Within-host pathogen dynamics: Some ecological and evolutionary consequences of transients, dispersal mode, and within-host spatial heterogeneity

Robert D. Holt and Michael Barfield

Abstract. The ecology and evolution of infectious disease occur at multiple spatial scales. In this paper, we explore some consequences of transient dynamics of pathogens within individual hosts. If infected hosts die quickly, relative to internal equilibration in pathogen dynamics, within-host transients may influence between-host transmission and spread. We develop a formulation for characterizing the overall growth rate of an infectious disease, which includes both within-host dynamics and between-host transmission, when the disease is sufficiently rare that the supply of available hosts can be viewed as a constant. This formulation is analogous to the familiar Euler equation in age-structured demography. We suggest that the pathogen growth rate estimated this way may be a better measure of pathogen fitness than is $R_0$. We point out that even simple models of within-host pathogen dynamics can have phases in which numbers overshoot the final equilibrium, and that such phases may influence pathogen evolution. We touch on the potential importance of within-host spatial heterogeneities in pathogen dynamics, and suggest that an interesting question for future work is understanding the interplay of spatial structure and transient dynamics in the within-host infection process.

1. Introduction

The ecology of infectious disease plays out in arenas at vastly different spatial scales. Traditional epidemiology focuses on between-host infection dynamics, either between individual hosts within well-mixed host populations, or among spatially segregated populations [1]. In recent years, there has been increasing attention given to an important arena of infection embedded within the host population scale, namely that of within-host infection dynamics (e.g., [2, 3, 7, 12, 13, 15, 16, 17, 20]). Following successful infection by a virus, bacterium, or fungus, a population of the pathogen is established within an individual host, which in effect is a “patch” being colonized by that pathogen. As in any colonization, there is a phase of population growth before an equilibrium (if any) becomes established in that host. This transient phase has two broad implications for disease dynamics, which will provide the interwoven themes for this chapter.

We would like to thank the University of Florida Foundation and NIH (grant 5 R01 GM60792-04) for support.
First, the transient dynamics of the within-host infection process may contribute significantly to the spread of the infection in the host population as whole. General ecological theory has long been dominated by a focus on the asymptotic behavior of systems (e.g., characterizing equilibria) as the determinant of natural population and community dynamics. There is a growing recognition that an understanding of transient dynamics may be crucial for understanding ecological systems [9]. Epidemiological models that assume host individuals are either susceptible or infected (or immune, etc.) in effect assume that the transient phase of within-host establishment dynamics can be ignored. One of our goals will be to relax this assumption and explore some consequences of explicitly tracking the size of the within-host pathogen population during each bout of infection.

Second, it may be a gross oversimplification to simply track total pathogen numbers within a host. Considerable heterogeneity in pathogen numbers and even genotype exists among tissues within individual hosts (e.g. [23]). Spatial heterogeneity and patterns of connectivity can in turn have a strong influence on the kind of transient dynamics observed in within-host infection [18]. For instance, the initial site of infection will often be distinct from the sites within which the greatest population growth is expected to occur. This implies that there will be a lag observed in the growth rate of the pathogen, following initial infection, reflecting the internal spread of the pathogen population into tissues other than that provided by the original site of entry.

In the first part of this paper, we discuss some simple aspects of transient within-host dynamics, and how this bears on between-host transmission dynamics. We then turn to some more complex issues involving transient conditions, first for a spatially unstructured host, and then briefly for one with multiple spatial compartments, and touch on how transient dynamics influence between-host transmission dynamics and pathogen evolution.

Figure 1 displays four patterns of within-host dynamics, for four different scenarios (three of which correspond to special cases of model (3.1) discussed below; note that in Figure 1B, C, and D, abundances are shown on a log scale). In Figure 1A, we portray within-host dynamics implicit in the usual SI model (e.g., [1]), in which individuals are simply divided into infected and non-infected classes (the latter may be further subdivided, for instance into susceptible vs. immune), and all infected individuals are assumed equivalent (relative to their capacity to infect other hosts). In effect, this assumes that following infection, the population of pathogens in a host immediately reaches some kind of carrying capacity, where the within-host pathogen population stays until the host either dies or recovers.

In Figure 1B, we depict another scenario, in which all the dynamics within the host are transient. Here, we assume that the initial infective propagule is tiny, relative to the size of the pathogen population that could be produced inside the host. The pathogen simply grows exponentially, until the host dies, either because of the effect of the pathogen itself on the host, or for other reasons. In either case, the death of the host eliminates the local population of the pathogen entirely. There is a sense in which the entire within-host dynamic in this case is a transient, since there is no local equilibrium reached before the host is removed. (This picture could also apply to a scenario in which the host mounts an effective, abrupt immune defense; below, we mainly consider cases involving host death.)
Figure 1C is intermediate between A and B. Here, again there is a phase of exponential growth, but the pathogen approaches an internal equilibrium, before host death. If the host lives sufficiently long after infection, then the pattern of Figure 1A (the usual assumption of SI and SIR models) may be a reasonable approximation to the actual within-host dynamics.

Finally, in Figure 1D, there is a more complex pathogen dynamic within the host. In the example shown, there is an initial rapid exponential growth. Susceptible host cells are used up, and slowly replenished by the host; this generates a predator-prey-like dynamic leading to a series of damped oscillations, in which the maximal pathogen population achieved greatly exceeds the final, equilibrial population. Comparable transient dynamics can emerge in a wide range of models (e.g., incorporating within-host spatial structure, or immune responses).

In all cases, the within-host dynamic of the pathogen occurs within a timeframe set by host death. In general, the magnitude of this timeframe should itself
depend upon the pathogen’s own dynamics. For instance, host death rates might increase with pathogen load, or pathogen growth rate. If the host dies prior to the establishment of a within-host equilibrium, then analyzing transients becomes crucial to understanding the overall host-pathogen system.

2. Shedding or Bursting

To link models for within-host dynamics to the dynamics of the pathogen in the entire host population, we have to relate within-host pathogen population size to transmission between this focal, infected host individual, and other susceptible hosts. There are several ways one can imagine such linkage to occur mechanistically. As idealized extremes, we will contrast “shedding” and “bursting.”

“Shedding” occurs when pathogens are emitted from an infected host throughout the course of infection. In this case, the degree of infectivity of the infected host in contact with another, susceptible host individual might scale with the instantaneous population size of the pathogen in the host. One issue that arises here is that shedding pathogens can act as a loss term for the microbial population within the host, which then should grow more slowly because of the “drain” of shed pathogens to the external environment. If disease transmission occurs via an environmental pool of infective propagules, then allocation of within-host pathogen growth to emission of such propagules should reduce the rate of continued growth within the focal host itself. (More broadly, the shedding may impact only a portion of the pathogen population, when there is internal spatial structuring and only certain tissues are sites of emission; see below.)

“Bursting” would instead apply to a system in which no pathogens are emitted until the host dies. This might pertain to some disease systems in which the host has to be consumed for pathogen transmission to occur (e.g., prions in bovine spongiform encephalopathy), and closely matches the life cycles of many baculoviruses infecting insects. Such a pathogen abstractly resembles an insect parasitoid, in that host death is required for successful reproduction. The degree of infection of other host individuals should then scale with the final pathogen population size within a host, at the time of that host’s death.

We assume below that the host is a free-living organism. However, it is worth noting that the same equations can also describe dynamics within a multicellular host individual, with the reinterpretation that the “host” in the models would then be a single cell that is either healthy or infected by a parasite, for example, a virus. Some types of virus tend to be continuously released from an infected cell until it dies, which is equivalent to the shedding model below (assumed for instance by [17]). Other types of virus tend to remain within the infected cell until there are a sufficient number of them to cause the cell to break open, releasing the virions to attack other host cells (corresponding to the burst model; see [11]). The chapter by Kelly [10] shows that this distinction does not alter the evolutionary generation time of the virus. However, it could affect the overall reproductive rate of different clones, and so the ideas presented below could have implications for the evolution of virulence within viral populations inhabiting single hosts (we thank John Kelly for this insight).

In our initial model, we assume that the pathogen is a microparasite (a virus, bacterium, or fungus) which infects a host starting with a single infective particle (we will ignore stochastic factors), establishing a local population. Clearly, ignoring
stochastic factors is not reasonable for the initial phases of infections starting from a single particle. However, the basic equations below apply if it is assumed that the initial infection is by an inoculum consisting of a fixed number of infectious particles (by simply rescaling the function that relates pathogen density to host mortality, and the constant that relates free pathogens to new infections). Also, if infection is by a single particle, then given the initial growth rate in the host, the probability of loss due to stochastic factors could be calculated and incorporated into the new infection probability parameter.

As noted above, the pathogen can spread from the infected host individual in one of two ways. It can be shed from the host continuously at a per-capita rate $s$, until the host dies, at which time all remaining pathogens in the host are lost. Alternatively, all the pathogens may be retained in the host until the host dies, at which time all pathogens are released (in a burst). In either case, the host mortality ($m$) may be a function of the pathogen level (denoted $V$) in the host. (An alternative mechanistic assumption might be that mortality scales with the rate of pathogen growth, $dV/dt$, a quantity which should be related to the instantaneous rate of conversion of host resources into pathogen resources.) In some cases, for simplicity we will assume a linear relationship between pathogen load and host mortality, so that $m(V) = a + bV$, where $a$ is the intrinsic mortality rate and $b$ gives the increase in mortality due to each additional pathogen. If the pathogen level is time-varying, the host mortality rate also varies over time. The probability that an individual host survives from the time of initial infection at time 0 to time $t$ is given by

$$P_s = \exp \left\{ - \int_0^t m[V(\tau)] \, d\tau \right\}.$$

We assume that each individual pathogen, once released (either by shedding or in a burst), has a probability $c$ of infecting another host. If it does not successfully infect another host, following release, the pathogen is assumed to die. In general, the quantity $c$ will be a function of the density of susceptible hosts. For now, we will assume that $c$ stays constant.

### 2.1. Shedding.

With pathogen shedding, the shedding rate $s$ is the fraction of the host pathogens released per unit time, so the rate at which pathogens are released is $sV$. Since a fraction $c$ of these are assumed to infect another host, the rate of new infections is $csV$. As a deliberate simplification of within-host pathogen dynamics, in some cases we assume that the pathogen tends to grow exponentially (as in Figure 1B), up to the point of host death, with intrinsic growth rate $r$. With this assumption, the rate of growth of pathogens in the focal individual host is reduced by the shedding rate $s$, so that starting from a single pathogen at time 0, the pathogen level (as long as the host is alive) is given by

$$V(t) = \exp \left\{ (r - s)t \right\}.$$

Given density dependence in the pathogen (e.g., because host resources are used up), or defensive responses, more complex forms of $V(t)$ are likely.

The expected value of the rate of new infections from a host at time $t$ is the product of the probability that the host is alive at time $t$ and the rate of new infections assuming the host is alive at $t$, which is

$$(2.1) \quad csV(t) \exp \left\{ - \int_0^t m[V(\tau)] \, d\tau \right\}.$$
The expected value of the total number of secondary infections is the integral of this rate over all time:

\[ \int_0^\infty c s V(t) \exp \left\{ - \int_0^t m[V(\tau)] \, d\tau \right\} \, dt. \]

If the number of pathogens is constant at \( \hat{V} \) (as in Figure 1A), the total number of pathogens released that infect other hosts, generally known as the basic reproduction number or “rate” \( R_0 \) is simply \( c s \hat{V} / m(\hat{V}) \). For any time-varying \( V \), to make further progress, we need to make an assumption about how changes in pathogen load translate into changes in host mortality.

For instance, as a limiting case we can assume that uninfected hosts only die due to infection. If the mortality rate \( m \) is proportional to \( V \), so that \( m(V) = b V \), then the above expression becomes

\[ \frac{-c s}{b} \int_0^\infty -b V(t) \exp \left\{ - \int_0^t b V(\tau) \, d\tau \right\} \, dt. \]

Note that the integrand is the derivative of the exponential term (which is the probability of host survival), so the integral is that term, evaluated at the limits. At the lower limit, this is 1, and at the upper limit, this is the probability of survival for an infinite time after infection, which we assume is 0 [this is true for any \( V(t) \) that does not eventually go extinct within the living host, such as those shown in Figure 1B–D]. Taking into account the appropriate signs at the two limits, the total expected number of secondary infections produced per primary infection is simply \( R_0 = c s / b \).

Note that this result does not depend upon the details of within-host dynamics. A pathogen population within a host can reach a high level, at which point it will be shedding many pathogens into the environment, but such a pathogen will also tend to kill off its host quickly. Conversely, if the pathogen stays at a low level, it sheds at a low rate, but its host lives much longer. The two effects cancel out, given our linear assumption relating host mortality to pathogen load.

If the pathogen grows exponentially, the condition that host survival eventually goes to 0 limits \( s \) to be no greater than \( r \), the intrinsic rate of within-host pathogen growth. (If this is not true, the pathogen goes extinct within the host because it is shed more rapidly than it can grow; the realized rate of growth is \( r - s \)). Comparing different pathogen strains, the one with the greatest \( R_0 \) is the one with the greatest value of the ratio \( c s / b \), so the maximal value is approximately \( cr / b \).

However, as other authors have noted (e.g., [6]), \( R_0 \) is not typically the best measure of pathogen fitness. One reason for this is that in considering infection of novel hosts, one must consider the tertiary infections spawned by each secondary infection, the quarternary infections then generated, and so on (other reasons have to do with the potential for co-infection, which we are assuming does not occur). Another measure of fitness might thus be the total growth rate of the pathogen population, summed over all infected hosts, a fitness measure which is called the Malthusian parameter [5]. In the next few paragraphs we derive an expression for the Malthusian parameter of a pathogen, taking into account both within-host dynamics and between-host transmission.

We can derive a measure for the total growth rate of the pathogen in the host population, if we assume that the number of susceptible hosts does not change over the time scale in question (e.g., because the number is large, or replenishment
occurs rapidly. We assume that after a sufficient time has passed since an initial infection, the total amount of pathogen in the population will grow exponentially at a rate of \( r' \), so that the total pathogen at time \( T \) from an initial infection at time 0 for sufficiently large \( T \) has the form \( A \exp\{r'T\} \) (the quantity \( A \) reflects initial transients). To estimate \( r' \), we use expression (2.1) above, which gives the expected rate of secondary infections produced at time \( t \) from the initial infected host. Each of these secondary infections should likewise result in an exponentially growing pathogen population, of the form \( A \exp\{r'(T - t)\} \) (since we assume that the susceptible host density remains constant, the only difference between secondary and initial infection is the start time). The total pathogen population at time \( T \) from secondary infections in the interval \((t, t + dt)\) from an individual infected at time \( 0 \) is therefore

\[
(2.2) \quad csV(t) \exp\left\{-\int_0^t m[V(\tau)] \, d\tau\right\} A \exp\{r'(T - t)\} \, dt.
\]

The total pathogen population at time \( T \) resulting from the initial host is this quantity integrated over all time (or at least until the time at which the rate of new infections from this host becomes negligible), which we set equal to \( A \exp\{r'T\} \), since this is the assumed ultimate form of the total pathogen population from the initial host. After dividing both sides by \( A \exp\{r'T\} \), we obtain

\[
(2.3) \quad \int_0^\infty csV(t) \exp\left\{-\int_0^t m[V(\tau)] \, d\tau\right\} \exp\{-r't\} \, dt = 1.
\]

Given assumptions about the functional forms of within-host pathogen dynamics, and the mortality term, this equation can be solved numerically for \( r' \). [In the Appendix, we provide some numerical results verifying that this expression for asymptotic growth rate is correct, assuming within-host pathogen dynamics are exponential up to the point of host death. In general, \( r' < r \).]

Note that our derivation of this implicit expression for total pathogen population growth (which incorporates both within-host dynamics, and between-host transmission) is analogous to the familiar derivation of Euler’s equation in age-structured demography (e.g., [19]; see [4] and [14] for other applications of the Euler’s equation in pathogen population dynamics). In place of age-specific mortality, we have a mortality term that depends upon pathogen load; in place of age-specific fecundity, we have an effective rate of transmission per infected host, which also depends upon pathogen load (as well as shedding rate).

If the pathogen level is fixed at \( \hat{V} \), then the mortality is fixed at \( m(\hat{V}) \), and (2.3) can be solved for \( r' = csV - m(\hat{V}) \). Recall that in this case \( R_0 = csV/m(\hat{V}) \), so \( r' = m(\hat{V})[R_0 - 1] \). For a given \( R_0 > 1 \), higher mortality gives a higher growth rate, since the same number of pathogens is released, but they are released sooner.

If the host mortality rate in the absence of the pathogen is \( a \), we can write the mortality rate as \( m(V) = a + f(V) \). Then equation (2.3) above becomes

\[
\int_0^\infty csV(t) \exp\left\{-\int_0^t f[V(\tau)] \, d\tau\right\} \exp\{-(r' + a)t\} \, dt = 1.
\]

Since \( a \) appears only in the term in which it is added to \( r' \), we can solve the equation for \( a = 0 \) to get \( r'_0 \), and then find \( r' \) using \( r'_0 = r' + a \), or \( r' = r'_0 - a \). Therefore, the background host mortality rate \( a \) directly reduces \( r' \).
As a limiting case, if we assume that host mortality is independent of the pathogen population within the host \( f(V) = 0 \), and that there is exponential within-host growth, then \( V(t) = \exp\{(r - s)t\} \) and the above equation can be solved, giving \( r' = r - s(1 - c) - a \). In this case, the maximum \( r' \) is obtained by the minimal shedding of the pathogen. With constant mortality rate, all hosts that are alive have the same expected mortality, and the same pathogen growth rate. Therefore, unless \( c = 1 \), there is a penalty associated with being shed (the possibility of not finding a new host) but no benefit. In this case, the expected number of secondary infections from an infected host is \( cs/(a + s - r) \) (unless the denominator is negative or 0, in which case it is infinite, since in that case the pathogen population size grows faster than the probability of host death). Assume that \( r < a \) (e.g., a “slow virus”), and that we are comparing pathogen clones which may vary in their shedding rate. It is interesting to note that in this limiting case, the pathogen that generates the highest number of secondary infections per primary infection (over the lifetime of that infection, viz. \( R_0 \)) is the one that has the lowest growth rate (as measured by its Malthusian parameter), in the host population as a whole.

The effect of varying the other parameters can be found by numerically solving the above equation. One question of interest may be how the rate of increase depends upon shedding rate. We earlier saw that \( R_0 \) increases linearly with \( s \) (with linear mortality). If \( R_0 \) is used as a measure of pathogen fitness, then more shedding is better. However, the result of numerical explorations shows that often \( r' \) increases with increasing \( s \) for low values of \( s \), but eventually peaks, and then decreases with further increase in \( s \) (Figure 2A). The reason is basically that a pathogen remaining in the host can multiply, but faces the risk that the host will die, after which all pathogens in the host also die. However, a pathogen being shed also faces a risk of death. If there is too low a shedding rate, the pathogen population rises quickly and increases the host mortality rate, and thus many pathogens are lost at host death. Too high a shedding rate and the host lives longer, but the pathogen level increases more slowly within each host.

Thus, an intermediate rate of shedding may lead to the maximal overall rate of increase of the pathogen in the host population. Note that this arises because of the interwoven dynamics of within-host pathogen growth, and between-host infection, and not because of any assumption about tradeoffs among parameters in the model. If pathogen clones differ in their shedding rate, but are otherwise identical, the one with an intermediate rate of shedding will have the greatest growth rate in the host population.

Surprisingly, increasing the shedding rate can lead to an increase in the average time to release of pathogens (with exponential pathogen growth). With low shedding, the pathogen population rises quickly and kills the host early, so that the average time of shedding is short. With high shedding, the pathogen population grows more slowly and so the host lives longer. Since the pathogen population is growing (if \( r > s \)), the number of pathogens shed per unit time increases as long as the host is alive. Therefore, the average time of pathogen release tends to increase with higher shedding. Somewhat counter-intuitively, with a decrease in the probability of infecting a new host (\( c \)), the overall growth rate \( r' \) decreases, as expected, but the shedding rate at which \( r' \) is maximized can increase, because increased shedding means each pathogen population within a host on average spends more
A. $a = 0$, $b = 0.01$, $c = 0.1$.

B. $a = 0$, $b = 0.001$ and intrinsic growth rate $r = 1$.

C. $a = 0$, $b = 0.01$, $c = 1$, $r = 1$.

Figure 2. Pathogen population growth rate for shedding model with linear mortality rate. Note there is an intermediate optimum in A and B, but not C.
time in the protective environment of the host, before killing its host. In effect, increased shedding provides a mechanism for decreased virulence, so this result is consistent with familiar models for the evolution of virulence. Figure 2B shows a numerical example of this effect, assuming a linear relationship between pathogen load and host mortality.

For these results on pathogen demography to be translated directly into assertions about the likely direction of pathogen evolution requires additional assumptions. For instance, if there are distinct pathogen clones, and each host has a single such clone (i.e., there is no admixture of pathogen clones within individual hosts), then relative growth rates can translate into relative Malthusian fitnesses.

Although there is often an optimum shedding rate, if \( c = 1 \) (i.e., each shed pathogen infects a new host), then \( r' \) monotonically increases with increasing \( s \) (see Figure 2C for an example). In this case, each shed pathogen finds a new host, so there is no longer any pathogen mortality associated with being shed. Therefore, the pathogen does better if shedding is higher, since this minimizes the chance of being in the host when it dies. The growth rate \( r' \) with \( c = 1 \) and high shedding approaches the within-host pathogen growth rate \( r \).

Finally, the quantity \( c \) is likely to change over time, not least because as the infection spreads in the host population, the density of susceptibles will decline. Incorporating the demographic impact of infection on host numbers is required for a full ESS analysis of evolution in the pathogen.

An approximate expression for \( r' \) is the logarithm of \( R_0 \) divided by the average time from host infection to secondary infection. This expression is

\[
\frac{R_0 \ln[R_0]}{\int_0^\infty tcsV(t) \exp \left\{- \int_0^t m[V(\tau)] d\tau \right\} dt}
\]

where

\[
R_0 = \int_0^\infty csV(t) \exp \left\{- \int_0^t m[V(\tau)] d\tau \right\} dt.
\]

This expression would be exact if all pathogens were released at the average release time. The closer the release times cluster around the mean, the better this approximation becomes. It is not good for constant virus loads, since in this case release times have an exponential distribution, which has a high variance. The approximation is much better with exponential pathogen increase and pathogen-dependent mortality.

### 2.2. Bursting.

The other model of pathogen spread has all pathogens retained until the death of the host, at which time a fraction \( c \) of the released virions infect a new host. With bursting, the rate at which new infections are generated is the product of the death rate, the pathogen number at the time of death, and the probability that a pathogen infects a new host (\( c \)). The death rate at time \( t \) is the probability that the host has survived to time \( t \), multiplied by the mortality rate. Therefore, the new infection rate with bursting is

\[
cV(t)m[V(t)] \exp \left\{- \int_0^t m[V(\tau)] d\tau \right\}.
\]

If we assume exponential pathogen growth within the host, \( V(t) = \exp\{rt\} \) with bursting, since there is no shedding to reduce the within-host growth rate. The
equation for population growth rate of the pathogen can be found in a similar manner to that described above for shedding:

\[
(2.4) \quad \int_0^\infty cV(t)m[V(t)] \exp \left\{ - \int_0^t m[V(\tau)] \, d\tau \right\} \exp\{-r't\} \, dt = 1.
\]

Note that in contrast to expression (2.3), the “fecundity” analogue now includes host mortality. With bursting, and within-host exponential growth, the overall pathogen population growth rate is an increasing function of the within-host pathogen growth rate \( r \) and infection rate \( c \), and is equal to \( r \) when \( c = 1 \). For \( c < 1 \), it is a decreasing function of the mortality parameters, since pathogen growth proceeds at its maximum rate as long as the host is alive, and host death leads to a loss of some pathogens. In a more realistic scenario in which pathogen numbers saturate or decrease at some point in the infection, host mortality is likely to be beneficial to the pathogen at or before the time the pathogen level approaches its maximum.

With exponential pathogen growth, if mortality is independent of pathogen level \((m = a)\), then the above equation can be solved for \( r' = r - (1 - c)a \). If the host instead dies when the pathogen reaches a threshold level \( V_T \), then the total pathogen population grows at rate \( r \) for a time \( \ln(V_T)/r \) [the time it takes pathogen in a host to increase from the initial propagule size (scaled to unity) to the threshold level], after which host death occurs. Only a fraction \( c \) of the released pathogens finds a new host. Therefore, over one cycle (and all hosts will be synchronized if we start from one host), the pathogen population increases by \( cV_T \) in time \( \ln(V_T)/r \). The equivalent \( r' \) is the logarithm of the increase divided by the time between infection and host death, leading to

\[
r' = \frac{r \ln(V_T)}{\ln(V_T)} = r \left( 1 + \frac{\ln(c)}{\ln(V_T)} \right)
\]

where the last entry in the bracketed term is negative, since \( c < 1 \).

We have thus far focused on the consequences of assuming that host death occurs prior to the equilibration of the pathogen within the host, but we have largely assumed a very simple model for within-host growth: exponential growth following infection. In the following sections, we will provide some explorations of other aspects of transient dynamics. Note that expressions (2.1) through (2.4) above do not depend upon explicit assumptions about within-host pathogen dynamics (e.g., exponential growth), provided all pathogens in the host are potentially available for shedding or release at host death. One way to generalize the model would be to allow the shedding rate itself to depend upon pathogen population size. Another route of generalization, which we touch on below, is to have a spatial structure for the pathogen within individual hosts, so that only a fraction of the total population is available for export to the external environment.

3. Within-host transient dynamics

3.1. Single-compartment, single-cell-type model. We now consider a particular class of pathogens - viruses. Because viruses must replicate within host cells, and are transmitted via free virion particles, one must monitor levels within the host of healthy cells, infected cells, and free virions. A simple dynamic model
with the virus attacking only one cell type within one tissue or organ (e.g., \([11]\)) is
\[
\frac{dn}{dt} = \lambda - \mu n - \betaqn,
\]
\[
\frac{dn^*}{dt} = \betaqn - \mu^*n^*,
\]
\[
\frac{dq}{dt} = \nu\mu^*n^* - \mu'q - \betaqn
\]
(3.1)
where \(n\) is the number of uninfected cells, \(n^*\) is the number of infected cells and \(q\) is the number of free virions. Uninfected cells are input at a constant rate \(\lambda\), die at a per-capita rate of \(\mu\), and are infected at rate \(\beta\). Infected cells die at rate \(\mu^*\), and release \(\nu\) virions at death. Virions are cleared at rate \(\mu'\). The model is quite simple, yet illustrates some general points about transient dynamics.

### 3.2. Increase in virus when rare.
If there is no infection, then there are no infected cells or free virions, and the number of uninfected cells should equilibrate at \(\hat{n}_0 = \lambda/\mu\). Given that uninfected cells are at this equilibrium, the basic reproduction ratio (the average number of free virions produced within an infected host per free virion, or the number of infected cells produced per infected cell, when virus is rare) is
\[
R_0 = \frac{\beta\lambda\nu}{\beta\lambda + \mu'\nu} = \frac{\beta\hat{n}_0\nu}{\beta\hat{n}_0 + \mu'}.
\]
This is derived as follows. Virions either infect cells (at rate \(\beta\hat{n}_0 = \beta\lambda/\mu\) if virus is rare), which on death release \(\nu\) virions each, or are cleared (at rate \(\mu'\)), producing no virions. It is assumed that both infection and clearance of virions are Poisson processes. If so, then the probability that a virion infects a cell before it is cleared is \((\beta\lambda/\mu)/(\beta\lambda/\mu + \mu') = \beta\lambda/(\beta\lambda + \mu'\nu)\), which is the rate of infection divided by the sum of the two rates. The expected value of the number of virions produced by a free virion is the probability of infection multiplied by the number of free virions produced by an infection \((\nu)\), plus the probability of clearance multiplied by the number of free virions produced in this case (which is 0, so this second term drops out). Therefore, the expected value of the number of free virions produced by a free virion is \(\beta\lambda\nu/(\beta\lambda + \mu'\nu)\), which is the basic reproductive ratio, \(R_0\). (A similar argument can be used projecting from infected cell to infected cell.)

For the virus to increase when rare, \(R_0\) must exceed 1, which implies that \(\beta(\nu - 1)\hat{n}_0 > \mu'\). If this condition is met, virus introduced at a low level will increase exponentially (after a transient period during which infected cells and virions reach their steady-state ratio, and until the virus reaches high enough levels to affect the uninfected cell number). The rate of exponential increase of the virions and infected cells is the dominant eigenvalue of the Jacobian of the infected cell and virion equations above, evaluated at the no-infection equilibrium. The Jacobian is
\[
\begin{bmatrix}
-\mu^* & \beta\hat{n}_0 \\
\nu\mu^* & -\beta\hat{n}_0 - \mu'
\end{bmatrix}
\]
and the dominant eigenvalue is
\[
s_d = \sqrt{\left(\frac{\mu^* - \mu' - \beta\hat{n}_0}{2}\right)^2 + \beta\hat{n}_0\nu\mu^*} - \frac{\mu^* + \mu' + \beta\hat{n}_0}{2}.
\]
(3.2)
(The non-dominant eigenvalue \((s_n)\) is the same except for a negative sign before the first term.)
During the initial phase of the infection, infected cells and virions will reach a steady-state ratio which can be found from the eigenvector corresponding to the dominant eigenvalue, and both then increase exponentially at the same rate (until they start to have an effect on uninfected cell number). The ratio of virions to infected cells during this initial phase of increase is

\[
\left(\frac{\mu^* - \mu' - \beta \hat{n}_0}{2} + \frac{\beta \hat{n}_0 \nu \mu^* - \mu^* - \mu' - \beta \hat{n}_0}{2}\right) \cdot \frac{1}{\beta \hat{n}_0}.
\]

Virions and infected cell will not, in general, have this ratio at the start of the infection. However, the initial transient response (the part of the transient response occurring when the virus is rare) for either virions or infected cells will have the form \(A \exp\{s_d t\} + B \exp\{s_n t\}\). The first term is the exponential increase at the dominant eigenvalue. This is reached as soon as the second term, due to the non-dominant eigenvalue, becomes negligible. Therefore, the non-dominant eigenvalue (which is always negative for this system) determines how fast the virions and infected cells reach an exponential increase, with the ratio between them then given by the above expression. If uninfected cell number drops slowly enough, the virion and infected cell numbers could adjust their growth rates and ratio according to the above expressions with the actual uninfected cell numbers substituted for the no-virus uninfected cell numbers. However, as the uninfected cell number drops, the magnitude of the non-dominant eigenvalue drops, so that the speed with which the ratio of virions to infected cells adjusts drops. Often, the uninfected cell number drops too fast for this approximation to hold.

### 3.3. Equilibria

The above dynamical system has a stable equilibrium, provided the virus can increase when rare. In this case, the eventual equilibrium is

\[
\hat{n} = \frac{\mu'}{\beta (\nu - 1)} \approx \frac{\mu'}{\beta \nu},
\]

\[
\hat{n}^* = \frac{\lambda}{\mu} - \frac{\mu'}{\beta \mu^*(\nu - 1)} \approx \frac{\lambda}{\mu^*},
\]

\[
\hat{q} = \frac{\lambda (\nu - 1) - \mu}{\mu'} \approx \frac{\lambda \nu}{\mu'}.
\]

[The approximate equilibria are obtained by assuming that at the equilibrium the vast majority of uninfected cells become infected before they die, and that the vast majority of virions are cleared before they infect a cell; see [11].]

If there is a trade-off between burst size \(\nu\) and infectivity \(\beta\), so that as burst size increases infectivity decreases (an infected cell can produce either many virions that are poor at infecting new cells, or fewer virions that are each better at infection), then this trade-off may have different consequences for equilibrium virion numbers, infected cell numbers and initial growth rate. The equilibrium virion number is \(\hat{q} = (\lambda (\nu - 1) - \mu / \beta) / \mu'\). The effect of varying burst size and infectivity are in opposite directions, but the effect of burst size is likely to be larger. If \(\beta \nu\) is constant = \(d\), then virion equilibrium abundance increases with increasing \(\nu\) as long as \(\lambda / \mu' > \mu / d\). Substituting \(\beta \nu\) for \(d\) and rearranging, this condition is \(\beta \nu \lambda / \mu > \mu'\), while the condition for virus to increase when rare is \(\beta (\nu - 1) \lambda / \mu > \mu'\). Hence, as long as the virus can increase when rare, virion abundance will increase with increasing \(\nu\). The infected cell equilibrium is \(\hat{n}^* = \lambda / \mu^* - \mu' / (\beta \mu^* \nu - \beta \mu^*)\).
In this case, with the above trade-off, increasing $\beta$ decreases the magnitude of the denominator of the second term, which increases the magnitude of this term, which decreases infected cells. So again, increasing $\nu$ would increase infected cell numbers. However, this would likely have a smaller effect than the effect on virions. The direction of the effect on the initial rate of increase is in the same direction.

3.4. Transient response - initial overshoot. The transient response of the system of equations given in (3.1) often contains significant oscillations (Figure 1D). Particularly important is the initial overshoot of the equilibrium. The tendency to overshoot depends on the relative values of the uninfected and infected cell death rates and virion clearance rate. If there is an overshoot, the magnitude is also mostly determined by these quantities, although the other parameters have some effect also on the magnitude of the overshoot (we assume in the results below that $\lambda$ is proportional to $\mu$, so the initial uninfected cell number is the same). If the uninfected cell death rate ($\mu$) is lower than both the infected cell death rate ($\mu^*$) and virion clearance rate ($\mu'$), all three quantities (uninfected and infected cells, and virions) tend to overshoot their respective equilibria (uninfected cell numbers decrease as the virus increases, so when it overshoots its equilibrium, it becomes less than its equilibrial value). If $\mu^* > \mu, \mu'$, then only the infected cell number will usually overshoot. If both are true ($\mu^* > \mu > \mu'$), then all three quantities tend to overshoot. The infected cell death rate would normally be higher than the uninfected cell death rate. This generally results in overshoot of all quantities.

Figure 3 shows the maximum virion abundance divided by the equilibrium virion abundance (a measure of the overshoot) when the virus increases from low levels, for a virion clearance rate of 1 (time and other rates can be scaled by virion clearance rate, so this assumption does not lead to a loss in generality). The surface is at 1 (i.e., no overshoot) for uninfected cell death rate of 1 or more, since in this region we have $\mu \geq \mu'$. It is also 1 on and to the left of the $\mu = \mu^*$ diagonal, since in this region $\mu \geq \mu^*$. The non-unity values at the right and back of the Figure represent overshoot of the virion number, which increases as $\mu'/\mu$ increases, with the value of $\mu$ being the more important. The overshoot of uninfected cell number has the same form (the equilibrium over the minimum value, since uninfected cells actually undershoot), but slightly lower values where overshoot occurs (maximum difference of 10% in the right corner where the overshoot reaches the maximum values). Infected cell number overshoot is similar for low $\mu^*$ but much higher for high $\mu^*$. It also is equal to 1 only to the left of the $\mu = \mu^*$ diagonal.

These patterns describe within-host dynamics. Note that if virions released from the infected host are the means of transmission of the infection to novel hosts, then we could link the above model (3.1) to our earlier treatment of pathogen population growth over the entire host population. To do that for the shedding scenario, we would need to explicitly include a loss term due to shedding (which could be added to the virion clearance rate). For instance, a clone with a higher $\mu^*$ has a higher initial exponential growth rate in its host, and if the host dies while the dynamics are in this transient phase, it should also have an overall higher growth rate in the host population as a whole (whether or not the mode of transmission is via virions, or infected cells). But if the pathogen reaches a dynamic equilibrium in the host, and the host is long-lived, this parameter becomes irrelevant to the number of virions present. And if between-host transmission is via infected cells, at equilibrium the abundances of those cells are inversely related to their death rates.
Figure 3. Magnitude of initial overshoot of virion abundance, relative to its equilibrium value, for $\beta = 0.0001$, $\nu = 50$, $\mu' = 1$, and $\lambda = 10000\mu$ (so the initial number of uninfected cells is 10000). There is no overshoot unless the infected cell death rate and virion clearance rate are both higher than the uninfected cell death rate.

[see (3.3)]. So whether or not evolution favors viruses which rapidly kill host cells, or disfavors such viruses, may depend on whether or not within-host dynamics is largely transient, or at equilibrium, over the lifetime of the infected host.

As a step towards this evolutionary question, one can couple within-host dynamics to between-host transmission using the machinery derived above, to generate the growth rate of the pathogen in the host population. Figure 4 shows two examples of pathogen population growth rate as a function of shedding rate calculated using equations (3.1) and some approximations based on them. The dotted lines were calculated numerically using the dynamics in equations (3.1), with the shedding rate added to the virion clearance rate parameter $\mu'$, (since shedding represents a loss of virions). This corresponds to Figure 1D, with the full intrahost dynamics incorporated into calculating $r'$. The solid lines were calculated assuming that the virion level was constant at the virion equilibrium level of equations (3.1) (with shedding rate added to $\mu'$; note that this corresponds to the assumption of Figure 1A). The dashed lines were calculated assuming that the virions increased exponentially at the rates given by the dominant eigenvalues given in equation (3.2) (with shedding rate added to $\mu'$; this matches the scenario of Figure 1B).

In both panels the host death rate is assumed to be directly proportional to the virion level. As we demonstrated above, this means that $R_0 = cs/b$, independent of the virion dynamics, as long as host death eventually occurs. We chose $c = 10b$ in both cases, so that $R_0 = 1$ at $s = 0.1$. Therefore, all curves have $r' = 0$ at this
A. $a = 0$, $b = 10^{-14}$, $c = 10^{-13}$.

B. $a = 0$, $b = 10^{-13}$, $c = 10^{-12}$.

**Figure 4.** Pathogen population growth rate for shedding model with linear mortality rate and virus dynamics calculated using equations (3.1) (dotted curve), assuming virion level constant at equilibrium virion level from those equations (solid curve) or assuming a virion level exponentially increasing at the dominant eigenvalue in equation (3.2). The examples depicted show that at times one can accurately assume either a simple exponential pattern for within-host dynamics, or a rapid approach to a within-host equilibrium. Parameters for virus model as in Figure 1D.
Figure 5. Pathogen population growth rate for a bursting model with virus-independent mortality and virus dynamics calculated using equations (3.1) (parameters for virus model as in Figure 1D), as a function of mortality rate ($a$). Other parameters are $b = 0$, $c = 10^{-10}$. Solid lines give the calculated population growth rate. Solid circles represent an approximation based on an estimated probability of release at the peak abundance during the transient (see text); in this example, the approximation is quite accurate.

shedding rate. In Figure 4A, the death rate parameter $b$ is set low enough that the host has a fair probability of surviving the initial transient, after which the virion level approaches the equilibrium. Therefore, the dotted line using the full dynamics of (3.1) is close to the results using the equilibrium virion levels. In Figure 4B, the death rate parameter is increased an order of magnitude, which makes it unlikely that a host could survive the initial transient. Therefore, in this case, the numerical results with the virion dynamics described by equations (3.1) are close to the results produced assuming exponential increase.

For the burst model, the existence of transients could potentially have a large impact on the number of virions present at the time of host death. If infected hosts are short-lived, relative to the transient phase of infection dynamics, then small differences in the expected time of host death could have large impacts on the likelihood of viral transmission. For instance, in Figure 1D, viral abundance plummets several orders of magnitude over just a few time steps, after the initial peak abundance. If the pathogen could determine the time of host death, so as to maximize its own fitness, it is clear that this death should occur no later than the initial peak in virion abundance. Early host death in this example both capitalizes on the internal transient overshoot, and permits rapid release of the pathogen.

Figure 5 shows the virus population growth rate as a function of host death rate, for the burst model with a fixed host death rate, as a function of that death rate (in this example, for illustrative purposes we assume that the pathogen is
avirulent, as measured by effects on host mortality rates). For the example shown, the growth rate peaks at about 0.01, which gives a mean time of death of 100 time units. This is approximately the time that the virion level peaks in the virus model (the virion dynamics used in Figure 5 are shown in Figure 1D). If the death rate is higher, then individuals tend to die before the peak and so fewer virions are released. If the death rate is lower, then individuals tend to live longer and again release fewer virions (because the virus settles into an equilibrium within-host abundance less than observed during the peak of the transient phase). In the latter case, the virions also tend to be released later, which further depresses the growth rate.

The growth rate in this case depends largely on the host death rate at the time of the peak virion abundance, which is at time 112 in this example. Since the time of death has an exponential distribution with mean $1/a$, the probability density for death at this time is $a \exp\{-112a\}$. The expected number of secondary infections due to virions released at this time will be equal to the product of this value, the infection constant $c$, the peak virion density, and the effective duration of the peak (to convert the probability density above to a probability of death). If this peak release were the only release, then the population growth rate $r'$ would be equal to the natural logarithm of this number of secondary infections divided by the time of the peak (112). The only quantity that has to be estimated is the effective duration of the peak. For illustration, this was approximated by dividing the integral of the virion density from the start of the infection to the time of the first relative minimum by the peak density. The solid circles in Figure 5 show the results of this calculation. The agreement is very good, indicating that variation in the probability of host mortality at the peak virion time can account for the shape of the curve.

This model thus suggests that there may be a nonlinear relationship between host mortality rates and the overall rate of spread of a pathogen in a host population, when the pathogen relies upon host death for infective propagules to be released, and when there are strong transient effects in within-host dynamics. In the example shown (corresponding to Figure 1D), an intermediate death rate of the host catches the pathogen population during a substantial overshoot of numbers, above the potential within-host equilibrium. Again, the existence of an optimal mortality rate for the host, with respect to maximizing the overall rate of growth of the pathogen, does not reflect a tradeoff in the usual sense, but rather emerges from the interplay of within-host and among-host dynamics.

3.5. Two or more compartments. Transient dynamics can also be important for more complex hosts with internal structure. Here, we briefly consider a more general virus model that includes multiple compartments and cell types (see also [10]). The variables now are the number of virions in each compartment $i(q_i)$, and the number of uninfected ($n_{ji}$) and infected ($n^c_{ji}$) cells of each type $j$ and compartment $i$. We assume that cells and virions can move between compartments.
The system is described by the following equations.

$$\frac{dq_i}{dt} = \sum_j \mu_j^* \nu_j n_{ji}^* - \mu_i' q_i + \sum_j \beta_j n_{ji} q_i + \sum_{k \neq i} m_{ki} q_k - \sum_{k \neq i} m_{ik} q_i,$$

$$\frac{dn_{ji}}{dt} = \lambda_{ji} - \mu_{ji} n_{ji} - \beta_{ji} n_{ji} q_i + \sum_{k \neq j} M_{ki} n_{jk} - \sum_{k \neq j} M_{ik} n_{ji},$$

$$\frac{dn_{ji}^*}{dt} = \beta_{ji} n_{ji} q_i - \mu_j^* n_{ji}^* + \sum_{k \neq j} M_{ki}^* n_{jk} - \sum_{k \neq j} M_{ik}^* n_{ji}^*$$

which include the same parameters as the simpler model, with the subscripts $j$ indicating the cell type, and $i$ the compartment. The new parameters are migration rates: $m$ for virions, $M$ for uninfected cells and $M^*$ for infected cells. The first subscript on the migration parameters denotes the source compartment, and the second the destination. Migration rates are assumed to be independent of cell type.

Analysis of this set of equations is difficult even with one cell type and two compartments. Some initial steps towards such an analysis are reported in [18], and elsewhere we intend to explore spatially structured hosts in more detail. Rather than explore the above model in full, here we merely touch on some limiting cases.

First, consider what happens when movement of virions and cells between compartments is slow, compared to dynamics within compartments. During the initial exponential growth phase, the overall growth rate of the infection is approximately just the growth rate expected in the compartment with the highest growth rate [18]. Thus, all else being equal, the compartment where the virus growth rate is highest will dominate the initial transient dynamics of the system. For instance, the compartment with the highest value for $\beta$ or $\mu^*$ will also permit the most rapid initial growth rate of the pathogen. If the host dies while the pathogen is still in this transient phase of initial increase, then differences between pathogen clones in these habitat-specific parameters will translate into differences in Malthusian fitness. (Quantitatively describing these effects would require a more complex analogue of expressions (2.3) and (2.4), permitting for instance different shedding rates from different compartments, and different compartment-specific impacts of the pathogen on host mortality.) If the pathogen settles into a within-host equilibrium, then at low movement rates the density of virions and infected cells should approximate that given by equations (3.3). At equilibrium, the local density of virions does not depend on infected cell death rates, and it only weakly depends on infection rates of healthy host cells, so within-host heterogeneity in these parameters is not likely to influence overall transmission to new hosts.

At high movement rates, in this model if there is heterogeneity in parameters among compartments, the growth rate is depressed below the maximal possible. Moreover, one finds that the potential for dramatic overshoots in the transient phase as the infection moves towards a within-host equilibrium can be reduced (for details see [18]). As with the single-compartment case discussed above, a full treatment would need to couple these within-host transient dynamics with among-host transmission. We intend to address this theme in a future contribution.
4. Conclusions

In conclusion, we have argued that a consideration of transient dynamics of within-host infection may be important in understanding the epidemiology of infectious diseases. We have presented a formulation for characterizing the overall growth rate of an infectious disease, which includes both within-host dynamics and between-host transmission, when the disease is sufficiently rare that the supply of available hosts can be viewed as a constant. We have shown that even simple models of within-host pathogen dynamics can have phases in which numbers overshoot the final equilibrium, and suggest that such phases may be of selective importance in pathogen evolution. The distinction between shedding and bursting in pathogen transmission is relevant to pathogen dynamics within individual hosts, as well as to pathogen dynamics in host populations as a whole. Finally, we have noted that many hosts have substantial internal spatial heterogeneities, and that such heterogeneities can influence the magnitude and character of transient dynamics. There is a growing appreciation in general ecology of the importance of transient dynamics and spatial processes [8, 9, 22]. In like manner, we suggest that the interplay of transient dynamics and internal spatial structure should provide a fruitful avenue for future empirical and theoretical studies of infectious disease dynamics.

5. Appendix

We wrote a program to simulate the shedding process, so as to assess expression (2.3) in the main text. The growth of the pathogen numbers within each host was assumed to be deterministic (assumed to start at 1 at infection and to increase exponentially at a rate of $r - s$). Host mortality and new host infections were assumed to be Poisson processes which occurred randomly at the rates determined by the equations in the text. Host mortality was assumed to be a linear function of pathogen number. Small, fixed time steps were used. For each time step, the pathogen population of each host was determined using $V(T) = \exp\{(r - s)(T - t)\}$, where $T$ is the time at the start of the time step and $t$ is the time at which the host became infected (the only information we needed to keep about each host). The pathogen level of each host was used to determine its mortality rate, from which the probability of mortality was determined for that time step (the product of mortality rate and time step). A random number generator was used to determine whether death occurred. If so, the host was deleted. If not, the rate of production of secondary infections from that host (and corresponding probability) was determined, and another random number generator used to determine if a secondary infection occurred (the time step was small enough so that there was a very small probability of more than one secondary infection in a step). If so, a new infected host was created with the current time as the time of infection. Initially, there was a single host infected with a single pathogen, and the total pathogen population size was determined at each time step. This process was repeated 400 times and the results averaged at each time. The population size initially increased at a rate close to $r$ (since the virus numbers at the beginning are dominated by the first host). Later, the growth rate dropped, and eventually reached exponential growth at the predicted rate. Figure 6 shows an example. Note that after an initial phase, the lines on the log plot are parallel (the solid line is a simulation run, and the dashed line is the exponential increase predicted, using the rate of increase calculated from equation (2.3), and starting from time 0). In this example, the
Figure 6. Simulated shedding scenario with parameters $a = 0$, $b = 0.01$, $c = 0.1$, $s = 0.5$ and $r = 1$. The solid line is the simulation results. The dashed line is $\exp\{r't\}$ with the value of $r'$ calculated from equation (2.3) in the main text. The assumption used in deriving (2.3) is that eventually $V(t) = A\exp\{r't\}$. Since the lines are parallel (on a logarithmic scale) after an initial period, the assumption of eventual exponential growth is clearly valid for this example.

prediction from equation (2.3) does describe quite well the asymptotic growth rate of the pathogen population.

Acknowledgements

We thank Erin Taylor for editorial assistance, and John Kelly for very useful suggestions.

References


[10] J. K. Kelly, Evolutionary and dynamic models of infection with internal host structure, this volume.


Department of Zoology, University of Florida, 223 Bartram Hall, PO Box 118525, Gainesville, FL 32611.

E-mail address: rdholt@zoo.ufl.edu

E-mail address: mjbarfield@ufl.edu