



Cite this article: Dibble CJ, O'Dea EB, Park AW, Drake JM. 2016 Waiting time to infectious disease emergence. *J. R. Soc. Interface* **13**: 20160540.

<http://dx.doi.org/10.1098/rsif.2016.0540>

Received: 6 July 2016

Accepted: 22 September 2016

Subject Category:

Life Sciences – Mathematics interface

Subject Areas:

computational biology

Keywords:

bifurcation delay, critical slowing down, early-warning systems, emerging infectious disease

Author for correspondence:

Christopher J. Dibble


e-mail: dibble.christopherj@gmail.com

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.fig-share.c.3500364>.

Waiting time to infectious disease emergence

Christopher J. Dibble¹, Eamon B. O'Dea^{1,2}, Andrew W. Park^{1,2}
and John M. Drake^{1,2}

¹Odum School of Ecology, and ²Center for the Ecology of Infectious Diseases, University of Georgia, 140 East Green Street, Athens, GA 30602-2202 USA

 CJD, 0000-0002-0303-7330; EBO, 0000-0003-4748-683X; AWP, 0000-0003-4080-7274; JMD, 0000-0003-4646-1235

Emerging diseases must make a transition from stuttering chains of transmission to sustained chains of transmission, but this critical transition need not coincide with the system becoming supercritical. That is, the introduction of infection to a supercritical system results in a significant fraction of the population becoming infected only with a certain probability. Understanding the waiting time to the first major outbreak of an emerging disease is then more complicated than determining when the system becomes supercritical. We treat emergence as a dynamic bifurcation, and use the concept of bifurcation delay to understand the time to emergence after a system becomes supercritical. Specifically, we consider an SIR model with a time-varying transmission term and random infections originating from outside the population. We derive an analytic density function for the delay times and find it to be, in general, in agreement with stochastic simulations. We find the key parameters to be the rate of introduction of infection and the rate of change of the basic reproductive ratio. These findings aid our understanding of real emergence events, and can be incorporated into early-warning systems aimed at forecasting disease risk.

1. Introduction

The emergence and re-emergence of infectious diseases is an enormous concern to global public health. Nearly 70% of emerging human diseases have zoonotic origins [1–3], including SARS [4], MERS [5], AIDS [6,7] and Ebola [8]. Similarly, the re-emergence of previously controlled or eradicated diseases, including vaccine-preventable diseases such as measles [9] and pertussis [10], continues to be a health and economic burden [11]. Owing to factors including reductions in vaccine uptake [12], pathogen evolution [13] or changes in host demography [14], emerging and re-emerging diseases make a transition from self-limiting, stuttering chains of infection (which inevitably go extinct), to sustained chains of human-to-human transmission (major outbreaks or epidemics). This critical transition becomes possible when the average number of infections caused by a single infectious individual in an entirely susceptible population (called R_0) becomes greater than 1—at this point the population is often referred to as supercritical [15,16]. Forecasting this critical transition is the goal of early-warning systems [17–19] of disease emergence [12]. In addition to knowing when a population will become supercritical [12], accurate forecasting also requires estimating the time of the first major outbreak or epidemic of a newly emerging or re-emerging disease. Crucially, major outbreaks do not happen immediately after $R_0 = 1$. Rather, they occur with some non-zero delay—for which we do not yet have an adequate theory.

Waiting times between criticality and emergence imply that host populations may be at risk of a major outbreak for an extended period of time, despite experiencing only small clusters of cases (figure 1). In this way, a substantial waiting time hides the true state of a system, inhibiting our ability to intervene in response to a

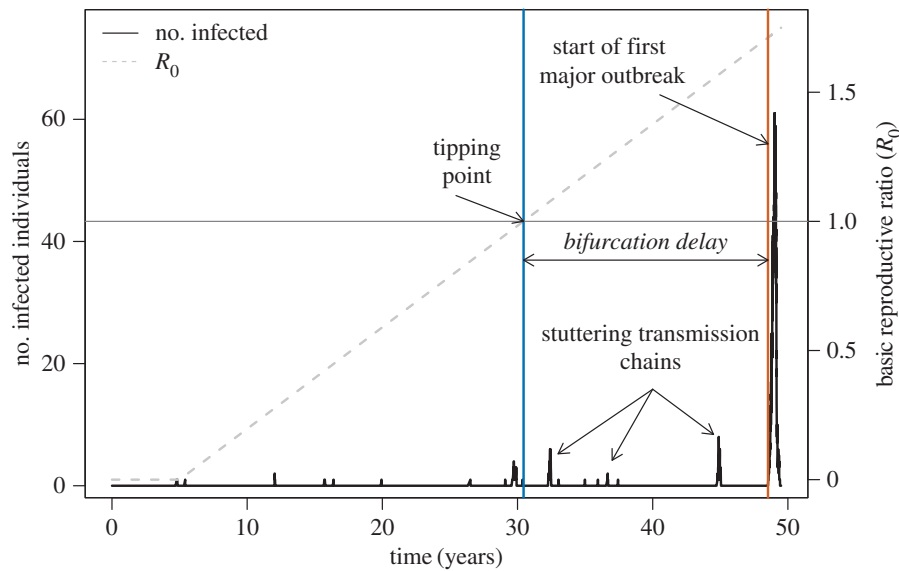


Figure 1. Bifurcation delay is the time between the tipping point at $R_0 = 1$ (intersection of grey horizontal line and blue vertical line) and the beginning of an actual disease epidemic (orange vertical line, definition in text). This simulated time series is a stochastic realization of the hinge SIR model, where the transmission rate (β) begins to increase at a set time (here, $t = 5$ years) at rate α . This example contains multiple ‘sparks’ after the transition to supercriticality that do not lead to epidemics, demonstrating that bifurcation delay is not simply the waiting time for a spark to occur after $R_0 > 1$. Other parameters for this simulation: rate of increase in transmission, $\alpha = 0.001 \text{ yr}^{-1}$, initial population size (N_0) = 1000, recovery rate, $\gamma = 365/14 \text{ yr}^{-1}$ (corresponding to a 14 day infectious period), spontaneous infection rate, $\xi = 0.00067 \text{ yr}^{-1}$, birth and death rate, $\mu = 1/60 \text{ yr}^{-1}$. (Online version in colour.)

large, but obscured, threat. Furthermore, long waiting times in disease re-emergence could mean that reversing the driver of a critical transition to its pre-outbreak levels might not push a pathogen’s R_0 below 1. For instance, if reductions in vaccine uptake facilitate disease outbreaks (e.g. measles [20–22]), increasing coverage to vaccination levels that existed immediately prior to an outbreak may not reduce R_0 below 1, because the transition to $R_0 > 1$ could have occurred well before the first major outbreak. An improved understanding of the factors determining the waiting time to a major disease outbreak will facilitate epidemic forecasting.

To better relate the transition to disease emergence and the timing of major outbreaks and epidemics, we developed a model incorporating a time-varying transmission term and the acquisition of infections from outside the population. For this model, we derive an approximate density function of waiting times. For validation, this approximation is compared with stochastic simulations across a wide range of notional life-history parameters. Together, these results show that waiting times can be described by a two-dimensional distribution encapsulating the rate of infections from outside the population and the rate of change of R_0 . We discuss implications of our results for early-warning systems of infectious disease emergence.

2. Material and methods

2.1 Main model and concepts

Let X denote the number of susceptible individuals in the population; Y , the number of infective individuals; Z , the number of recovered individuals; β , the transmission rate; γ , the individual recovery rate; μ , the individual birth and death rate (assumed here to be equal) and ξ , the rate at which infections are spontaneously acquired from outside the population. The transmission rate is 0 until time s , at which point it increases linearly at a small rate α . That is, for $t > s$, $\beta = \alpha t$. Our main model

Table 1. Events and rate laws in our stochastic model.

event	$(\Delta X, \Delta Y, \Delta Z)$	rate
birth	$(1, 0, 0)$	$\mu(X + Y + Z)$
death of S	$(-1, 0, 0)$	μX
death of I	$(0, -1, 0)$	μY
death of R	$(0, 0, -1)$	μZ
transmission	$(-1, 1, 0)$	βXY
recovery	$(0, -1, 1)$	γY
sparking	$(-1, 1, 0)$	ξX

is the continuous-time Markov chain with the time-dependent rates given in table 1.

For our model, we consider only the case that $(1 \gg \xi > 0)$. This case is of interest because in most disease emergence scenarios, there is a small influx of infectious individuals from outside the focal population—for example, due to zoonotic spillover events [23] or periodic importations of a disease from an endemic location [21]. Because of this influx, the expected number infected is necessarily greater than zero. However, the order of the expected number of cases changes qualitatively as the basic reproductive ratio passes through 1. The basic reproductive ratio is the expected number of new infections caused by an infected individual in an otherwise susceptible population, which for our model equals $R_0 = \beta(N - 1)/(\gamma + \mu)$, where $N = X + Y + Z$. For $R_0 < 1$, the expected number infected is of order ξN , whereas for $R_0 > 1$ the expected number is of order N . In other words, the expected number infected is approximately $1/\xi$ times larger when $R_0 > 1$ compared with when $R_0 < 1$. The transition from one order of magnitude to another may be referred to as an imperfect bifurcation, and the delay between this event and the time when $R_0 = 1$ is known as a *bifurcation delay*. Our goal is to find the model parameters that largely control this delay. As we shall see, they turn out to be

the rate of infection from outside of the population, which we call the *sparkling rate*, and the rate of change of R_0 , which we call the *sweep rate*.

The results of this work are based on stochastic models for two reasons. First, an analysis of deterministic bifurcation delay in a model similar to a deterministic SIR model is already available [24]. Second, the number of infectious individuals prior to disease emergence may be small, and thus it may not be appropriate to neglect stochastic effects. We study our stochastic model and extensions of it using both simulation and analysis.

2.2. Simulation

Our simulation model extends our main model by including additional sources of demographic stochasticity. We make three modelling decisions to allow demographic stochasticity to have a substantial effect on delay times. First, the rate of birth in our model is always equal to the rate of death. Thus, the population size has no positive steady state, and the variance of the population size across multiple realizations of the model with the same initial population size increases over time. Second, the timing (s) of the increase in the transmission rate was set to year 100 to allow the variance to grow before the approach to the critical point even began. Third, the transmission rate is density dependent so that R_0 varies with fluctuations in population size (table 3). Although frequency-dependent transmission is the better supported model for many diseases of interest [25,26], the direct effect of demographic stochasticity on density-dependent transmission makes it more likely to cause behaviour different from our analytic model. As the purpose of the simulation was to investigate such behaviour, and essentially 'stress test' the analytic results, we present results based on density-dependent transmission in the main text. A comparison of these results with those of a frequency-dependent model is included in the electronic supplementary material.

To calculate bifurcation delays in our simulation models, we first simulated trajectories of the number infected using Gillespie's direct method [27]. These simulations are not exact because the algorithm only updates the time-dependent rates at the time when state-variables are updated, but for our parameters updates should happen frequently enough that the simulations remain highly accurate. We parametrized models based roughly on the human lifespan, with a range of life-history parameters that correspond to several disease systems of emergence concern (table 2). To explore the range of possible solutions to the stochastic model, we ran 1000 stochastic simulations for each combination of parameter values. Each simulation ended at the first major outbreak, which occurred when three criteria were satisfied: (A) the size of the susceptible population was reduced below N/R_0 (after $\sqrt{N}(R_0 - 1) > 1$, as it is difficult to distinguish epidemics from small outbreaks when R_0 is only just above 1), (B) at least 15% of the population was in the recovered class, and (C) at least 2.5% of the population was infected. These criteria ensured that the disease had surpassed the long-term, deterministic endemic equilibrium number of susceptible individuals (N/R_0) (i.e. a major outbreak occurred, A), that a substantial portion of the population had been exposed to the disease (B) and that the population was experiencing an active outbreak (C). We quantified bifurcation delay as the time between the deterministic tipping point of the system (calculated as the expected time when $R_0 = 1$, based on the initial population size N_0 and time-varying transmission rate β , table 3) and the last time point with 0 infectious individuals before the actual bifurcation (e.g. figure 1). To determine whether longer delays make outbreaks larger when they eventually occur, we also recorded the peak epidemic prevalence in each simulation.

Table 2. Parameters used for stochastic SIR simulations.

parameter	meaning	set or range of values
μ	host lifespan ⁻¹ , birth rate, death rate	1/60 yr ⁻¹
γ	recovery rate	[365/5 yr ⁻¹ , 365/41 yr ⁻¹]
β	transmission rate	if $t < s$, 0; otherwise $\alpha(t - s)$ individual ⁻¹ yr ⁻¹
N_0	initial population size	[100 individuals, 10 000 individuals]
ξ	spontaneous infection rate	{1/1500 yr ⁻¹ , 1/5000 yr ⁻¹ , 1/10 000 yr ⁻¹ }
α	rate of increase in transmission rate	[0.0005 yr ⁻¹ , 0.01 yr ⁻¹]

Table 3. Key calculations for hinge SIR model.

value	calculation
expected time for system to become supercritical	$t_c = (\mu + \gamma)/(N_0\alpha) + s$
time of bifurcation	$t_{\text{bif}} =$ last time before epidemic when no infectives present
bifurcation delay	$t_{\text{bif}} - t_c$
basic reproductive ratio	$R_0(t) = \beta(t)N(t)/(\gamma + \mu)$
maximum infection prevalence	$\max(Y(t)/N(t))$

2.3. Analysis

Our analytic model explores the effect on delay of the degree of heterogeneity in secondary cases. The degree of heterogeneity of secondary cases summarizes the consequences of heterogeneities in susceptibility, infectivity, infectious periods and contact rates. For the sake of tractability, our analytic model does not allow for demographic stochasticity in the same manner as our simulations. Our main approach is to suppose that each imported infection initiates a chain of infections which may be modelled as a branching process. The delay time is considered to be the time before the initiation of a branching process that does not go extinct. This time is a random variable that may be modelled like the time to death in a survival analysis. In the following, we derive a survival function for this time that is the basis for equations for the mean and median times given in Results.

A branching process model can be used to calculate a probability of a major outbreak that accounts for the level of heterogeneity of a given disease system. The branching process we use models the number of cases in each generation as the sum of the random number of secondary cases that are caused by each case in the previous generation. Following previous authors [28], we suppose that the secondary cases are drawn from a negative binomial distribution. The theory of branching processes then tells us that when the expected number of secondary cases is greater than 1, the probability of extinction is the smallest root of the equation

$$s = \left[1 + \frac{R_0(1-s)}{k}\right]^{-k}, \quad (2.1)$$

where R_0 is the mean and k the overdispersion parameter of the negative binomial distribution. The parameter k allows us to tune the level of heterogeneity in the distribution of secondary cases. As $k \rightarrow \infty$, the negative binomial distribution approaches a Poisson distribution, and as $k \rightarrow 0$ the distribution becomes more overdispersed. For the case $k = 1$, we recover the equation that results from the assumption that cases occur at constant rate over the duration of an exponentially distributed infectious period [29]. Although equation (2.1) does not, in general, provide a closed-form equation for the extinction probability, numerically finding the smallest root of equation (2.1) is trivial. The probability of a major outbreak follows immediately as the complement of the extinction probability.

To provide a more readily understandable equation for the extinction probability, we introduce the approximation $\Pr(\text{extinction}) \approx 1/R_0^p$, where p is chosen based on k . Electronic supplementary material, figure S1 shows the accuracy of this approximation for range of k and p . We can see that for $k < 1$, $p \approx k$ often has a maximum error of less than 0.1. As expected for k and p both equal to one, the calculated error is small and reflects rounding error. For $k > 4$, $p \approx 2$ typically leads to a maximum error of less than 0.05. This approximation facilitates the following analysis.

Given a model for the parameters of the branching process, many equations describing the distribution of delay times follow readily from standard survival analysis. We first consider the model in which $R_0 = 1$ at time zero and then increases linearly. For this model, the hazard function λ for a major outbreak is given by

$$\lambda(t) = a[1 - (1 + bt)^{-p}], \quad (2.2)$$

where a is the rate at which infections occur due to contact with other populations and b is the rate of change of R_0 with respect to time. The parameters a and b are the sparking and sweep rates mentioned earlier. Our hazard function employs the previous approximation by letting the probability of a major outbreak, given a spark, be equal to $1 - R_0(t)^{-p}$. The cumulative hazard is given by

$$\int_0^T \lambda(t) dt = \begin{cases} aT - \frac{a(1 + bT)^{1-p} - 1}{1-p}, & p \neq 1 \\ aT - \frac{a}{b} \log(1 + bT), & p = 1. \end{cases} \quad (2.3)$$

The survival function S computes the probability that a major outbreak has not happened before a given time, and it follows

$$S(t) = \exp\left(-\int_0^t \lambda(u) du\right). \quad (2.4)$$

The probability density function for the delay times is the function $-dS/dt$. The mean delay time $E(T)$ is given by $E(T) = \int_0^\infty S(t) dt$. The quantile function, which provides that time at which the probability of emergence having occurred equals F , may be found by solving $1 - F = S(q)$ for q . The median delay time is equal to q when $F = 1/2$.

2.4. Comparison of analytic and simulated delay distributions

To assess the agreement between analytic and simulation results, we calculated the Kullback–Leibler (KL) divergence [30] between analytic and simulated distributions of bifurcation delay for all parameter combinations ($N = 7200$, with 1000 simulations generating each simulated distribution). KL divergence quantifies the amount of information lost (i.e. entropy) when approximating one distribution with another [30,31]. The KL divergence of the distribution with density p from that with density q is defined as $D_{\text{KL}} = \int_{-\infty}^{\infty} p(x) \log_2(p(x)/q(x)) dx$. We calculated KL divergence using the entropy R package [32]. For the results presented, we used the analytic density function

for p and the density of simulated delays for q . Our results are quantitatively very similar using the reverse formulation.

The KL divergence values are measured in bits. Higher values indicate more information lost via the approximation, and thus more dissimilarity between analytic and simulated distributions. In our case, it quantifies the extent to which difference in the simulation and analytic models lead to different distributions of delays.

The KL divergence of course depends on the parameters of the models. We used recursive partitioning to create a decision tree that split KL-divergence scores according to a least-squares criterion, iteratively minimizing the remaining variance below each node of the decision tree [33]. Thus, recursive partitioning allowed us to identify and visualize the parameter combinations that generated the largest discrepancies between analytic and simulated delays.

3. Results

For our analytic model, we obtained explicit equations for the mean and median delays for various levels of heterogeneity as well as results for the sensitivity of the delays for all levels of heterogeneity. Recall that the two parameters of our derived density function for delays are the sparking rate a , which is the rate at which infections are introduced from outside of the population, and the sweep rate b , which is the rate of change of R_0 . The parameter p indicates the degree of heterogeneity in secondary cases. For the case $p = 1$, which corresponds to the heterogeneity caused by the assumption of exponentially distributed infectious period in our simulation model, the mean follows

$$E(t) = e^{a/b} \left(\frac{a}{b}\right)^{-a/b} a^{-1} \Gamma\left(\frac{a}{b} + 1, \frac{a}{b}\right), \quad (3.1)$$

where Γ denotes the upper incomplete gamma function and the median is equal to

$$\frac{-W(-1/(2^{b/a}e)) - 1}{b}, \quad (3.2)$$

where W denotes the non-principal branch of the Lambert W function. The median is equal to

$$\frac{b \log 2 + \sqrt{ab \log 2}}{ab}, \quad (3.3)$$

when $p = 0.5$ and

$$\frac{b \log 2 + \sqrt{(b \log 2)^2 + 4 ab \log 2}}{2ab}, \quad (3.4)$$

when $p = 2$. Recall that $p = 2$ corresponds to a Poisson distribution of secondary cases, whereas $p = 0.5$ corresponds to a highly overdispersed distribution of secondary cases. In the electronic supplementary material, we show that for all p the sensitivity of the mean and median to changes in either the sparking rate or the sweep rate depends only on the quotient of those two parameters. For relatively high sparking rates, the mean and median have a similar level of sensitivity to both parameters. As the sparking rate becomes relatively lower, the sweep rate becomes less important because for a broad range of rates the probability of an epidemic will be close to one by the time the first spark occurs. The level of heterogeneity in the distribution of secondary cases determines how quickly this transition happens by way of its effect on the relationship between R_0 and the probability of an epidemic. Overall, the analytic results provide a rather simple picture of the waiting times. However, in the

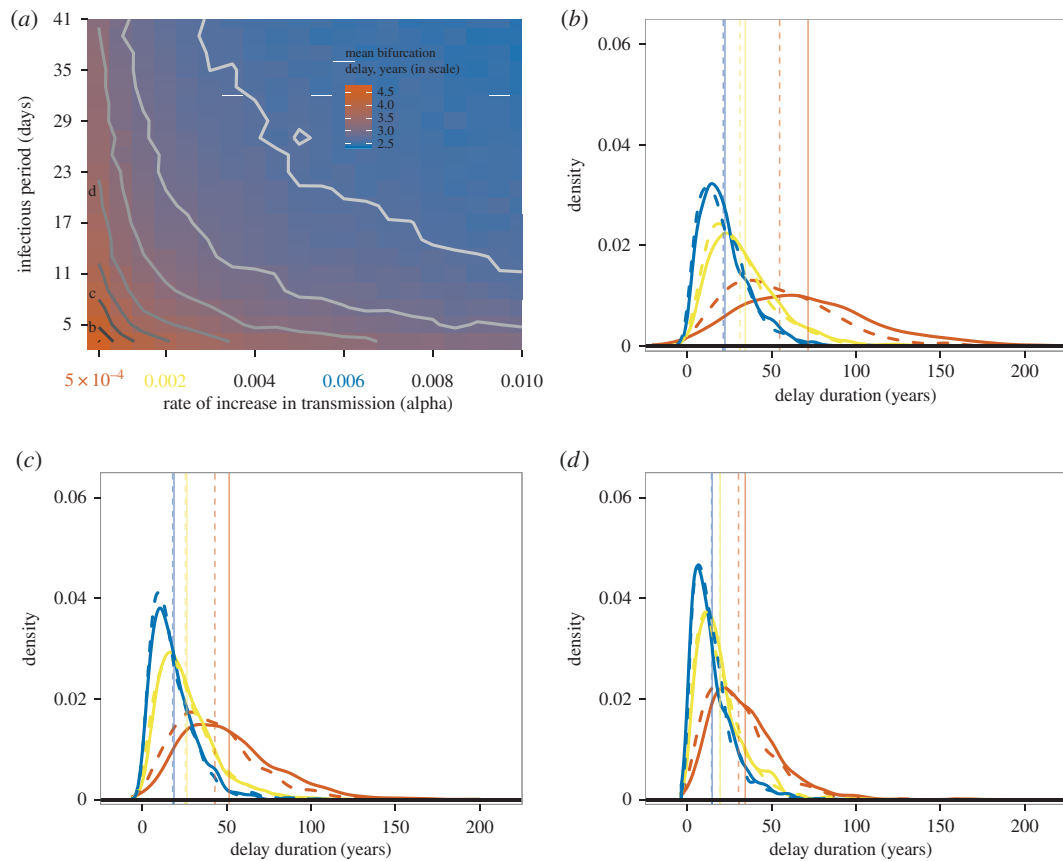


Figure 2. (a) Mean bifurcation delay (in years) decreases with increasing rate of change in the transmission rate (α) and increasing mean infectious period (i.e. inverse recovery rate, γ), as indicated by the colour of the heatmap and the contour lines (darker = more delay). Letters b–d indicate three mean different infectious periods detailed in subsequent panels. (b–d) The distribution of bifurcation delay from 1000 stochastic simulations (solid lines) and expected distributions using the analytic results (dashed lines) at three rates of change in the transmission rate ($\alpha = 0.0005 \text{ yr}^{-1}$, red; $\alpha = 0.002 \text{ yr}^{-1}$, yellow and $\alpha = 0.006 \text{ yr}^{-1}$, blue). Each panel represents a different mean infectious period ($b = 5$ days, $c = 9$ days, $d = 21$ days). Vertical lines show the mean delay for the each value of α , for both the observed (solid line) and expected (dashed line) delay distributions. Other life-history parameters for these examples: $N_0 = 1000$ and $\xi = 0.0001 \text{ yr}^{-1}$. (Online version in colour.)

electronic supplementary material, we also show that the picture can be rather different if R_0 changes according to a step function instead of increasing linearly. For example, the median delay has the lowest proportional sensitivity to the sparking rate when the median delay just exceeds the time of the jump in R_0 . In the following, we explore the extent to which the picture changes due to the elements included in our more complicated simulation models.

With data from stochastic simulations, we found mean bifurcation delay to decrease with increasing values of (i) sparking rate, (ii) population size, (iii) rate of increase in transmission, and (iv) infectious period (figure 2). These qualitative effects were not surprising, although the results highlight the importance of particular combinations of life-history traits. For instance, at a given value for the transmission rate's rate of change (α), different mean infectious periods ($1/\gamma$) can lead to very different predicted and observed distributions of bifurcation delay (figure 2b–d) because short infectious periods (i.e. high recovery rates) lead to lower sweep rates for a given value of α . The sparking rate and the sweep rate parameters of the analytic model are thus also important parameters of the simulated delay times. This parametrization also leads to a strong correlation of the delays of the frequency- and density-dependent models, although the delays of the frequency-dependent model are typically a little smaller (electronic supplementary material,

figure S2). Note that the sweep rate is independent of population size in the frequency-dependent model, so the relationship between life-history parameters and delay is highly dependent on the type of transmission.

Closer examination of the sensitivity of the median simulated delay to changes in the sweep rate revealed both similarities and differences with our analytic model. This sensitivity may be quantified as the slope of a least-squares line in the plot of $\log(\text{median delay})$ versus $\log(\text{sweep rate})$. Such a slope is an estimate of the elasticity of the median with respect to the sweep rate. Consistent with the analytic results, increasing the sparking rate a increases the slope of the negative relationship between median simulation delays and the sweep rate (figure 3 and table 4). However, for large a the elasticity for the simulated delays is much more negative than -0.5 , which is the minimum possible value for our analytic model (electronic supplementary material). This discrepancy is due to simulations tending to produce more delay than expected, which is in large part a consequence of the stringent conditions for an epidemic in our simulations.

In general, predictions based on theory derived assuming large, fixed population sizes are representative of stochastic simulations (figure 2; electronic supplementary material, figure S3). Recursive partitioning shows that approximately 78% of the variation in KL divergence can be accounted for by the sparking (a) and sweep (b) rates. Differences in

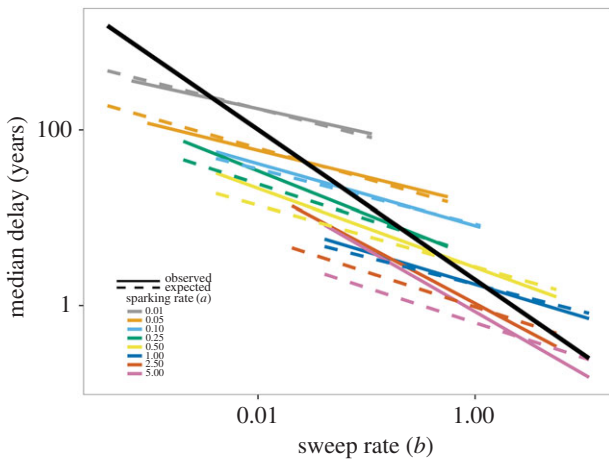


Figure 3. The negative relationship between bifurcation delay and the sweep rate of a system depends on the rate of sparking. The solid black line shows the overall relationship from a linear model of $\log(\text{observed delay})$ by $\log(\text{sweep rate})$. Solid colour lines represent the observed median bifurcation delay from model simulations, and dashed colour lines represent the expected median bifurcation delay based on analytic results. Colours show the effect of different levels of the sparking rate, a , on the relationship between delay and the sweep rate. Slopes and intercepts can be found in table 4. (Online version in colour.)

predicted and observed delay distributions were greatest with slow sweep rates, fast recovery rates, high sparking rates and small population sizes (electronic supplementary material, figure S3).

For a given combination of host and pathogen life-history parameters, longer delays give rise to larger outbreaks, because R_0 continues to increase ever-higher above the critical value of 1. Thus, a measles outbreak that occurs, by chance, 1 year after the population becomes supercritical will be larger, on average, than a measles outbreak that occurs six months after the population becomes supercritical (assuming the same underlying parameters). The slope of this relationship between delay and prevalence *within* a life-history parametrization depends on how quickly R_0 changes relative to the sparking rate (i.e. b/a , figure 4). Note that the relationship between delay and prevalence does not necessarily hold *across* parameter combinations (i.e. when comparing different diseases; figure 4). Systems with shorter delays can have higher peak prevalence if they exhibit a longer infectious period, for instance, and thus a faster change in R_0 over time (which itself contributes to differences in mean delay).

4. Discussion

Forecasting outbreaks of newly emerging and re-emerging diseases is difficult. Adding to this difficulty is the fact that major outbreaks of emerging diseases need not occur immediately after they become a possibility. Rather, they can exhibit bifurcation delay, the duration of which has until now remained largely undescribed. We have shown that even for models that include substantial demographic stochasticity, this delay is well described by a simple survival function with two parameters. One is the rate of transmission from outside of the population, which we call the sparking rate, and the other is the rate of change of R_0 , which we call the sweep rate. Further, the sensitivity of the delay to changes in these parameters depends only on their quotient.

We next explain the intuition that indicates the importance of these parameters.

We found that the greater the sparking rate, the smaller the waiting time for a major disease outbreak (figure 3). This result follows Bartlett [34], who showed that very low sparking rates could generate long time lags between recurrent outbreaks. In dynamic bifurcations more generally, the addition of stochasticity or noise to slow passages through bifurcation points reduces bifurcation delay [35,36] by hastening the departure from an unstable equilibrium. While our results are in agreement, we also note that ‘noise’ in our system is strictly required for the critical transition to disease emergence. Without any infections from outside the population, a supercritical host population will never experience an outbreak or epidemic.

We also found that bifurcation delay depends on the sweep rate. This result is comparable with the phenomenon in Bartlett’s model [34] where the time to a recurrent outbreak depends on both the rate of influx of susceptibles and R_0 . With slow sweep rates, the addition of an infectious individual to a population in the interval following the time when $R_0 = 1$ h as a low probability of leading to a major outbreak [37]. Naturally, the expected time to the major outbreak increases in this case, and bifurcation delay in disease emergence is not simply the interval of time preceding a sparking event.

Our theoretical and simulation results captured the expected negative relationship between the duration of bifurcation delay and the sweep rate of a system [36,38]. However, we also found that the slope and intercept describing the relationship between the duration of bifurcation delay and the sweep rate depended on the sparking rate. The duration of bifurcation delay depends more on the sparking rate when the sparking rate is low (i.e. outbreaks are spark-limited), and more on the sweep rate when the sparking rate is high (i.e. driver-limited) (figure 3, see also electronic supplementary material). The degree of spark- or driver-limitation of a system reflects the size of disease outbreaks (figure 4). It may also determine the ability of early-warning systems based on critical slowing down to detect the tipping point at $R_0 = 1$ because a driver-limited system would likely experience a series of minor outbreaks before a major outbreak. The near-critical dynamics of the minor outbreaks could trigger an early warning.

There are a number of practical consequences of bifurcation delay in disease emergence. First, delayed emergence means host populations can be supercritical for extended periods of time without experiencing a major outbreak (e.g. figure 1). Stuttering chains of transmission do not necessarily indicate, then, that $R_0 < 1$. Without accounting for bifurcation delay, we might underestimate the likelihood of a major epidemic when transmission appears limited. This phenomenon is exhibited in many of our simulations, and is likely to occur in generic emergence time series with small clusters of cases followed by large, unexpected outbreaks. A major difficulty for disease risk assessment lies in determining whether a system with apparent subcritical transmission is truly subcritical, or actually supercritical and experiencing a delayed bifurcation.

Additionally, bifurcation delay in disease re-emergence can obscure the timing of the system becoming supercritical, and thus the level of a given driver that targeted control measures might try to achieve to ‘roll-back’ disease risk. For instance, slow reductions in vaccine uptake that contribute to a major outbreak of a vaccine-preventable disease are also subject to bifurcation delay [12]. Owing to the delay, control measures

Table 4. Scaling relationships between bifurcation delay and the sweep rate of the system. The left-most column is the linear slope of the relationship between $\log(\text{delay})$ and $\log(\text{sweep rate}, b)$ for all simulations across all levels of the sparking rate a (e.g. solid black line in figure 3). All other columns show the slope of $\log(\text{delay})$ by $\log(\text{sweep rate})$ for each level of the sparking rate, a (e.g. colour lines in figure 3). Observed values represent results from individual stochastic simulations, and expected values for the mean and median delay arise from our analysis based on a branching process model. We fitted general linear models to each to obtain estimates of explained variance, R^2 .

observed delay		observed median delay			expected median delay			observed mean delay			expected mean delay		
intercept	sparking rate (a)	intercept	slope ($\log(b)$)	slope s.e.	intercept	slope ($\log(b)$)	slope s.e.	intercept	slope ($\log(b)$)	slope s.e.	intercept	slope ($\log(b)$)	slope s.e.
0.54	0.010	3.847	-0.289	0.008	3.743	-0.306	0.002	3.600	-0.331	0.021	4.124	-0.262	0.002
slope	0.025	3.055	-0.336	0.011	3.095	-0.310	0.003	2.893	-0.352	0.029	3.433	-0.267	0.003
-0.852	0.050	2.652	-0.311	0.009	2.528	-0.345	0.002	2.318	-0.404	0.024	2.825	-0.307	0.003
slope s.e.	0.067	2.847	-0.203	0.011	2.103	-0.408	0.003	1.790	-0.479	0.031	2.350	-0.380	0.003
0.003	0.100	2.096	-0.352	0.011	2.135	-0.313	0.003	1.953	-0.377	0.029	2.413	-0.270	0.003
	0.125	1.837	-0.405	0.011	1.764	-0.399	0.003	1.698	-0.417	0.030	2.001	-0.368	0.003
	0.167	1.684	-0.401	0.011	1.553	-0.410	0.003	1.410	-0.454	0.030	1.773	-0.383	0.003
	0.250	1.258	-0.493	0.011	1.345	-0.400	0.003	1.178	-0.489	0.030	1.559	-0.370	0.003
	0.330	1.060	-0.520	0.011	1.142	-0.412	0.003	0.944	-0.522	0.029	1.345	-0.383	0.003
	0.500	1.002	-0.450	0.009	1.020	-0.349	0.002	0.884	-0.477	0.024	1.228	-0.312	0.003
	0.670	0.596	-0.613	0.011	0.734	-0.412	0.003	0.547	-0.603	0.029	0.917	-0.383	0.003
	1.000	0.557	-0.373	0.011	0.553	-0.313	0.003	0.478	-0.384	0.029	0.733	-0.270	0.003
	2.500	0.057	-0.658	0.011	-0.037	-0.400	0.003	0.003	-0.714	0.029	0.109	-0.370	0.003
	3.330	-0.043	-0.749	0.011	-0.213	-0.412	0.003	-0.113	-0.813	0.029	-0.076	-0.383	0.003
	5.000	-0.162	-0.713	0.011	-0.453	-0.400	0.003	-0.158	-0.736	0.029	-0.327	-0.370	0.003
	6.670	-0.169	-0.824	0.011	-0.620	-0.412	0.003	-0.193	-0.814	0.030	-0.504	-0.383	0.003
R^2 (adj)		0.6624	0.9921		0.9995			0.7776			0.9993		

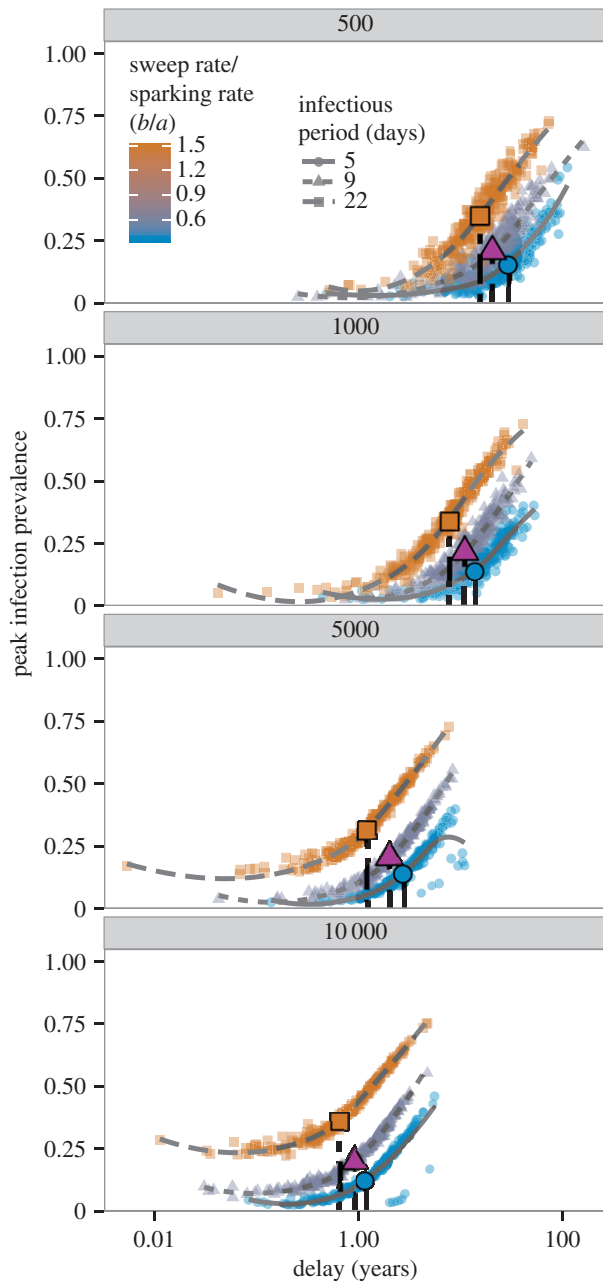


Figure 4. Bifurcation delay can lead to epidemics with greater peak prevalence, although this effect depends on the life-history parameters of the system. Panels show peak epidemic prevalence by stochastic realizations of bifurcation delay (x -axis on log-scale). Symbols and line types represent three different infectious periods, approximating three important childhood diseases (5 days for measles, 9 days for chickenpox, 22 days for pertussis). Lines are loess smooths added for visual clarity. Colour represents the ratio of b to a , which describes how quickly R_0 changes relative to how frequently imported infections occur in a system. Symbols are positioned at the median of the x - and y -axes for each combination of life-history parameters. Panels show four different initial population sizes ($N_0 = 500, 1000, 5000$ and $10\,000$). The relationship between delay and peak infection prevalence depends on how quickly R_0 increases, which is governed by the sweep rate, mean infectious period and population size. (Online version in colour.)

that increase vaccination coverage to what it was immediately before an outbreak may not reduce R_0 below 1 because R_0 may have exceeded 1 substantially earlier than the beginning of the outbreak. Further, separation in the timing of disease (re)emergence and changes in factors permitting emergence can compromise detection of statistical association between epidemic occurrence and underlying causes.

While the effects of sparking (a) and the probability of an epidemic given a spark (which is governed by the sweep rate, b) on the duration of bifurcation delay are relatively intuitive, it is worth reiterating the importance of the terms that go into each. For instance, the absolute number of infectious imports for a given system may depend on population size (as in pre-vaccine era measles [39]), which would have to be accounted for in any estimate of the sparking rate. On the other hand, unless the populations being modelled are very small (e.g. a school or village), frequency-dependent transmission is probably a better assumption for many diseases [25,26] and so the sweep rate may be largely independent of population size. Nevertheless, the same rate of change in the transmission rate for two different diseases could still be consistent with different sweep rates and thus different expectations of delay, and even different sizes of eventual disease outbreaks. For example, differences in the infectious period of childhood diseases such as measles and pertussis could mean that a constant 1% reduction in vaccine uptake per year would result in earlier, larger pertussis outbreaks, because a longer infectious period (all else being equal) generally increases R_0 (figure 4).

5. Conclusion

Moving forward, efforts relying on the identification of critical slowing down to forecast the emergence of infectious diseases may incorporate our findings about bifurcation delay [12]. The rate of critical slowing down and the sweep rate of a system are closely related, and indicators of critical slowing down may be informative of the sweep rate. With this information and an estimate of the sparking rate from observed case data, future early-warning systems may be able to not only generate forecasts of when R_0 exceeds 1, but also produce distributions of likely waiting times to major outbreaks or epidemics (e.g. figure 2*b–d*). This approach could also be used in hindsight, in order to better assess the timing of when R_0 exceeded 1 to set the stage for previous epidemics. Our goal is that this combined framework will produce model-independent forecasts that provide actionable information about the state of an emerging disease, particularly during the transition from subcritical dynamics to sustained transmission among hosts.

Data accessibility. Simulation data have been archived at the Dryad Digital Repository doi:10.5061/dryad.dp4kb [40].

Authors' contributions. J.M.D., C.J.D. and A.W.P. conceived of the study; J.M.D., C.J.D. and E.B.O.D. constructed stochastic models and E.B.O.D. derived the analytic results. C.J.D. wrote the initial drafts of the manuscript, and all authors contributed greatly to revisions. All authors approve the publication of this manuscript in its current form.

Competing interests. We declare we have no competing interests.

Funding. This research was funded by the National Institutes of Health grant no. U01GM110744.

Acknowledgements. We wish to thank Toby Brett and three anonymous reviewers for comments and suggestions that improved the clarity and analysis of the manuscript, and Bogdan Epureanu and Shiyang Chen for helpful discussion.

Disclaimer. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

References

- Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008 Global trends in emerging infectious diseases. *Nature* **451**, 990–993. (doi:10.1038/nature06536)
- Wolfe ND, Dunavan CP, Diamond J. 2007 Origins of major human infectious diseases. *Nature* **447**, 279–283. (doi:10.1038/nature05775)
- Lloyd-Smith JO, George D, Pepin KM, Pitzer VE, Pulliam JRC, Dobson AP, Hudson PJ, Grenfell BT. 2009 Epidemic dynamics at the human-animal interface. *Science* **326**, 1362–1367. (doi:10.1126/science.1177345)
- Anderson RM, Fraser C, Ghani AC, Donnelly CA, Riley S, Ferguson NM, Leung GM, Lam TH, Hedley AJ. 2004 Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic. *Phil. Trans. R. Soc. Lond. B* **359**, 1091–1105. (doi:10.1098/rstb.2004.1490)
- Khalafalla AI, Lu X, Al-Mubarak AIA, Dalab AHS, Al-Busadah KAS, Erdman DD. 2015 MERS-CoV in upper respiratory tract and lungs of dromedary camels, Saudi Arabia, 2013–2014. *Emerg. Infect. Dis.* **21**, 1153–1158. (doi:10.3201/eid2107.150070)
- Gao F *et al.* 1992 Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. *Nature* **358**, 495–499. (doi:10.1038/358495a0)
- Hahn BH. 2000 AIDS as a zoonosis: scientific and public health implications. *Science* **287**, 607–614. (doi:10.1126/science.287.5453.607)
- Morens DM, Folkers GK, Fauci AS. 2004 The challenge of emerging and re-emerging infectious diseases. *Nature* **430**, 242–249. (doi:10.1038/nature02759)
- King A, Varughese P, Serres GD, Tipples GA, Waters J. 2004 Measles elimination in Canada. *J. Infect. Dis.* **189**, S236–S242. (doi:10.1086/378499)
- Rohani P, Drake JM. 2011 The decline and resurgence of pertussis in the US. *Epidemics* **3**, 183–188. (doi:10.1016/j.epidem.2011.10.001)
- Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. 2006 Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747–1757. (doi:10.1016/S0140-6736(06)68770-9)
- O'Regan SM, Drake JM. 2013 Theory of early warning signals of disease emergence and leading indicators of elimination. *Theor. Ecol.* **6**, 333–357. (doi:10.1007/s12080-013-0185-5)
- Park AW, Daly JM, Lewis NS, Smith DJ, Wood JLN, Grenfell BT. 2009 Quantifying the impact of immune escape on transmission dynamics of influenza. *Science* **326**, 726–728. (doi:10.1126/science.1175980)
- Morse SS. 1995 Factors in the emergence of infectious diseases. *Emerg. Infect. Dis.* **1**, 7–15. (doi:10.3201/eid0101.950102)
- Blumberg S, Lloyd-Smith JO. 2013 Inference of RO and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Comput. Biol.* **9**, e1002993. (doi:10.1371/journal.pcbi.1002993)
- Lee SS, Wong NS. 2015 Probable transmission chains of Middle East respiratory syndrome coronavirus and the multiple generations of secondary infection in South Korea. *Int. J. Infect. Dis.* **38**, 65–67. (doi:10.1016/j.ijid.2015.07.014)
- Scheffer M *et al.* 2009 Early-warning signals for critical transitions. *Nature* **461**, 53–59. (doi:10.1038/nature08227)
- Scheffer M *et al.* 2012 Anticipating critical transitions. *Science* **338**, 344–348. (doi:10.1126/science.1225244)
- Dakos V *et al.* 2012 Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE* **7**, e41010. (doi:10.1371/journal.pone.0041010)
- Blumberg S, Enanoria WTA, Lloyd-Smith JO, Lietman TM, Porco TC. 2014 Identifying postelimination trends for the introduction and transmissibility of measles in the United States. *Am. J. Epidemiol.* **179**, 1375–1382. (doi:10.1093/aje/kwu068)
- Blumberg S, Worden L, Enanoria W, Ackley S, Deiner M, Liu F, Gao D, Lietman T, Porco T. 2015 Assessing measles transmission in the United States following a large outbreak in California. *PLoS Curr.* (doi:10.1371/currents.outbreaks.b497624d7043b1aecbfbd3dfda3e344a)
- de Quadros CA, Hersh BS, Nogueira AC, Carrasco PA, da Silveira CM. 1998 Measles eradication: experience in the Americas. *Bull. World Health Organ.* **76**(Suppl. 2), 47–52.
- Gardner LM, MacIntyre C. 2014 Unanswered questions about the Middle East respiratory syndrome coronavirus (MERS-CoV). *BMC Res. Notes* **7**, 358. (doi:10.1186/1756-0500-7-358)
- Erneux T, Mandel P. 1986 Imperfect bifurcation with a slowly-varying control parameter. *SIAM J. Appl. Math.* **46**, 1–15. (doi:10.1137/0146001)
- Ferrari MJ, Perkins SE, Pomeroy LW, Bjørnstad ON. 2011 Pathogens, social networks, and the paradox of transmission scaling. *Interdiscip. Perspect. Infect. Dis.* **1**, 10. (doi:10.1155/2011/267049)
- de Jong MCM, Diekmann O, Heesterbeek H. 1995 How does transmission of infection depend on population size. In *Epidemic models: their structure and relation to data* (ed. D Mollison). Cambridge, UK: Cambridge University Press.
- Gillespie DT. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–404. (doi:10.1016/0021-9991(76)90041-3)
- Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359. (doi:10.1038/nature04153)
- Diekmann O, de Jong MCM, Metz JAJ. 1998 A deterministic epidemic model taking account of repeated contacts between the same individuals. *J. Appl. Probab.* **35**, 448–462. (doi:10.1239/jap/1032192860)
- Kullback S, Leibler RA. 1951 On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86. (doi:10.1214/aoms/1177729694)
- Burnham KP, Anderson DR (eds). 2004 *Model selection and multimodel inference*. New York, NY: Springer.
- Hausser J, Strimer K. 2014 *Entropy: estimation of entropy, mutual information and related quantities. R package version 1.2.1*. See <http://strimmerlab.org>.
- Therneau T, Atkinson B, Ripley B. 2014 *rpart: recursive partitioning and regression trees. R package version 4.1-8*. See <https://cran.r-project.org>.
- Bartlett MS. 1956 Deterministic and stochastic models for recurrent epidemics. In *Contributions to biology and problems of health*, pp. 81–109. Berkeley, CA: University of California Press.
- Kuske R. 1999 Probability densities for noisy delay bifurcations. *J. Stat. Phys.* **96**, 797–816. (doi:10.1023/A:1004658609270)
- Berglund N, Gentz B. 2006 *Noise-induced phenomena in slow-fast dynamical systems*. London, UK: Springer.
- Keeling MJ, Rohani P. 2008 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
- Tredicce JR, Lippi GL, Mandel P, Charasse B, Chevalier A, Picqué B. 2004 Critical slowing down at a bifurcation. *Am. J. Phys.* **72**, 799–809. (doi:10.1119/1.1688783)
- Bjørnstad ON, Finkenstädt BF, Grenfell BT. 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* **72**, 169–184. (doi:10.1890/0012-9615(2002)072[0169:DOMEEES]2.0.CO;2)
- Dibble C, O'Dea EB, Park AW, Drake JM. 2016 Data from: Waiting time to infectious disease emergence. Dryad Digital Repository. (doi:10.5061/dryad.dp4kb)