CrossMark

R_0: Host Longevity Matters

L. M. Viljoen[1] · L. Hemerik[2] · J. Molenaar[2] ⬤

Abstract The basic reproduction ratio, R_0, is a fundamental concept in epidemiology. It is defined as the total number of secondary infections brought on by a single primary infection, in a totally susceptible population. The value of R_0 indicates whether a starting epidemic reaches a considerable part of the population and causes a lot of damage, or whether it remains restricted to a relatively small number of individuals. To calculate R_0 one has to evaluate an integral that ranges over the duration of the infection of the host. This duration is, of course, limited by remaining host longevity. So, R_0 depends on remaining host longevity and in this paper we show that for long-lived hosts this aspect may not be ignored for long-lasting infections. We investigate in particular how this epidemiological measure of pathogen fitness depends on host longevity. For our analyses we adopt and combine a generic within- and between-host model from the literature. To find the optimal strategy for a pathogen from an evolutionary point of view, we focus on the indicator R_0^{opt}, i.e., the optimum of R_0 as a function of its replication and mutation rates. These are the within-host parameters that the pathogen has at its disposal to optimize its strategy. We show that R_0^{opt} is highly influenced by remaining host longevity in combination with the contact rate between hosts in a susceptible population. In addition, these

✉ J. Molenaar
    jaap.molenaar@wur.nl

    L. M. Viljoen
    lourensv@mbdcs.co.za

[1] Department of Mathematics and Applied Mathematics, North West University, Potchefstroom, North West, South Africa

[2] Biometris, Department of Mathematical and Statistical Methods, Wageningen University, Wageningen, The Netherlands

two parameters determine whether a killer-like or a milker-like strategy is optimal for a given pathogen. In the killer-like strategy the pathogen has a high rate of reproduction within the host in a short time span causing a relatively short disease, whereas in the milker-like strategy the pathogen multiplies relatively slowly, producing a continuous small amount of offspring over time with a small effect on host health. The present research allows for the determination of a bifurcation line in the plane of host longevity versus contact rate that forms the boundary between the milker-like and killer-like regions. This plot shows that for short remaining host longevities the killer-like strategy is optimal, whereas for very long remaining host longevities the milker-like strategy is advantageous. For in-between values of host longevity, the contact rate determines which of both strategies is optimal.

1 Introduction

In general, successful parasites engage in a dynamic co-evolutionary interaction with their host population. As such, they do not eradicate their hosts. Instead, they live in some kind of stable hostility, resulting in an arms race between parasite and host (Haraguchi and Sasaki 1996). For vertebrate hosts and their obligatory directly transmitted pathogenic microbes this arms race is taking place between the hosts' immune system on the one hand, and the complex genetic systems that the pathogens develop on the other hand. The main tools pathogens have at their disposal are antigenic diversification and variation in replication rates. These two mechanisms are essential to the pathogen's continued struggle to evade the host's immune system that is adapting to control the proliferation of infectious organisms (Deitsch et al. 2009).

The clash of the above-mentioned antigenic players can lead to different optimal strategies for the pathogen. To achieve ecological success (i.e., persistence within a host population), pathogens require mechanisms both for survival within hosts and transmission between hosts. Deitsch et al. (2009) consider some of these strategies and in particular pay attention to the mechanisms of antigenic variation adopted by pathogens to avoid eradication by the hosts' immune system in order to maintain persistent infections. Herewith, they ensure the feasibility of their transmission to new hosts. Two main strategies combining within- and between-host dynamics are the so-called 'milker-like' and 'killer-like' strategies (van Baalen and Sabelis 1995). The 'milker-like' strategy relates to a pathogen replicating slowly within the host in an attempt to damage the host minimally. In contrast, the 'killer-like' strategy refers to a pathogen replicating very quickly and making the host ill. The latter strategy puts all emphasis on a high release rate of pathogens.

In most epidemiological models, the main tool used to identify the ecological success of a pathogen is the basic reproduction ratio, R_0 (Diekmann et al. 1990; Heesterbeek 2002). This measure of pathogen fitness is widely applied, because it

provides information regardless of the specific characteristics of the infection under consideration, such as prevalence, virulence change, host switch, and control escape. R_0 is defined as the expected number of secondary infections arising from a single individual during its entire infectious period, in a population of susceptible hosts. See, e.g., Anderson and May (1995), Diekmann and Heesterbeek (2000), Keeling and Grenfell (2000) and Heffernan et al. (2005). An important conclusion from their work is that if $R_0 < 1$, the infection does not disappear immediately, but subsequent generations of infected organisms are smaller and smaller in size. Alternatively, $R_0 > 1$ indicates a possible outbreak of the disease (van den Driessche and Watmough 2002).

Traditionally, R_0 was incorporated in studies that relate to between-host dynamics. However, to account for the complex genetic systems that pathogens have developed, nowadays much attention is given to within-host competition between strains of a given pathogen, as opposed to the implicit assumption that hosts are exploited by a single clone of pathogens (see e.g., Anderson and May 1979; Ewald 1983). When combined with the more traditional population-level approach, this relatively new approach to infectious disease modelling is very promising to gain better understanding of the arms race described above. Consequently, when using systems in which within- and between host processes are modelled simultaneously, R_0 depends on a combination of host-specific and pathogen-related characteristics. One of these characteristics is the remaining longevity of the host.

Increased computational capabilities brought on by technological advances has allowed to consider mathematical models that combine within- and between-host interactions as opposed to viewing the two systems in isolation (see e.g., Bhattacharya et al. 2014; Numfor et al. 2014). Identifying and analysing both host- and pathogen-specific characteristics that influence the spread of an infection in a host-population, is nowadays a topic of intense research. Traditionally, in this field much emphasis has been placed on factors such as force of infection, transmission probabilities, contact neighbourhoods, etc. Of late, the different transmission mechanisms also receive more and more attention. E.g., Rohani et al. (2009), and Brooks-Pollock et al. (2014) consider the impact of environmental transmission mechanisms on control efforts related to various microorganisms. Heinzmann et al. (2011) utilise density based models to investigate environmental factors and evaluate intervention programs. Alexander et al. (2009) evaluate various treatment strategies for influenza infection—with R_0 as threshold parameter - and the optimal timing of treatment campaigns.

In the basic definition of R_0 given in Eq. (9) below, the integral is taken over the entire infectious period. See, e.g., Diekmann and Heesterbeek (2000) and Anderson and May (1995). Although the upper boundary in this integral is taken as ∞, the effective length of integration is determined by the transmission rate $q(t)$ in the integrand, which is defined in Eq. (4). This transmission rate vanishes as soon as the infection dies out. This may be caused by recovery but also by the death of the host. The first is automatically included via (4), but the latter must be introduced by setting the upper boundary at a finite value. We remark that in simulation studies the upper boundary in the integral must always be set at a finite duration, as done by, e.g., Lange and Ferguson (2009), since numerically

the integration cannot exceed to ∞. However, as we show in this paper, the value used for the upper boundary may strongly determine the model predictions and can be decisive whether the milker-like or the killer-like strategy is optimal. The reason for this is not hard to understand: a milker-like strategy can only last for the specific chosen value of host longevity. In the literature some studies consider host longevity. Examples are a study of the evolution of virulence under the trade-off between transmission probability and host longevity (Sigmund et al. 2002), and a study of the effects of shortened host life span on the evolution of parasite life history and virulence (Nidelet et al. 2009).

The specific relationship between R_0, as a measure of pathogen fitness in a host population, and host longevity has not yet received attention. Therefore, we explicitly study the effect of a finite host longevity on R_0. To investigate this relationship we take an existing model from the literature with no extra mortality due to the infection. In our simulations we use dimensionless versions of these models and by applying numerical procedures that are computationally very efficient, we are able to perform extensive parameter scans. Taking the (remaining) host longevity variable enables us to draw new conclusions about how host longevity shapes the maximum pathogen fitness strategies.

2 Methods

We first discuss the modelling concepts underlying the present study. The model used is taken from Lange and Ferguson (2009). The numerical implementation is optimized to allow for many and long computer simulations. We started with replicating their results (see Fig. A1 in Appendix A) to check for computational correctness. Although the modelling principles used in our analysis are thus not new, we prefer to present them in this paper in a self-contained way to avoid unnecessary reference to the literature.

In the next section we describe the modelling of the dynamical evolution of an infection in a host. This *within-host model* represents how in an infected host the number of pathogens evolves in time. Because we allow mutations, the model may deal with a variable number of pathogen strains, each with its own dynamics, depending on their interactions with the immune system of the host and the availability of resources.

Next, we describe the model that represents how an infection spreads in a population. This *between-host model* incorporates the effect of one infected host in a direct and an indirect way: it may infect its direct neighbours directly, and it may infect its neighbours via other neighbours that were already infected by this host at an earlier time. To that end, we include in an averaged way the structure of the relational network of the population.

Eventually we come to the essence of the paper by coupling the two models and considering the R_0 of the coupled system. R_0 depends on host longevity and it is this dependency that gives rise to remarkable insights, as shown in the Sect. 3.

2.1 Within-Host Model

To describe how a pathogen replicates within a host and how the immune system of the host reacts in response to the presence of the pathogen, we use a model with a variable number of strains. The idea is as follows:

We start the invasion in the host with one specific strain. This strain replicates and builds up a viral load, and at the same time evokes an immune response. When replicating a mutation may take place with probability μ. Most mutations will not lead to a new strain type. The fraction of mutations that is 'successful' and generates a new strain type is equal to parameter δ. In the simulations this process of replication and mutation is modelled using a Poisson process. After the first successful mutation a second strain type comes into existence, with its own dynamics. These two strains may mutate, but may also die and disappear from the system. The consequence is that the number of strains n is stochastically varying. The immune system of the host may adapt to recognize new strains. The specific mortality rate of strain i due to the immune system of the host is weighted by the difference between between the new strain and the strains that are already present. As a quantative measure for this difference we use the so-called Hamming distance, which is explained below.

The state variables in the within-host model are: V_i, the number of pathogens of strain i, also referred to as pathogen load i ; X_i, the adaptive immunity of the host specific to strain i; and C, the resource level, i.e., the number of target cells available for any strain to multiply.

The dynamics of pathogen load V_i is governed by the following differential equation:

$$\frac{dV_i}{dt} \equiv [\text{pathogen growth}] - [\text{natural death}] - [\text{death due to acquired host immunity}]$$

$$= (1 - \mu)\,\rho\,V_i\,\frac{v_1 C}{v + v_1 C} - \psi\,V_i - \sigma\,V_i \sum_{k=1}^{n} y(\varrho_{ik})\,X_k. \tag{1}$$

where μ is the probability that a mutation occurs, ρ the replication rate of the pathogen, v_1 a conversion factor: it is the number of pathogens that may stem from one unit of resource, and v the total viral load in the host, $v \equiv \sum_{i=1}^{n} V_i$. Equation (1) links pathogen replication to two inhibiting mechanisms: host immunity and resource limitation. Pathogen growth is restricted via a Monod function that shows saturation behaviour if $v \gg v_1 C$. This term links the pathogen load V_i to the resource level C. The last two terms in (1) represent the clearance of pathogen: $\psi\,V_i$ gives the natural clearance of pathogens from a host, independent of any immune response, while the term with σ represents the decrease in pathogen load due to immunity acquired by the host. Here, the latter term includes cross immunity via $\sum_{k=1}^{n} y(\varrho_{ik})\,X_k$, which models via the Hamming distance ϱ_{ik} (see below) how *similar* strain i is to any other strain k that is currently in circulation.

In the computer simulations a strain is represented by a string of 5 elements (representing loci) and each element may attain 3 different values (representing alleles). Altogether this allows for 243 possible, different strains.

We use the Hamming distance ϱ_{ik} as a metric that relates similarities between antigenic variants, where ϱ_{ik} is defined as the fraction of loci at which strains i and k differ. The relative Hamming metric is widely used in the literature to measure antigenic variation. For other possible choices, see e.g. Cai et al. (2012), Plotkin et al. (2002), and Neher et al. (2016). The degree of cross-immunity is incorporated through the function $y(\varrho_{ik})$, where $y(\varrho_{ik}) = 1 - (1 - \chi)\varrho_{ik}$ if $\varrho_{ik} \leq 1/(1 - \chi)$, and $y(\varrho_{ik}) = 0$ otherwise. Parameter χ thus regulates the degree of cross-immunity with $\chi = 0$ in case of total cross immunity. Increasing χ values correspond with less and less cross immunity.

The dynamics of the adaptive immune response X_i is governed by

$$\frac{dX_i}{dt} \equiv [\text{decline of immunity}] + [\text{acquisition of immunity}]$$
$$= \xi\,(x_0 - X_i) + \zeta\,X_i\,\frac{V_i}{\eta + V_i}. \tag{2}$$

In (2), the first term describes the decline of immunity with x_0 the minimum immune level. Note that in general $X_i > x_0$. The acquisition of immunity depends on the load of strain i via a Monod function.

The dynamics of the resource level C is modelled as:

$$\frac{dC}{dt} \equiv [\text{limited exponential growth}] - [\text{use by present pathogens}]$$
$$= \kappa\,(C_0 - C) - \frac{\rho}{\nu_1} \sum_{i=1}^{n} V_i\,\frac{\nu_1 C}{\nu + \nu_1 C}. \tag{3}$$

In any host, resource is being replenished at a rate κ, with C_0 a maximum value for C, chosen to represent a realistic number of target cells. Note that in general $C < C_0$.

As in an infected host each pathogen strain i grows at a rate proportional to $V_i\,(\nu_1 C/(\nu + \nu_1 C),)$ the resource is depleted proportionally to the sum of all these growth terms.

Because we allow the generation of new strains, the number of pathogen strains is variable, and so is the number of differential equations in our model. If at some moment in time we have n different strains, the model consists at that moment of $2n + 1$ differential equations. We solve differential equations (1), (2), and (3) to obtain the time evolution of $V_i(t), X_i(t), i = 1, \ldots, n$, and $C(t)$ until a new successful mutation takes place. At that moment we replace n by $n + 1$ and extend the set of model equations with two new equations. It may also happen that we have to decrease n with 1, namely when one of the pathogen strains goes extinct. This happens when $V_i \leq \nu_0$.

The model in Eqs. (1–3) yields state variables that vary greatly in magnitude, with total pathogen load reaching levels of $\nu \approx 10^{11}$, while adaptive immunity is saturated at $\eta = 10^5$. In contrast with these high values, most parameter values (see Table 1) are of order 1. To allow us to investigate the relationships between parameters and the influence of specific parameters on the dynamics of the system, it is therefore essential to

Table 1 Parameters and variables of the within host model

	Description	Default value
Parameter		
ρ	Replication rate of pathogen	$\rho \in [3, 8]$
δ	fraction of mutations that are 'successful'	$\delta \in [10^{-9}, 10^{-3}]$
μ	Probability that a replication leads to a mutation	0.1
v_0	Initial and minimum pathogen load	10
v_1	Conversion factor from resource to pathogens	10^3
C_0	Initial/maximum resource level	10^8
ψ	Clearance rate of pathogen	0.25/day
σ	Immune-related clearance rate	10^{-3}/day
η	Critical load of saturated immune response	10^5
ξ	Decline of immunity	0.3/day
x_0	Initial/minimum immunity	1
ζ	Growth of immunity	0.8/day
κ	Replenishment of resources	1/day
χ	Degree of cross-immunity	0.6
Variable		
$V_i(t)$	Number of pathogens of strain i	
$X_i(t)$	Adaptive immunity against strain i	
$C(t)$	Number of target cells available for a strain to multiply	
$v(t)$	Total pathogen load	$\sum_{i=1}^{n} V_i(t)$

rewrite the model in dimensionless form. For the non-dimensional version we refer to Appendix B. The descriptions of the parameters and state variables of the within-host model are listed in Table 1.

2.2 Between-Host Model

As for the transmission of the pathogen between hosts, our model is based on the use of a network structure as investigated in Keeling (1999). Let us focus on one initially infected host connected to $N - 1$ susceptible, but still healthy neighbours. As an example, a sub-network with $N = 7$ nodes is sketched in Fig. 1.

In this part of the network the number S of susceptible nodes is thus initially $S(0) = N - 1$. $S(t)$ will decrease in time, because the infected host may transfer its infection to its neighbours. The transmission rate q will depend on the pathogen load in the infected host, and thus on its age. It is modelled as

$$q(t) = \beta \left[1 - \exp \left(-\frac{v(t)}{v_T} \right) \right].
\tag{4}$$

So, $q(t)$ depends on the between-host parameter β (transmission rate) and the dynamic within-host total pathogen load, $v(t)$. Parameter v_T is the infectiousness
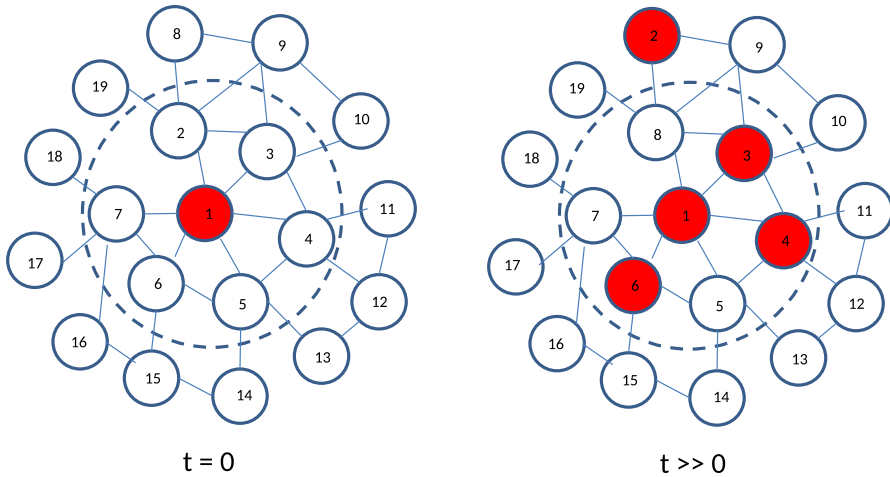
Fig. 1 Left: initial network of a susceptible population with one infected host (node 1). This infected host is surrounded by 6 direct neighbours (within the dashed circle), so for the local subnetwork we have $N = 7$. Right: the same part of the network after some time. Some of the neighbours of node 1 are now also infected. The infection of node 4 could be caused by its link with node 1 (direct infection), but could also stem from its contact with node 3 (secondary infection). Note that some of the infected neighbours in the local neighbourhood of node 1 may be in the meantime replaced by new susceptible ones. In the present case this happened to node 2, which has exchanged position with node 8

threshold. If $v \gg v_T$, q will no longer depend on v and converge to β. The transmission rate β is defined as $\beta \equiv \alpha \gamma / N$ with α the contact rate and γ the transmission probability per contact ($\gamma \in (0, 1)$).

It remains to model how $S(t)$ evolves in time. This is done in an averaging way. In our sub-network the total number N of nodes is conserved and the number I of infected nodes and the number S of susceptible nodes sum up to the total number of nodes, $I + S = N$. At the initial time $t = 0$, we can simply write

$$\frac{dI}{dt}(t) = q(t) S(t), \tag{5}$$

because only direct infections may occur. However, at times $t > 0$ the infestation of a susceptible could not only be direct, but also secondary, i.e., due to other neighbours that in the meantime got infected by the originally infected host. When these neighbours were infested $\tau < t$ times ago, this happened with a chance proportional to the transmission rate $q(t - \tau)$ at that moment. This leads to the extended integro-differential equation

$$\frac{dI}{dt}(t) = S(t) \left[q(t) + \phi \int_0^t \frac{dI}{dt}(\tau) \, q(t - \tau) \, d\tau \right]. \tag{6}$$

Here, ϕ defines the cliquishness of the network, i.e. the proportion of neighbours of a node who are neighbours of each other. The integral represents the secondary

infections caused by new infectives up to time t and takes into account the changing transmission rates resulting from time dependent pathogen loads.

Using that $dI/dt = -dS/dt$, we can rewrite this in the form

$$\frac{dS}{dt}(t) = -S(t)\left[q(t) - \phi\left(q * \frac{dS}{dt}\right)\right],\tag{7}$$

with the convolution of two functions $f(t)$ and $g(t)$ defined in the usual way as $(f * g)(t) \equiv \int_0^t f(\tau)\, g(t - \tau)\, d\tau$.

As a last modelling step we assume that the infected nodes in our sub network are replaced at a rate ω by susceptible nodes from the environment in a constantly ongoing exchange process. That process is in a quasi- steady state, so that N remains conserved. In (7) this effect is incorporated by replacing dS / dt with $dS/dt - \omega I$. Eventually we then arrive at

$$\frac{dS}{dt} = \omega\,(N - S) - S\left[q + \phi\, q * \left(\omega(N - S) - \frac{dS}{dt}\right)\right].\tag{8}$$

The parameters of the between-host model, together with descriptions, are listed in Table 2.

2.3 Basic Reproduction Ratio R_0

The basic reproduction ratio, R_0, is a fundamental concept in epidemiology. It is defined as the total number of secondary infections brought on by a single primary infection, in a virgin, susceptible population. The value of R_0 indicates whether a single infection may lead to an epidemic or will probably die out soon. According to Diekmann and Heesterbeek (2000), R_0 is given by the integral

$$R_0 = \int_0^\infty S(t)\, q(t)\, dt.\tag{9}$$

Table 2 Parameters and variables of the between host model

	Description	Default value
Parameter		
α	Contact rate between hosts	$\alpha \in [0, 10]$
ω	Replacement rate of hosts	0.001
γ	Probability of transmission	0.2
β	Transmission rate	$\beta = \alpha\gamma/N$
ϕ	Cliquishness	0.75
v_T	Infectiousness threshold	10^8
ϖ	Expected maximum pathogen load	10^{11}
N	Neighbourhood size	20
Variable		
S	Number of susceptible hosts	
$q(t)$	Transmission rate between hosts	$q = \beta\left(1 - e^{-v(t)/v_T}\right)$

Here, $q(t)$, defined in (4), is the transmission rate from the infected host to the surrounding susceptibles. This rate depends on the pathogen load in the infected host, which is obtained by evaluating the within-host model. $S(t)$ is the number of susceptibles in the neighbourhood of the primary infective. It is obtained as a solution of the between-host model in (7) or (8).

The upper boundary in the integral in (9) is ∞. In computer simulations one can only work with a finite upper boundary. In Lange and Ferguson (2009) the authors interpret the upper bound as the maximum duration of infection and they take as a cut off point a fixed value of 2 years. This value is in many practical cases sufficiently long. In general however, this may introduce inaccurate results, since in practice the maximum duration of an infection is limited by the host's (remaining) longevity. That is why we take in this paper a finite upper boundary D_{max} and use the definition

$$R_0 = \int_0^{D_{max}} S(t)\, q(t)\, dt. \tag{10}$$

It is clear that $S(t)$ and $q(t)$, and thus also R_0, depend in quite a complex way on the model parameters summarized in Tables 1 and 2. If our only goal would be to show how R_0 depends on D_{max}, it would suffice to fix all parameter values, and to vary D_{max}. However, our aim is more intricate: we want to look at strategies that pathogens may develop under different circumstances. For this we assume that the pathogen attempts to optimize R_0 in the course of time. For this purpose, the pathogen has two parameters at its disposal, namely the replication rate ρ, and the fraction of successful mutations δ. To find the optimal strategy that a pathogen will develop in an evolutionary process, we optimize R_0 with respect to these two within-host parameters. As for the between-host dynamics, the dominant parameter is the contact rate α. This parameter is not under control of the pathogen. Consequently, we consider $R_0 = R_0(\rho, \delta; \alpha, D_{max})$ and we are in particular interested in the dependence of R_0 on the contact rate α and host longevity D_{max}. That's why we optimize over ρ and δ:

$$R_0^{opt}(\alpha, D_{max}) = \max_{(\rho, \delta)} R_0(\rho, \delta; \alpha, D_{max}). \tag{11}$$

R_0^{opt} measures the 'success' of a pathogen when adjusting its replication and mutation parameters ρ and δ.

The optimization of Eq. (11) is done by calculating $R_0(\rho, \delta; \alpha, D_{max})$ on a rectangular grid in the (ρ, δ) plane and the optimum is simply found by comparison of all grid points.

Solving the within-host system of the system of differential equations is done numerically by using a fourth order Runge Kutta method. Subsequently, Eq. (11) is maximized using the well-known optimisation method of Levenberg and Marquardt.

3 Results

We combined and implemented the within- and between-host models in Eqs. (1–8) and, as a check, first recalculated the results as reported in Lange and Ferguson (2009). We found the same results; they are given in Appendix A and B. Next, we focussed at finding how the evolutionary optimal R_0^{opt} depends on contact rate α between hosts and host longevity D_{max}.

To ensure realistic within-host model dynamics that allows for an infection to become endemic in a host, it suffices to use the ranges $\rho \in [3, 8]$ and $\delta \in [10^{-9}, 10^{-3}]$. For the contact rate we take as a realistic range $\alpha \in [0, 10]$. The maximum of R_0 over (ρ, δ) for fixed values of α and D_{max} is found by calculating R_0 on a grid in the (ρ, δ)-plane. For this we need to evaluate the within- and between-host models for many (ρ, δ)-pairs. This time consuming procedure yields a so-called fitness landscape. An example is given in Fig. A1 of Appendix A, where we not only present R_0, but also the cumulative pathogen load and the duration of the infection, as functions of (ρ, δ).

In Fig. 2 we show R_0^{opt} as a function of α, for the values $D_{max} = 500, 1000, 1200,$ 1500, and 2000 days. The results in Fig. 2 exhibit remarkable behaviour. If D_{max} is relatively short ($D_{max} = 500$) days, the dependence of R_0^{opt} on α is nearly linear over the full interval that we tested. This implies that, when the contact rate increases, the success of the pathogen in spreading itself in a host population goes up in a way proportional to the number of contacts, as is to be expected.

However, for high values of D_{max} ($D_{max} \in (1200-2000)$ days), the linear dependence on α is shown on a much smaller interval. R_0^{opt} tends to increase fast to a high value of about 32 as soon as α exceeds 1, and increases from there on only very slowly as a function of α. In the limiting case of $D_{max} \to \infty$, R_0^{opt} immediately attains the value of 32 as soon as $\alpha > 0$, and shows a linear dependence on α for all α, but with a very small slope.

When observing the results in Fig. 2, the question arises for which combination of pathogen replication rate ρ and mutation rate μ, R_0 attains its maximum, and how
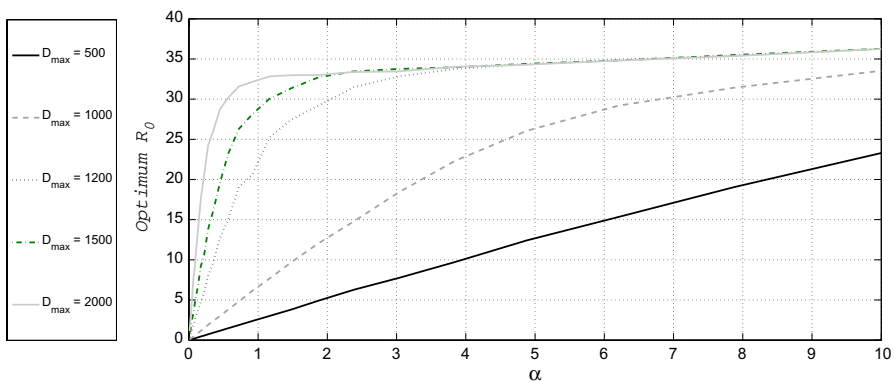

Fig. 2 R_0^{opt} as a function of the contact rate between hosts (α), for various values of host longevity D_{max}

this combination changes when host longevity D_{max} and/or between hosts contact rate α vary. This question is answered via the information in Fig. 3.

When $D_{max} \leq 1000$ days (≈ 3 years), we observe in Fig. 3 two completely distinct R_0^{opt} regions: first, a region centred at about $\rho = 3$ and $\delta = 10^{-3}$ (low replication and high diversity, the 'milker-like' region, and second, a region centred at about $\rho = 8$ and $\delta = 10^{-6}$ (high replication and intermediate diversity, the 'killer-like' region). Low contact rates α are associated with the killer-like region and high rates with the milker-like region. This is to be expected: if contacts are rare, the pathogen should optimize the transmission rate of the disease by increasing the viral load during contact, because this is its only option to survive. The viral load is high in the killer-like regime and low(er) in the milker-like regime. A striking observation is that the jump between the regions occurs at a critical value of α, which we denote by α_c. On its turn, this critical value depends on D_{max}. In Fig. 4 we show how α_c behaves as a function of D_{max}. The curve in Fig. 4 divides this plane into two regions with completely distinct pathogen behaviour. Above the curve the 'milker-like' behaviour (replication rate ρ low and mutation rate δ high) is optimal for the pathogen and below the curve 'killer-like' behaviour (replication rate ρ high and mutation rate δ intermediate).

Another remarkable detail observed in Fig. 4 is that α_c vanishes if D_{max} is bigger than some threshold, which occurs for our parameter values (taken from Lange and Ferguson 2009) at $D_{max} = 1425$ days. Above that threshold the killer-like region disappears and all optimum R_0 values are attained in the milker-like region. This implies that for D_{max} larger than this threshold the system always ultimately converges to approximately the same 'optimal' values for ρ and δ, namely $\rho \approx 3$ and $\delta \approx 10^{-3}$. It should be noticed that this effect will only be seen if the epidemics gets
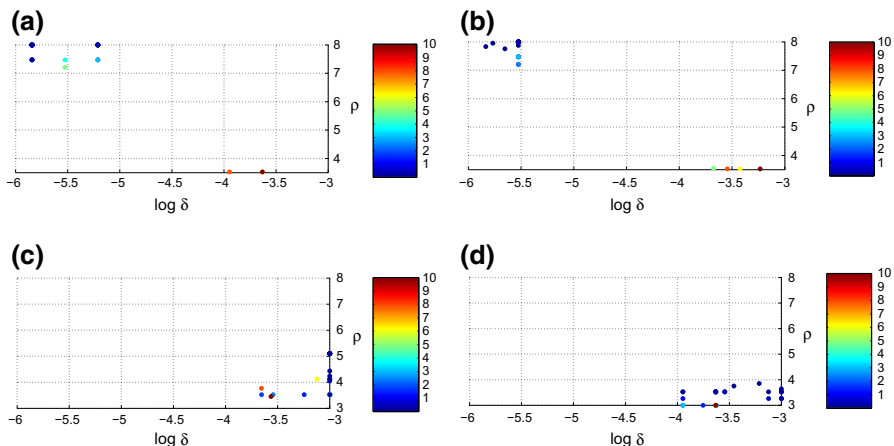
Fig. 3 Position of R_0^{opt} in the (ρ, δ) plane, with ρ the replication rate of pathogen and δ the mutation rate of pathogen within the host. This position depends on the between-host contact rate α. The values of $\alpha \in (0, 10)$ are indicated with colours and colour bars. The position of R_0^{opt} also depends on remaining host longevity D_{max}. Results are shown for 4 values of D_{max}: **a** $D_{max} = 500$, **b** $D_{max} = 1000$, **c** $D_{max} = 1500$, and **d** $D_{max} = 2000$
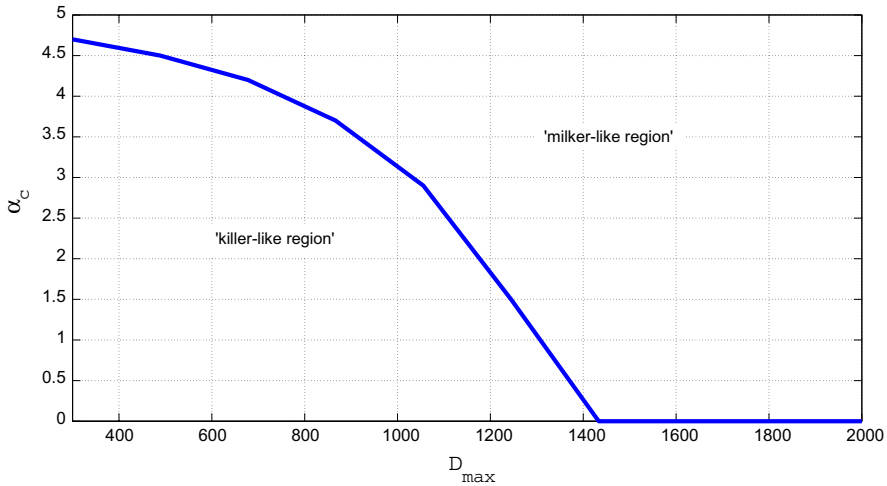
Fig. 4 Behaviour of the critical value α_c as a function of D_{max}. The curve divides the (D_{max}, α) plane into two regions. Below the curve the optimal strategy for the pathogen is the 'killer-like' behaviour, above the curve the 'milker-like' behaviour is optimal

enough time to adjust its parameters, since that is the philosophy underlying the introduction of R_0^{opt} in (11). In this limiting case, R_0^{opt} still increases linearly with α, but quite slowly. It should be noted that the milker-like strategy is the one that eventually will show up if one waits long enough. In reality some diseases are still in the region of transient behaviour and thus on their way to a milker-like regime, but now still displaying killer-like behaviour.

4 Discussion

A lot of studies relate pathogen success or failure to R_0 by investigating how changes in the parameters governing model dynamics lead to changes in R_0. The commonly considered factors, like force of infection, transmission probabilities and contact neighbourhoods are important. This certainly leads to more insight pertaining to the success or failure of infectious disease progression. Here, we have shown that the factor "remaining longevity of the host", also influences the basic reproduction number of an epidemic and thus can essentially determine whether an outbreak is to be expected or not. We used a specific model from the literature to investigate the relation between R_0 and remaining host longevity D_{max}. The within-host model used allows for variation in the diversity of the infectious agent, as well as for different rates of replication. Coupling it with the population-level model, we are able to show how the optimal strategy for the pathogen depends on the parameters. It should be noted that these optimal strategies will develop in the end and that during the transient phase suboptimal strategies could be observed. We found that optimal strategies depend heavily on D_{max}, as is clearly shown in Fig. 2. We conclude that the

value of D_{max} should be chosen in correspondence with the specific host-pathogen system under consideration. The value of D_{max} may vary from very short, e.g., in the cases of fly or mosquito, to very long, e.g., for humans. One of the consequences of this insight is that the results given in Lange and Ferguson (2009), from which we adopted the modelling ideas, are valid for systems in which the hosts live at most 2 years after the start of the infection, because they fixed D_{max} at that period, assuming that it would not change their conclusions. The present work shows that varying (remaining) host longevity can change conclusions drastically: in our long run simulations to determine the optimal R_0 (Eq. 11), one of the three different regimes that Lange and Ferguson (2009) distinguished never pops up as the optimum strategy.

When expected host longevity is relatively short, we find that optimum between-host R_0 only emerges in two clearly distinct regions in the plane spanned by the two relevant pathogen characteristics, namely δ, ther probability of a mutation to be successful, and ρ, the replication rate. These two regions could be associated with the well-known 'milker-like' (low replication rate ρ) and 'killer-like' (high ρ) strategies. It should be noted that longevity may vary over sub-populations. For example, for humans longevity really hinges on the food availability in the environment and it ranges in WHO's (2016) report between 50.1 and 83.7 years. This might also be the case for other species.

For given host longevity D_{max}, the choice between the milker-like and killer-like strategies depends on the contact rate α. Figures 3 and 4 show that there exists a critical value α_c where the transition from killer-like to milker-like as the best strategy occurs. Figure 4 also shows that if D_{max} exceeds some threshold value, the milker-like strategy is always optimal, independent of the contact rate α. For the parameter setting used throughout this paper this threshold is approximately 4 years. This value itself is not the most important discovery, but the achieved insight is that such a threshold exists and can be quite easily calculated. If we combine the information contained in Figs. 3 and 4, we not only conclude that for D_{max} above this threshold the milker-like strategy is optimal, but also that R_0 in this regime linearly increases with the contact rate α, although the slope of this increase is very small.

Our analysis suggests that in long-living hosts like humans most infectious diseases will finally attain a milker-like strategy. That does not imply that all human diseases are already of the milker-like type, but diseases that are already for a long time among us tend to be of milker-like type. Counterexamples are the diseases, so-called zoonotic diseases, that jump from animals to the human population. Examples like bird influenza and Ebola show that diseases that come from organisms with a shorter longevity. For example, bird influenza (de Jong et al. 2009) and fruit bats (Ebola, WHO 2017) are highly virulent (killer-like type).

It should be noted that both an individual based model or a physiologically structured population model incorporating a contact network, the within-host dynamics of the pathogen, and the stochastic longevity of hosts, might be able to show the impact of host longevity on R_0 in a more rigorous way. However, this kind of modellling is beyond the scope of the present analysis.

The presented results show that host longevity really matters when pathogen fitness is investigated. When using R_0 as a measure of fitness, D_{max} should be taken

equal to the expected remaining host longevity. The optimum infection strategies can only be reliably predicted if one uses appropriate estimates for host longevity.

5 Supporting Information

Appendix A This appendix gives some typical results obtained through numerical simulations. We show typical types of within-host behaviour, as well as the fitness landscapes that result from the combination of the two models.

Appendix B In this section of the supporting documentation, we show how the within- and between-host models, given in the Methods section, have been made dimensionless and how they have been solved numerically.

References

Alexander ME, Dietrich SM, Hua Y, Moghadas SM (2009) A comparative evaluation of modelling strategies for the effect of treatment and host interactions on the spread of drug resistance. J Theor Biol 259(2):253–263. https://doi.org/10.1016/j.jtbi.2009.03.029

Anderson RM, May RM (1979) Population biology of infectious diseases: part I. Nature 280:361–367

Anderson RM, May RM (1995) Infectious diseases of humans. Oxford Science Publications, Oxford

Bhattacharya S, Martcheva M, Li XZ (2014) A predator–prey-disease model with immune response in infected prey. J Math Anal Appl 411(1):297–313. https://doi.org/10.1016/j.jmaa.2013.09.031

Brooks-Pollock E, Roberts GO, Keeling MJ (2014) A dynamic model of bovine tuberculosis spread and control in Great Britain. Nature 511:228–231 **(Brunham et al. (1993) or Brunman et al. (1993))**

Cai Z, Zhang T, Wan XF (2012) Antigenic distance measurements for seasonal influenza vaccine selection. Vaccine. https://doi.org/10.1016/j.vaccine.2011.10.051

de Jong MCM, Stegeman A, van der Goot J, Koch G (2009) Intra- and interspecies transmission of H7N7 highly pathogenic avian influenza virus during the avian influenza epidemic in the Netherlands in 2003. Rev Sci Tech Off Int Epizoot 28:333–340

Deitsch KW, Lukehart SA, Stringer JR (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat Rev Microbiol 7:493–503. https://doi.org/10.1038/nrmicro2145

Diekmann O, Heesterbeek JAP (2000) Mathematical epidemiology of infectious diseases. Wiley, New York. ISBN 0-471-49241-8

Diekmann O, Heesterbeek JAP, Metz JAJ (1990) On the definition and the computation of the basic reproduction ratio Ro in models for infectious diseases in heterogeneous populations. J Math Biol 28:365–382

Ewald PW (1983) Host-parasite relations, vectors, and the evolution of disease severity. Ann Rev Ecol Syst 14:465–85

Haraguchi Y, Sasaki A (1996) Host–parasite arms race in mutation modifications: indefinite escalation despite a heavy load? J Theor Biol 183(2):121–137

Heesterbeek JAP (2002) A brief history of R_0 and a recipe for its calculation. Acta Biotheor 50:189–204

Heffernan JM, Smith RJ, Wahl LM (2005) Perspectives on the basic reproductive ratio. J R Soc Interface 2(4):281–293

Heinzmann D, Barbour AD, Torgerson PR (2011) A mechanistic individual-based two-host interaction model for the transmission of a parasitic disease. Int J Biomath 4:4436–460. https://doi.org/10.1142/S1793524511001313

Keeling MJ (1999) The effects of local spatial structure on epidemiological invasions. Proc R Soc Lond B 266:859–867

Keeling MJ, Grenfell BT (2000) Individual-based perspectives on R_0. J Theor Biol 203(1):51–61

Lange A, Ferguson NM (2009) Antigenic diversity, transmission mechanisms, and the evolution of pathogens. PLoS Comput Biol 5(10):e1000536

Neher RA, Bedford T, Daniels RS, Russel CA, Shraiman BI (2016) Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. PNAS 113(12):E1701–E1709

Nidelet T, Koella JC, Kaltz O (2009) Effects of shortened host life span on the evolution of parasite life history and virulence in a microbial host–parasite system. BMC Evol Biol 9:65. https://doi.org/10.1186/1471-2148-9-65

Numfor E, Bhattacharya S, Lenhart S, Martcheva M (2014) Optimal control in coupled within-host and between-host models. Math Model Nat Phenom 9(4):171–203. https://doi.org/10.1051/mmnp/20149411

Plotkin JB, Dushoff J, Levin SA (2002) Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. PNAS 99(9):6263–6268

Rohani P, Breban R, Stallknecht DE, Drake JM (2009) Environmental transmission of low pathogenicity avian influenza viruses and its implications for pathogen invasion. PNAS 106(25):10365–10369. https://doi.org/10.1073/pnas.0809026106

Sigmund K, Sabelis MW, Dieckmann U, Metz JAJ (2002) Adaptive dynamics of pathogen-host interactions. In: Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (eds) Adaptive dynamics of infectious diseases: in pursuit of virulence management. Cambridge University Press, International Institute for Applied Systems Analysis, Cambridge, Laxenburg, pp 39–59

van Baalen M, Sabelis MW (1995) The milker–killer dilemma in spatially structured predator–prey interactions source. Oikos 74:391–400

van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180:29–48

WHO (2016) Life expectancy increased by 5 years since 2000, but health inequalities persist. In World Health Statistics: Monitoring Health for the SDG's, p 7. ISBN 978-92-4-156526-4

WHO (2017) Ebola virus disease. http://www.who.int/mediacentre/factsheets/fs103/en/. Accessed 31 Jan 2018