vol. 181, no. 2 the american naturalist february 2013

E-Article

Distinguishing between Indirect and Direct Modes of Transmission Using Epidemiological Time Series

Michael H. Cortez[1,*] and Joshua S. Weitz[2]

1. School of Biology and School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332; 2. School of Biology and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

Submitted April 27, 2012; Accepted August 28, 2012; Electronically published January 8, 2013

Online enhancement: appendix PDF.

abstract: Pathogen transmission can involve direct and/or indirect pathways. Using theoretical models, in this study we ask, "do directly and indirectly transmitted pathogens yield different population-level epidemiological dynamics?" and "can the transmission pathway be inferred from population-level epidemiological data?" Our approach involves comparing the continuous-time dynamics of a class of compartmental epidemiological models with direct versus environmentally mediated indirect transmission pathways. Combing analytical theory and numerical simulations we show that models with direct and indirect transmission can produce quantitatively similar time series when the pathogen cannot reproduce in the environment, particularly when the environmental pathogen dynamics are fast. We apply these results to a previous study on chronic wasting disease and show that identifying the transmission pathway is more difficult than previously acknowledged. Our analysis and simulations also yield conditions under which numerical differences can potentially identify the transmission route in oscillating endemic systems and systems where the environmental pathogen dynamics are not fast. This work begins to identify how differences in the transmission pathway can result in quantitatively different epidemiological dynamics and how those differences can be used to identify the transmission pathway from population level time series.

Keywords: indirect transmission, direct transmission, chronic wasting disease, disease dynamics.

Introduction

Pathogen transmission can involve direct or indirect pathways. Direct transmission occurs via contacts between infectious and susceptible conspecifics. Indirect transmission occurs when susceptible hosts are exposed to the pathogen in the environment via free-living pathogens or other infected host species. Identification of the transmission route is important because it can facilitate improved prevention and control of infectious diseases; for example, in foot and mouth disease and tuberculosis in livestock (Brennan et al. 2008) and influenza (Spicknall et al. 2009), *Vibrio cholerae* (Kaper et al. 1995), viral hepatitis A (Ajelli et al. 2008), and *Giardia lamblia* in humans (Wolfe 1992). Knowledge of the transmission pathway can yield understanding about the ecological effects of pathogens in natural communities, for example, plague in prairie dogs (Webb et al. 2006), black band disease in corals (Aeby and Santavy 2006), avian influenza (Breban et al. 2009), chronic wasting disease in deer (Miller et al. 2006), amphibian ranavirus in salamanders (Brunner et al. 2007), and *Mycoplasma gallisepticum* in house finches (Dhondt et al. 2005). Knowledge of the transmission pathway is also important for understanding the evolution of ecologically important pathogen traits (e.g., virulence evolution; Day 2001; Roche et al. 2011).

While transmission routes for many pathogens lie between these two extremes (e.g., airborne-transmitted pathogens like *Mycobacterium tuberculosis*; Siegel et al. 2007), in theoretical studies pathogen transmission is often partitioned into direct and indirect components or pathways (Miller et al. 2006; Roche et al. 2009; Spicknall et al. 2009; Tien and Earn 2010; Roche et al. 2011). When only a single transmission pathway is assumed, it may be unclear which class of models is more appropriate. In some cases, the true transmission pathway is unknown, and in others, multiple transmission pathways may be available but the pathway typically utilized by the pathogen is unknown. For example, *Escherichia coli* persists in the environment, but its mode of transmission is not always known and can depend on the host species (LeJeune et al. 2004; Cornick and VuKhac 2008; Arthur et al. 2009).

One approach used to investigate the contribution different transmission pathways have on the epidemiological dynamics of a system involves simulating differential equation based models and comparing the output to epidemic time series data. In this approach, the dynamics of candidate direct, indirect, and mixed transmission models are

* Corresponding author; e-mail: michael.cortez@biology.gatech.edu.

Am. Nat. 2013. Vol. 181, pp. E43–E52.
0003-0147/2013/18102-53819$15.00.
DOI: 10.1086/668826

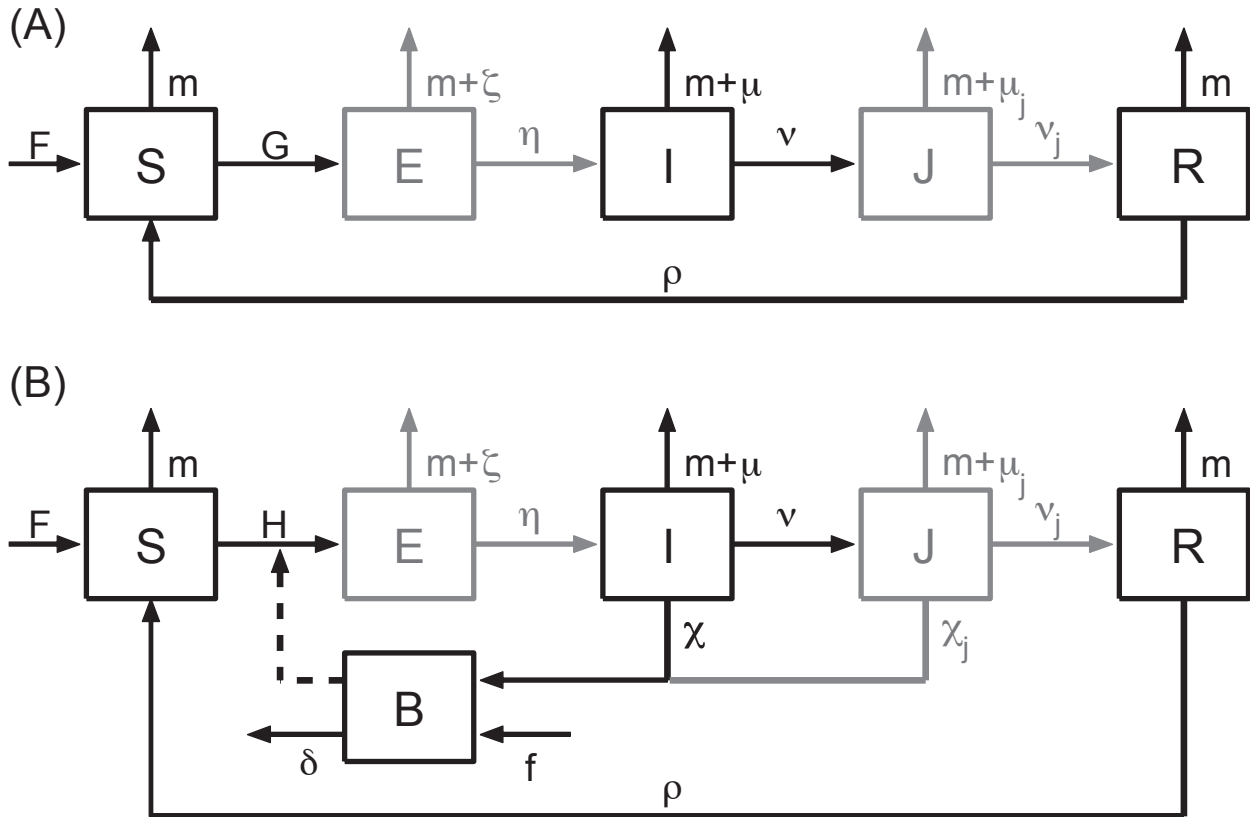

Figure 1: Model diagrams for direct (*A*) and indirect (*B*) transmission pathogen systems. Hosts can be susceptible (S), exposed (E), infectious (I, J), or recovered (R). The density of the pathogen in the environment is denoted by B. Black denotes direct and indirect models with a single infectious class and no exposed class. Gray denotes the addition of exposed or multiple infectious classes.

compared to time series data using statistical methods (like the Akaike Information Criteria). The relative fits of the models and their fitted parameters are then used as evidence for the relative contributions different transmission pathways have on the epidemiological dynamics of the system. This comparative approach has been used (albeit with a model that incorporates stochasticity) to determine the ratio of asymptomatic cases and the duration of immunity for cholera (King et al. 2008). That same study also explored if the seasonality of cholera epidemics was driven by interactions between *V. cholerae* and phage in the environment. This approach has also been used to study the impact of waterborne indirect transmission and direct transmission on the disease dynamics of avian influenza in wild birds (Roche et al. 2009) and *Cryptosporidium parvum* in humans (Brookhart et al. 2002).

In principle, numerical differences between model fits can allow one to statistically differentiate between models and identify the transmission pathway. However, if the models can exhibit numerically similar dynamics, then in some cases it may not be possible to identify the transmission pathway from epidemiological time series. As an example of this arising in practice, we revisit a study on chronic wasting disease that fit direct and indirect transmission models to mortality data on mule deer (Miller et al. 2006). While the authors concluded that the data and model fits supported an indirect transmission pathway (Miller et al. 2006), our analysis shows that a direct transmission model and an indirect transmission model that exhibit equivalent dynamics explain the data equally well.

In this study, using theoretical models, we ask, "do directly and indirectly transmitted pathogens yield different population-level epidemiological dynamics?" and "can the transmission pathway can be inferred from population-level epidemiological data?" We argue that in some cases inferences made about the transmission pathway from time series data may be more challenging than previously recognized. In particular, population level epidemiological time series can provide less information than previously thought about the transmission pathway because the population-level dynamics of those systems can be similar or in some cases identical. Our arguments are based on the

analysis of a class of compartmental epidemiological models. As an initial case, we consider deterministic ordinary differential equation (ODE) models with a single transmission pathway (fig. 1). Furthermore, we focus on indirect transmission models where the pathogen cannot reproduce in the environment. We analyze these models and explore when their continuous-time dynamics are numerically identical and when they quantitatively differ.

Our approach is the following. We first show analytically that direct and indirect transmission models can produce identical host time series when the environmental pathogen load dynamics are fast. Thus, under these conditions one cannot distinguish between models. As a demonstration of this issue arising in practice, we revisit a previous study of chronic wasting disease in mule deer (Miller et al. 2006). We then compare the simulations of direct and indirect transmission models and identify conditions under which numerical differences arise. We find that in some cases our models can produce similar time series, suggesting that identifiability issues can arise even when there is not a separation of timescales. However, we also identify conditions under which numerical differences do arise and these differences could be leveraged when using statistical comparisons of model fits. Finally, we explore when phase relations in endemically oscillating systems differ between direct and indirect transmission models.

Methods

Direct Transmission Model

We begin with a direct transmission system where S, I, and R, denote the densities of the susceptible, infectious, and recovered host classes, respectively (fig. 1*A*). Note that we only consider horizontally transmitted pathogens. The direct transmission system is

$$\begin{aligned}\frac{dS}{dt} &= F(S, I, R) - G(S, I, R) + \rho R - mS,\\ \frac{dI}{dt} &= G(S, I, R) - \nu I - \mu I - mI,\\ \frac{dR}{dt} &= \nu I - \rho R - mR,\end{aligned} \tag{1}$$

where G is the transmission rate, m is the natural mortality rate, ν is the recovery rate, μ is pathogen mortality rate, and ρ is the rate at which recovered individuals lose immunity. The function F is the growth rate of the population, and we assume all new individuals enter through the susceptible class. We assume that G is an increasing function of S and I and a nonincreasing function of R; for example, $G = \beta SI$ or $G = \beta SI/(S + I + R)$.

Indirect Transmission Model

We compare system (1) to an environmentally mediated indirect transmission system where infectious individuals excrete the pathogen into the environment and susceptible individuals become infected through contact with the contaminated environment (fig. 1*B*). The pathogen environmental load is denoted by B, and we assume that the pathogen cannot reproduce in the environment. The indirect transmission system is

$$\begin{aligned}\frac{dS}{dt} &= F(S, I, R) - H(S, I, R, B) + \rho R - mS,\\ \frac{dI}{dt} &= H(S, I, R, B) - \nu I - \mu I - mI,\\ \frac{dR}{dt} &= \nu I - \rho R - mR,\\ \frac{dB}{dt} &= \chi I - \delta B.\end{aligned} \tag{2}$$

Here, H is the infection rate due to contact with the pathogen in the environment, χ is the per capita excretion rate of the pathogen, and δ is the pathogen per capita mortality or decay rate. We assume that H is an increasing function of S and B and a nonincreasing function of I and R, for example, mass action transmission $H = \alpha SB$ or frequency-dependent transmission $H = \alpha SB/(S + I + R)$ (Stilianakis and Drossinos 2010).

Modeling Assumptions and Background

The level of generality of systems (1) and (2) includes systems where transmission is density dependent or frequency dependent, where immunity is permanent (SIR systems) or can be lost (SIS and SIRS systems), and where recovery is not possible (SI systems). These models and their stochastic variants are widely used in the epidemiological literature, but we note that many pathogens utilize both transmission pathways and that the dynamics of mixed transmission systems have been studied previously (Brookhart et al. 2002; Sauvage et al. 2003; Ajelli et al. 2008; Roche et al. 2009; Tien and Earn 2010). In addition, we assume that the causative agent of the disease cannot reproduce in the environment. While this is true for some diseases (e.g., prions or viruses), there are many systems where pathogen reproduction in the environment is important (e.g., *Vibrio cholerae*; Jensen et al. 2006; Hove-Musekwa et al. 2011). Finally, we assume that the loss of pathogen in the environment due to uptake by susceptible individuals is negligible, but other studies have considered models that relax this assumption (Li et al. 2009).

In our numerical examples, we will typically use $F = 0$,

$G(S, I, R) = \beta SI$ and $H(S, I, R, B) = \alpha SB$. Nonetheless, our analytical results in sections "Time Series Are Identical When Environmental Dynamics Are Fast" and "Differences in the Phase Lags of Oscillating Endemic Systems" hold for any functions F, G, and H that satisfy our assumptions. In appendix A, available online, we extend systems (1) and (2) to include exposed and multiple infectious classes. We reference the results pertaining to those models throughout.

Results

Time Series Are Identical When Environmental Dynamics Are Fast

We first ask if there exist conditions under which systems (1) and (2) produce numerically identical time series. Direct and indirect transmissions systems can yield identical time series when the environmental load dynamics of system (2) are much faster than the epidemiological dynamics of the host. Biologically, this arises when the pathogen is shed in large quantities by infectious individuals and the pathogen has a short expected life or degradation time in the environment (χ and δ are large).

Under these conditions, the dB/dt dynamics are much faster than the dS/dt, dI/dt, and dR/dt dynamics in system (2) This results in a fast-slow system (an area known as singular perturbation theory; Arnold et al. (1995) with a separation of timescales between the environmental and epidemiological dynamics of the system. Mathematically, in this fast-slow case, the environmental load is approximated by the solution to $dB/dt = 0$ in system (2); see appendix B, available online, for details. Solving yields $B(t) = \chi I(t)/\delta$, implying that environmental load tracks the infectious population. Substitution into system (2) yields

$$\begin{aligned}
\frac{dS}{dt} &= F(S, I, R) - H\left(S, \frac{\chi}{\delta} I\right) + \rho R - mS,\\
\frac{dI}{dt} &= H\left(S, \frac{\chi}{\delta} I\right) - \nu I - \mu I - mI, \qquad (3)\\
\frac{dR}{dt} &= \nu I - \rho R - mR.
\end{aligned}$$

If the transmission terms G and H in systems (1) and (2) have similar functional forms, then χ and δ can be chosen such that the time series of the host dynamics in systems (1) and (2) are identical. This equivalence has been shown previously for a particular mixed transmission model (Tien and Earn 2010). Note that technically systems (1) and (2) are approximately equal in the fast-slow limit (see app. B). While mathematically this means that small numerical differences do exist between the time series of the two systems, because of the separation of timescales, these numerical differences are much smaller than the error in experimental measurements or small stochastic fluctuations.

The above equivalence depends strongly on the transmission functional forms G and H. For example, consider systems (1) and (2) with density-dependent transmission rates $G_1(S, I) = \beta SI$ and $H_1(S, B) = \alpha SB$, respectively. When χ and δ are large (implying fast environmental dynamics), the above implies that the indirect transmission rate will become $H_1 = \alpha\chi SI/\delta$. When $\alpha = \beta\delta/\chi$, H_1 simplifies to $H_1 = \beta SI$ and the direct and indirect transmission models yield equivalent host time series (see fig. S1, available online). Now consider the frequency-dependent direct transmission rate, $G_2(S, I) = \beta SI/N$ where $N = S + I + R$ is the total population size. If N varies in time, then the indirect transmission model with $H_1(S, B) = \alpha SB$ will not yield identical dynamics when the environmental dynamics are fast. In contrast, consider system (2) with the frequency-dependent indirect transmission rate $H_2(S, B) = \alpha SB/N$ (see Stilianakis and Drossinos 2010 and app. E, available online, for a description). Following the above, when the environmental dynamics are fast, the direct transmission model with $G_2(S, I)$ and the indirect transmission model with $H_2(S, B)$ will have equivalent dynamics when $\beta = \alpha\delta/\chi$. Hence, via fast-slow systems theory, there does exist at least one indirect transmission model that yields equivalent dynamics when the environmental dynamics are fast. However, the time series from a given direct transmission system cannot be exactly reproduced by all indirect transmission systems.

The above conclusions hold for any system with a single infectious class when the pathogen cannot reproduce in the environment. For example, direct and indirect transmission systems with an exposed class (e.g., SEIR) produce equivalent dynamics when the environmental dynamics are fast; see appendix B. These conclusions do not always hold for systems with multiple infectious classes. Indirect and direct transmission models are not equivalent in the fast-slow limit when there are multiple infectious classes and transmission is nonlinear; see appendix B. Nonlinear functional forms have been proposed and used previously in the literature (Anderson and May 1978; Liu et al. 1987; Diekmann and Kretzschmar 1991; Hochberg 1991; Heesterbeek and Metz 1993; Briggs and Godfray 1995; Barlow 2000; McCallum et al. 2001). Thus, host heterogeneity with respect to shedding rates and disease progression can prevent this identifiability issue from arising.

Application to Chronic Wasting Disease

As an example of how the above issues can arise in practice, we apply our theory to the Miller et al. (2006) study on

the prion disease, chronic wasting disease (CWD), in mule deer (*Odocoileus hemionus*). In that study, three direct (SI, SEI, and SI_1I_2 with two infectious classes), two indirect (iSI and iSEI), and one mixed transmission models were fit using maximum likelihood techniques to observed cumulative mortality data from two epidemics in captive mule deer populations. Model fits were compared using the Akaike Information Criterion (AIC); see appendix D, available online, for details. In all models, the prions could not reproduce in the environment, and infected individuals could not recover from the disease. Of the six fitted models, the iSI model was best supported by the data (reported AIC value closest to zero), followed by the SI and iSEI models. The reported parameter and statistical values for these three models are in table 1. The apparent statistical support for the iSI model led the authors to conclude that CWD is most likely indirectly transmitted (Miller et al. 2006).

Two key observations can be made about the reported values in table 1. First, for the reported parameter values, the environmental load dynamics of the iSEI model are much faster than the host dynamics. Numerically, this is seen by simulating the fitted iSEI system where χ and δ are unaltered and where χ and δ are 100 times larger. Due to the separation of timescales, the dynamics remain unchanged; see figure S3, available online. For the particular transmission functions used in the Miller et al. (2006) study, the separation of timescales in the fitted iSEI model implies that there exists parameter values for the SEI model that reproduce the host dynamics of the fitted iSEI model; see appendix D for details. In particular, setting the direct transmission coefficient to $\beta = 0.66$ in the SEI model and holding all other parameters constant yields an SEI model, denoted by SEI*, that produces host dynamics identical to those of the fitted iSEI model (fig. 2).

Second, the fitted iSEI model fits the data the best (reported log likelihood closest to zero in table 1). Since the SEI* and fitted iSEI host dynamics are identical, the SEI* model fits the data just as well. After accounting for the two fewer parameters in the SEI* model, we find that the SEI* model has the lowest AIC value in table 1; see appendix D for details. Thus, the SEI* model has the best fit and the most support out of the six models. How should we interpret this finding? The SEI* and fitted iSEI models exhibit equivalent host dynamics, have the same parameter values except for those relating to transmission, and give estimates for R_0 (the basic reproductive number) consistent with the long-term survival of the disease in natural populations (Miller et al. 2006). Thus, we cannot reject the iSEI model, nor can we reject the SEI* model. Our results suggest that it is premature to conclude whether CWD is (i) primarily transmitted directly, (ii) primarily transmitted indirectly via prions with a short survival time in the environment, or (iii) transmitted via both modes of transmission.

Comparisons of Time Series When Environmental Dynamics Are Not Fast

Numerical Comparisons of Disease Models with Direct versus Indirect Transmission. The previous section shows how certain indirect transmission models can exactly reproduce the time series from a subset of direct transmission models. In addition, when the environmental load dynamics are fast, direct transmission models can exactly reproduce the time series from indirect transmission models. However, it is not clear whether direct transmission models can reproduce the time series from indirect transmission models when the environmental load dynamics are not fast. Hence, as an alternative, we compared the simulated time series of three particular pairs of direct and indirect transmission models in the following way. First, we generated time series with the indirect transmission model. Then we simulated the direct transmission model and searched for parameters that produced time series that best fit the indirect transmission time series. This fitting of direct transmission models to indirect transmission time series cap-

Table 1: Reported and computed parameter values and statistics for fitted transmission models of chronic wasting disease from Miller et al. (2006)

	Parameters[a]						Statistics[b]		
Model	β_1	β_2	μ	α	χ	δ	ln $\mathcal{L}$	AIC	Source
iSI	...	.787	.567	...	.111	2.55	−30.61	84.5	Miller et al. 2006
SI	.035	...	.481	...	...	...	−36.21	86.9	Miller et al. 2006
iSEI	...	1.13	.546	.486	.332	5.66	−28.94	87.0	Miller et al. 2006
SEI*	.66	...	.546	.486	...	...	−28.94	76.3	Current study

[a] Parameters are direct transmission coefficient (individual^{-1}yr^{-1}), indirect transmission coefficient (mass^{-1}yr^{-1}), disease induced mortality rate (yr^{-1}), proportional latent time, excretion rate (yr^{-1}), and pathogen death rate (yr^{-1}), respectively.

[b] Statistics are log likelihood and the Akaike Information Criteria (AIC) for small samples, respectively.

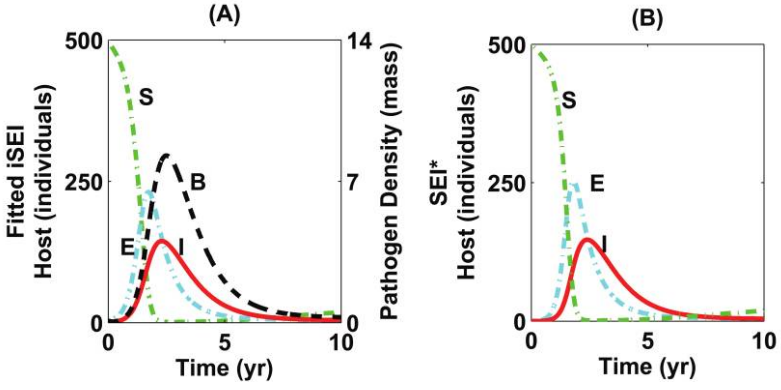

Figure 2: Numerical simulations of an indirect transmission model fitted to mortality data from two epidemics of chronic wasting disease in mule deer in Miller et al. (2006) and a direct transmission model that exactly reproduces the host epidemiological dynamics of the fitted indirect transmission model. Numbers of susceptible (green dash-dot), exposed (cyan dash-dot), and infectious individuals (solid red) and environmental load (dashed black) time series for the iSEI model (*A*) and the SEI model (*B*) from the Miller et al. (2006) study. Parameters in *A* are the reported best-fitting parameters. Parameters in *B* were calculated using the fast-slow theory in section "Time Series Are Identical When Environmental Dynamics Are Fast"; for details see main text and appendix D, available online. *A*, Model is $dS/dt = a - S(\beta_1 I + m)$, $dE/dt = \beta_1 SI - E(\mu/\alpha + m)$, and $dI/dt = \mu L/\alpha - I[\mu/(1-\alpha) + m]$. Initial conditions are $S = 500$, $E = 0$, $I = 1$, and $B = 0.1$. *B*, Model is $dS/dt = a - S(\beta_2 B + m)$, $dE/dt = \beta_2 SB - E(\mu/\alpha + m)$, $dI/dt = \mu L/\alpha - I[\mu/(1-\alpha) + m]$, and $dB/dt = \chi I - \delta B$. Initial conditions are $S = 500$, $E = 0$, $I = 1$, and $B = 0$. In both systems, $a = 10$ year^{-1} is the immigration rate. All other parameters are defined in table 1.

tures how well direct transmission models can reproduce the dynamics of indirect transmission models.

The pairs of transmission terms (G, H) used in our analysis were $(\beta SI, \alpha SB)$, $(\beta SI/N, \alpha SB/N)$, and $(\beta SI/[1 + m_1 I], \alpha SB/[1 + m_2 B])$; see appendix E for a description of the functional forms. For each model comparison, we generated 5,000 parameter sets using Latin Hypercube sampling for the indirect transmission model. For each parameter set, the direct transmission model was fit to the indirect transmission time series by minimizing the maximum pointwise difference between the infectious class time series from the two models. We chose this metric so as to focus on large deviations between time series; see appendix E for details. When fitting the direct transmission models, only the transmission coefficient β, the initial infectious class size $I(0)$, and, when appropriate, the parameter h_2 were allowed to vary. All other parameters were fixed at those used in the indirect transmission model. Fixing the other parameters constrains the model fitting in a way akin to having highly informative prior estimates of life-history parameters.

Representative examples of the numerical fitting for the model pair $(\beta SI, \alpha SB)$ are presented in figures S4 and S5, available online. Parameters chosen in figures S4 and S5 correspond to epidemic and endemic systems, respectively. Figures 3*A* and 4*A* illustrate how the peak height of the time series (circles) and the quality of the fits of the direct transmission model (relative error, squares) depend on the speed of the environmental dynamics. When the environmental dynamics are fast (small timescale values), the maximum prevalence of the time series is relatively large and the differences between the time series are negligible. As expected, the direct and indirect transmission time series are identical in the limit of fast environmental dynamics (figs. 3*B*, 4*B*).

As the speed of the environmental dynamics decreases, the measured error between the direct and indirect transmission time series increases. When the environmental dynamics are comparable to or slower than the epidemiological dynamics, the time series can remain quantitatively similar, for example, figure 3*C*. However, in many cases large numerical differences arise between time series. The kinds of numerical differences depend on the epidemiological dynamics of the system. In epidemic systems, slow environmental dynamics yield small epidemics (compare scales in fig. 3*C* and 3*D*) and direct transmission models poorly fit the width of the epidemic peak, thus large differences between time series arise (fig. 3*D*). In endemic systems, direct transmission models cannot fit both the initial rise in infections and the endemic level of the infection (fig. 4*C* and 4*D*).

We find the same relationship between model fits and the speed of the environmental dynamics for the other model pairs. Furthermore, the relationship is driven by the same kinds of numerical differences between the direct and indirect transmission time series. The results are summarized in figures S4 and S5. These simulations suggest that time series corresponding to direct and indirect transmission models can have quantitative differences between them when environmental dynamics are not fast. In some

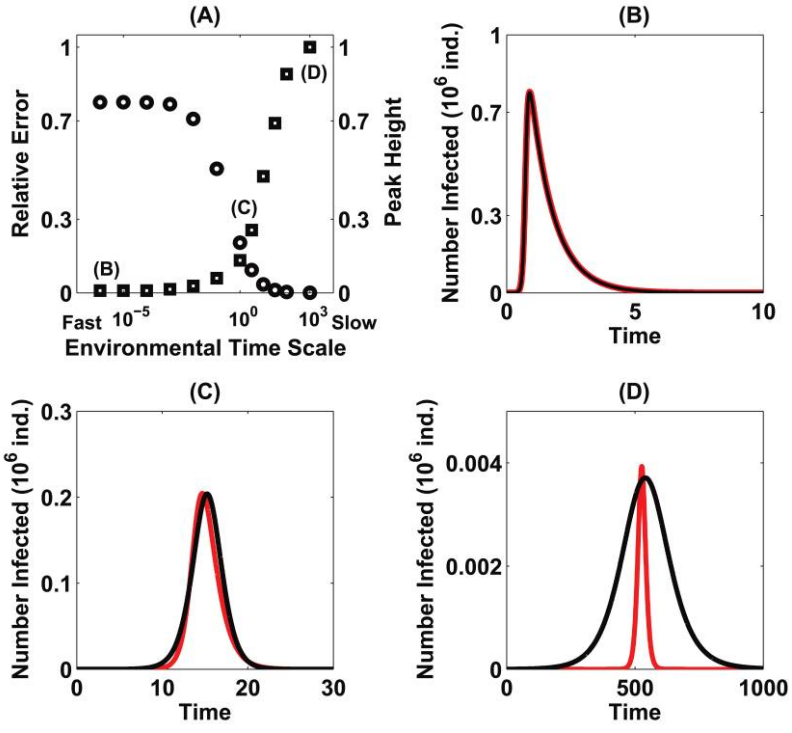

Figure 3: Numerical comparisons of epidemic time series for different speeds of the environmental load dynamics. *A*, Error (squares) between indirect transmission and the best-fitting direct transmission time series relative to the peak height of the infectious class (circles). The abscissa is the speed of the environmental dynamics relative to the speed of the environmental dynamics (ε) on a log axis. Small (large) values indicate faster (slower) environmental dynamics. Letters denote time scale in *B–D*. *B–D*, Examples of indirect transmission time series (black) and the best-fitting direct transmission time series (red) for fast ($\varepsilon = 10^{-4}$), comparable speed ($\varepsilon = 1$), and slow ($\varepsilon = 10^4$) environmental dynamics, respectively. Initial host population size is 10^6 individuals. Parameter values and functions are $F = 0$, $G(S, I, R) = \beta SI$, $H(S, I, R, B) = \alpha SB$, $\alpha = 2$, $m = \rho = 0$, $\mu = 0.1$, $\nu = 1$, $\chi = 1/\varepsilon$, and $\delta = 0.1/\varepsilon$. In *B–D*, (β, ε) is given by $(19.12, 10^{-6})$, $(2.54, 1)$, and $(1.21, 10^2)$, respectively.

cases, these differences can be substantial enough to suggest an indirect pathway of transmission.

The Effects of Incorporating Prior Information about Life-History Parameters. Here we interpret our results in the context of identifying transmission routes via fits of deterministic models to experimental epidemic time series data. Note that our arguments only apply to data with observational noise.

Consider the case where parameters are estimated via model fits by maximizing the posterior probability of a set of parameters given the data. Let $P_{\text{ind}}(\beta_{\text{ind}}, \gamma|x)$ be the posterior probability of a set of parameters given the data for an indirect transmission model. Here x is the data; β_{ind} are the parameters associated with indirect transmission, pathogen excretion and pathogen mortality; and γ are all other parameters. Similarly, let $P_{\text{dir}}(\beta_{\text{dir}}, \gamma|x)$ be the posterior probability for a direct transmission systems where β_{dir} are the parameters associated with direct transmission, pathogen excretion, and pathogen mortality. In a Bayesian framework, $P_{\text{ind}}(\beta_{\text{ind}}, \gamma|x)$ is proportional to the product $\mathcal{L}_{\text{ind}}(x|\beta_{\text{ind}}, \gamma)p_{\text{ind}}(\beta_{\text{ind}}, \gamma)$, where $\mathcal{L}_{\text{ind}}$ is the likelihood of the data given the parameters and p_{ind} is the prior distribution of the parameters. Similarly, $P_{\text{dir}}(\beta_{\text{dir}}, \gamma|x)$ is proportional to the product $\mathcal{L}_{\text{dir}}(x|\beta_{\text{dir}}, \gamma)p_{\text{dir}}(\beta_{\text{dir}}, \gamma)$, where $\mathcal{L}_{\text{dir}}$ is the likelihood of the data given the parameters and p_{dir} is the prior distribution of the parameters. Our results from simulations suggest that direct and indirect transmission models can produce quantitatively similar time series in some cases. In the cases where the best-fitting direct and indirect models produce quantitatively similar time series, the likelihood values for the models will be similar; that is, $\mathcal{L}_{\text{ind}}(x|\beta_{\text{ind}}, \gamma) \approx \mathcal{L}_{\text{dir}}(x|\beta_{\text{dir}}, \gamma)$. This implies that differences in the posterior probabilities will be driven by prior information about the parameters.

However, it is important to note that prior information will not always allow one to distinguish between models. If little information is known about the transmission mechanism and parameters related to transmission, then

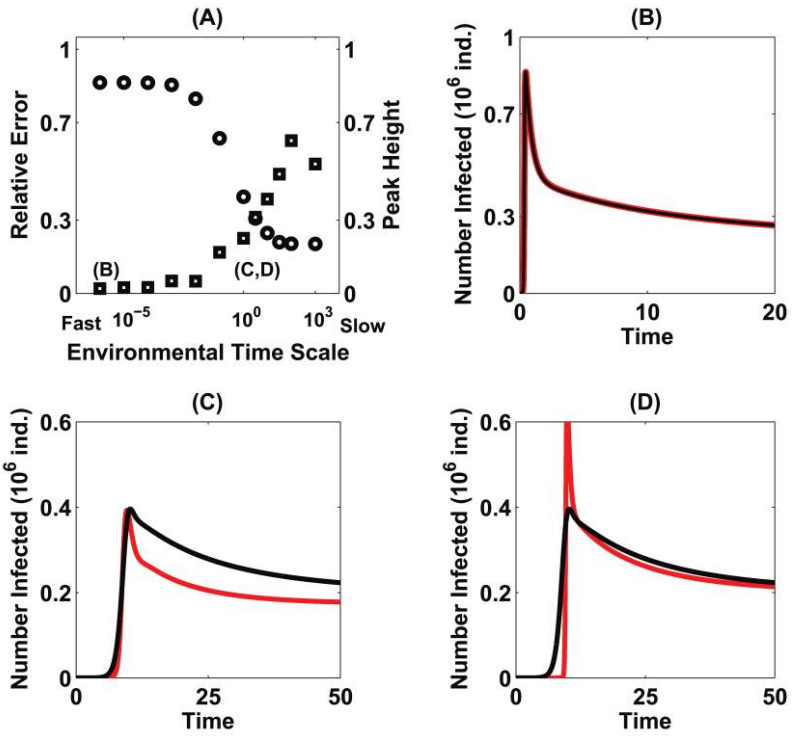

Figure 4: Numerical comparisons of endemic time series for different speeds of the environmental load dynamics. *A*, Error (squares) between indirect transmission and the best-fitting direct transmission time series relative to the peak height of the infectious class (circles). The abscissa is the speed of the environmental dynamics relative to the speed of the environmental dynamics (ε) on a log axis. Small (large) values indicate faster (slower) environmental dynamics. Letters denote time scale for *B–D*. *B*, *C*, Examples of indirect transmission time series (black) and the best-fitting direct transmission time series (red) for fast ($\varepsilon = 10^{-4}$) and comparable speed ($\varepsilon = 1$) environmental dynamics, respectively. *D*, Direct transmission time series (red) that fits the endemic state but not the initial peak of the indirect transmission time series (black). Initial host population size is 10^6 individuals. Parameter values and functions are $F = rN(1 - N/K)$, $G(S, I, R) = \beta SI$, $H(S, I, R, B) = \alpha SB$, $r = 0.1$, $K = 10^6$, $\alpha = 4$, $\rho = 0$, $m = 0.01$, $\mu = 0.1$, $\nu = 1$, $\chi = 1/\varepsilon$, and $\delta = 0.1/\varepsilon$. In *B–D*, (β, ε) is given by $(36.4, 10^{-6})$, $(4.08, 1)$, and $(12.37, 11)$, respectively.

the priors will only depend on the other parameters, that is, $p_{\text{ind}}(\beta_{\text{ind}}, \gamma) \sim p_{\text{ind}}(\gamma)$ and $p_{\text{dir}}(\beta_{\text{dir}}, \gamma) \sim p_{\text{dir}}(\gamma)$. In the simulations discussed in the previous section, in some cases direct transmission models were able to fit indirect transmission time series while holding parameters unrelated to transmission constant. When fitting models to experimental data this can result in cases where the best-fitting direct and indirect models will have similar likelihood values and the same (limited) prior information, that is, $\mathcal{L}_{\text{ind}}(x|\beta_{\text{ind}}, \gamma) \approx \mathcal{L}_{\text{dir}}(x|\beta_{\text{dir}}, \gamma)$ and $p_{\text{ind}}(\gamma) = p_{\text{dir}}(\gamma)$. As a consequence, the posterior probabilities for both models will be similar, that is, $P_{\text{ind}}(\beta_{\text{ind}}, \gamma|x) \approx P_{\text{dir}}(\beta_{\text{dir}}, \gamma|x)$. In such cases, this implies that one would have to know the transmission pathway in order to infer it from fitted ODE models. In total, our results suggest that in some cases it may be difficult to identify the transmission pathway when fitting models to time series data, even when prior information about life history parameters unrelated to transmission is available.

Differences in the Phase Lags of Oscillating Endemic Systems

Here we explore how the phase relations in endemic systems undergoing endogenous oscillations (as opposed to externally forced systems) can help identify the transmission pathway. In particular, we explore if the transmission pathway affects the phase differences. Oscillations in direct transmission models have been studied previously and arise as a consequence of nonlinear contact rates (Anderson and May 1978; Liu et al. 1986, 1987). Oscillations driven by host heterogeneity in indirect transmission systems have also been studied (e.g., Dwyer et al. 2000), though we do not address that case here.

Our work follows from Bulmer (1975) and is based on the analysis of small amplitude oscillations; see appendix C, available online, for details. We focus on this case because the phase relations of small-amplitude oscillations can be calculated analytically and approximate the phase

relations of larger oscillations. In this section, we assume that infectious individuals in the direct transmission system also excrete a noninfectious pathogen into the environment. This allows us to explore how the phase differences between the susceptible class and the environmental load depend on the transmission pathway. Mathematically this is done by appending the environmental dynamics equation dB/dt in system (2) to system (1). In addition to our analytical analysis, we also searched parameter space for numerical examples of our predicted phase relations. Our analytical and numerical results are summarized in tables C1 and C2, available online. All of the derivations and a discussion of our numerical simulations are contained in appendix C. Note that our results also apply to systems where the pathogen can reproduce in the environment (e.g., oscillations of *Vibrio cholerae*; Jensen et al. 2006).

First, consider systems (1) and (2) with a single infectious class. In the direct transmission model (1) the infectious population lags behind the susceptible population by less than a quarter of the period (fig. 5*A*, 5*B*). In contrast, in the indirect transmission model (2) that lag can be greater than a quarter of the period (fig. 5*C*, 5*D*). Thus, in models with a single infected class, phase differences between $S(t)$ and $I(t)$ help identify an indirect transmission pathway. In principle, the phase differences between the susceptible class and the environmental pathogen can also differ. The phase difference between $S(t)$ and $B(t)$ must be less than half of the period in direct transmission systems while phase differences greater than half of the period are possible in indirect transmission systems. However, we have not observed lags greater than half of the period in numerical simulations; see appendix C for a discussion. Our lack of numerical evidence suggests that only phase differences between $S(t)$ and $I(t)$ will likely yield evidence of an indirect pathway.

Now consider systems (A1) and (A3) from appendix A with an exposed class. Our theoretical analysis does not predict that any phase relations will differ between models. However, our numerical simulations suggest that differences do arise in the lag between $S(t)$ and $B(t)$. In simulations of indirect transmission models, the lag between $S(t)$ and $B(t)$ is only observed to be greater than half of the period in small regions of parameter space near Hopf bifurcation curves. In contrast, in direct transmission models, lags greater than half of the period are commonly observed; for example, see figure S2. When comparing systems with a single exposed and infectious class, our numerical simulations suggest that phase lags greater than half of the period between the susceptible population and the environmental load should be considered strong evidence for a direct transmission pathway.

Finally, consider systems (A2) and (A4) from appendix A with multiple infectious classes. Because the structure of the infectious classes may be unknown or it may be difficult to distinguish between individuals of different infectious classes, we focus on the dynamics of the total infectious population, $I_T(t) = I_1(t) + \ldots + I_n(t)$. Depending on the number of infectious classes, the phase lag between $S(t)$ and $I_T(t)$ can be of any size in both systems. Furthermore, any phase relation can exist between $I_T(t)$ and $B(t)$. For example, assume one infectious stage excretes significantly more pathogen than any other, that is, a supershedder class. If the last infectious class is the super-

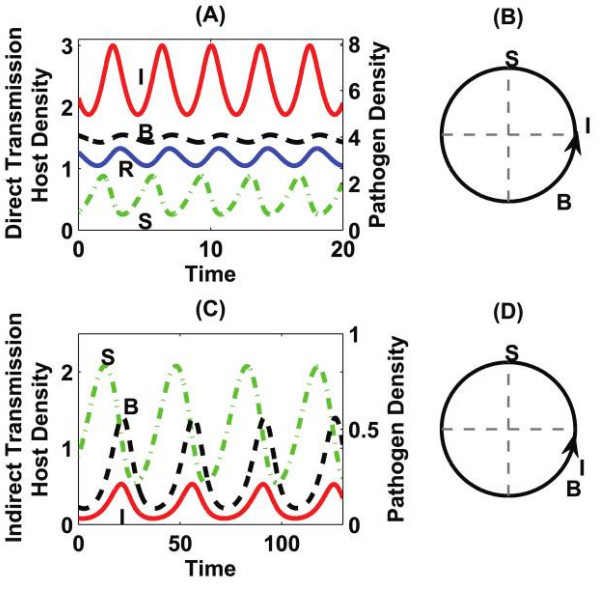

Figure 5: Differences in the phase relations of the direct transmission model (1) (*A*, *B*) and the indirect transmission model (2) (*C*, *D*). *A*, *C*, Susceptible (green dash-dot), infectious (solid red), recovered (solid blue), and environmental load (dashed black) time series. In *C*, individuals can recover, but there is no recovered class. *B*, *C*, Circle represents one period of the cycles, and the arrow denotes counterclockwise rotation. *S*, *I*, and *B* denote when the susceptible, infectious, and environmental load time series reach their peak during the cycle. The infectious class lags behind the susceptible class by less than one-quarter of the period in the direct transmission model. That lag can be up to half of the period in the indirect transmission model. Host density, environmental load, and time are dimensionless. *A*, *B*, Parameters and functions are $F(\cdot) = r_S S[1 - k_S(S + I + R)]$, $G(S, I, R) = \beta SI/(1 + S)$, $f(B) = 0$, $\beta = 3$, $m = 0.4$, $\mu = 0.1$, $\nu = 0.5$, $\rho = 0.6$, $r_S = 4.15$, $k_S = 0.025$, $\chi = 0.5$, and $\delta = 0.3$. Initial conditions are $S = 0.5$, $I = 1$, $R = 0.5$, and $B = 0$. *C*, *D*, Host equations are $dS/dt = F(S, I) - H(S, I, R, B) + \rho I - mS$ and $dI/dt = H(S, I, R, B) - (\rho + \mu + m)I$. Parameters and functions are $F(\cdot) = r_S S[1 - k_S(S + I + R)]$, $H(\cdot) = \alpha SB/(1 + S)$, $f(B) = 0$, $\alpha = 1.6$, $m = 0$, $\mu = 0.6$, $\rho = 0.3$, $r_S = 0.2$, $k_S = 0.25$, $\chi = 1.5$, and $\delta = 1.4$. The transmission forms are analogous to type 2 functional responses and account for the handling time of the pathogen; see also appendix B, available online. Initial conditions are $S = 1$, $I = 0.1$, and $B = 0.2$.

shedder class, then the environmental load can lag behind the total infectious population by any amount. Alternatively, if the first infectious class is the superspreader class, then the environmental load oscillations can potentially precede those of the total infectious population. Our analysis suggests that these relations can give important information on what stages of infection excrete the most pathogen but do not necessarily help identify the mode of transmission.

Discussion

In this study we explored differences between time series generated by a class of direct and environmentally mediated indirect transmission models. Our analysis shows that in some cases direct and indirect transmission models can produce numerically similar dynamics (figs. 3, 4), especially when the environmental dynamics of the system are fast (fig. S1). Our results also identify cases in which the models yield numerically different time series. In particular, when the environmental dynamics of the system are slow, direct transmission models do not accurately capture the infectious peak characteristics and endemic equilibrium values of indirect transmission time series (figs. 3*D*, 4*C*). Furthermore, quantitatively different phase relations can arise in oscillatory endemic systems (fig. 5). Here we discuss what implications these results have on identifying the transmission pathway from epidemiological time series data.

Our results suggest a few reasons why difficulties may arise when inferring the transmission pathway from fits of deterministic continuous-time models to epidemiological time series. First, our numerical simulations suggest that in some cases our direct and indirect transmission models can produce numerically similar epidemiological time series, especially when the environmental dynamics are fast. Depending on the level of observational noise in the data, this similarity in output may pose difficulties when trying to infer the transmission route from model fits. This identifiability issue has also been studied in systems with multiple transmission pathways where the environmental dynamics are fast (Tien and Earn 2010). Thus, depending on the transmission functional forms in the mixed transmission model, it may not be possible to identify the contribution each route has on disease transmission.

Second, our numerical simulations suggest that prior information about life-history parameters may not aid in identifying the transmission route. In our simulations, indirect transmission time series could be fit by direct transmission models by varying only parameters related to transmission. However, transmission, excretion, and pathogen mortality rates are often the most poorly characterized life-history parameters. Thus, prior estimates of life-history parameters unrelated to transmission may not necessarily aid in model selection. Third, differences in model fits to data do not necessarily imply that a pathogen utilizes a particular pathway. For example in the Miller et al. (2006) study, the authors observed that their candidate models had different fits to the data and concluded that the data supported an indirect transmission pathway. However, there must exist a direct transmission model that fits the data equally well because the inferred environmental dynamics of the best-fitting indirect transmission model are fast. Thus, it is important to be cautious about the interpretation of model fits based on numerical optimization routines in cases where identifiability issues can arise.

Our results do suggest that features of epidemic time series can help identify the transmission pathway when the environmental dynamics are not fast. In our numerical simulations of epidemic systems, we found that direct transmission models could not capture both the peak height and the peak width of the infectious class. Similarly, in our numerical simulations of endemic systems, we found that direct transmission models could not capture both the initial increase and the endemic level of infection. These numerical differences suggest that statistical comparisons of model fits to time series can potentially identify an indirect route of transmission. However, it should be noted that additional complications arise when the environmental dynamics are slow. In particular, epidemics tend to be small in size because excretion and pathogen mortality rates are small (fig. 3*D*) and numerical differences between endemic time series rely on capturing both the initial increase and the endemic level of infection (fig. 4*D*).

Phase relations (i.e., the lag between peaks) in oscillating endemic systems can also yield evidence supporting a particular route of transmission (fig. 5). These differences suggest that statistical comparisons of model fits could potentially identify the transmission pathway. It is important to note that our work does not address cycles driven by host heterogeneity, which has been shown to be important in insect systems (Dwyer et al. 2000). Our analysis also assumes a single transmission pathway and endogenous oscillations. In practice it can be difficult to identify if oscillations are driven by external factors and the study of phase relations in externally forced systems is an important future area of research.

Our results also demonstrate how environmental load data can aid in distinguishing between transmission pathways. At a minimum, environmental load data will yield better informed statistical comparisons between models. In addition, in some cases, direct and indirect models can only be distinguished using environmental load time series. For example, while the infectious class time series in

figure S1*B* and S1*C* are identical, the relative timing of the $I(t)$ and $B(t)$ peaks differ because the environmental dynamics are not fast in the direct transmission model. Further investigation of how environmental measurements assist model selection is warranted.

While our work indicates that identifiability issues can arise in epidemiological models, the generality of our results is limited by the class of functions we considered and suggests areas of future research. First, our class of models does not include systems where the environmental load decreases due to uptake by the host (e.g., Li et al. 2009). Nonetheless, we expect identifiability issues to arise when the environmental dynamics are fast. For example, in the fast environmental dynamics limit, an indirect transmission model with $H = \alpha SB$ and $dB/dt = \chi I - \delta B - H$ will be equivalent to a direct transmission model with a transmission rate of the form $G = \beta SI/(h + S)$; see appendix B. Second, our class of models does not include pathogen reproduction in the environment (e.g., Jensen et al. 2006; King et al. 2008; Joh et al. 2009). Based on our analysis of a particular model in appendix B, we do not expect indirect transmission models with pathogen reproduction to be equivalent to direct transmission models in most natural systems. However, it is unclear whether direct and indirect transmission models can produce quantitatively similar time series.

Third, there may exist quantitative differences beyond those we considered. Additional mechanistic information about epidemiological systems could help distinguish between direct and indirect transmission systems (e.g., bistability due to a minimal infectious dose of the pathogen, as in Joh et al. 2009). Fourth, the results derived about our deterministic models may not generalize to stochastic models. For example, stochastic models were used in the King et al. (2008) study to model cholera epidemics. In that study, via simulations of stochastic direct and indirect transmission models, the authors concluded that the seasonality of cholera epidemics was not driven by interactions between *Vibrio cholerae* and phage in the environment. We have shown that in some cases identifiability issues can arise in the underlying deterministic dynamics of stochastic models. However, statistical comparisons of time series generated from models with process noise may yield signatures of the transmission pathway. Understanding when stochasticity provides traction despite the equivalence of the underlying deterministic model is an important area of future research.

In total, our analysis suggests that the transmission pathway can be inferred from time series data under some biological conditions. Thus, the effects of the transmission pathway can be observed in epidemiological dynamics at the population level. However, our analysis also shows that there exists conditions under which the identify of the transmission pathway is masked at the population level. While additional work is needed in areas beyond our particular class of models, this work begins to identify how and when the transmission model can be inferred from time series data and how the transmission pathway utilized by a pathogen can influence the epidemiological dynamics of the system.

Acknowledgments

We thank J. Dushoff and J. Keen for discussions and suggestions about the manuscript. We thank two anonymous reviewers for helpful comments on the manuscript. This work was supported by the Defense Advanced Research Projects Agency under grant HR0011-09-1-0055 and a grant from the James S. McDonnell Foundation. Joshua S. Weitz holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund.

Literature Cited

Aeby, G. S., and D. L. Santavy. 2006. Factors affecting susceptibility of the coral *Montastraea faveolata* to black-band disease. Marine Ecology Progress Series 318:103–110.

Ajelli, M., M. Iannelli, P. Manfredi, and M. L. C. degli Atti. 2008. Basic mathematical models for the temporal dynamics of HAV in medium-endemicity Italian areas. Vaccine 26:1697–1707.

Anderson, R. M., and R. M. May. 1978. Regulation and stability of host-parasite population interactions. I. Regulatory processes. Journal of Animal Ecology 47:219–247.

Arnold, L., C. K. R. T. Jones, K. Mischaikow, and G. Raugel. 1995. Geometric singular perturbation theory. Pages 44–118 *in* R. Johnson, ed. Dynamical systems. Vol. 1609. Springer, Berlin.

Arthur, T. M., J. E. Keen, J. M. Bosilevac, D. M. Brichta-Harhay, N. Kalchayanand, S. D. Shackelford, T. L. Wheeler, X. Nou, and M. Koohmaraie. 2009. Longitudinal study of *Escherichia coli* O157:H7 in beef cattle feedlot and role of high-level shedders in hide contamination. Applied and Environmental Microbiology 75:6515–6523.

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

Breban, R., J. M. Drake, D. E. Stallknecht, and P. Rohani. 2009. The role of environmental transmission in recurrent avian influenza epidemics. PLoS Computational Biology 5:e1000346.

Brennan, M. L., R. Kemp, and R. M. Christley. 2008. Direct and indirect contacts between cattle farms in north-west England. Preventive Veterinary Medicine 84:242–260.

Briggs, C. J., and H. C. J. Godfray. 1995. The dynamics of insect-pathogen interactions in stage-structured populations. American Naturalist 145:855–887.

Brookhart, M. A., A. E. Hubbard, M. J. van der Laan, J. M. Colford Jr., and J. N. S. Eisenberg. 2002. Statistical estimation of parameters in a disease transmission model: analysis of a *Cryptosporidium* outbreak. Statistics in Medicine 21:3627–3638.

Brunner, J. L., D. M. Schock, and J. P. Collins. 2007. Transmission

dynamics of the amphibian ranavirus *Ambystoma tigrinum* virus. Diseases of Aquatic Organisms 77:87–95.

Bulmer, M. G. 1975. Phase relations in the ten-year cycle. Journal of Animal Ecology 44:609–621.

Cornick, N. A., and H. VuKhac. 2008. Indirect transmission of *Escherichia coli* O157:H7 occurs readily among swine but not among sheep. Applied and Environmental Microbiology 74:2488–2491.

Day, T. 2001. Parasite transmission modes and the evolution of virulence. Evolution 45:2389–2400.

Dhondt, A. A., S. Altizer, E. G. Cooch, A. K. Davis, A. Dobson, M J. L. Driscoll, B. K. Hartup, et al. 2005. Dynamics of a novel pathogen in an avian host *Mycoplasmal conjunctivitis* in house finches. Acta Tropica 94:77–93.

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

Dwyer, G., J. Dushoff, J. S. Elkinton, and S. A. Levin. 2000. Pathogen-driven outbreaks in forest defoliators revisited: building models from experimental data. American Naturalist 156:105–120.

Heesterbeek, J. A. P., and J. A. J. Metz. 1993. The saturating contact rate in marriage and epidemic models. Journal of Mathematica Biology 29:539–570.

Hochberg, M. E. 1991. Non-linear transmission rates and the dynamics of infectious diseases. Journal of Theoretical Biology 153: 301–321.

Hove-Musekwa, S. D., F. Nyabadza, C. Chiyaka, P. Das, A. Tripathi, and Z. Mukandavire. 2011. Modelling and analysis of the effects of malnutrition in the spread of cholera. Mathematical and Computer Modelling 53:1583–1595.

Jensen, M. A., S. M. Faruque, J. J. Mekalanos, and B. R. Levin. 2006 Modeling the role of bacteriophage in the control of cholera outbreaks. Proceedings of the National Academy of Sciences of the USA 103:4652–4657.

Joh, R. I., H. Wang, H. Weiss, and J. S. Weitz. 2009. Dynamics of indirectly transmitted infectious diseases with immunologica threshold. Bulletin of Mathematical Biology 71:845–862.

Kaper, J. B., J. G. Morris Jr., and M. M. Levine. 1995. Cholera. Clinical Microbiological Review 8:48–86.

King, A., E. L. Ionides, M. Pascual, and M. J. Bouma. 2008. Inapparent infections and cholera dynamics. Nature 454:877–880.

LeJeune, J. T., T. E. Besser, D. H. Rice, J. L. Berg, R. P. Stilborn, anc D. D. Hancock. 2004. Longitudinal study of fecal shedding of *Escherichia coli* O157:H7 in feedlot cattle: predominance and persistence of specific colonal types despite massive cattle population turnover. Applied and Environmental Microbiology 70:377–384.

Li, S., J. N. S. Eisenberg, I. H. Spicknall, and J. S. Koopman. 2009. Dynamics and control of infections transmitted from person to person through the environment. American Journal of Epidemiology 170:257–265.

Liu, W., H. W. Hethcote, and S. A. Levin. 1987. Dynamical behavior of epidemiological models with nonlinear incidence rates. Journal of Mathematical Biology 25:359–380.

Liu, W., S. A. Levin, and Y. Iwasa. 1986. Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. Journal of Mathematical Biology 23:187–204.

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

Miller, M. W., N. T. Hobbs, and S. J. Tavener. 2006. Dynamics of prion disease transmission in mule deer. Ecological Applications 16:2208–2214.

Roche, B., J. M. Drake, and P. Rohani. 2011. The curse of the pharaoh revisited: evolutionary bi-stability in environmentally transmitted pathogens. Ecology Letters 14:569–575.

Roche, B., C. Lebarbenchon, M. Gauthier-Clerc, C. Chang, F. Thomas, F. Renaud, S. van der Werf, and J. Guegan. 2009. Water-borne transmission drives avian influenza dynamics in wild birds: the case of the 2005–2006 epidemics in the Camargue area. Infection, Genetics and Evolution 9:800–805.

Sauvage, F., M. Langlais, N. G. Yoccoz, and D. Pontier. 2003. Modelling hantavirus in fluctuating populations of bank voles: the role of indirect transmission on virus persistence. Journal of Animal Ecology 72:1–13.

Siegel, J. D., E. Rhinehart, M. Jackson, L. Chiarello, and the Health Care Infection Control Practices Advisory Committee. 2007. 2007 guideline for isolation precautions: preventing transmission of infectious agents in health care settings. American Journal of Infection Control 35:S65–S164.

Spicknall, I. H., J. S. Koopman, M. Nicas, J. M. Pujol, S. Li, and J. N. S. Eisenberg. 2009. Informing optimal environmental influenza interventions: how the host, agent, and environment alter dominant routes of transmission. PLoS Computational Biology 6: e1000969.

Stilianakis, N. I., and Y. Drossinos. 2010. Dynamics of infectious disease transmission by inhalable respiratory droplets. Journal of the Royal Society Interface 7:1355–1366.

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

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

Wolfe, M. S. 1992. Giardiasis. Clinical Microbiology Reviews 5:93–100.

Associate Editor: Pejman Rohani
Editor: Judith L. Bronstein