

Research

Covariation between the physiological and behavioral components of pathogen transmission: host heterogeneity determines epidemic outcomes

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

L. A. White (<http://orcid.org/0000-0003-3852-5374>) (whit1951@umn.edu), Dept of Ecology, Evolution and Behavior, Univ. of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, St. Paul, MN 55108, USA. – *J. D. Forester* (<http://orcid.org/0000-0002-5392-9556>), Dept of Fisheries, Wildlife and Conservation Biology, Univ. of Minnesota, St. Paul, MN, USA. – *M. E. Craft* (<http://orcid.org/0000-0001-5333-8513>), Veterinary Population Medicine, Univ. of Minnesota, St. Paul, MN, USA.

Oikos

127: 538–552, 2018

doi: 10.1111/oik.04527

Subject Editor: David Chalcraft

Editor-in-Chief: Dries Bonte

Accepted 10 October 2017

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.

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 and Caillaud 2011), 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 et al. 1997, Beldomenico and Begon 2010, Gibson et al. 2016), and variability in infectiousness has also been observed, particularly when concomitant infections are present (Cattadori et al. 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 versus 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 and 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 Supplementary material Appendix 1 for a schematic representing transmission and definitions of key terms). Thus, β can further be broken down into the behavioral (β_b) 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 (Supplementary material Appendix 1), 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 (Gear 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, Berdoy et al. 2000) or indirectly through sickness behaviors that reduce host activity levels such as fever, lethargy, and limited foraging (Adelman et al. 2014, Welicky and Sikkell 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 and Codeço 2009, Funk et al. 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. 2005). In disease ecology, however, the focus has largely been on variability in contact rate (β_b) (VanderWaal and Ezenwa 2016, White et al. 2017a), 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 super-spreading has not been fully elucidated (Hawley and Altizer 2011, VanderWaal and 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 and 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 et al. 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. 2005, Meyers 2007). Network models are a tool that can capture individual variability in the number and rate of contacts (β_b). 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. 2005).

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 and Holme 2013, White et al. 2017a). 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 and Meyers 2007, Bansal et al. 2010). However, the implications of using static versus 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 and Holme 2013). In a theoretical framework, utilizing a static network for a dynamic system was found to overestimate epidemic predictions (Fefferman and Ng 2007, Masuda and 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. 2017a).

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. 2005), our results suggest that both heterogeneity in physiology and subsequent covariation of physiology with contact rate could powerfully influence epidemic dynamics.

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 and 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 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

Table 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

contacts, and disease-induced behavioral changes (Krivitsky and 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 and 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 et al. 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 1).

Models were constructed in R (ver. 3.3.2, <www.r-project.org>) using self-written modules in the ‘EpiModel’ package (ver. 3.4.0, <www.epimodel.org/>) (Jenness et al. 2016a). The EpiModel package 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 ‘statnet’ 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 the Supplementary material Appendix 2, and all code and simulation data for the manuscript are

available from the Dryad Digital Repository (White et al. 2017b).

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 versus contact rate; 2) infectiousness versus contact rate; and 3) infection status versus contact rate (Supplementary material Appendix 1 Fig. A2). 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) (Supplementary material Appendix 1 Fig. A2a, blue line), and a negative covariation scenario where physiology (β_p) negatively covaries with contact rate (β_c) (Supplementary material Appendix 1 Fig. A2a, 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 (Supplementary material Appendix 1 Fig. A2b; Table 2). These behavioral phenotypes can be thought of as corresponding roughly to spectrums of individual personality (e.g. shy versus 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 et al. 2014).

Table 2. Experimental design for sections ‘Susceptibility versus contact rate’ and ‘Infectiousness versus contact rate’.

Type of covariation	No. 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	<i>unif</i> {0,2}
	175	4	4	<i>unif</i> {0,2}
	175	4	4	<i>unif</i> {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

For susceptibility versus contact rate and infectiousness versus 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. Supplementary material Appendix 1 Fig. A2b), and those with a ‘high’ separation would have three separate groups with mean degrees of 0, 4 and 8 (Table 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).

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).

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 et al. 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 (Supplementary material Appendix 1). We vary transmission efficiency, τ , in our factorial design (Table 1) and discuss the modifications for the final transmission probability for each specific experimental scenario below.

Susceptibility versus 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:

$$\mathbb{P}(T) = 1 - (1 - \min\{1, \tau \times s\})^\alpha$$

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 $\mathbb{P}(T) = \min\{1, \tau \times s\}$. 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 versus 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:

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

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 ($\kappa = 0$), the pathogen could not propagate further.

Infection status versus 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 and Sikkell 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 3). It is worth noting that in the above scenarios (‘Suscep-

Table 3. Experimental design for section ‘Infection status versus contact rate’.

Type of covariation	Mean degree pre-exposure (β_e)	Mean degree post-infection (β_e)	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

tibility versus contact rate’ and ‘Infectiousness versus contact rate’) it is possible for a secondary correlation to result between contact rate and infection status. For example, in the positive covariation scenario for susceptibility versus 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_e) 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_e) upon infection, but individuals had heterogeneity in infectiousness (individuals were randomly assigned an infectiousness of $\kappa = 1$ 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 ($\kappa = 1$) (Table 3). Unlike the first two scenarios tested (above), 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 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 et al. 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 ‘Infectiousness versus contact rate’ where infectiousness influences the action rate in the exponent of the final transmission probability $\mathbb{P}(T) = 1 - (1 - \tau)^{\alpha \kappa}$; and 2) a form where infectiousness (κ) directly modifies the probability of infection (‘Susceptibility versus contact rate’) so that the final transmission probability was equal to: $\mathbb{P}(T) = 1 - (1 - \min\{1, \tau \times \kappa\})^\alpha$, which simplifies to $\mathbb{P}(T) = \min\{1, \tau \times \kappa\}$ 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 (ver. 1.2-1) in R (Elzhov et al. 2016). We fit each individual epidemic simulation to the logistic growth equation: $I(t) = \frac{K}{1 + b \times e^{-\beta K t}}$ 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 the Supplementary material Appendix 3). We assigned values of I_0 appropriate to each simulation ($I_0 = 1$ for susceptibility or infectiousness versus contact rate and $I_0 = 2$ for infection status versus 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, 2009). Using the ‘cforest’ function in the ‘party’ package in R (ver. 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 permuted. Stated another way, higher mean decrease in accuracy scores indicate a greater importance in model prediction. For susceptibility versus contact rate and infectiousness versus 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 versus 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: <<http://dx.doi.org/10.5061/dryad.8t201>> (White et al. 2017b).

Results

Susceptibility versus 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 (Fig. 1, 2).

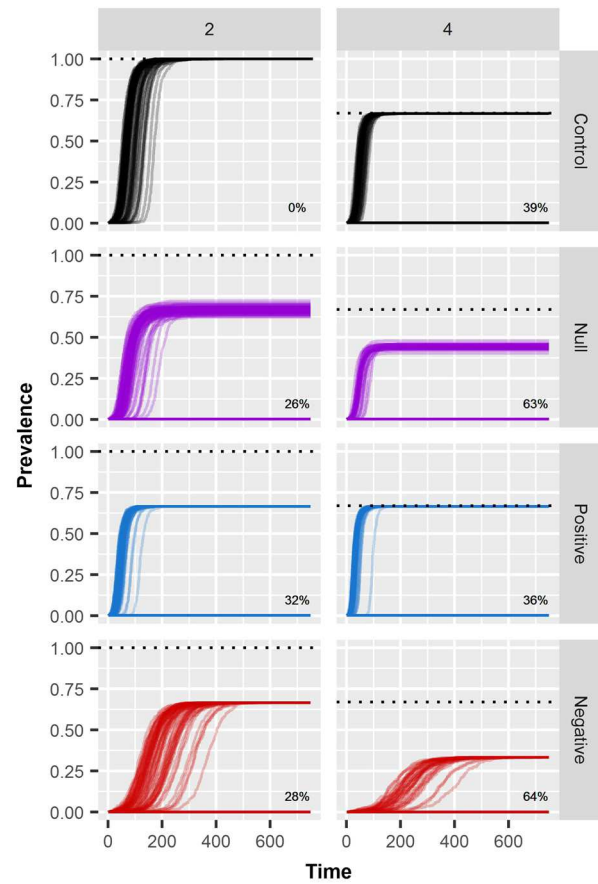


Figure 1. Time course of simulated epidemics for susceptibility versus contact rate for the lowest transmission efficiency tested of $\tau = 0.025$. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in susceptibility, 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 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 $\beta_c = 0$, limiting maximum prevalence to 0.66. Time courses for the corresponding medium ($\tau = 0.25$) and high ($\tau = 0.5$) transmission efficiencies are available in the Supplementary material Appendix 4 Fig. A3–A4.

This finding was consistent across differences in mean degree and for different transmission efficiencies (Supplementary material Appendix 4 Fig. A3–A4). 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 ($\beta_c = 0$) for networks constructed with higher variation in contact rate (i.e. mean degree difference = 4). Notably, positive covariation counteracted this observed difference

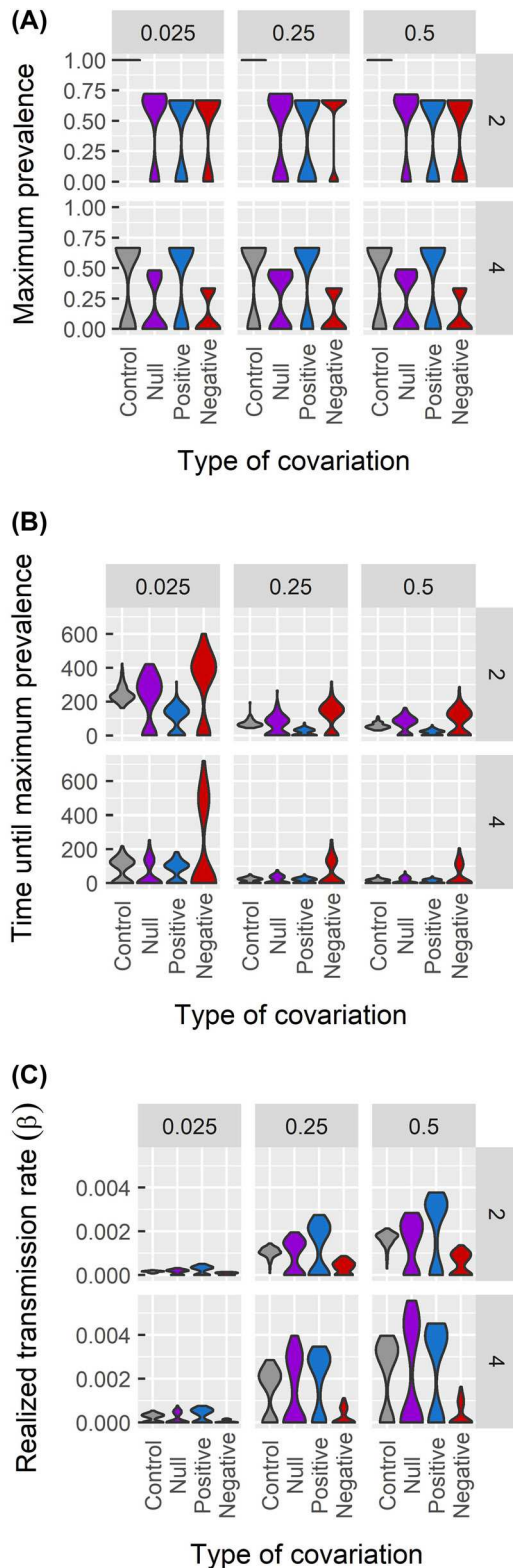


Figure 2. For susceptibility versus contact rate, violin plots depicting for all covariation types: (A) the maximum prevalence reached in 750 time steps; (B) the time it takes to reach that maximum prevalence; and (C) the realized transmission rate (β), which describes the rate of epidemic spread. The 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).

in maximum prevalence between the control case and other covariation types, and this effect was consistent across infection probabilities (Fig. 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 (Fig. 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 (Fig. 2C). Negative covariation continued to substantially suppress the realized β even at higher values of transmission efficiency.

Infectiousness versus contact rate

Variability in infectiousness (null case) increased variability in epidemic outcomes (Fig. 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 transmission efficiencies (Supplementary material Appendix 4 Fig. A5–A6). As with susceptibility versus contact rate, a larger simulated variation in contact rate within the population also decreased the maximum prevalence, even in the control case (Fig. 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 (Fig. 4B). Similar to the results for susceptibility versus 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 (Fig. 4C).

Infection status versus contact rate

For infection status versus contact rate, we tested two ways that infectiousness might play into the final transmission probability ('Infection status versus contact rate: disease-induced behavioral changes'), but results were consistent across these two different formulations. As with susceptibility versus contact rate and infectiousness versus contact rate, simply including heterogeneity in physiology increased variability in epidemic outcome (Fig. 5; compare control versus null cases). Reduction in contact rate upon infection (negative covariation) drastically reduced the maximum prevalence reached within 750 time steps (Fig. 5, 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

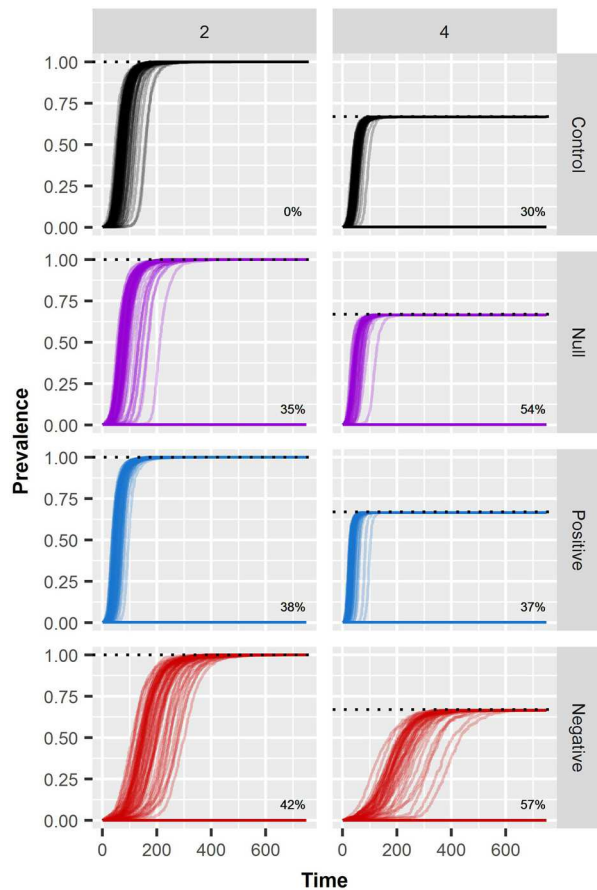


Figure 3. Time course of simulated epidemics for infectiousness versus contact rate for the lowest transmission efficiency tested of $\tau = 0.025$. Columns correspond to the difference in mean degree tested, and rows correspond to the mechanism of covariation: control (no variability in infectiousness, 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 $\beta_c = 0$, limiting maximum prevalence to 0.66. Time courses for the corresponding medium ($\tau = 0.25$) and high ($\tau = 0.5$) transmission efficiencies are available in the Supplementary material Appendix 4 Fig. A5–A6.

(Fig. 5, 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 (Fig. 6B). This is likely because the control, null and positive covariation cases all saturated very quickly at higher transmission efficiencies (Supplementary material Appendix 4 Fig. A7–A8). Consistent with susceptibility versus contact rate and infectiousness versus 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 (Fig. 6C).

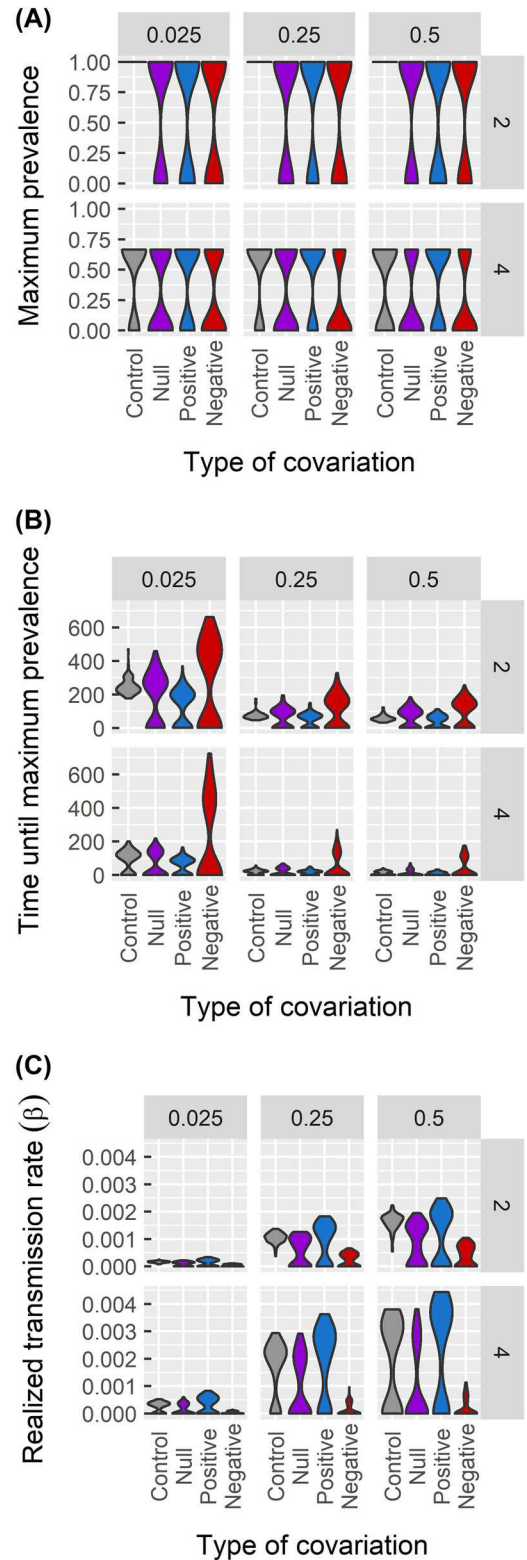


Figure 4. For infectiousness versus contact rate, violin plots depicting for all covariation types: (A) the maximum prevalence reached in 750 time steps; (B) the time it takes to reach that maximum prevalence; and (C) the realized transmission rate (β), which describes the rate of epidemic spread. The 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).

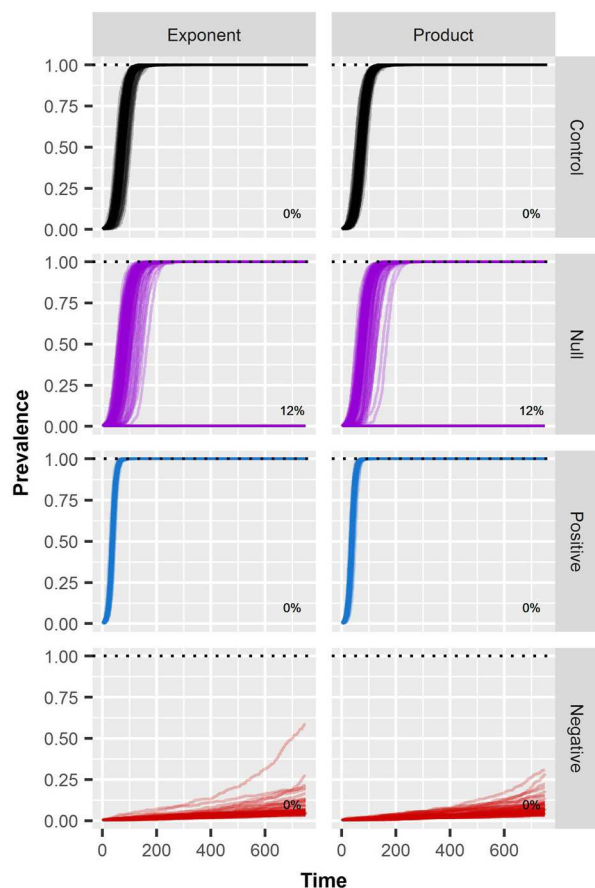


Figure 5. Time course of simulated epidemics for infection status versus contact rate for the lowest transmission efficiency tested of $\tau = 0.025$. Columns correspond to how infectiousness was modelled (either in the exponent or the product of the final transmission probability), and rows correspond 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 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. Time courses for the corresponding medium ($\tau = 0.25$) and high ($\tau = 0.5$) transmission efficiencies are available in the Supplementary material Appendix 4 Fig. A7–A8.

Random forest results

Variable importance scores for maximum prevalence indicate that the physiological phenotype of the index case had the highest importance for susceptibility versus contact rate and infectiousness versus contact rate; this was followed in importance by separation in mean degree, and then type of covariation (Table 4). Transmission efficiency had a negligible mean decrease in accuracy for both mechanisms in predicting maximum prevalence. For infection status versus 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 versus contact rate and infectiousness versus 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 versus contact rate (Table 4).

For the realized transmission rate (β), pathogen transmission efficiency was an informative predictor for all three experiments (Table 4). For both susceptibility versus contact rate and infectiousness versus 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 versus 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 versus contact rate.

Discussion

Accounting for contact heterogeneity has been shown to dramatically alter disease predictions (Keeling and 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 versus 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 versus contact rate and infectiousness

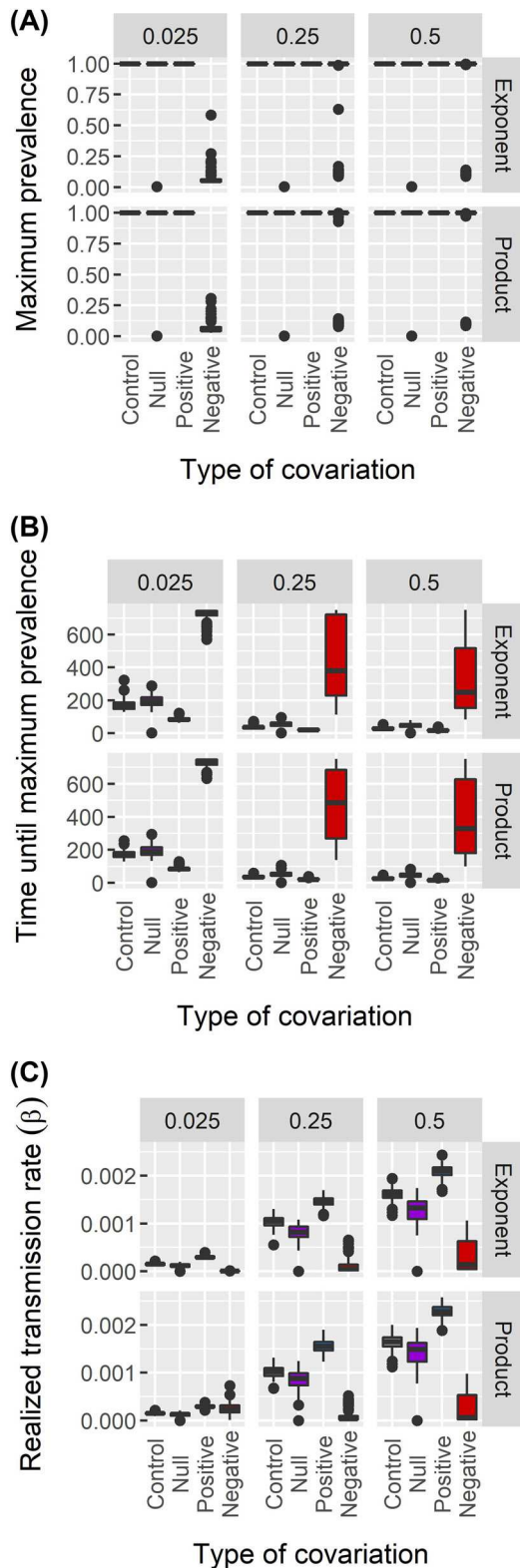


Figure 6. For infection status versus contact rate, box and whisker plots depicting for all covariation types: (A) the maximum prevalence reached in 750 time steps; (B) the time it takes to reach that maximum prevalence; and (C) the realized transmission rate (β), which describes the rate of epidemic spread. The columns

versus 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 (Fig. 2C, 4C, 6C).

Trends for time until maximum prevalence and the intrinsic rate of increase were consistent for susceptibility versus contact rate and infectiousness versus contact rate. Across parameter sets, infectiousness versus 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 (Fig. 2C, 4C, 6C), but overall, played a negligible role in explaining maximum prevalence, especially for susceptibility versus contact rate and infectiousness versus contact rate. For infection status versus 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

Figure 6. Continued

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.

Table 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 versus contact rate	Infectiousness versus contact rate	Infection status versus 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)	9.339×10^{-08}	1.570×10^{-07}	–
	Covariation	7.404×10^{-07}	5.623×10^{-07}	5.0133×10^{-07}
	Transmission efficiency (τ)	1.104×10^{-06}	6.402×10^{-07}	6.043×10^{-07}
	Physiology of index case (β_p : s or κ)	1.139×10^{-06}	6.497×10^{-07}	–
	Form of infectiousness (exponent or product)	–	–	5.436×10^{-09}

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 and Sikkell 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) versus 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), wild-life host–pathogen systems with existing dynamic contact network and individual physiological data are rare (Craft and Caillaud 2011, White et al. 2017a). 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 and 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 sick-

ness 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 and 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 and 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 and 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.

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 versus contact rate and infectiousness versus 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 versus 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.

Acknowledgements – Funding – LAW was funded by the National Science Foundation (GRFP-00039202 and DEB-1701069). MEC 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 Univ. of Minnesota for providing resources that contributed to the research results reported within this paper (<www.msi.umn.edu>).

References

- Adelman, J. S. et al. 2014. Using remote biomonitoring to understand heterogeneity in immune-responses and disease-dynamics in small, free-living animals. – *Integr. Comp. Biol.* 54: 377–386.
- Adelman, J. S. et al. 2015. Feeder use predicts both acquisition and transmission of a contagious pathogen in a North American songbird. – *Proc. R. Soc. B* 282: 20151429.
- Anderson, R. and May, R. 1991. *Infectious diseases of humans*. – Oxford Univ. Press.
- Anderson, R. et al. 1986. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. – *IMA J. Math. Appl. Med. Biol.* 3: 229–263.
- Bansal, S. et al. 2010. The dynamic nature of contact networks in infectious disease epidemiology. – *J. Biol. Dyn.* 4: 478–489.
- Barron, D. et al. 2015. Behavioral competence: how host behaviors can interact to influence parasite transmission risk. – *Curr. Opin. Behav. Sci.* 6: 35–40.
- Begon, M. et al. 2002. A clarification of transmission terms in host–microparasite models: numbers, densities and areas. – *Epidemiol. Infect.* 129: 147–153.
- Beldomenico, P. M. and Begon, M. 2010. Disease spread, susceptibility and infection intensity: vicious circles? – *Trends Ecol. Evol.* 25: 21–27.
- Berdoy, M. et al. 2000. Fatal attraction in rats infected with *Toxoplasma gondii*. – *Proc. R. Soc. B* 267: 1591–1594.

- Blonder, B. et al. 2012. Temporal dynamics and network analysis. – *Methods Ecol. Evol.* 3: 958–972.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- Cattadori, I. M. et al. 2007. Variation in host susceptibility and infectiousness generated by co-infection: the myxoma-*Trichostrongylus retortaeformis* case in wild rabbits. – *J. R. Soc. Interface* 4: 831–40.
- Chen, S. et al. 2014. Highly dynamic animal contact network and implications on disease transmission. – *Sci. Rep.* 4: 4472.
- Clay, C. A. et al. 2009. Contact heterogeneity in deer mice: implications for Sin Nombre virus transmission. – *Proc. R. Soc. B* 276: 1305–1312.
- Coelho, F. C. and Codeço, C. T. 2009. Dynamic modeling of vaccinating behavior as a function of individual beliefs. – *PLoS Comput. Biol.* 5: e1000425.
- Cosgrove, M. K. et al. 2012. Modeling vaccination and targeted removal of white-tailed deer in Michigan for bovine tuberculosis control. – *Wildl. Soc. Bull.* 36: 676–684.
- Craft, M. E. and Caillaud, D. 2011. Network models: an underutilized tool in wildlife epidemiology? – *Interdiscip. Perspect. Infect. Dis.* 2011: 676949.
- Croft, D. P. et al. 2011. Effect of gyrodactylid ectoparasites on host behaviour and social network structure in guppies *Poecilia reticulata*. – *Behav. Ecol. Sociobiol.* 65: 2219–2227.
- Cutler, D. R. et al. 2007. Random forests for classification in ecology. – *Ecology* 88: 2783–2792.
- Dwyer, G. et al. 1997. Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. – *Am. Nat.* 150: 685–707.
- Eisinger, D. and Thulke, H. 2008. Spatial pattern formation facilitates eradication of infectious diseases. – *J. Appl. Ecol.* 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 ver. 1.2-1. <<https://CRAN.R-project.org/package=minpack.lm>>.
- Ezenwa, V. O. et al. 2016. Host behaviour – parasite feedback: an essential link between animal behaviour and disease ecology. – *Proc. R. Soc. B* 283: 20153078.
- Fefferman, N. and Ng, K. 2007. How disease models in static networks can fail to approximate disease in dynamic networks. – *Phys. Rev. E* 76: 31919.
- Funk, S. et al. 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. – *J. R. Soc. Interface* 7: 1247–1256.
- Gibson, A. K. et al. 2016. Fine-scale spatial covariation between infection prevalence and susceptibility in a natural population. – *Am. Nat.* 188: 1–14.
- Godfrey, S. S. et al. 2009. Network structure and parasite transmission in a group living lizard, the gidgee skink, *Egernia stokesii*. – *Behav. Ecol. Sociobiol.* 63: 1045–1056.
- Grear, D. A. et al. 2009. Does elevated testosterone result in increased exposure and transmission of parasites? – *Ecol. Lett.* 12: 528–537.
- Hanneke, S. et al. 2010. Discrete temporal models of social networks. – *Electron. J. Stat.* 4: 585–605.
- Hawley, D. M. and Altizer, S. M. 2011. Disease ecology meets ecological immunology: understanding the links between organismal immunity and infection dynamics in natural populations. – *Funct. Ecol.* 25: 48–60.
- Hawley, D. M. et al. 2011. Does animal behavior underlie covariation between hosts' exposure to infectious agents and susceptibility to infection? Implications for disease dynamics. – *Integr. Comp. Biol.* 51: 528–539.
- Haydon, D. T. et al. 2006. Low-coverage vaccination strategies for the conservation of endangered species. – *Nature* 443: 692–695.
- Hirsch, B. T. et al. 2013. Raccoon social networks and the potential for disease transmission. – *PLoS One* 8: e75830.
- Hothorn, T. et al. 2006. Survival ensembles. – *Biostatistics* 7: 355–373.
- Jenness, S. M. et al. 2016a. EpiModel: mathematical modeling of infectious disease. – R package ver. 1.2.7. <<https://CRAN.R-project.org/package=EpiModel>>.
- Jenness, S. M. et al. 2016b. Impact of the Centers for Disease Control's HIV preexposure prophylaxis guidelines for men who have sex with men in the United States. – *J. Infect. Dis.* 214: 1800–1807.
- Keiser, C. N. et al. 2016. Personality composition alters the transmission of cuticular bacteria in social groups. – *Biol. Lett.* 12: 20160297.
- Keeling, M. J. and Eames, K. T. D. 2005. Networks and epidemic models. – *J. R. Soc. Interface* 2: 295–307.
- Koolhaas, J. M. 2008. Coping style and immunity in animals: making sense of individual variation. – *Brain. Behav. Immun.* 22: 662–667.
- Krause, J. et al. 2013. Reality mining of animal social systems. – *Trends Ecol. Evol.* 28: 541–51.
- Krivitsky, P. N. and Handcock, M. S. 2014. A separable model for dynamic networks. – *J. R. Stat. Soc. Ser. B Stat. Methodol.* 76: 29–46.
- Kyvsgaard, N. C. et al. 2007. Simulating transmission and control of *Taenia solium* infections using a Reed–Frost stochastic model. – *Int. J. Parasitol.* 37: 547–558.
- Lloyd-Smith, J. O. et al. 2005. Superspreading and the effect of individual variation on disease emergence. – *Nature* 438: 355–359.
- Lopes, P. C. et al. 2016. Infection-induced behavioural changes reduce connectivity and the potential for disease spread in wild mice contact networks. – *Sci. Rep.* 6: 31790.
- MacIntosh, A. J. J. et al. 2012. Monkeys in the middle: parasite transmission through the social network of a wild primate. – *PLoS One* 7: e51144.
- Masuda, N. and Holme, P. 2013. Predicting and controlling infectious disease epidemics using temporal networks. – *F1000Prime Rep.* 5: 6.
- May, R. M. et al. 1988. The transmission dynamics of human immunodeficiency virus (HIV). – *Phil. Trans. R. Soc. B* 321: 565–607.
- McCallum, H. et al. 2017. Breaking beta: deconstructing the parasite transmission function. – *Phil. Trans. R. Soc. B* 372: 20160084.
- Meloni, S. et al. 2011. Modeling human mobility responses to the large-scale spreading of infectious diseases. – *Sci. Rep.* 1: 62.
- Meyers, L. A. 2007. Contact network epidemiology: bond percolation applied to infectious disease prediction and control. – *Am. Math. Soc.* 44: 63–86.
- Natoli, E. et al. 2005. Bold attitude makes male urban feral domestic cats more vulnerable to feline immunodeficiency virus. – *Neurosci. Biobehav. Rev.* 29: 151–157.
- Paull, S. H. et al. 2012. From superspreaders to disease hotspots: linking transmission across hosts and space. – *Front. Ecol. Environ.* 10: 75–82.

- Perkins, S. E. et al. 2009. Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. – *J. Anim. Ecol.* 78: 1015–1022.
- Restif, O. et al. 2012. Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. – *Ecol. Lett.* 15: 1083–1094.
- Robins, G. et al. 2007. An introduction to exponential random graph (p*) models for social networks. – *Soc. Networks* 29: 173–191.
- Springer, A. et al. 2017. Dynamic versus static social networks in models of parasite transmission: predicting *Cryptosporidium* spread in wild lemurs. – *J. Anim. Ecol.* 86: 419–433.
- Stehlé, J. et al. 2011. Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. – *BMC Med.* 9: 87.
- Strobl, C. et al. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. – *BMC Bioinf.* 8: 25.
- Strobl, C. et al. 2009. Party on! A new, conditional variable-importance measure for random forests available in the party package. – *R J.* 1: 14–17.
- Tompkins, D. M. et al. 2011. Wildlife diseases: from individuals to ecosystems. – *J. Anim. Ecol.* 80: 19–38.
- VanderWaal, K. L. and Ezenwa, V. O. 2016. Heterogeneity in pathogen transmission: mechanisms and methodology. – *Funct. Ecol.* 30: 1606–1622.
- Verdolin, J. L. et al. 2014. Key players and hierarchical organization of prairie dog social networks. – *Ecol. Complex.* 19: 140–147.
- Volz, E. and Meyers, L. A. 2007. Susceptible–infected–recovered epidemics in dynamic contact networks. – *Proc. R. Soc. B* 274: 2925–2934.
- Welicky, R. L. and Sikkell, P. C. 2015. Decreased movement related to parasite infection in a diel migratory coral reef fish. – *Behav. Ecol. Sociobiol.* 69: 1437–1446.
- White, J. W. et al. 2014. Ecologists should not use statistical significance tests to interpret simulation model results. – *Oikos* 123: 385–388.
- White, L. A. et al. 2017a. Using contact networks to explore mechanisms of parasite transmission in wildlife. – *Biol. Rev.* 92: 389–409.
- White, L. A. et al. 2017b. Data from: Covariation between the physiological and behavioral components of pathogen transmission: host heterogeneity determines epidemic outcomes. Dryad Digital Repository <<http://dx.doi.org/10.5061/dryad.8t201>>.
- Woolhouse, M. E. et al. 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. – *Proc. Natl Acad. Sci. USA* 94: 338–342.
- Zohdy, S. et al. 2012. Mapping the social network: tracking lice in a wild primate (*Microcebus rufus*) population to infer social contacts and vector potential. – *BMC Ecol.* 12: 4.

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