (β_1).

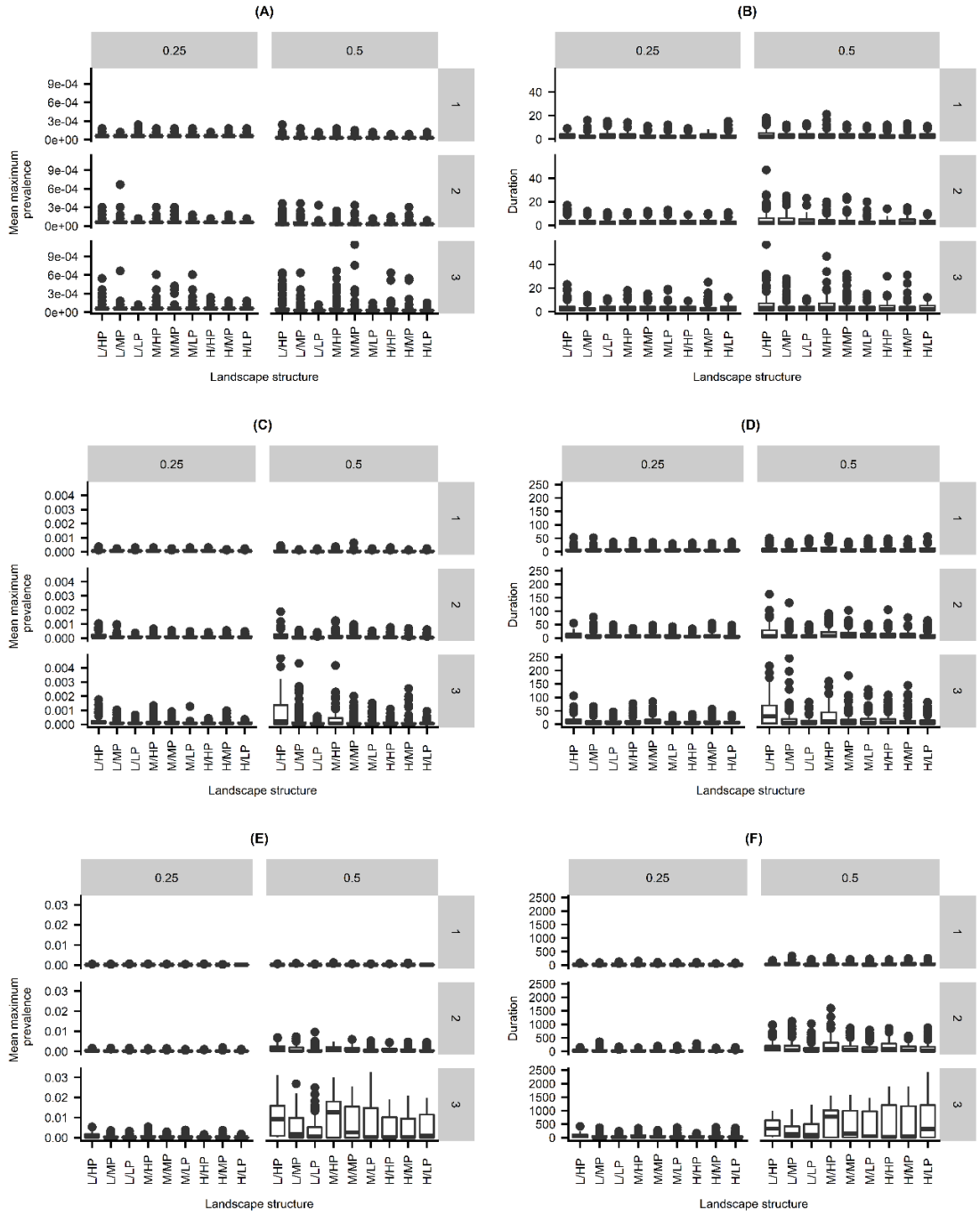


Figure G.6. For the RSF of $\boldsymbol{\beta_1 = 6, \beta_2 = 2,\ \beta_3 = -0.5}$, box plots of (*A, C, E*) mean maximum prevalence and (*B, D, F*) duration for the recovery rate (γ) of 0.4 (*A, B*), 0.2 (*C, D*), and 0.1 (*E*, *F*). Columns correspond to low and high densities and rows correspond to the perceptual range. Landscape structure

abbreviations take the form of proportion available habitat(p)/patchiness (Hurst exponent, H): “L” = low, “M” =medium, “H” = high proportion available habitat. “LP” =low patchiness, “MP” =medium patchiness, and “HP” =high patchiness.

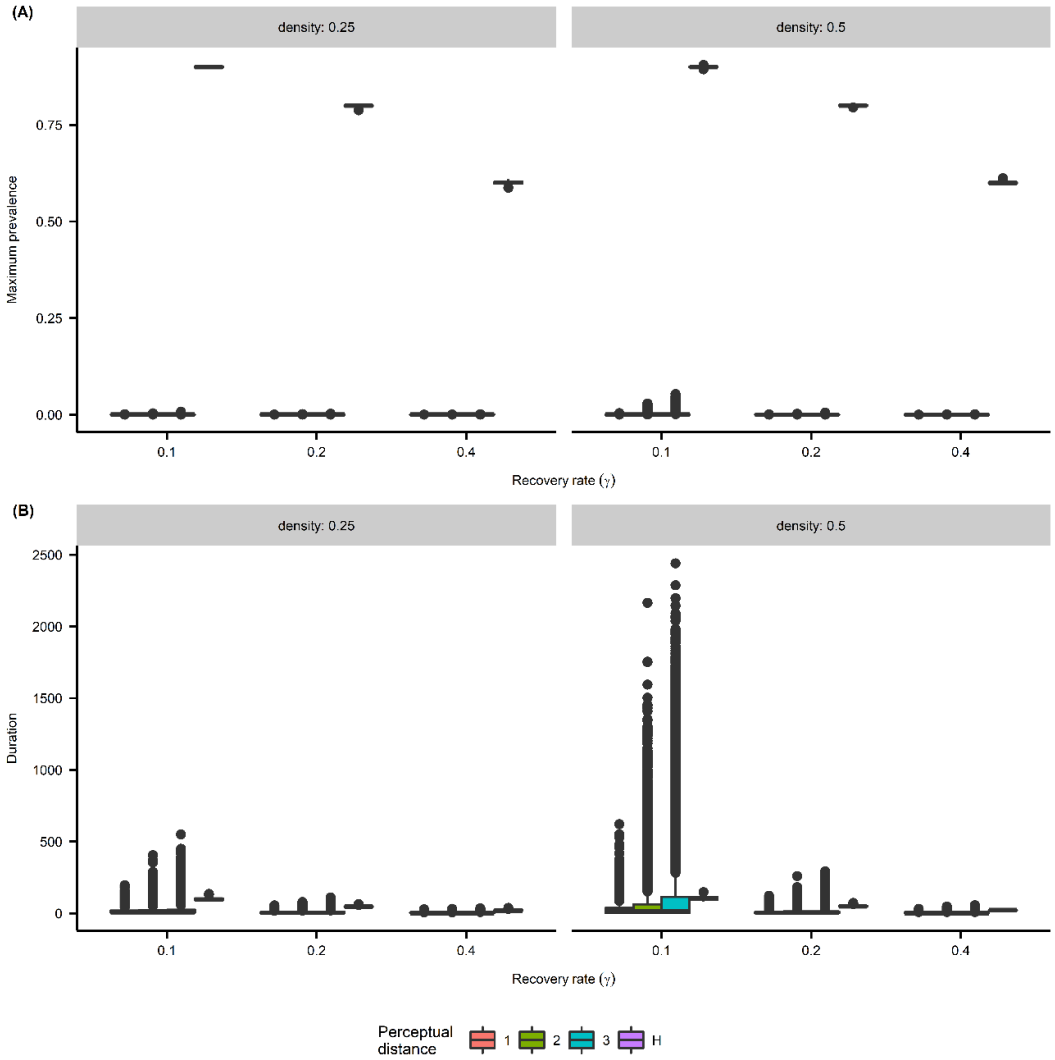


Figure G.7. Simulation results for stochastic SIR model that assumed density-dependent transmission function and homogenous mixing of conspecifics. (*A*) Maximum prevalence reached for each recovery rate tested (x-axis) for both conspecific densities (columns). (*B*) Epidemic duration for each recovery rate tested (x-axis) for both conspecific densities (columns). For the legend, "H" refers to the homogeneous mixing model.

Expanded Methods

Selection of transmission *probability* (β)

To help limit the number of parameter combinations tested, we represented the epidemiological parameter space as three combinations of recovery rate and a single transmission probability—equivalent to R_0 values of roughly 0.5, 1, and 1.5 (Table 4.1). These combinations of transmission probability and recovery rate provided a parameter space with at least some successful outbreaks, outbreaks that did not always saturate (reach all individuals in the population) regardless of the landscape structure, and outbreaks that were computationally feasible (i.e., finishing in less than 1000 timesteps). For much lower transmission probabilities (e.g., β = 0.02 or 0.002) with the same recovery rates, we did not observe enough successful transmission events (especially in our lower conspecific density treatment, d=0.25) to be able to differentiate the effects of landscape or movement. This trend is observable for our lowest conspecific density treatment and fastest recovery rate in Appendix G Figure G.3 *E* and *F*. While our results are likely robust for transmission probabilities of similar orders of magnitude, we would expect diminishing effects of movement behavior and landscape structure for less successful ($\beta < 0.02$) and highly successful pathogens ($\beta > 0.5$) with the same combinations of recovery rate.

Calibrating resource selection functions (RSFs)

We governed individual movement choices using the resource selection function, $\boldsymbol{w_j}$, for a given cell, $\boldsymbol{j}$, with the form: $w_j = \exp(\beta_1 \cdot R_j + \beta_2 \cdot N_j + \beta_3 \cdot N_j^2)$

where R_j is the resource value, and N_j is the number of conspecifics in cell, *j*.

Then, the probability that an individual would choose cell, *j*, over other cells in the neighborhood (including the current cell) is given by:

$$P_j = \frac{\exp(\beta_1 \cdot R_j + \beta_2 \cdot N_j + \beta_3 \cdot N_j^2)}{\sum_{i=1}^{n} \exp(\beta_1 \cdot R_i + \beta_2 \cdot N_i + \beta_3 \cdot N_i^2)}$$

Where the numerator represents the RSF of cell, *j*, and the denominator represents the sum of the RSFs of all *n* cells in the individual's neighborhood.

Calibrating strength of selection for resources

When holding the number of conspecifics (*N*) constant, what is the probability that an individual selects a cell that has a resource value of $R + \Delta R$ relative to a landscape with baseline resource values of *R*?

R	$R + \Delta R$	*R*
R	*R*	*R*
R	*R*	*R*

Probability of selecting a cell with $R+\Delta R$ resource value in neighborhood of nine cells (eight neighboring cells and the current cell):

$$P(R + \Delta R) = \frac{\exp(\beta_1 \cdot (R + \Delta R) + \beta_2 \cdot N + \beta_3 \cdot N^2)}{8\exp(\beta_1 \cdot R + \beta_2 \cdot N + \beta_3 \cdot N^2) + \exp(\beta_1 \cdot (R + \Delta R) + \beta_2 \cdot N + \beta_3 \cdot N^2)}$$

If *N* is constant across each cell in the neighborhood then this expression simplifies to:

$$P(R + \Delta R) = \frac{\exp(\beta_1 \cdot (R + \Delta R))}{8\exp(\beta_1 \cdot R) + \exp(\beta_1 \cdot (R + \Delta R))}$$

Table G.1. Calibration of β_1 values. β_1 values are calibrated so that holding conspecific density constant, the probability of selecting a cell with a resource value of 1 in a neighborhood of $R = 0$ should be: zero ($\beta_1 = 0$, i.e., random), biased ($\beta_1 = 3$, $P(R + \Delta R) = 0.72$) or almost completely deterministic ($\beta_1 = 6$; $P(R + \Delta R) \approx 1$).

β_1	Description	Probability of selecting cell with resource value of $R + \Delta R$ when $R = 0$ and $\Delta R = 1$	Probability of selecting cell with resource value of R when $R = 0$ and $\Delta R = 1$
0	Random	$1/9 \approx 0.11$	$1/9 \approx 0.11$
3	Biased	0.715	0.036
6	Approaching deterministic	0.981	0.002

Calibrating strength of selection for conspecifics

For a landscape with uniform resource values (R), number of conspecifics per cell (N), and resource selection function parameters ($\beta_1, \beta_2, \beta_3$), what is the probability of selecting a cell within the neighborhood that has $N + \Delta N$ conspecifics?

N	*N + ΔN*	*N*
N	*N*	*N*
N	*N*	*N*

The probability of selecting one of the eight cells with N conspecifics in the neighborhood of nine cells (including the current cell) is equal to:

$$P(N) = \frac{\exp(\beta_1 \cdot R + \beta_2 \cdot N + \beta_3 \cdot N^2)}{8\exp(\beta_1 \cdot R + \beta_2 \cdot N + \beta_3 \cdot N^2) + \exp(\beta_1 \cdot R + \beta_2 \cdot (N + \Delta N) + \beta_3 \cdot (N + \Delta N)^2)}$$

And the probability of selecting the cell with $N + \Delta N$ conspecifics:

$$P(N + \Delta N)$$

$$= \frac{\exp(\beta_1 \cdot R + \beta_2 \cdot (N + \Delta N) + \beta_3 \cdot (N + \Delta N)^2)}{8\exp(\beta_1 \cdot R + \beta_2 \cdot N + \beta_3 \cdot N^2) + \exp(\beta_1 \cdot R + \beta_2 \cdot (N + \Delta N) + \beta_3 \cdot (N + \Delta N)^2)}$$

β_2 and β_3 values were calibrated to exhibit three biologically feasible responses to changes in conspecific numbers across a landscape: (*i*) conspecific avoidance (Appendix *G* Figure G.8*B*); (*ii*) random (Appendix G Figure G.8*A*); (*iii*) attraction to conspecifics (Appendix G Figure G.8 *C-F*). This last scenario could reflect the idea that after a certain point depletion of resources would outweigh the benefit of conspecifics signaling a high-quality resource patch.

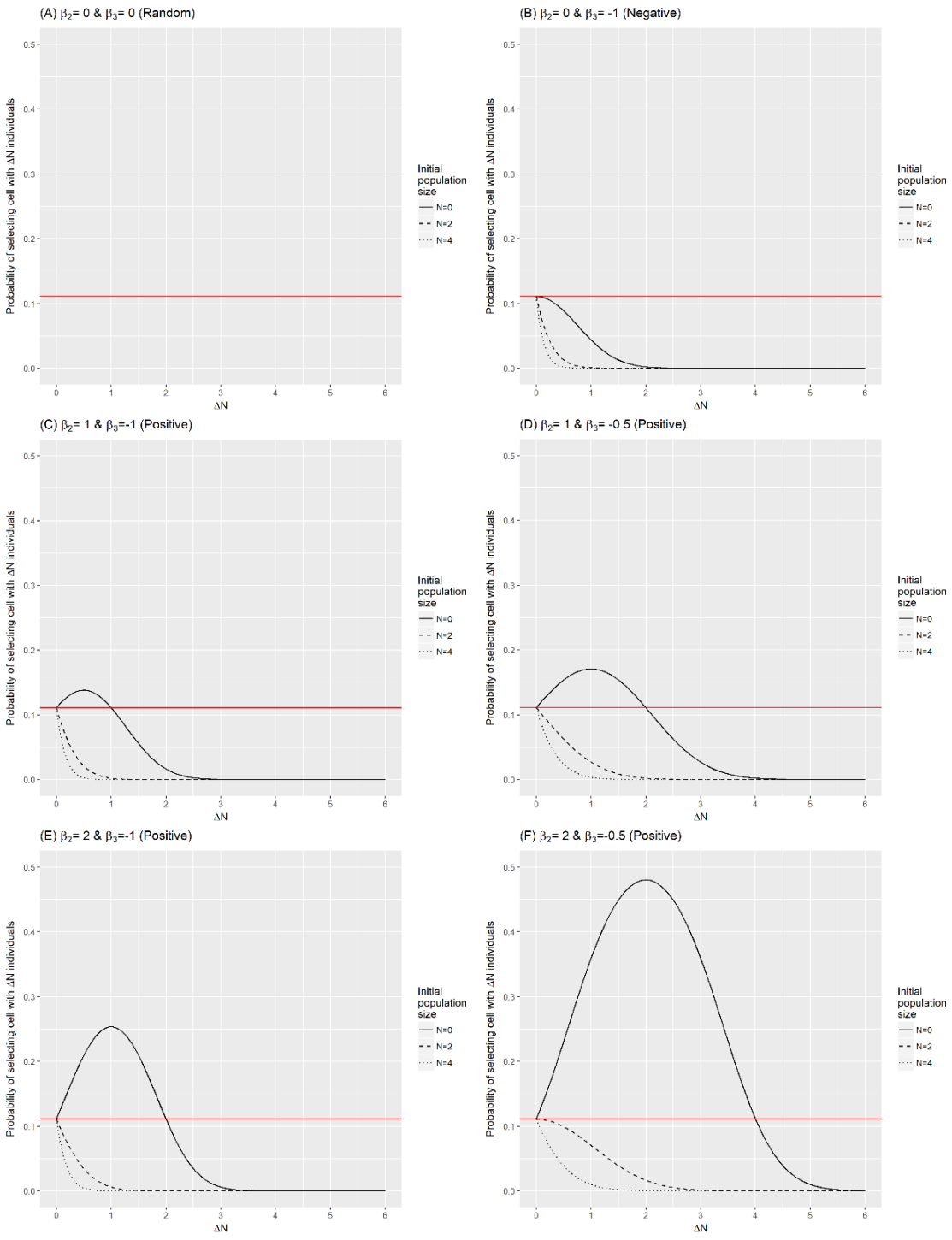


Figure G.8. Depiction of effects of conspecific presence on probability of an animal selecting a given cell within its neighborhood. Calculated for a neighborhood of nine cells. Δ*N* refers to the difference in the

number of conspecifics between other cells in the neighborhood and the cell of interest. Red line corresponds to expected probability based on random choice $(1/9 \approx 0.111)$.

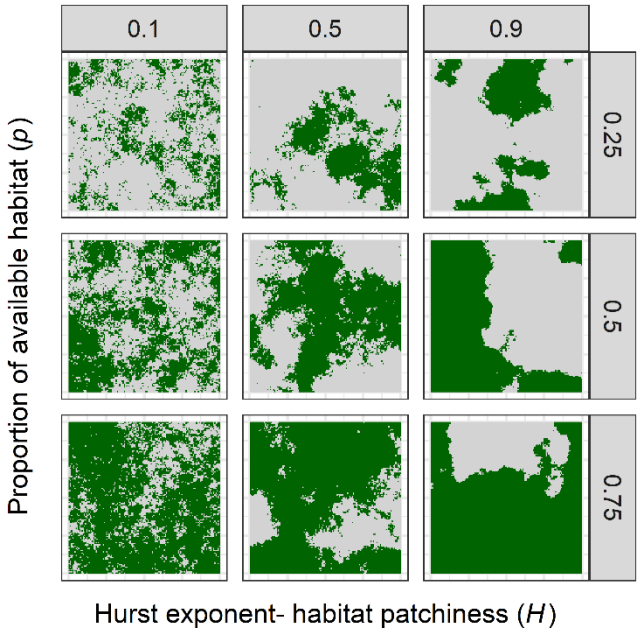


Figure G.9. Binary, neutral theoretical landscapes generated with the mid-point displacement algorithm. Landscapes varied in the proportion of available habitat (*p*; y-axis) and clustering of habitat (*H*; x-axis) (Table 4.1). Good quality habitat is shown in green (resource quality = 1), and poor-quality habitat is shown in grey (resource quality= 0). As the Hurst exponent (*H*) decreases, habitat patchiness increases. The size of the landscape was held constant at 257 x 257 units (e.g., grid cells). Spatial units are arbitrary.

Evaluating equivalent homogeneous mixing model

Comparable deterministic ordinary differential equations used to evaluate equivalent, homogeneous mixing, SIR model:

$$\frac{dS}{dt} = -\beta \cdot S \cdot I$$

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

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