University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Borovkov, K;Day, R;Rice, T

Title:
High host density favors greater virulence: A model of parasite-host dynamics based on multi-type branching processes

Date:
2013-05-01

Citation:
Borovkov, K., Day, R. & Rice, T. (2013). High host density favors greater virulence: A model of parasite-host dynamics based on multi-type branching processes. Journal of Mathematical Biology, 66 (6), pp.1123-1153. https://doi.org/10.1007/s00285-012-0526-9.

Persistent Link:
https://hdl.handle.net/11343/268669

High host density favors greater virulence: a model of parasite-host dynamics based on multi-type branching processes

K. Borovkov,* R. Day,† T. Rice‡

March 21, 2012

Abstract

We use a multitype continuous time Markov branching process model to describe the dynamics of the spread of parasites of two types that can mutate into each other in a common host population. While most mathematical models for the virulence of infectious diseases focus on the interplay between the dynamics of host populations and the optimal characteristics for the success of the pathogen, our model focuses on how pathogen characteristics may change at the start of an epidemic, before the density of susceptible hosts decline. We envisage animal husbandry situations where hosts are at very high density and epidemics are curtailed before host densities are much reduced. The use of three pathogen characteristics: lethality, transmissibility and mutability allows us to investigate the interplay of these in relation to host density. We provide some numerical illustrations and discuss the effects of the size of the enclosure containing the host population on the encounter rate in our model that plays the key role in determining what pathogen type will eventually prevail. We also present a multistage extension of the model to situations where there are several populations and parasites can be transmitted from one of them to another. We conclude that animal husbandry situations with high stock densities will lead to very rapid increases in virulence, where virulent strains are either more transmissible or favoured by mutation. Further the process is affected by the nature of the farm enclosures.

1 Introduction

There is an extensive literature on mathematical modeling of infectious diseases. Both the rate of spread and the virulence of pathogens are important, and both the dynamics

*Department of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia; e-mail: borovkov@unimelb.edu.au

†Department of Zoology, The University of Melbourne, Parkville 3010, Australia; e-mail: r.day@unimelb.edu.au

‡Department of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia; e-mail: t.rice@ms.unimelb.edu.au

of disease spread and the evolution of virulence in parasite-host systems have attracted much attention. The first mathematical model of disease dynamics apparently goes back to the 1760 paper by D. Bernoulli [15] (see also pp. 2–6 in [21]). A major step in further development of deterministic modelling was the "threshold theorem" [51] showing that the initial frequency of susceptibles must exceed a certain threshold value to give rise to an epidemic. The first stochastic epidemiological models appeared in the late 1920s [58], applying the "law of mass action" suggested in [15] to probabilistic description of the epidemics. Stochastic models became much more popular after Bartlett [10] formulated a model for the general stochastic epidemic by analogy with the deterministic model from [51]. Anderson and May [4] introduced the first model of evolutionary change in pathogen virulence in 1982, which explained the decline in virulence of myxomatosis, introduced to control Australian rabbits, in terms of optimal transmission of the parasite between hosts. Subsequent work, notably [29, 59, 53, 54, 34], has expanded on this classic paper, to cover the effects of a vector of the pathogen, vaccination of hosts, and various other issues. For more extensive surveys of the field and literature reviews, we refer the reader to the monographic literature [5, 12, 21, 6, 63] and also to papers [3, 27, 46, 28, 25].

In natural systems, the density of host organisms declines when a pathogen produces an epidemic. Almost all the models referred to above have considered such natural populations of hosts, where the the density of susceptible hosts is reduced by the pathogen. They show that as host density is reduced, less virulent strains of pathogen become more viable, such that infected hosts live longer, and thus tend to pass the pathogen to more susceptible new hosts, maintaining the epidemic. While hosts will have longer generation times than their pathogens, they too, will evolve to be more resistant, further reducing the virulence of the pathogen for these hosts. On this basis one might expect pathogens to be highly virulent only when they have recently switched hosts, such as the SARS coronavirus that was probably transferred to humans from animals such as Himalayan civet cats [40] or when the plague bacterium, *Yersina pestis*, has been transferred to humans from rodents [72]. Various reasons, such as low host resistance, have been suggested to explain high initial virulence of transferred pathogens [18].

Most models have used a game-theoretic approach (see e.g. [4, 22, 31]) and focused on the equilibrium level of both host density and pathogen virulence in natural host populations infected by a pathogen. But a number of papers have noted that the initial epidemic phase of pathogen infection may select for high virulence [54, 31, 25, 19, 18]. In particular, [24, 25] present a general theory of evolution of virulence that is capable of predicting both the short- and long-term evolution of virulence.

In animal husbandry as opposed to natural populations, epidemics of new pathogens are not left to reduce the densities of the stock. Instead, antibiotics are applied, hosts are destroyed, or other measures are used in an attempt to eliminate the pathogen, once it is recognised that a serious pathogen epidemic is spreading. If such pathogens persist, they must do so through resistant stages and some mechanism of transfer to new enclosures of domestic stock or crops, where the initial epidemic stage is repeated. The initial stages of each epidemic will typically occur at very high host density. This

is the situation we wish to model.

The main objective of the present paper is to provide quantitative and also qualitative (cf. the discussion of the above-mentioned threshold theorem and other findings from [51] on pp. 11 and 29 in [21]) insights into the dynamics of pathogen populations during the initial stages of epidemics, assuming that in this period the virulence of the pathogens can change due to mutation. More specifically, we are interested in analyzing the effects that host density may have on the virulence of the pathogen.

In animal husbandry, stock are commonly maintained at very high densities. Thus the question arises as to whether high host densities are likely to lead to selection for more virulent strains of pathogens during the initial stages of epidemics. This question has been investigated only recently [25, 18]. Both these studies use models of the epidemiological dynamics of natural host populations and the evolution of the pathogen, incorporating an assumed tradeoff between virulence (the extra mortality due to infection) and transmissibility (the probability of transmission from an infected to a susceptible host in contact with it). We propose a novel approach to this question.

The mathematical models we discuss in the present paper incorporate a continuous time Markovian multitype branching process. Branching processes first appeared in epidemiological context in the mid-1950s: one can mention here the paper [16] and also refer to [11, 50, 70]. For further literature review, we again refer the reader to the above-listed monographic and survey literature, while some examples of applications of multitype branching processes in epidemiology can be found in [13, 62, 71, 45, 1].

Branching processes are relatively simple and well-understood mathematical models; for treatments of the theory of branching processes, see e.g. [47, 8], a recent addition to the monographic literature on branching processes prepared specifically for biological audience being [41]. At the same time, the processes were shown to be good approximations to the general stochastic epidemic models at the initial stage (when the total population size is large and the initial number of infectives is small) and also at the final stages of the epidemics (see e.g. [57, 9, 46]). As our main objective is not to model the detailed dynamics and describe the whole history of the epidemic itself, but rather to suggest a model for changes in the pathogen's population composition depending on the density of the host population, we will restrict ourselves to working with the simpler branching process models.

One can also envisage an extension of our simple analysis to the general stochastic models as the same mechanism will certainly work for the latter as well. However, such extensions will be much less tractable analytically and may lead to no closed-form answers.

Moreover, the most critical stage of an epidemic is the initial one, when it is basically determined if there will be a large-scale event or the epidemic will die out. And as branching processes are good approximations to the general stochastic epidemic models at the initial stage, the threshold analysis aimed to determine if the "basic reproductive number" (defined roughly as the expected number of secondary cases produced by one infected individual) is greater than one (which implies the danger of an outburst or persistence of endemic levels) can be carried out using those models (see e.g. Chapters 6 and 8 in [63]).

In our analysis, we will look at supercritical two-type branching processes (so that the basic reproductive number will be greater than one: we are interested in what happens when there exists a threat of epidemics) and then look at the behaviour of the ratio of the sizes of the subpopulations in the process (representing two versions of the pathogen, that can mutate into each other). This quantity can be used to determine which of the two types will become dominant in the population over time.

As we are interested in stock or crop systems of high host density (not regulated by the parasite, cf. [54], we assume natural mortality of the hosts is zero and that host density is effectively not changed by the infection during the period of interest. We define the lethality of the pathogen as the mean survival time of an infected host (which corresponds to the usual definition of virulence as the extra mortality due to infection [22, 25, 18], and the transmissibility as the probability of an infected host infecting a susceptible one if a contact occurs (which corresponds to the transmission rate per susceptible host in many published models [18]). For recent discussions of the interplay between these two characteristics, see [38, 56]. The theory of a tradeoff between these postulates that high virulence (which involves higher host exploitation) requires high transmission, but this theory is still debated [2, 18].

In the paper, we use our approach to model the dynamics of a host/parasite population where parasites can be of one of two types that can differ in their lethality and transmissibility. We do not assume there is any necessary relation between these, as in most models (see [18]). The underlying simple continuous time Markov model of a two-type branching process is presented in Section 2. The analysis of the model and derivation of the dynamics for the mean functions are given in Section 3 and used in Section 4 to establish the eventual composition of the pathogen population. It turns out that our model does show changes in the overall parasite's lethality in response to increased density of the hosts. This is not dependent on any relationship between the lethality and transmissibility of a pathogen strain, as assumed in many previous models.

One of the key parameters of our model is the contact rate for hosts, which clearly depends on the structure of the host population and, provided that the hosts populate a certain region and are moving within it, on the size of that region and the hosts speed and character of movement. One way of specifying contact rates is to assume that hosts occupy sites in a contact network, and a number of publications have discussed spatially structured host-parasite dynamics basing on that standard assumption. Approaches used include the simulation of probabilistic cellular automata (see e.g. [48]), the derivation of modified versions of SIS/SIRS differential equations [55] and so-called correlation dynamics (see e.g. [68] and references therein).

However, using contact networks does not seem appropriate in the situations we envisage, where branching process models may be applicable, as the dynamics assumes host populations that are constantly mixing. Moreover, instead of stipulating contact rates (by means of contact networks), for such populations, the natural question that arises is how the contact rate will be determined by more basic parameters such as hosts speed, the character of their movement etc. In Section 5, assuming chaotic independent movements of the hosts, we present a few remarks on how the change in the size of the

enclosure a given population of hosts inhabits can affect the "encounter rate" for the hosts — the key parameter of the model describing the "effective density" in the host population.

Finally, in Section 6 we consider a multi-stage modification of our model that can be used to analyse farm or aquaculture situations in which there are many enclosures or tanks of animals, and an outbreak of an infectious disease occurs in one of them. We assume that the pathogen might at some stage be transmitted to the next enclosure, where the epidemic process starts anew etc., and discuss possible scenarios of the development of the epidemic in the farm. Section 7 contains a few final remarks concerning biological interpretation of our results.

2 Description of the branching process model

Assume that we have a large population of hosts that can be infected by parasites of one of two types that will be denoted by T_1 and T_2. The pathogen types can differ in both their lethality and transmission rate. The numbers of infected hosts at time t are represented by the vector

$$\boldsymbol{Z}(t) = (Z_1(t), Z_2(t)), \quad t \geq 0,$$

$Z_i(t)$ being the time t number of hosts infected with the type T_i pathogen (for simplicity, at any given time, any given host can be infected with one type of the pathogen only). We do not keep track of the number of hosts that remain uninfected (susceptible), assuming instead that this number will remain large enough during the time period for which our mathematical model is intended, so that the dynamics of the process $\{\boldsymbol{Z}(t),\, t \geq 0\}$ to be described do not change over time.

A general description of the model is as follows. We assume that each infected individual lives a random time (which will tend to be shorter when one is infected with the "more lethal" of the two pathogen types). During its lifetime, an infected host can encounter susceptible hosts and, with a probability depending on the type of the pathogen it carries, transmit the parasite to them. The rate of such (random) encounters will be specified by a special parameter that we can vary in order to model changes in the density of the host population.

Finally, we also allow the pathogens to mutate, so that when a host originally infected with T_1 encounters a susceptible host, the latter can become infected with T_2-type parasites (and the other way around).

Now we will present a formal mathematical model. First recall that the exponential distribution with rate (or intensity) $\alpha > 0$ has density of the form

$$p(t) = \alpha e^{-\alpha t}, \qquad t > 0, \tag{1}$$

with mean $1/\alpha$, and plays a special role in probability theory due to its unique memoryless property that makes the distribution ubiquitous in the theory of Markov processes. Namely, a random variable $\tau > 0$ modelling, say, the time of the first encounter of a given infected host with a healthy one, has this property if, for any $s, t > 0$, the

conditional probability of the event $\{\tau > s + t\}$ given that $\tau > s$ coincides with the probability of $\{\tau > t\}$:

$$\Pr(\tau > s + t \,|\, \tau > s) \equiv \frac{\Pr(\tau > s + t,\, \tau > s)}{\Pr(\tau > s)} \equiv \frac{\Pr(\tau > s + t)}{\Pr(\tau > s)} = \Pr(\tau > t). \quad (2)$$

In words, if, at time s we know that there has been no such encounter, then the conditional distribution (given that information) of the residual random time $\tau - s$ till the encounter will be the same as the original distribution of τ. It is obvious that if τ has density (1) then

$$\Pr(\tau > t) = \int_t^\infty p(s)ds = e^{-\alpha t}, \qquad t > 0, \quad (3)$$

and so the property (2) is clearly satisfied.

An equivalent formulation of the property can be given in terms of the distribution's hazard rate $r_\tau(s)$ that quantifies the probability that, given there has been no encounter up to time s, there will be one "immediately afterwards": in case a random variable τ has a continuous density p_τ, the hazard rate is defined by

$$r_\tau(s) = \frac{p_\tau(s)}{\Pr(\tau > s)} \equiv -\frac{d}{ds} \ln \Pr(\tau > s), \qquad s > 0, \quad (4)$$

and, as $h \downarrow 0$,

$$\Pr(\tau \leq s + h |\, \tau > s) = r_\tau(s)h + o(h), \qquad s > 0, \quad (5)$$

where, as usual, $o(h)$ denotes a quantity that vanishes faster than h: $o(h)/h \to 0$.

It is obvious from (3) and (4) that the hazard rate of a distribution is constant if and only if it is exponential (in that case, the hazard rate simply equals the distribution's rate). In applications, one often uses exponentially distributed random variables to model times between successive events of a particular kind and also lifetimes. This is because, on the one hand, such assumptions make sense from the modelling view point (in a large population, meeting a new individual during a time interval $[t, t+h]$, $h > 0$, can scarcely depend on one's "life history" prior to time t) and, on the other hand, as the resulting models are usually Markovian, so that the powerful machinery of the theory of Markov processes is applicable.

Our basic model assumptions are as follows:

(a) The initial population contains z_i hosts infected with parasites of type T_i, $i = 1, 2$.

(b) A host infected with type T_i parasites lives a random time which is exponentially distributed with parameter $\alpha_i > 0$, $i = 1, 2$. Clearly, the pathogen with a higher rate α_i will be the more lethal one, as the mean lifetime in that case will be lower.

(c) Any infected host can encounter susceptible ones. The time till the first encounter of a given infected host (of any type) with a susceptible host is a random variable exponentially distributed with rate λ. It is clear from the above discussion of the

properties of exponential distributions that, at any time t, the residual times till encounters of the current infected hosts with susceptible hosts are all exponential with the same rate λ. A similar observation applies to the residual lifetimes; this ensures that the process $\{\boldsymbol{Z}(t)\}$ will be Markovian: given the current state of the process, its future evolution does not depend on the past one.

In a modification of the model, one can assume that, at time t, the time till the first encounter of a given infected host (of any type) with a susceptible host has a hazard rate $\lambda(t)$ which can depend on t. This will enable one to model changes in the density of the host population that occur over time (the higher λ, the more often are encounters, which corresponds to higher host density situations). The process will remain Markovian, but will become time-inhomogeneous.

(d) At any encounter with susceptible hosts, a T_i-infected host meets only one susceptible host (there can be several such encounters during the host's lifetime). At each such instance, the T_i-infected host transmits the parasites to the susceptible one with probability β_i (so that, with the complement probability $1 - \beta_i$, the encounter will have no consequences for the susceptible host).

(e) Mutations $T_1 \leftrightarrow T_2$ are possible. A host infected with T_1-type pathogens will remain such till its end, but, when transmitting pathogens to a susceptible host during an encounter, the newly infected host will carry T_2-type pathogens with a probability μ_1. Likewise, μ_2 denotes the probability that a successful transmission of parasites from a T_2-infected host to a susceptible one resulted in making the latter T_1-infected.

(f) All the above-mentioned random times (lifetimes, times till the first encounter) are independent of each other.

Of course, the above assumptions oversimplify the real biological processes. There are likely to be several or many strains of pathogen, and the probability of mutation from one to another will vary. We suggest that simplifying the system to two strains is likely to retain the same key dynamic features. Further, the distributions of times until events occur are likely to be approximately exponential as argued earlier, and we do not see any reasons why host survival times and encounter rates should depend in any way on each other. Note that "encounters" between hosts are simply occasions where a pathogen can be transferred between hosts. They do not have to involve physical contact. Thus a pathogen transferred by aerosols might be transferred between pigs that are isolated in neighbouring stalls.

The diagram in Fig. 1 illustrates the "physical" meaning of our assumptions. The two horizontal segments represent the mean lifetimes of hosts infected by the pathogens of our two types: the longer segment (of length $1/\alpha_1$) corresponds to T_1 which we assume less lethal by stipulating that $\alpha_2 > \alpha_1$, whereas the shorter one (of length $1/\alpha_2$) corresponds to T_2.

When the host population density is relatively low (say, represented by the value $\lambda = \lambda''$, as depicted), the lifetime of T_2-infected hosts will be too short compared to

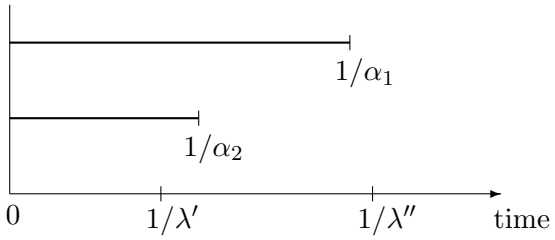


Figure 1: Mean lifetimes of infected hosts and mean times to encounter.

the mean time $1/\lambda''$ between encounters to give them a good opportunity to encounter susceptibles and hence further propagate in the host population. One may expect that this will result in the eventual prevalence of T_1 pathogens who have better chance of being transmitted as they live longer. However, if the host population density increases (say, to $\lambda = \lambda'$, as depicted), then the more lethal type T_2 may have frequent enough encounters which, combined with its higher transmissibility, can lead to its eventual prevalence. As we will see in Section 3, the above argument is confirmed by mathematical analysis.

Assumptions (a)–(f) imply that our process $\{\boldsymbol{Z}(t)\}$ is actually a two-type time homogeneous Markov branching process in continuous time, see e.g. Section V.7 in [8]. That is, $\{\boldsymbol{Z}(t)\}$ describes the dynamics of a population consisting of individuals (or "particles") of two types, 1 and 2. The transitions of different particles in the process are assumed to be independent. A particle of type i lives for an exponentially random time with rate a_i. At the end of its life it disappears. It can either

(i) simply disappear (in terms of our modelling assumptions above, this means that a given T_i infected host died having never encountered a susceptible host), or

(ii) produce one particle of the same type (meaning: there was an encounter, but no transmission occurred); we think of the "newly produced" particle of type i as just the "old" infected host of type T_i who keeps living — note that, due to the memoryless property of the exponential distribution, such an identification of a "new" particle with the "old" one is in agreement with our assumptions (a)–(f) (so that the life of one T_i-infected host can actually be represented by a succession of several type i particles, for which reason we do not use T_i to denote the type of particles in the branching process), or

(iii) produce two particles of the same type i (meaning: there was an encounter and successful transmission, but no mutation; one of the "newly produced" particles is actually the original host, the other represents a newly infected — with the same T_i-type pathogen — host), or

(iv) produce two particles of different types (meaning: there was an encounter and successful transmission and mutation; one of the "newly produced" particles is actually the original host, the other represents a newly infected — with the pathogen of the other type — host).

To see that both sets of assumptions (a)–(f) and (i)–(iv) describe the same dynamics of $\{\boldsymbol{Z}(t)\}$, it suffices to note that in both cases we deal with time-homogeneous continuous time jump Markov processes (which follows from the exponentiality assumptions) and then verify that, choosing suitable parameter values for the second model, one can obtain the same transition rates as for the first one.

To do that, we first observe that assumptions (i)–(iv) imply the branching property which means that, for any $s \geq 0$, given $\boldsymbol{Z}(s) = (z_1, z_2)$, the future $\{\boldsymbol{Z}(s+t),\, t > 0\}$ of the process will follow the same probability laws as that of the sum of z_1 independent copies of $\boldsymbol{Z}(t)$ starting at time 0 with a single particle of type 1 and z_2 independent copies of $\boldsymbol{Z}(t)$ starting at time 0 with a single particle of type 2. It is clear that the first model (specified by (a)–(f)) has the same property.

Moreover, the branching property implies that to completely describe the evolution of the process, it suffices to specify transition probabilities

$$p_i^{(h)}(j_1, j_2) = \Pr\big(\boldsymbol{Z}(h) = (j_1, j_2) |\, \boldsymbol{Z}(0) = \boldsymbol{e}_i\big)$$

from the basic initial states

$$\boldsymbol{e}_1 = (1, 0) \qquad \text{and} \qquad \boldsymbol{e}_2 = (0, 1)$$

for arbitrary small time increments h (for more detail, see Chapter V in [8]). It is obvious that the probability of having more than one transition during a small time interval $(0, h)$ will be $o(h)$, so we just need to consider where a single transition can take the process from a basic state $\boldsymbol{e}_i$ according to assumptions (a)–(f) and show that the transitions will have the same rates as for a process specified by (i)–(iv) (for a suitable choice of parameter values).

Suppose that, in our branching process, particles of type i have exponentially distributed random lifetimes with rates $a_i = \alpha_i + \lambda$, $i = 1, 2$. Moreover, at the end of a type i particle's life, it produces a random number of children (possibly of both types) according to the offspring distributions $q_i(j_1, j_2) = \Pr\big(\text{a particle of type } i \text{ gives birth to } j_1 \text{ particles of type 1 and } j_2 \text{ ones of type 2}\big)$ given in Table 1.

Then, due to independence and (5),

$$\begin{aligned} p_1^{(h)}(0,0) &= \Pr\big(\text{initial type 1 particle dies in } (0,h), \text{ no children produced}\big) \\ &= \Pr\big(\text{initial type 1 particle dies in } (0,h)\big) \times q_1(0,0) \\ &= ((\alpha_1 + \lambda)h + o(h)) \times \frac{\alpha_1}{\lambda + \alpha_1} = \alpha_1 h + o(h), \end{aligned}$$

which is clearly the same as the probability for a single T_1-infected host (from the first model) to die during $(0, h)$.

Likewise,

$$\begin{aligned} p_1^{(h)}(2,0) &= \Pr\big(\text{initial type 1 particle dies in } (0,h), \text{ producing 2 children of type 1}\big) \\ &= \Pr\big(\text{initial type 1 particle dies in } (0,h)\big) \times q_1(2,0) \\ &= ((\alpha_1 + \lambda)h + o(h)) \times \frac{(1-\mu_1)\beta_1\lambda}{\lambda + \alpha_1} = (1-\mu_1)\beta_1\lambda h + o(h). \end{aligned}$$

(j_1, j_2)	$q_1(j_1, j_2)$	$q_2(j_1, j_2)$
$(0,0)$	$\dfrac{\alpha_1}{\lambda+\alpha_1}$	$\dfrac{\alpha_2}{\lambda+\alpha_2}$
$(1,0)$	$\dfrac{(1-\beta_1)\lambda}{\lambda+\alpha_1}$	0
$(2,0)$	$\dfrac{(1-\mu_1)\beta_1\lambda}{\lambda+\alpha_1}$	0
$(1,1)$	$\dfrac{\mu_1\beta_1\lambda}{\lambda+\alpha_1}$	$\dfrac{\mu_2\beta_2\lambda}{\lambda+\alpha_2}$
$(0,2)$	0	$\dfrac{(1-\mu_2)\beta_2\lambda}{\lambda+\alpha_2}$
$(0,1)$	0	$\dfrac{(1-\beta_2)\lambda}{\lambda+\alpha_2}$

Table 1: Postulated offspring distributions in the branching process model (i)–(iv). These distributions result in transition probabilities as specified in (a)–(f).

The corresponding transition in the first model is as follows: a T_1-infected host did not die during $(0,h)$, but met a susceptible; the pathogen was transmitted, no mutation occurred. Due to independence, the probability of this will be

$$(1+O(h)) \times (\lambda h + o(h)) \times \beta_1 \times (1-\mu_1) = (1-\mu_1)\beta_1\lambda h + o(h).$$

Virtually the same argument shows that

$$\begin{aligned} p_1^{(h)}(1,1) &= \Pr\big(\text{initial type 1 particle dies in } (0,h), \text{ producing one child of each type}\big) \\ &= \Pr\big(\text{initial type 1 particle dies in } (0,h)\big) \times q_1(1,1) \\ &= ((\alpha_1+\lambda)h + o(h)) \times \frac{\mu_1\beta_1\lambda}{\lambda+\alpha_1} = \mu_1\beta_1\lambda h + o(h), \end{aligned}$$

coincides with the probability for a T_1-infected host to meet a susceptible and transmit the mutated form of the pathogen.

Finally, given $\boldsymbol{Z}(0) = \boldsymbol{e}_1$, the most likely state for $\boldsymbol{Z}(h)$ is $\boldsymbol{e}_1$. In our branching process this occurs when either the original particle lives through the time interval (with probability $1-(\alpha_1+\lambda)h+o(h)$) or it dies prior to h producing a single type 1 particle (of which the probability is $((\alpha_1+\lambda)h+o(h)) \times (1-\beta_1)\lambda/(\lambda+\alpha_1) = (1-\beta_1)\lambda h + o(h)$), so that

$$p_1^{(h)}(1,0) = 1-(\alpha_1+\lambda)h + (1-\beta_1)\lambda h + o(h) = 1-(\alpha_1+\beta_1\lambda)h + o(h).$$

In the first model, to get to this state, the initial T_1-infected host survives for h time units and either has no encounters with susceptibles or has one, but without transmission of a pathogen. The probability of this is

$$\begin{aligned} &(1-\alpha_1 h + o(h)) \times [(1-\lambda h + o(h)) + ((1-\beta_1)\lambda h + o(h))] \\ &\qquad = (1-\alpha_1 h + o(h)) \times (1-\beta_1\lambda h + o(h)) = 1-(\alpha_1+\beta_1\lambda)h + o(h), \end{aligned}$$

(j_1, j_2)	$p_1^{(h)}(j_1, j_2)$	$p_2^{(h)}(j_1, j_2)$
$(0,0)$	$\alpha_1 h$	$\alpha_2 h$
$(1,0)$	$1-(\alpha_1+\beta_1\lambda)h$	0
$(2,0)$	$(1-\mu_1)\beta_1\lambda h$	0
$(1,1)$	$\mu_1\beta_1\lambda h$	$\mu_2\beta_2\lambda h$
$(0,2)$	0	$(1-\mu_2)\beta_2\lambda h$
$(0,1)$	0	$1-(\alpha_2+\beta_2\lambda)h$

Table 2: Transition probabilities for small h values ($o(h)$ terms omitted).

the same value as for the branching process model.

Of course, we could also work out the probability for $(1,0)$ just as the complement probability

$$p_1^{(h)}(1,0) = 1 - \big(p_1^{(h)}(0,0) + p_1^{(h)}(2,0) + p_1^{(h)}(1,1)\big) + o(h),$$

but the presented argument demonstrates the difference between the interpretation of the elements of the two models (T_i-infected hosts in the first model and type i particles in the second one are not the same, as we noted earlier).

The same calculations are applicable in the case of the initial state $\boldsymbol{e}_2$, which leads us to the transition probabilities presented in Table 2 (for $h \to 0$, the additive terms $o(h)$ being omitted for brevity).

3 The dynamics of the means

Consider the matrix of mean values $\boldsymbol{M}(t) = (M_{ij}(t))$, where

$$M_{ij}(t) = \mathbb{E}(Z_j(t) |\, \boldsymbol{Z}(0) = \boldsymbol{e}_i), \qquad i,j = 1,2,$$

is the expected number of type j particles present in the process at time t given that the process started at time 0 with a single particle of type i. Clearly,

$$\boldsymbol{M}(0) = \boldsymbol{I} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the identity matrix, and, using the branching and Markov properties of $\{\boldsymbol{Z}(t)\}$, it is easy to demonstrate that $\{\boldsymbol{M}(t),\, t \geq 0\}$ possesses the operator semigroup property:

$$\boldsymbol{M}(s+t) = \boldsymbol{M}(s)\boldsymbol{M}(t) \qquad \text{for any} \quad s,t \geq 0. \tag{6}$$

Indeed, given that $\boldsymbol{Z}(0) = (z_1, z_2)$, the value of $\boldsymbol{Z}(s)$ is just the sum of z_1 independent copies of $\boldsymbol{Z}(s)$ starting at time 0 with a single particle of type 1 and z_2 independent

copies of $\boldsymbol{Z}(s)$ starting at time 0 with a single particle of type 2. As the process is time-homogeneous, we infer that

$$\mathbb{E}\big(\boldsymbol{Z}(s+t)\big|\,\boldsymbol{Z}(s)=(z_1,z_2)\big)=\sum_{i=1}^{2} z_i\mathbb{E}(\boldsymbol{Z}(t)|\,\boldsymbol{Z}(0)=\boldsymbol{e}_i)=(z_1,z_2)\boldsymbol{M}(t). \tag{7}$$

From here, using the Markov property and the double expectation law for conditional expectations, we have

$$\begin{aligned}\mathbb{E}\big(\boldsymbol{Z}(s+t)\big|\,\boldsymbol{Z}(0)=\boldsymbol{e}_i\big)&=\mathbb{E}\big[\mathbb{E}(\boldsymbol{Z}(s+t)\big|\,\boldsymbol{Z}(s))\,\big|\,\boldsymbol{Z}(0)=\boldsymbol{e}_i\big]\\&=\mathbb{E}\big[\boldsymbol{Z}(s)\boldsymbol{M}(t)\big|\,\boldsymbol{Z}(0)=\boldsymbol{e}_i\big]=(M_{i1}(s),M_{i2}(s))\boldsymbol{M}(t),\end{aligned}$$

which is equivalent to (6).

Relation (6) implies (cf. p. 202 in [8]) that one has the matrix exponential representation

$$\boldsymbol{M}(t)=e^{t\boldsymbol{A}}\equiv\sum_{k=0}^{\infty}\frac{t^k\boldsymbol{A}^k}{k!},\qquad t\geq 0, \tag{8}$$

where $\boldsymbol{A}^0=\boldsymbol{I}$ and

$$\boldsymbol{A}=\lim_{h\downarrow 0}\frac{1}{h}(\boldsymbol{M}(t)-\boldsymbol{I}) \tag{9}$$

is the so-called infinitesimal generator of $\{\boldsymbol{M}(t)\}$. Indeed, from (6) one has $\boldsymbol{M}(t+h)-\boldsymbol{M}(t)=(\boldsymbol{M}(h)-\boldsymbol{I})\boldsymbol{M}(t)$, so that (9) implies that $\dfrac{d}{dt}\boldsymbol{M}(t)=\boldsymbol{A}\boldsymbol{M}(t)$, $\boldsymbol{M}(0)=\boldsymbol{I}$, for which (8) is clearly a solution, as seen from its term-wise differentiation.

Evaluating matrix exponentials is rather straightforward: it basically reduces to calculating the values of the function on the matrix' spectrum (for more detail on functions of matrices, see e.g. Chaper V in [35]). If, say, $\boldsymbol{A}$ is diagonalisable, so that there exists an invertible matrix $\boldsymbol{Q}$ (with inverse $\boldsymbol{Q}^{-1}$: $\boldsymbol{Q}^{-1}\boldsymbol{Q}=\boldsymbol{Q}\boldsymbol{Q}^{-1}=\boldsymbol{I}$) such that

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{D}\equiv\operatorname{diag}\{\sigma_+,\sigma_-\}\equiv\begin{bmatrix}\sigma_+ & 0\\ 0 & \sigma_-\end{bmatrix} \tag{10}$$

for some $\sigma_\pm$ (which is the case in our situation, as we will see below), then clearly $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1}$,

$$\boldsymbol{A}^2=(\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1})^2=\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1}=\boldsymbol{Q}\boldsymbol{D}^2\boldsymbol{Q}^{-1}=\boldsymbol{Q}\operatorname{diag}\{\sigma_+^2,\sigma_-^2\}\boldsymbol{Q}^{-1}$$

and so on: $\boldsymbol{A}^k=\boldsymbol{Q}\boldsymbol{D}^k\boldsymbol{Q}^{-1}=\boldsymbol{Q}\operatorname{diag}\{\sigma_+^k,\sigma_-^k\}\boldsymbol{Q}^{-1}$. We obtain

$$e^{t\boldsymbol{A}}=\sum_{k=0}^{\infty}\frac{t^k(\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1})^k}{k!}=\boldsymbol{Q}\sum_{k=0}^{\infty}\frac{t^k\operatorname{diag}\{\sigma_+^k,\sigma_-^k\}}{k!}\boldsymbol{Q}^{-1}=\boldsymbol{Q}\operatorname{diag}\{e^{t\sigma_+},e^{t\sigma_-}\}\boldsymbol{Q}^{-1}. \tag{11}$$

Thus, to derive the dynamics of the means, we just have to compute the generator $\boldsymbol{A}$, which can easily be done using Table 2. Indeed, we infer from the table that,

as $h \to 0$,

$$\begin{aligned} M_{11}(h) &= 0 \times \alpha_1 h + 1 \times [(1 - (\alpha_1 + \beta_1\lambda)h) + \mu_1\beta_1\lambda h] + 2 \times (1 - \mu_1)\beta_1\lambda h + o(h) \\ &= 1 + ((1 - \mu_1)\beta_1\lambda - \alpha_1)h + o(h); \\ M_{12}(h) &= 0 \times \alpha_1 h + 1 \times \mu_1\beta_1\lambda h + o(h) = \mu_1\beta_1\lambda h + o(h) \end{aligned}$$

and similarly for $M_{2i}(h)$. From here and (9) we immediately obtain

$$\boldsymbol{A} = \begin{bmatrix} \gamma_1 & \delta_1 \\ \delta_2 & \gamma_2 \end{bmatrix}, \qquad \gamma_k = (1 - \mu_k)\beta_k\lambda - \alpha_k, \quad \delta_k = \mu_k\beta_k\lambda, \quad k = 1, 2.$$

Now solving the characteristic equation $\det(\boldsymbol{A} - \sigma\boldsymbol{I}) = 0$ for σ we find the eigenvalues $\sigma_\pm$ of $\boldsymbol{A}$ given by

$$\sigma_\pm = \frac{1}{2}\big(\gamma_1 + \gamma_2 \pm \Delta\big), \qquad \Delta = \sqrt{(\gamma_1 - \gamma_2)^2 + 4\delta_1\delta_2}$$

(cf. similar calculations of the threshold parameter for a somewhat different two-type model in Section 8.4 of [63]), with the respective (right) eigenvectors

$$\boldsymbol{u}_\pm = \begin{bmatrix} u_\pm \\ 1 \end{bmatrix}, \qquad u_\pm = \frac{\gamma_1 - \gamma_2 \pm \Delta}{2\delta_2} \gtrless 0.$$

The eigenvalues $\sigma_\pm$ are clearly different from each other (in any case, this is guaranteed by the fact that σ_+ is the Perron-Frobenius root for the quasi-positive matrix $\boldsymbol{A}$, cf. Section A.8 in [67]), which ensures that $\boldsymbol{A}$ is diagonalizable and we can take the transformation matrix $\boldsymbol{Q}$ from (10) to be given by

$$\boldsymbol{Q} = (\boldsymbol{u}_+, \boldsymbol{u}_-) \equiv \begin{bmatrix} u_+ & u_- \\ 1 & 1 \end{bmatrix}, \qquad \boldsymbol{Q}^{-1} = \frac{1}{u_+ - u_-}\begin{bmatrix} 1 & -u_- \\ -1 & u_+ \end{bmatrix}.$$

Hence (8) and (11) imply that

$$\begin{aligned} \boldsymbol{M}(t) &= \frac{1}{u_+ - u_-}\begin{bmatrix} u_+ & u_- \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{\sigma_+ t} & 0 \\ 0 & e^{\sigma_- t} \end{bmatrix}\begin{bmatrix} 1 & -u_- \\ -1 & u_+ \end{bmatrix} \\ &= \frac{1}{u_+ - u_-}\begin{bmatrix} u_+e^{\sigma_+ t} - u_-e^{\sigma_- t} & u_+u_-(e^{\sigma_- t} - e^{\sigma_+ t}) \\ e^{\sigma_+ t} - e^{\sigma_- t} & u_+e^{\sigma_- t} - u_-e^{\sigma_+ t} \end{bmatrix}, \qquad t \geq 0. \end{aligned} \tag{12}$$

4 The asymptotic composition of the pathogen types

As $\sigma_+ > \sigma_-$, it is clear from (12) that σ_+ is the so-called Malthusian parameter of the branching process that determines the long-term behaviour of the branching process means. Note that this "long-term behaviour" still corresponds to the initial stage of the epidemic that we are aiming to model (even more so since the means of the branching population "settle" in that asymptotic composition exponentially fast, as shown below).

As we said in Section 1, the most interesting for us is the case of supercritical processes for which $\sigma_+ > 0$, implying unbounded exponential growth of the population (unless it becomes extinct at a pretty early stage). Otherwise, the process $\{\boldsymbol{Z}(t)\}$ would be doomed to die out very soon, so that no epidemic would arise.

It is clear that, in the supercritical case, the ratio of the time t expected number of type 2 particles to that of type 1 particles will be given by

$$R_1(t) \equiv \frac{M_{12}(t)}{M_{11}(t)} = \frac{u_+u_-(e^{\sigma_- t} - e^{\sigma_+ t})}{u_+e^{\sigma_+ t} - u_-e^{\sigma_- t}} = |u_-|\biggl[1 - \Bigl(1 - \frac{u_-}{u_+} + o(1)\Bigr)e^{-\Delta t}\biggr], \quad t \to \infty,$$

if the process starts with a single type 1 particle, and by

$$R_2(t) \equiv \frac{M_{22}(t)}{M_{21}(t)} = \frac{u_+e^{\sigma_- t} - u_-e^{\sigma_+ t}}{e^{\sigma_+ t} - e^{\sigma_- t}} = |u_-|\biggl[1 + \Bigl(1 - \frac{u_+}{u_-} + o(1)\Bigr)e^{-\Delta t}\biggr], \quad t \to \infty,$$

provided that it started with a type 2 particle (recall that $u_- < 0$ and $\sigma_+ - \sigma_- = \Delta > 0$).

Therefore, using (7), we see that, regardless of the initial state (z_1, z_2), the eventual ratio of the mean number of T_2-infected hosts to that for T_1-infected hosts will be one and the same quantity

$$R \equiv \lim_{t\to\infty} \frac{\mathbb{E}\bigl(Z_2(t)|\, \boldsymbol{Z}(0) = (z_1, z_2)\bigr)}{\mathbb{E}\bigl(Z_1(t)|\, \boldsymbol{Z}(0) = (z_1, z_2)\bigr)} = |u_-| \equiv \frac{\gamma_2 - \gamma_1 + \Delta}{2\delta_2},$$

which is a well-known fact from the theory of multitype branching processes (it follows, for instance, from Theorem 1 on p. 185 in [8], see also p. 203, *ibid.*). The convergence rate is clearly exponential: the remainder term decays as $e^{-\Delta t}$. Note also the obvious facts that $R_1(0) = 0$, $R_2(0) = \infty$ and that $R_1(t)$ ($R_2(t)$) is an increasing (decreasing) function of t (so that always $R_1(t) < R_2(t)$).

Thus the single value $R = R(\alpha_1, \alpha_2, \beta_1, \beta_2, \mu_1, \mu_2, \lambda)$ completely specifies the eventual balance of the mean numbers of individuals of different types in our supercritical process (whatever the initial values). This reflects a much deeper result on the long-term behaviour of $\{\boldsymbol{Z}(t),\, t \ge 0\}$ — namely, the fact that, with probability one, the scaled vector $e^{-\sigma_+ t}\boldsymbol{Z}(t)$ will converge, as $t \to \infty$, to a non-trivial random vector whose distribution is concentrated on the ray $\{r\boldsymbol{v},\, r \ge 0\}$ collinear to the (positive) left eigenvector $\boldsymbol{v}$ of $\boldsymbol{A}$ corresponding to the Perron-Frobenius eigenvalue σ_+ (see e.g. Theorem 2 on p. 206 in [8] and references therein). This implies that convergence to R holds not only for the ratio of the means, but for the random variables $Z_2(t)/Z_1(t)$ as well: if we denote by A the event $\{\boldsymbol{Z}(t) \ne \boldsymbol{0} \text{ for all } t > 0\}$, then

$$\lim_{t\to\infty} \frac{Z_2(t)}{Z_1(t)} = R \quad \text{on} \quad A$$

(up to an event of probability zero). In words, this means that either the branching process becomes extinct in finite time or the sizes of the subpopulations of individuals of the two types grow unboundedly in such a way that their ratio tends to R.

Observe also that the above shows how fast the composition of the population will change if the encounter rate λ switches to another value. Suppose that the initial value

is λ'. As we saw, after some (exponentially short) time, the balance of types in the process will establish around the value $R(\lambda') \equiv R(\alpha_1, \alpha_2, \beta_1, \beta_2, \mu_1, \mu_2, \lambda')$. Now if the value of λ quickly changes to λ'', then the population will re-establish balance at a new level $R(\lambda'')$ — again exponentially fast, with the rate characterized by the new value of Δ (provided, of course, that the process will still be supercritical, i.e. $\sigma_+ > 0$ for the new value of λ).

As we discussed earlier, the increase in the encounter rate λ ought to be beneficial for the parasite type with higher lethality as that increases its chances to spread in the host population. Indeed, we have

$$\frac{\partial R}{\partial \lambda} = \frac{(\alpha_2 - \alpha_1)(\gamma_2 - \gamma_1 + \Delta)}{2\delta_2 \Delta} > 0$$

since $\alpha_2 > \alpha_1$ by assumption and it is obvious that $|\gamma_2 - \gamma_1| < \Delta$. Fig. 2 displays the dependence of R on λ varying in (0,20), for different levels of the lethality α_2 (left pane) and mutation rate μ_2 (on the right) of the type 2 pathogen.

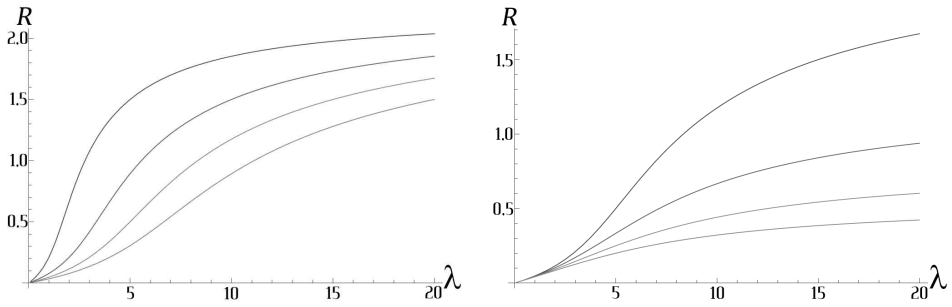


Figure 2: The plots of R as a function of $\lambda \in (0, 20)$. For the fixed common values $\alpha_1 = 0.5$, $\mu_1 = 0.2$, $\beta_1 = 0.3$ and $\beta_2 = 0.6$, the left pane displays the plots of R for four different lethality levels $\alpha_2 = 1, 1.5, 2$ and 2.5 for fixed $\mu_2 = 0.2$ (the lower the value of α_2, the higher the curve), whereas the right one shows the plots for different mutation rates $\mu_2 = 0.2, 0.3, 0.4$ and 0.6 (the higher the mutation rate, the lower the curve), for fixed $\alpha_2 = 2$.

In this example, our model corresponds to many previous models in that transmissibility is greater for the more virulent pathogen. The model shows the expected increase in the frequency of more lethal parasites in a host population when the density of the hosts increases. Observe that the threshold value $R = 1$ (after which type 2 pathogen dominates the population) may play no critical role: our model is a crude approximation for the initial stage of an epidemic only, so $R/(R+1)$ will just give the proportion of the carriers of type 2 parasite in the population of infected hosts at the end of that stage.

It is important to note that it is the "splitting" of the single "virulence" characteristic of the parasite into two (lethality and transmissibility) that made such a capability possible: if, say, there is no difference in lethality ($\alpha_1 = \alpha_2$) then, as a simple algebraic calculation shows, the value of R does not depend on density λ. The last observation we could have actually made earlier, as it follows from the model construction.

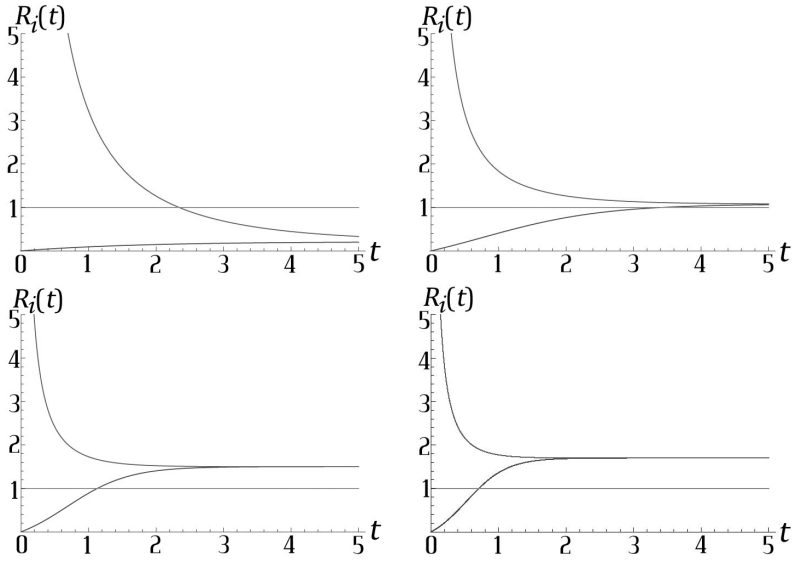


Figure 3: Convergence of $R_i(t)$ to R as $t \to \infty$. For a common set of parameter values, the plots display the behaviour of $R_i(t)$, $i = 1, 2$, for $\lambda = 2, 6, 10$ and 14 (from left to right, top to bottom), with the respective values $R \approx 0.210$, 1.076, 1.500 and 1.699. In all the cases, the process is supercritical ($\sigma_+ > 0$).

Figure 3 illustrates the stated exponentially fast convergence of the ratios $R_i(t)$ to a common limit R in four situations that have common parameter values $\alpha_1 = 0.5$, $\alpha_2 = 1.5$, $\mu_1 = \mu_2 = 0.2$, $\beta_1 = 0.3$, $\beta_2 = 0.6$, but different encounter rates. The plot in the top left corner displays the graphs of $R_1(t) < R_2(t)$ in the case when the encounter rate $\lambda = 2$ is small enough to allow type 1 parasite to dominate ($R \approx 0.210$). On the top right plot, we see that, for $\lambda = 6$, there establishes a rough balance ($R \approx 1.076$), whereas on the plots in the second raw we see type 2 parasites to gain dominance pretty fast (which is due to higher values of $\Delta = \sigma_+ - \sigma_-$), with the limiting values $R \approx 1.500$ and 1.699, resp.

The character of the dependence of R on the transmission probabilities is illustrated in Fig. 4. For four different values of the encounter rate ($\lambda = 2, 6, 10$ and 14), the figure shows the plots of R as a function of (β_1, β_2), restricted to the regions where the process is supercritical (i.e. $\sigma_+ > 0$). The values of the other parameters are $\alpha_1 = 0.5$, $\alpha_2 = 1.5$, $\mu_1 = \mu_2 = 0.2$. As one could expect, the value of R strongly depends on (β_1, β_2) and is an increasing function of β_2 and a decreasing one of β_1.

In all four cases presented in Fig. 4 the threshold value $R = 1$ is exceeded (that is, the more virulent pathogen becomes dominant) only when the transmissibility of type 2 pathogen is greater than that for type 1 ($\beta_2 > \beta_1$), so it may appear that that inequality is a necessary condition for type 2 to prevail. This, however, is not true: it turns out that a higher mutation rate from type 1 to type 2 can compensate for some lack of transmissibility. Figure 5 shows the plots of R as a function of the mutation probabilities $(\mu_1, \mu_2) \in (0, 1)^2$ for different encounter rates, all other parameters being fixed and common, with the transmission probability for type 1 being double that for type 2 ($\beta_1 = 2\beta_2 = 0.4$). On the left plot corresponding to $\lambda = 4$, the maximum

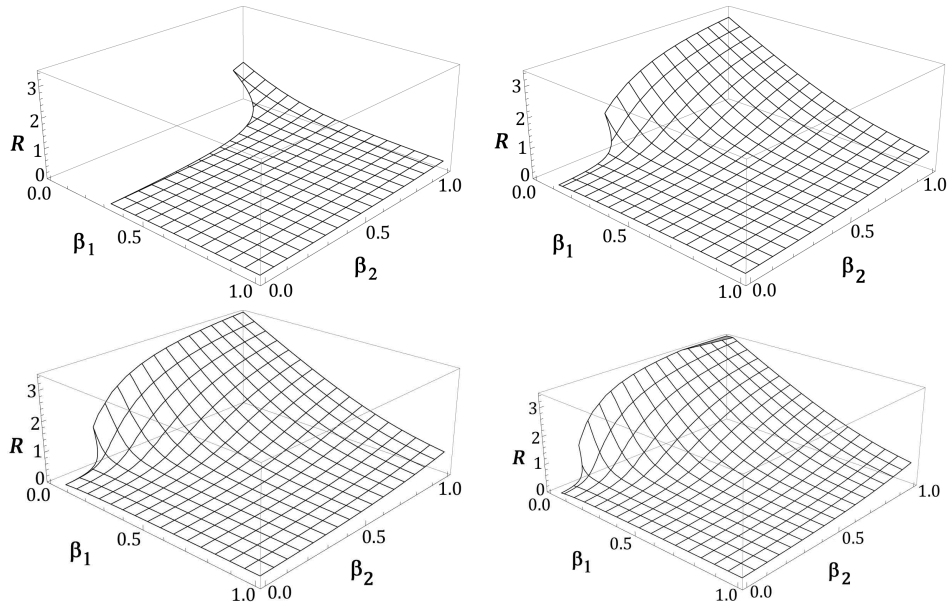


Figure 4: The plots of the limiting value R as a function of the transmissibility probabilities (β_1, β_2) in four cases: $\lambda = 2, 6, 10$ and 14 (from left to right, top to bottom), when all the other parameters of the model are common ($\alpha_1 = 0.5$, $\alpha_2 = 1.5$, $\mu_1 = \mu_2 = 0.2$). The plots are restricted to the regions where the respective processes are supercritical.

value of R barely exceeds 0.5, whereas on the right one, due to the increase in the encounter rate to $\lambda = 7$, not only the supercriticality region is much bigger, but also the maximum value is $R = 3$. We see type 2 parasite's domination in the region where the mutation rate μ_1 (from type 1) is high enough, while μ_2 is relatively small.

Finally, we turn to the dependence of R on the lethalities α_j. As one can easily see,

$$\frac{\partial R}{\partial \alpha_1} = \frac{1}{2\delta_2}\Big(1 - \frac{\gamma_1 - \gamma_2}{\Delta}\Big) > 0, \qquad \frac{\partial R}{\partial \alpha_2} = -\frac{\partial R}{\partial \alpha_1} < 0.$$

This is quite natural, as the increase in a pathogen type's lethality does not improve its chances to prevail when all other parameters of the model remain unchanged. The character of the dependence is illustrated in Fig. 6 showing R as a function of $(\alpha_1, \alpha_2) \in (0, 7)^2$, for $\lambda = 2, 6, 10$ and 14. Note how the encounter rate λ influences the size of the region where the process is supercritical (thus, for $\lambda = 2$ it shrinks to a narrow strip in the (α_1, α_2)-domain, corresponding to small values of α_1).

This is consistent with previous models of virulence evolution — it appears that they all would show that a more virulent pathogen strain would not be competitive if all its other parameters were equal to other strains. The "trade-off" relation between virulence and transmission rate in those models is necessary for more virulent pathogens to dominate. The recent model in [24] which also does not assume a tradeoff, produces a similar result.

We do not assume a trade-off between transmission and virulence (i.e. a function linking these characteristics), or that there is a genetic covariance between them. As noted in [24], virulence and transmissibility are affected by both the pathogen and its

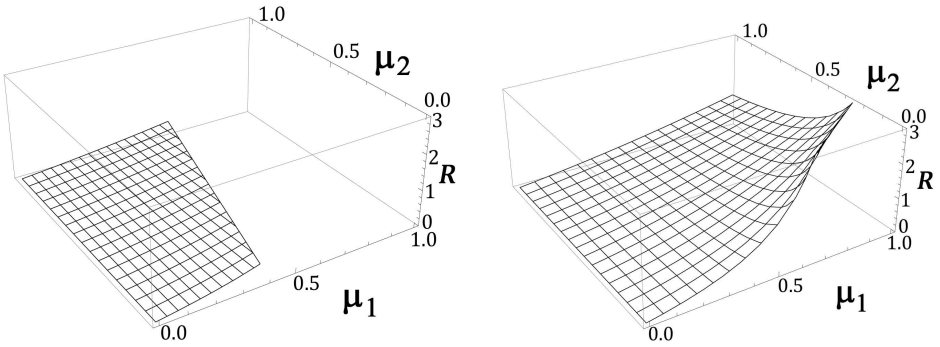


Figure 5: The plots of R as a function of the mutation probabilities for $\lambda = 4$ (the left pane) and $\lambda = 7$ (on the right), for fixed common values of all the other parameters of the model ($\alpha_1 = 1$, $\alpha_2 = 2$, $\beta_1 = 0.4$ and $\beta_2 = 0.2$. The plots are restricted to the regions where the respective processes are supercritical. Note the unusual orientation of the μ-coordinate axes (chosen so as to have a better view of the plots).

host. But our model suggests that more virulent strains can dominate only if higher transmissibility or mutation rate (or a combination thereof) favours them. It is this subset of more virulent pathogens that can dominate more easily at high host density.

5 On the effects of the change in enclosure size on the encounter rate

As we said, our main motivation was to model the effect that a change in "effective density" represented by the parameter λ (the rate at which infected hosts encounter susceptibles) can have on the ratio of the number of hosts infected, say, with T_2 pathogen to that for T_1-infected hosts. But how does the value of λ relate to physically measurable parameters of the modelled situation — for instance, the density of fish in an aquaculture tank (given the tank size and all other parameters of the model are fixed) or the size of the tank (given the number of fish in it is fixed)? How will λ change if one, for example, "squashes" the same host population to a "world" whose linear dimensions are twice as small as for the original one?

The answer to this type of questions will in general depend on what one assumes about the character of the hosts' movements (and of course, on the the pathogen transmission mechanism — but we will not address this aspect in our simple analysis in the section). One of the most popular models for "wandering particles" is the famous Brownian motion process $\{W(t), t \geq 0\}$ (see e.g. p.169 in [49]), which can be thought of as a continuous analog of a simple (symmetric) random walk. Recall that the Brownian motion is defined as a continuous time process with continuous trajectories that starts at zero at time $t = 0$ and has independent Gaussian increments: $W(t+h) - W(t) \sim N(0, h)$ for $t, h \geq 0$. One of the key properties of the process is its self-similarity: for any $a > 0$, one has

$$\{aW(t), t \geq 0\} \stackrel{d}{=} \{W(a^2 t), t \geq 0\}, \tag{13}$$

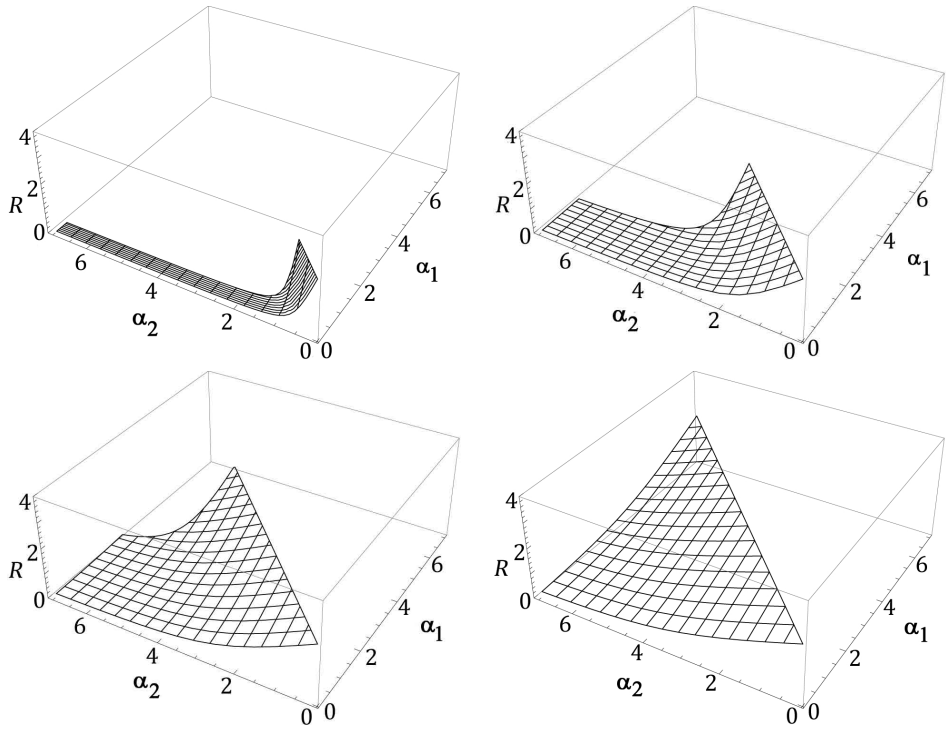


Figure 6: The plots of R as a function of the lethalities (α_1, α_2), in the same four cases as on Figs. 3, 4: $\lambda = 2, 6, 10$ and 14 (from left to right, top to bottom), when all the other parameters of the model are common ($\beta_1 = 0.3$, $\beta_2 = 0.6$, $\mu_1 = \mu_2 = 0.2$). The plots are restricted to the regions where the respective processes are supercritical and $\alpha_1 < \alpha_2$ (as we assumed). Note the unusual orientation of the α-coordinate axes (chosen so as to have a better view of the plots).

i.e. these two processes have the same distribution.

Since the total number of hosts is assumed to be large enough, encounter rates are mostly determined by the "local" characteristics (the density of the population and the dimensionality of the space) of the enclosure and will have little dependence on the "shape of the world". Therefore, for analysis purposes, we will assume in this section that hosts perform independent Brownian motions in a "simple world" S in one, two or three dimensions (starting at some "individual" initial points) and show how the encounter rates change when one "contracts" the original world without changing its shape, i.e. switches from S to the set $\varepsilon S = \{\varepsilon \boldsymbol{x} : \boldsymbol{x} \in S\}$. In this context, the dimensionality can actually be thought of as a crude description of the the shape of the world.

Suppose that there are N susceptibles in the population and that the movements of all the hosts are independent of each other. The value of the parameter $\lambda = \lambda_N$ gives the average number of encounters of a given infected host with susceptible ones per time unit (we assume that N is large enough so that the "conversion" of some susceptibles into infected hosts during our modelling "time horizon" does not significantly change the number N and hence the encounter rate λ). It is clear that $\lambda_N = N\lambda_1$ and,

moreover, that $\lambda_1 = 1/t^*$, where t^* is the mean time to encounter of our infected host with a given susceptible host. Thus the answer to the question on how the encounter rate will increase if the host density living in a "fixed world" increases is simple: it is just proportional to the number of hosts in the world. However, to understand the effect of the world size change (when we switch from S to εS) on the encounter rate $\lambda = \lambda(\varepsilon)$, we will have to analyze that effect on the mean time $t^* = t^*(\varepsilon)$.

The one-dimensional case. First we assume that $S = [0,1] \subset \mathbb{R}$. Our hosts move according to independent Brownian motions, reflecting from the boundaries 0 and 1 of the set S. This can be formalised by introducing the function

$$\varphi(x) := \begin{cases} x - \lfloor x \rfloor & \text{if } \lfloor x \rfloor \text{ is even,} \\ 1 - x + \lfloor x \rfloor & \text{if } \lfloor x \rfloor \text{ is odd,} \end{cases} \qquad \lfloor x \rfloor := \max\{k \in \mathbb{Z} : \, k \le x\},$$

and letting the location of our infected host to be given by $H_0 = H_0(t) = \varphi(H_0(0) + W_0(t))$ and that of a given susceptible one by $H_1 = H_1(t) = \varphi(H_1(0) + W_1(t))$, where $H_i(0) \in S$ are some fixed initial positions and W_i are independent standard Brownian motions, $i = 0, 1$. We say that the hosts have an encounter at time t if $H_0(t) = H_1(t)$.

Now consider the space εS, where our hosts move according to $\{\varepsilon H_i(t),\, t \ge 0\}$ and denote by $t^*(\varepsilon)$ the mean time to encounter of the hosts in this new world, $\varepsilon > 0$. It is obvious from the self-similarity property (13) that $t^*(\varepsilon) = \varepsilon^2 t^*(1)$ and therefore

$$\lambda(\varepsilon) = \varepsilon^{-2} \lambda(1).$$

Thus, if our hosts are confined to a one-dimensional world where they wander at random, the encounter rate displays inverse quadratic dependence on the world size: say, halving the "living space" will increase the encounter rate fourfold.

The two-dimensional case. To avoid dealing with any boundaries, we assume that our hosts live on the two-dimensional sphere

$$\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 : \, x^2 + y^2 + z^2 = 1\}$$

and that our hosts wander on it at random according to independent spherical Brownian motions $\{H_i(t),\, t \ge 0\}$ (see e.g. [49], Chapter 15, Section 13I), starting at some fixed distinct points $H_i(0)$, $i = 0, 1$.

In this case, $\Pr(H_0(t) \ne H_1(t),\, t \ge 0) = 1$, so we need to modify our definition of encounter. Fix a small enough $\delta > 0$ and define the encounter time of the hosts H_i as $\inf\{t > 0 : r(H_0(t), H_1(t)) = \delta\}$, where $r(\cdot, \cdot)$ is the geodesic distance on $\mathbb{S}^2$. Denote by $t^*_\delta = t^*_\delta(1)$ the mean value of this time (suppressing the dependence of the initial locations $H_i(0)$; to simplify the exposition, we deliberately make it somewhat sketchy).

Next we consider the "contracted world" $\varepsilon\mathbb{S}^2$, where our hosts wander according to $\{\varepsilon H_i(t),\, t \ge 0\}$, but the definition of encounter remains unchanged (the hosts should find themselves within distance δ of each other); the mean time to encounter for this case is denoted by $t^*_\delta(\varepsilon)$. Again using self-similarity, we can easily conclude that

$$t^*_\delta(\varepsilon) = \varepsilon^2 t^*_{\delta/\varepsilon}(1). \tag{14}$$

However, we want to relate $t^*_\delta(\varepsilon)$ to $t^*_\delta(1)$, so it remains to clarify the relationship between $t^*_{\delta/\varepsilon}(1)$ and $t^*_\delta(1)$.

It is obvious that $t^*_\eta = t^*_\eta(1)$ is a decreasing function of $\eta > 0$. As we are interested in situations where δ/ε is small (despite the small size of the "world", encounters are still relatively rare), it suffices to find the asymptotic behaviour of t^*_η as $\eta \to 0$. To do that, we first observe that analyzing the dynamics of the position of $H_0(t)$ relative to $H_1(t)$ shows that finding the mean time when the two points are first within distance η of each other is equivalent to finding the mean time a Brownian particle $H^*(t)$ (with an initial position at a distance $r(H_0(0), H_1(0))$ from the "North Pole" $P = (0, 0, 1) \in \mathbb{S}^2$ and local diffusion coefficient $\sqrt{2}$ times that for the original spherical Brownian processes) will need to get within distance η of P.

Denoting by $V(t)$ the projection of $H^*(t)$ on the z-axis, one can easily see from Itô's formula that $\{V(t),\, t \geq 0\}$ is a diffusion process with the state space $[-1, 1]$ governed by the stochastic differential equation

$$dV(t) = \mu(V(t))dt + \sigma(t)dW_0(t),$$

where W_0 is a standard univariate Brownian motion process and the drift and diffusion coefficients are given by

$$\mu(z) = -\frac{z}{2}, \qquad \sigma^2(z) = \frac{1 - z^2}{2},$$

respectively (see e.g. p.194 in [61]; for convenience we assumed that $H^*(t)$ follows a standard spherical Brownian motion on $\mathbb{S}^2$ which will have no adverse implications for the validity of our analysis). The geodesic distance from H^* to P is equal to η iff its projection on the z-axis equals $r := \cos\eta$, so that we need to find

$$v(z) = \mathbb{E}(\tau |\, V(0) = z), \qquad \text{where} \quad \tau = \inf\{t > 0 :\, V(t) = r\}.$$

This can easily be done using the standard technique of the method of differential equations (see e.g. Problem B on p.192 in [49]): the function $v(z)$ is the bounded solution to the equation

$$1 = -\mu(z)v'(z) - \frac{1}{2}\sigma^2(z)v''(z) \equiv \frac{1}{2}zv'(z) + \frac{1}{4}(z^2 - 1)v''(z), \qquad z \in (-1, r), \tag{15}$$

with the boundary condition $v(r) = 0$. Setting $u(z) := (z^2 - 1)/4$, we notice that $u'(z) = z/2$ and so (15) is equivalent to

$$1 = u'v' + uv'' = (uv')',$$

which means that $u(z)v'(z) = z + c_1$, and so

$$v'(z) = \frac{z + c_1}{u(z)} = \frac{4(z + c_1)}{z^2 - 1}.$$

Therefore the general solution to (15) is given by

$$v(z) = 2\big[(1+c_1)\ln(1-z) + (1-c_1)\ln(1+z)\big] + c_2,$$

which is bounded on $(-1, r)$ iff $c_1 = 1$. Now using the boundary condition at $z = r$ to find c_2 leads to

$$v(z) = 4\ln\frac{1-z}{1-r}, \qquad z \in [-1, r].$$

To find the asymptotic behaviour of $v(z)$ as $\eta \to 0$, we use $\cos\eta = 1 - \eta^2/2(1+o(1))$ to obtain that, for a fixed initial value z,

$$v(z) = 8|\ln\eta| + O(1).$$

This means that, for fixed initial positions of H_0 and H_1, we have $t^*_\eta = (c+o(1))|\ln\eta|$ as $\eta \to 0$. Together with (14) this yields, for small δ/ε,

$$t^*_\delta(\varepsilon) \approx \varepsilon^2 \frac{|\ln\delta| - |\ln\varepsilon|}{|\ln\delta|} t^*_\delta(1).$$

That is, the encounter rate behaves as

$$\lambda(\varepsilon) \approx \varepsilon^{-2} \frac{|\ln\delta|}{|\ln\delta| - |\ln\varepsilon|} \lambda(1).$$

In the two-dimensional case, it might be more natural to relate the encounter rate not to the linear dimensions of the enclosure, but rather to its area (proportional to ε^2). The above formula shows that, in this case, the encounter rate in a "shrinking" world still grows somewhat faster than the inverse proportion of its area, the latter giving the density of hosts.

The three-dimensional case. We still have (14), and an analysis similar to the one carried out in the two-dimensional case shows that now $t^*_\eta = (c + o(1))\eta^{-1}$ as $\eta \to 0$ (cf. p.195 in [61]), so that $t^*_\delta(\varepsilon) = \varepsilon^3 t^*_\delta(1)(1+o(1))$. That is,

$$\lambda(\varepsilon) = \varepsilon^{-3}\lambda(1)(1+o(1)).$$

We see that, in the three-dimensional case (assuming that the hosts wander according to three-dimensional Brownian motions), the encounter rate in a "shrinking" world is inversely proportional to its volume. That is, in this case the rate is proportional to the density of hosts.

To summarise the above analysis, we observe that the encounter rate $\lambda = \lambda(\varepsilon)$ grows rather fast when the linear dimensions (specified by the parameter ε) of the hosts' "world" diminish. In the three-dimensional case, λ is inversely proportional to the volume per host, in the two dimensional case it grows slightly faster than the reciprocal of the area per host, while in the one-dimensional case the growth rate of λ is inversely proportional to the square of the size of a host's share of the enclosure. This indicates that not only effective density per se, but also the shape of the enclosure can be an important factor leading to an epidemic. Thus the nature of the enclosures in which animals are kept can be an important factor in determining the progress and nature of an epidemic.

6 A multistage modification of the model

In this section we will consider an aggregate model for situations in which there are several populations of hosts that exist in originally isolated enclosures. The situation we model here could be viewed as a simple instance of a meta-population model (we refer the reader to [42] as an extensive source on metapopulation modelling). But in our scenario there is no feedback from later enclosures to earlier ones, and we do not consider any dynamics of the hosts. Thus we find it more useful to consider it as a chain of epidemic episodes.

In intensive animal husbandry, there are often many pens, paddocks or sheds enclosing animals at a farm or many fish tanks at an aquaculture facility. We consider the scenario where initially, one of the enclosures becomes infected with a single type pathogen. This can give rise to a "local epidemic" in the infected enclosure, which can be modelled using our processes from Section 2 (assuming that we have supercriticality: $\sigma_+ > 0$). The original pathogen may also mutate to become more or less lethal. We will initially assume that it may mutate to a more lethal type 2 pathogen ($\alpha_2 = r\alpha$ for some $r > 1$, $\alpha_1 = \alpha$). That new pathogen type can also have different transmissibility and mutation rate, but, to make our model as simple as possible, we will assume for the time being that it differs from type 1 pathogen in lethality only, all other parameters being unchanged. Denoting them simply by β and μ, we see that the Malthusian parameter for that process is given by

$$\sigma_+(\alpha_1, \alpha_2) = \frac{1}{2}\left(2(1-\mu)\beta\lambda - \alpha_1 - \alpha_2 + \sqrt{(\alpha_1 - \alpha_2)^2 + 4\mu^2\beta^2\lambda^2}\right). \tag{16}$$

At some point, the infection might be transmitted to the next enclosure (say, by a worker in a farm situation). We assume that this occurs after the epidemic has gone through the initial stage, and so the ratio of the numbers of hosts infected with different pathogen type can be assumed equal to ρ_0, where $\rho_k \equiv R(r^k\alpha, r^{k+1}\alpha, \beta, \beta, \mu, \mu, \lambda)$. The transmitted pathogen is chosen at random, so that the probability of transmitting the one with lethality α (denote this event by A) is $1/(1+\rho_0)$, while the one with lethality $r\alpha$ is transmitted with probability $\rho_0/(1+\rho_0)$. This, in turn, may lead to a local epidemic in the new enclosure: we again assume the possibility of mutation to a more lethal pathogen (so that now we will have $\alpha_1 = \alpha$, $\alpha_2 = r\alpha$ if the event A occurred, and $\alpha_1 = r\alpha$, $\alpha_2 = r^2\alpha$ otherwise), and to have an epidemic we again need $\sigma_+ = \sigma_+(\alpha_1, \alpha_2) > 0$ (now for the new set of parameters α_1, α_2). Once the epidemic has established itself in the second enclosure (and the balance of pathogen types has stabilized around the respective R-value), the next step is the transmission of the disease (by means of a random mechanism of the same type as in the first instance) to the next enclosure, and so on.

Scenarios of this type have been encountered often where once a disease is recognized in a herd, animals in the infected enclosure are removed or killed, but the disease is subsequently found in other herds, for example the spread of Foot and mouth disease among herds in Taiwan, which was related to herd size and the number of herds in a province [36]. Of course, biosecurity measures are intended to prevent such transmission between enclosures, but often the need for diligence is learned after the event.

It is easily seen from (16) that

$$\sigma_+(0,0) > 0, \qquad \frac{\partial}{\partial\alpha}\sigma_+(\alpha, r\alpha) < 0, \qquad \lim_{\alpha\to\infty} \sigma_+(\alpha, r\alpha) = -\infty, \tag{17}$$

and also that $\dfrac{\partial}{\partial\alpha}R(\alpha, r\alpha) < 0$. Thus, if the lethality of the pathogen will keep increasing, the Malthusian parameter of the model will eventually drop below zero, and then the epidemic will collapse. More specifically, setting

$$k^* \equiv \inf\{k \geq 0 :\ \sigma_+(r^k\alpha, r^{k+1}\alpha) < 0\},$$

we see that

$$\sigma_+(r^k\alpha, r^{k+1}\alpha) < 0 \quad \text{ for all } k \geq k^*. \tag{18}$$

It is clear that the transition of the disease from enclosure to enclosure according to the above scheme is described by a discrete time Markov chain $\{X_n\}$, the "time" n having the meaning of the number of steps (i.e. enclosures infected), X_n representing the level of lethality of the pathogens in the nth infected enclosure: we set $X_n = k$ if $(\alpha_1, \alpha_2) = (r^k\alpha, r^{k+1}\alpha)$ in the enclosure. Thus the state space of the Markov chain is $\{0, 1, 2, \dots\}$ and the only nonzero entries in the transition matrix $\boldsymbol{P} = [p_{j,k}]_{j,k\geq 0}$ of the chain are

$$p_{k,k} \equiv \Pr(X_{n+1} = k |\, X_{n+1} = k) = \frac{1}{1+\rho_k},$$
$$p_{k,k+1} \equiv \Pr(X_{n+1} = k+1 |\, X_{n+1} = k) = \frac{\rho_k}{1+\rho_k},$$

$k = 0, 1, 2, \dots$

Further, in view of (18), one can assume that once the Markov chain $\{X_n\}$ has reached the state k^*, the epidemic becomes unsustainable, and hence there will be no further transmission of the disease to other enclosures. So we can truncate our state space to $\{0, 1, 2, \dots, k^*\}$, which results in a finite decomposable Markov chain with a single absorbing state k^*. Whatever the current state of the chain, at the next step it can either stay at it or move to the right, the transition probabilities forming the matrix

$$[q_{jk}] = \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{T} & \boldsymbol{r}^\top \\ \boldsymbol{0} & 1 \end{bmatrix},$$

where $\boldsymbol{T}$ is the $k^* \times k^*$ substochastic matrix formed by the first k^* rows and k^* columns of $\boldsymbol{P}$, $\boldsymbol{r} = (0, \dots, 0, p_{k^*-1,k^*}) \in \mathbb{R}_+^{k^*}$ and $^\top$ denotes transposition.

The (random) number of steps T the chain will need to reach the absorbing state is nothing else but the total number of enclosures that will be affected by the epidemic prior to its collapse. Using our model, we can easily find the distribution of T.

Indeed, using the standard approach to solving such problems (see e.g. p.80 in [52]), we note that as the state k^* is absorbing, we have $\Pr(T \leq n |\, X_0 = 0) = q^{(n)}_{0,k^*}$, where $q^{(n)}_{jk}$ are the n-step transition probabilities:

$$[q^{(n)}_{jk}] = \boldsymbol{Q}^n = \begin{bmatrix} \boldsymbol{T}^n & \boldsymbol{r}_n^\top \\ \boldsymbol{0} & 1 \end{bmatrix}, \qquad \boldsymbol{r}_n^\top = (\boldsymbol{I} + \boldsymbol{T} + \dots + \boldsymbol{T}^{n-1})\boldsymbol{r}^\top,$$

so that for the probability mass vector function $\boldsymbol{f}(n) = \{f_j(n),\, j = 0, \ldots, k^* - 1\}$, $f_j(n) = \Pr(T = n |\, X_0 = j)$, one obtains

$$\boldsymbol{f}(n) = \boldsymbol{r}_n - \boldsymbol{r}_{n-1} = \boldsymbol{r}(\boldsymbol{T}^{n-1})^\top, \qquad n \geq 1.$$

Of course, we are only interested in the first entry of the vector $\boldsymbol{f}(n)$.

To compute the mean and higher moments of T one can use the generating function

$$\boldsymbol{f}^*(z) \equiv \sum_{n=1}^{\infty} z^n \boldsymbol{f}(n) = \boldsymbol{r}z\big(\boldsymbol{I} - z\boldsymbol{T}^\top\big)^{-1}, \qquad |z| \leq 1.$$

In particular, since $\dfrac{d}{dz}\boldsymbol{f}^*(z) = \boldsymbol{r}\big(\boldsymbol{I} - z\boldsymbol{T}^\top\big)^{-1} + \boldsymbol{r}z\boldsymbol{T}^\top\big(\boldsymbol{I} - z\boldsymbol{T}^\top\big)^{-2}$, we find that

$$\mathbb{E}(T |\, X_0 = j) = \frac{d}{dz}\boldsymbol{f}^*(z)\bigg|_{z=1} = \boldsymbol{r}\big(\boldsymbol{I} - \boldsymbol{T}^\top\big)^{-2} = \boldsymbol{1}\big(\boldsymbol{I} - \boldsymbol{T}^\top\big)^{-1},$$

where the last equality follows from the obvious observation that $\boldsymbol{r}\big(\boldsymbol{I} - \boldsymbol{T}^\top\big)^{-1} = \boldsymbol{f}^*(1) = \boldsymbol{1} \equiv (1, \ldots, 1) \in \mathbb{R}_+^{k^*}$.

A possible objection to the above simple aggregate model is that pathogens will not always mutate to become more lethal. The model can be further generalized by allowing, within each enclosure, mutations of our pathogen not only in the direction of higher lethality, but also in the opposite direction. So we will first have to generalize our basic model from Section 2 to a three-type branching process, assuming that, if an enclosure is infected with a pathogen with lethality α, then the pathogen can mutate to ones with lethalities $r^{-1}\alpha$ and $r\alpha$, respectively, where, as before, $r > 1$ is a fixed number (all other parameters being assumed equal for the different types of pathogens). Mathematically, analyzing such processes is essentially equivalent to what we did in Sections 2–4, although all the closed-from expressions will become much more complicated, and so we will not give much technical detail for brevity's sake. The main assertions concerning the asymptotic behaviour of the branching process will remain true: there will exist a limiting balance of types in the supercritical case (denote the shares of the different pathogens by $\pi_j = \pi_j(\alpha)$, $j = 1, 2, 3$, $\sum_j \pi_j = 1$), which can be found from the generator $\boldsymbol{A}$ of the semigroup of the mean matrices, and the almost sure convergence of the process scaled by $e^{-\sigma_+ t}$ (as before, σ_+ denotes the Perron-Frobenius root of $\boldsymbol{A}$) to a limiting random vector will hold.

In the multi-stage model, we start with initial infection of one of the enclosures with a pathogen with lethality α. That leads to an epidemic (provided, of course, that $\sigma_+ > 0$) in which pathogens of three types will be present, with lethalities given by the vector $(\alpha_1, \alpha_2, \alpha_3) = (r^{-1}\alpha, \alpha, r\alpha)$. The next enclosure to be infected will receive a pathogen chosen at random from those present in the first infected enclosure, and so it will have lethality α_j with probability π_j, $j = 1, 2, 3$, and so on.

Observe that the triplets of lethalities $(\alpha_1, \alpha_2, \alpha_3)$ characterizing the pathogens present in a given enclosure in our system will all be of the form (x, rx, r^2x) for some $x > 0$, i.e. lying on a common ray L with the direction vector $(1, r, r^2)$. Therefore we

will again have a basically "univariate" Markov chain $\{X_n\}$ showing what pathogens can be present in different enclosures, $X_n = k$ meaning that the nth affected enclosure was initially infected with the pathogen of lethality $r^k\alpha$, $k \in \mathbb{Z}$ (and in this case there can also be pathogens with lethalities $r^{k-1}\alpha$ and $r^{k+1}\alpha$ in that enclosure), assuming that $X_0 = 0$ (as α is the lethality of the pathogen that was initially introduced into our system). The original state space for the process will be $\mathbb{Z} \equiv \{\dots, -1, 0, 1, \dots\}$, which is infinite in both directions. At each step, the value of the chain can remain unchanged (the interpretation being that the pathogen transmitted to the next enclosure had the same lethality as the one with which the epidemic started in the current one) or can either decrease or increase by one (that is, the transmitted pathogen would have lethality values equal to r^{-1} or r times the current one, respectively).

Further, , as before, one can show that $\frac{\partial}{\partial x}\sigma_+(x, rx, r^2x) < 0$, so that, moving along the ray L in the "positive direction", we will eventually enter the subcriticality region for the branching process, where the basic reproductive number will be less than one. Therefore, at this stage the epidemic will collapse, and hence we again can "truncate" the state space for $\{X_n\}$ — now to $\{\dots, -1, 0, 1, \dots, k^*\}$, where k^* is an absorbing state that has the same meaning as above.

The variety of behaviours that such a chain can display will be somewhat wider than for our first multi-stage model. The dynamics of the chain are determined by the behaviour of the mean step values

$$\mathbb{E}\big(X_{n+1} - X_n \big| X_n = k\big) = \pi_3(r^k\alpha) - \pi_1(r^k\alpha), \qquad k < k^*.$$

In particular, the parameters of the model can be such that the above quantities will be negative. Then absorbtion at k^* occurs with probability less than one, while on the complement event the chain will drift away in the negative direction, which corresponds to the disease "fading", when the pathogen's lethality vanishes, and so on.

7 Discussion

We conclude with remarks concerning possible biological interpretation of our results.

First, our model results are generally consistent with previous models of virulence, but we have focused on the initial stages of epidemics in an animal husbandry context, and used a different approach and assumptions. The mathematical models we have presented show that, at the beginning stages of any epidemics that arise in such situations, the density of hosts is an important factor in determining whether more or less lethal strains of the pathogen will predominate. Many authors have noted that selection for virulence differs during the initial (spreading) stage of an epidemic [31, 25, 19, 7]; and others have noted that high host densities are likely to lead to increased virulence [30, 34, 19].

Many previous models [30, 31, 23] have assumed that more rapid production of copies of the pathogen inside the host would both increase the likelihood of transmission to new hosts and also shorten the life of the host. There has been considerable criticism of this "tradeoff hypothesis", although there are limited data supporting it [26, 33, 18],

but it has now been realised [19] that it is simply not relevant to the initial stages of an epidemic. Nevertheless, our model shows that increased lethality alone will not lead to dominance by a pathogen strain — although more lethal strains can invade more easily at high densities and they will dominate more rapidly, provided they are favoured by some factor such as higher transmission or mutation rates.

Our models show that the ratio of infections by more and less lethal pathogens stabilizes very fast, so that even if measures to prevent further spread of the disease are put in place as soon as an outbreak is identified (such as elimination of the animals in an enclosure), the relative frequency of pathogen types is likely to have changed before action is taken. Thus a high density of animals on farms or mariculture facilities will rapidly lead to the dominance of more lethal strains of pathogens if these can enter the farm and mutations occur to produce more lethal variants. An example of this process has recently been described in the mariculture of fish in [64]. The problem has been recognized in intensive poultry production [65], and the identification of more virulent trains of Marek's disease in chickens [69] may also be an example of this process. It also seems possible that the advent of very virulent strains of Spanish Flu in August 1918 in US military camps, troop ships, and European disembarkation camps, some time after the start of the epidemic may be linked to high densities of soldiers in these camps and troop ships and the difficulty in implementing quarantine in these overcrowded camps [60], although there are other plausible explanations, such as overcrowded conditions leading to higher infection doses.

Thus the increasing density at which animals are kept (intensification of the industry) should be considered as a risk factor for the evolution of more virulent diseases. Various authors have pointed to this issue [17, 34, 7] but in many cases attention has been focused on the possibility that vaccination — especially where the vaccine reduces pathogen replication — selects for higher virulence. We contend that it is likely that the higher transmission of pathogens at high densities that allows more virulent pathogens (which may replicate faster) to become dominant in spite of such vaccines.

Assuming a chaotic pattern of movement of animals inside the enclosure where they are kept (as modelled by independent Brownian motion processes), we have modelled how the effective stock density affects their encounter rate (the key parameter of our model) and thus influences what pathogen type will predominate. The effect increases more rapidly with density where an enclosure is most appropriately measured as area (cattle feedlots, paddocks) than if the appropriate measure is a volume (fish tanks). We assumed constant mixing so that the rate of encounter with susceptible hosts is not materially changed in the first stages of the epidemic. These assumptions will limit the situations in which this result may apply to cases where animals mix freely and rapidly in the enclosure, but our result may also apply if there is a large area around an infected individual where others may become infected by a pathogen, because it is dispersed in air, water or over time (as a long-lived propagule [17] that can disperse or survive until restocking [19]).

In agricultural systems, humans have been maintaining animals and plants at very high densities for about 10,000 years, and humans themselves began living at high local densities in villages, towns and cities from about the same time [20]. Densities of both

humans and domesticated organisms have probably increased continually since then, but the densities at which some domestic animals are kept appear to have increased very rapidly in the last few decades (see e.g. [32]), and the recent rapid development of large scale commercial aquaculture involves the maintenance of very high densities of a whole new set of marine species that will have their own pathogens (see [44, 43, 14]).

The ratio of pathogen types will clearly also depend on the transmissibility of pathogen strains. We have focused on differences in lethality between pathogen strains, because pathogen strains with increased transmissibility would obviously become more prevalent. It is interesting to note that our model showed that, even if the transmissibility of pathogen of one type is lower than that for the other, the former pathogen can still prevail provided it is favoured by mutation. Increased density of animal hosts may itself increase transmissibility, due to stress on the hosts caused by crowding [66]. Increased transmissibility might in turn lead to the evolution of increased lethality [30].

Modern animal husbandry often involves a large number of separate enclosures, each containing a large number of animals at very high densities. Once a disease is detected in an enclosure, farmers would either use antibiotics, or sacrifice or remove the animals, but pathogens may be carried between enclosures by various mechanisms, depending on the type of pathogen and the biosecurity practices followed. A multistage version of our model for this situation suggests that if pathogens are transferred a number of times, then the evolution of more lethal pathogens may be very rapid, but the increase in lethality will eventually lead to the epidemic becoming unsustainable (hosts dying too fast to be able to transit the pathogen).

We suggest that the outcomes predicted by the mathematical models discussed in the present paper can carry important messages for animal husbandry, where there are strong commercial incentives to increase the densities of animals in enclosures to very high levels, and often very large numbers of enclosures are built in a single farm.

Acknowledgments. K. Borovkov was supported by the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS) and ARC Discovery Grant DP120102398. T. Rice was partially supported by MASCOS. The original motivation for this study arose from conversations of R. Day with Jeremy Prince and Harry Gorfine.

The authors are grateful to the referee whose thoughtful comments helped to improve the presentation of the paper.

References

[1] H.K. Alexander and T. Day. (2010) Risk factors for the evolutionary emergence of pathogens. *J. R. Soc. Interface,* 7, 1455–1474.

[2] S. Alizon, A. Hurford, N. Mideo and M. van Baalen. (2009) Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* 22, 245–259.

[3] L.J.S. Allen. (2008) An introduction to stochastic epidemic models. In: *Mathematical Epidemiology, Lecture Notes in Math.,* Springer, Berlin, 1945, 81–130.

[4] R.M. Anderson and R.M. May. (1982) Coevolution of hosts and parasites. *Parasitology,* 85, 411–426.

[5] R.M. Anderson and R.M. May. (1991) *Infectious Diseases of Humans: Dynamics and Control.* Oxford University Press, Oxford.

[6] H. Andersson and T. Britton. (2000) *Stochastic Epidemic Models and Their Statistical Analysis.* Lecture Notes in Statistics, 151. Springer, New York.

[7] J-B. André and M.E. Hochberg. (2005) Virulence evolution in emerging infectious diseases. *Evolution,* 59, 1406–1412.

[8] K.B. Athreya and P.E. Ney. (1972) *Branching Processes.* Springer, New York.

[9] F.G. Ball and P. Donelly. (1995) Strong approximations for epidemic models. *Stoch. Proc. Appl.* 55, 1–21.

[10] M.S. Bartlett. (1949) Some evolutionary stochastic processes. J. Roy. Statist. Soc. B 11, 211–229.

[11] M.S. Bartlett. (1955) *An Introduction to Stochastic Processes.* Cambridge Univeristy Press, Cambridge.

[12] N.G. Becker. (1989) *Analysis of Infectious Disease Data.* Chapman and Hall, London.

[13] N.G. Becker and I. Marschner. (1990) The effect of heterogeneity on the spread of disease. In: *Stochastic Processes in Epidemic Theory,* J.-P. Gabriel et al., eds. Lecture Notes in Biomathematics, 86, 90–103.

[14] O. Bergh. (2007) The dual myths of the healthy wild fish and the unhealthy farmed fish. *Dis. Aquat. Org.* 75, 159–164.

[15] D. Bernoulli. (1760) Essai d'une nouvelle analyse de la mortalité causée par la petit vérole et des advanteges de l'inoculation pour la prévenir. *Mém. Math. Phys. Acad. Roy. Sci., Paris,* 1–45.

[16] A.T. Bharucha-Reid. (1958) Comparison of populations whose growth can be described by a branching stochastic process. *Sankhyā,* 19, 1–14.

[17] P.M. Biggs. (1985) Infectious animal disease and its control. *Phil. Trans. Roy. Soc. London B,* 310, 259–274.

[18] B.M. Bolker, A Nanda and D Shah. (2010) Transient virulence of emerging pathogens. *J. R. Soc. Interface,* 7, 811–822.

[19] J.J. Bull and D. Ebert. (2008) Invasion thresholds and the evolution of non-equilibrium virulence. *Evol. Appl.* 1, 172–182.

[20] D. Crawford. (2007) Deadly companions: how microbes shaped our history. Oxford University Press, Oxford.

[21] D.J. Daley and J. Gani. (1999) *Epidemic Modelling: An Introduction.* Cambridge University Press, Cambridge.

[22] T. Day. (2002) On the evolution of virulence and the relationship between various measures of mortality. *Proc. R. Soc. Lond. B*, 269, 1317–1323.

[23] T. Day. (2003) Virulence evolution and the timing of disease life-history events. *Trends in Ecology and Evolution,* 18, 113-118.

[24] T. Day and S. Gandon. (2006) Insights from Price's equation into evolutionary epidemiology. In: *Disease Evolution: Models, Concepts, and Data Analyses.* Z. Feng, U. Dieckmann and S. A. Levin, eds. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 71, 23–43

[25] T. Day and S.R. Proulx. (2004) A general theory for the evolutionary dynamics of virlence. *Amer. Naturalist,* 163, E41–E63.

[26] J.C. de Roode, A.J. Yates and S. Altizer. (2008) Virulence — transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proc. Natl Acad. Sci. USA,* 105, 7489–7494.

[27] K. Dietz and D. Schenke. (1985) Mathematical models for infectious disease statistics. In: *A Celebration of Statistics,* A.C. Atkinson and S.E. Fienberg, eds. Springer, New York, 167–204.

[28] D. Ebert. (1999) The evolution and expression of parasite virulence. In: *Evolution in health and disease,* S.C. Stearns, ed. Oxford University Press, Oxford, 161–172.

[29] P.W. Ewald. (1983) Host-parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Syst.* 14: 465–485.

[30] P.W. Ewald, (1994) *Evolution of infectious disease.* Oxford Univ. Press, Oxford.

[31] S.A. Frank. (1996) Models of parasite virulence. *Quarterly Rev. Biol.* 71, 37–78.

[32] D. Fraser. (2005) *Animal welfare and the intensification of animal production: An alternative interpretation.* UN Food and Agriculture Organization, Rome.

[33] C. Fraser, T.D. Hollingsworth, R. Chapman, F. de Wolf and W.P. Hanage. (2007) Variation in HIV–1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl Acad. Sci. USA,* 104, 17441–17446.

[34] S. Gandon, M.J. Mackinnon, S. Nee and A.F. Read. (2001) Imperfect vaccines and the evolution of pathogen virulence. *Nature,* 414, 751–756.

[35] F.R. Gantmakher. (1989) *The theory of matrices.* Vol. 1. Chelsea, New York.

[36] G. Gerbier. (1999) *Effect of animal density on FMD spread.* European Commission for the Control of Foot-and-Mouth Disease Report. Research Group of the Standing Technical Committee, Maisons-Alfort.

[37] W.M. Getz and J.O. Lloyd-Smith. (2006) Basic methods for modeling the invasion and spread of contagious diseases. In: *Disease evolution, DIMACS Ser. Discrete Math. Theoret. Comput. Sci.,* Amer. Math. Soc., Providence, RI, 71, 87–109.

[38] M. Girvan1, D.S. Callaway, M.E.J. Newman, and S.H. Strogatz (2002), A simple model of epidemics with pathogen mutation. *Phys. Rev. E,* 65, 031915.

[39] M. Greenwood. (1931) On the statistical measure of infectiousness. *J. Hyg. Camb.* 31, 336–351.

[40] Y. Guan, B. J. Zheng, Y. Q. He et al. (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science,* 302, 276–278.

[41] P. Haccou, P. Jagers and V.A. Vatutin. (2005) *Branching Processes: Variation, Growth, and Extinction of Populations.* Cambridge University Press, Cambridge.

[42] I.A. Hanski and O.E. Gaggiotti, eds. (2004) *Ecology, Genetics and Evolution of Metapopulations.* Elsevier Academic Press, Burlington, MA.

[43] C.D. Harvell, K. Kim, J.M. Burkholder et al. (1999). Emerging marine diseases — climate links and anthropogenic factors. *Science,* 285, 1505–1510.

[44] D. Harvell, R. Aronson, N. Baron et al. (2004) The rising tide of ocean diseases: unsolved problems and research priorities. *Front. Ecol. Environ.* 2, 375–382.

[45] D. Heinzmann. (2009) Extinction times in multitype Markov branching processes *J. Appl. Probab.* 46, 296–307.

[46] H.W. Hethcote. (1994) A thousand and one epidemic models. In: *Lecture Notes in Biomathematics*, 100, 504-515.

[47] P. Jagers. (1975) *Branching Processes with Biological Applications.* Wiley, London.

[48] F. Jeltsch, M.S. Mller, V. Grimm, C. Wissel and R. Brandl. (1997) Pattern formation triggered by rare events: Lessons from the spread of rabies. *Proc. R. Soc. London B,* 264, 495-503.

[49] S. Karlin and H.M. Taylor. (1981) *A Second Course in Stochastic Processes.* Academic Press, New York.

[50] D. Kendall. (1956) Deterministic and stochastic epidemics in closed popultaions. In: *Proc. Third Berkeley Symp. Maths. Statist. Prob.* 4, 149–165.

[51] W.O. Kermack and A.G. McKendrick. (1927) A contribution to the mathematcial theory of epidemics. *Proc. R. Soc. Lond. A* 115, 700–721.

[52] M. Kijima. (1997) *Markov Processes for Stochastic Modeling.* Chapman & Hall, London.

[53] H. Knolle. (1989) Host density and the evolution of parasite virulence. *J. Theor. Biol.* 136, 199–207.

[54] R.E. Lenski and R.M. May. (1994) The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. *J. Theor. Biol.* 169, 253–265.

[55] S. Lion and M. Boots. (2010) Are parasites 'prudent' in space? *Ecol. Lett. 13,* 1245–1255.

[56] M. Lipsitch and E.R. Moxon. (1997) Virulence and transmissibility of pathogens: what is the relationship. *Trends Microbiol.* 5, 31–37.

[57] D. Ludwig. (1975) Qualitative behavior of stochastic epidemics. *Math. Biosci.* 23, 47–73.

[58] A.G. McKendrick. (1926) Applications of mathematics to medical problems. *Proc. Edin. Math. Soc.*, 14, 98–130.

[59] E. Massad. (1987) Transmission rates and the evolution of pathogenicity. *Evolution,* 41, 1127–1130.

[60] J.D. Mathews, J.M. Chesson, J.M. McCaw and J.McVernon. (2009) Understanding influenza transmission, immunity and pandemic threats. *Influenza Other Respi. Viruses,* 3(4), 143–149.

[61] P. Matthews. (1988) Covering problems for Brownian motion on spheres. *Ann. Prob.* 18, 189–199.

[62] R. Meester, J. de Koning, M.C.C de Jong and O. Diekmann. (2002) Modeling and real-time prediction of classical swine fever epidemics. *Biometrics,* 58, 178–184.

[63] C.T. Mode and C.K. Sleeman. (2000) *Stochastic processes in epidemiology.* World Scientific, Singapore.

[64] K. Pulkinnen, L.-R. Suomalainen, A.F. Read, D. Ebert, P. Rintamäki and E.T. Valtonen. (2010) Intensive fish farming and the evolution of pathogen virulence: the case of coumnaris disease in Finland. *Proc. Roy. Soc. B,* 277, 593–600.

[65] J. Slingenbergh, M. Glibert. (2010) Do old and new forms of poultry go together. In: *FAO Abstracts: Poultry production in the 21st century.* FAO, Rome, p. 47.

[66] S. Sniezko. (1974) The effects of environmental stress on outbreaks of infectious diseases of fishes. *J. Fish. Biol.,* 6, 197–208.

[67] H.R. Thieme. (2003) *Mathematics in population biology.* Princeton University Press, Princeton, NJ.

[68] M. van Baalen. (2002) Contact networks and the evolution of virulence. In: *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*, U. Dieckmann, J.A.J. Metz, M.W. Sabelis and K. Sigmund, eds. Cambridge University Press, 85-103.

[69] R.L. Witter. (1997) Increased virulence of Marek's disease virus field isolates. *Avian Diseases,* 41, 149–163.

[70] P. Whittle. (1955) The outcome of a stochastic epidemic — a note on Bailey's paper. *Biometrica*, 42, 116-122.

[71] A. Yates, R. Antia and R.R. Regoes. (2006) How do pathogen evolution and host heterogeneity interact in disease emergence? *Proc. R. Soc. Lond. B,* 273, 3075–3083.

[72] A. Yersin. (1894) La peste bubonique á Hong Kong. *Ann. Inst. Pasteur,* 8, 662–667.