Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Theoretical Biology 503 (2020) 110388

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Prophylactic host behaviour discourages pathogen exploitation

Evan Mitchell *, Geoff Wild

Department of Applied Mathematics, Western University, London, ON N6A 3K7, Canada

ARTICLE INFO

Article history:
Received 26 February 2020
Revised 13 May 2020
Accepted 19 June 2020
Available online 9 July 2020

Keywords:
Disease dynamics
Evolution
Game theory
Mathematical model

ABSTRACT

Much work has considered the evolution of pathogens, but little is known about how they respond to changes in host behaviour. We build a model of sublethal disease effects where hosts are able to choose to engage in prophylactic measures that reduce the likelihood of disease transmission. This choice is mediated by utility costs and benefits associated with prophylaxis, and the fraction of hosts engaged in prophylaxis is also affected by population dynamics. When prophylactic host behaviour occurs, we find that the level of pathogen host exploitation is reduced, by the action of selection, relative to the level that would otherwise be predicted in the absence of prophylaxis. Our work emphasizes the significance of the transmission-recovery trade-off faced by the pathogen and the ability of the pathogen to influence host prophylactic behaviour.

1. Introduction

Mathematical study of infectious diseases has a long history that predates even the well-known contributions of Ross (1916) and Kermack and McKendrick (1927). Mathematical study of the evolution of pathogens, however, is relatively recent. One avenue of inquiry has explored the ways in which natural selection shapes the virulent effects pathogens inflict on their hosts. Models have provided insight into the ways in which various factors—such as parasite reproductive rates (Bremermann and Pickering, 1983), host density (Knolle, 1989), relatedness among co-infecting pathogen strains (Frank, 1992), and multiple types of hosts (Gandon, 2004)—modulate the expressed level of virulence. A key prediction emerging from this body of work is that, in many cases, selection acts to maximize the basic reproduction number (Anderson and May, 1982; Day and Burns, 2003), i.e., the expected number of new infections caused by a single infective host (Heffernan et al., 2005). In particular, adaptive levels of pathogen virulence balance a trade-off between the average duration of infection and the rate of disease transmission (Alizon and Michalakis, 2015; Cressler et al., 2016). As disease transmission rates slow, then, standard theory predicts selection will respond by shaping virulence in a way that decreases the duration of infection.

Mathematics has also contributed to our understanding of the ways in which host traits impact the evolution of pathogen virulence. Here, models have explored the effects of co-evolution with innate host defences (Day and Burns, 2003; van Baalen, 1998), use of antibiotics (Reluga, 2005), vaccines and vaccination behaviour (Bauch and Earn, 2004; Murall et al., 2015), and other social factors related to hosts themselves (Bauch and Galvani, 2013). It is this last item—namely, the effect of host social behaviour on pathogen evolution—on which we focus our attention in this paper. Schaller (2011) discusses the idea of a "behavioural immune system" that complements the standard physiological immune system in humans. This behavioural immune system is comprised of various prophylactic measures, such as social distancing (e.g., avoiding handshakes when greeting) or improved personal hygiene (e.g., hand washing), that individuals may adopt to reduce the likelihood of infection (Schaller, 2011). Importantly, individuals can start and stop these behaviours as often as desired (e.g., as in Pharaon and Bauch (2018)), as opposed to measures like vaccination that are, in a sense, irreversible.

Unfortunately, little is known about how hosts' behavioural immune system impacts the evolution of pathogens. What work has been done in this area predicts that prophylactic behaviour exhibited by hosts can select for higher pathogen virulence, assuming that the perceived severity is higher for the more virulent strain and that the prophylactic measures are more effective against the less virulent strain (Pharaon and Bauch, 2018). This work, however, considers short-term evolutionary outcomes only, and does not consider the effects on host behavioural changes of factors outside of social learning. In particular, the model of host behavioural dynamics in Pharaon and Bauch (2018) does not account for all the ways in which host demographics and disease dynamics could affect the proportion of susceptible individuals engaging in prophylaxis. In order to assess additional risks posed by pathogen evolution, then, different models are required.

* Corresponding author.
E-mail address: emitch44@uwo.ca (E. Mitchell).

https://doi.org/10.1016/j.jtbi.2020.110388
0022-5193/

We devise a model that tightly couples changes in the host's behavioural immune system, host demographics, and disease dynamics, in a way that allows us to make predictions about the long-term evolution of pathogen host exploitation. We focus on pathogen exploitation instead of virulence as we are considering sublethal disease effects and low-risk host behaviours (e.g., hand washing). An important aspect of the host behaviours studied here is the relation between the utility cost associated with engaging in that behaviour, which may take many different forms, and the benefit in terms of a reduced likelihood of pathogen transmission. A simple example would be more frequent hand washing, where an individual engaging in this behaviour will benefit from a lower risk of contracting the disease at the cost of some extra time out of their day and having to spend more money on soap than they would otherwise. Another, more compelling, example is social distancing, which again reduces disease transmission but can have much more severe social and economic costs as can be seen in the current COVID-19 pandemic (Anderson et al., 2020). We find that prophylactic behaviour uniformly reduces the pathogen's exploitation below the level expected in the absence of such behaviour. Furthermore, we argue that the driving force behind our result is the modified nature of the transmission-recovery tradeoff faced by the pathogen as a result of the inclusion of host prophylactic behaviour.

2. Model

2.1. Disease dynamics

We begin with a version of an endemic SIR model of infectious-disease dynamics (Britton, 2003) modified in a way that separates a host population of total size N into two groups. Individuals in group $i = 0$ are those who do not take prophylactic measures that limit (but do not completely prevent) disease transmission, whereas those in group $i = 1$ do take such measures. Individuals in each group are further subdivided according to their disease status. Let S_i, I_i, and R_i denote the number of individuals in group i who are susceptible to infection, infective, and recovered from infection, respectively.

We assume that transmission is frequency-dependent. This is a common assumption when modelling sexually transmitted infections (STIs) since contacts between individuals in those cases are generally not the result of random mixing (Antonovics et al., 1995; McCallum et al., 2001). While we are not considering STIs here, some prophylactic behaviours may create similar contact patterns within the population. For example, in the case of social distancing, individuals intentionally limit their contact with others, and adding more people to the population may not greatly affect the average contact rate. This would make frequency-dependent transmission a more appropriate model.

Each individual encounters another at a fixed rate and, given that an encounter is between a susceptible and an infective, the likelihood of disease transmission depends on the groups to which individuals belong. If β_{ij} denotes the product of the probability of disease transmission from a j-infective to an i-susceptible and the per-capita encounter rate, then $S_i\beta_{ij}I_j/N$ gives us the total rate at which new infections are created in group i. Infective individuals recover at a fixed per-capita rate, γ, independent of their group. As a result of recovery, individuals are imbued with life-long immunity to future infection. Since we are considering sublethal disease effects, we do not include disease-related mortality (virulence) in our model.

Individuals can also switch groups. For now, we use the constants τ_{ij}, ϕ_{ij} and η_{ij} to represent the per-capita rates at which susceptible, infective, and recovered individuals, respectively, switch from group j to i. We expand on the details surrounding group switching later. We can summarize the description above using a system of differential equations. Scaling time so that the background death rate is unity (i.e., one time unit is equivalent to the average lifetime of an individual in the population), and matching birth and death rates, we get

$$\frac{dS_0}{dt} = N - S_0\frac{\beta_{00}}{N}I_0 - S_0\frac{\beta_{01}}{N}I_1 - S_0 + \tau_{01}S_1 - \tau_{10}S_0 \tag{1a}$$

$$\frac{dS_1}{dt} = -S_1\frac{\beta_{10}}{N}I_0 - S_1\frac{\beta_{11}}{N}I_1 - S_1 - \tau_{01}S_1 + \tau_{10}S_0 \tag{1b}$$

$$\frac{dI_0}{dt} = S_0\frac{\beta_{00}}{N}I_0 + S_0\frac{\beta_{01}}{N}I_1 - (1+\gamma)I_0 + \phi_{01}I_1 - \phi_{10}I_0 \tag{1c}$$

$$\frac{dI_1}{dt} = S_1\frac{\beta_{10}}{N}I_0 + S_1\frac{\beta_{11}}{N}I_1 - (1+\gamma)I_1 - \phi_{01}I_1 + \phi_{10}I_0 \tag{1d}$$

$$\frac{dR_0}{dt} = \gamma I_0 - R_0 + \eta_{01}R_1 - \eta_{10}R_0 \tag{1e}$$

$$\frac{dR_1}{dt} = \gamma I_1 - R_1 - \eta_{01}R_1 + \eta_{10}R_0. \tag{1f}$$

Here, we make the assumption that all newborns enter the S_0 compartment. While realistically we might expect a fraction to enter the S_1 compartment (e.g., through cultural vertical transmission of prophylactic behaviours), the presence of the switching terms allows individuals in the S_0 compartment to immediately move into the S_1 compartment if they choose to do so, justifying our choice of modelling births as entering only the S_0 compartment.

Note that the differential equations in (1) sum to zero, and so total population size N is constant. This, along with the fact that group membership among recovered individuals is of no consequence, allows us to omit (1e) and (1f). We now use $u_i = S_i/N$ and $v_i = I_i/N$ to denote the fraction of susceptible and infective individuals, respectively, in group i. Similarly, we use $u = u_0 + u_1$ and $v = v_0 + v_1$ to denote the total fraction of susceptible and infective individuals, respectively. The dynamics of u and v can then be modelled by the following set of differential equations:

$$\frac{du}{dt} = 1 - [y(\beta_{01}(1-x) + \beta_{11}x) + (1-y)(\beta_{00}(1-x) + \beta_{10}x)]uv - u \tag{2a}$$

$$\frac{dv}{dt} = [y(\beta_{01}(1-x) + \beta_{11}x) + (1-y)(\beta_{00}(1-x) + \beta_{10}x)]uv - (1+\gamma)v. \tag{2b}$$

We also track the fraction of susceptible and infected individuals taking prophylactic measures using $x = u_1/u$ and $y = v_1/v$, respectively. Given this definition of x, we can derive the following differential equation to describe how the proportion of susceptible individuals engaging in prophylaxis changes due to various factors:

$$\frac{dx}{dt} = -\overbrace{\frac{x}{u}}^{\text{births}} - \overbrace{\tau_{01}x + \tau_{10}(1-x)}^{\text{group switching}} + \overbrace{(\beta_{00}(1-y)v + \beta_{01}yv)(1-x)x}^{\text{infection of hosts not engaged in prophylaxis}} - \underbrace{(\beta_{10}(1-y)v + \beta_{11}yv)x(1-x)}_{\text{infection of hosts engaged in prophylaxis}}. \tag{2c}$$

The first term represents the fact that births increase the pool of individuals not engaging in prophylactic measures, which in turn reduces the relative proportion of susceptible individuals engaging in prophylactic measures. The τ_{ij} terms capture the effects of susceptible individuals switching between engaging and not engaging in prophylaxis. The final two terms correspond to infection. In one case, infection reduces the pool of individuals not engaging in prophylaxis and subsequently increases the relative proportion of individuals engaging in prophylaxis. Conversely, the proportion of susceptible individuals engaging in prophylaxis is directly reduced by infection of these individuals. Similarly, we can define the differential equation for y as:

$$\frac{dy}{dt} = -\overbrace{\phi_{01}y + \phi_{10}(1-y)}^{\text{group switching}} - \overbrace{(\beta_{00}(1-y) + \beta_{01}y)(1-x)uy}^{\text{infection of hosts not engaged in prophylaxis}} + \underbrace{(\beta_{10}(1-y) + \beta_{11}y)xu(1-y)}_{\text{infection of hosts engaged in prophylaxis}}. \tag{2d}$$

The first term represents the effects of infective individuals switching between engaging and not engaging in prophylaxis. The latter two terms correspond to infection in the same vein as in Eq. (2c).

Eqs. (2c) and (2d) capture dynamic features of the proportion of hosts engaged in prophylaxis not found in previous work. The switching terms in the modelling undertaken by Pharaon and Bauch (2018) account for utility costs and benefits of prophylaxis, but neglect changes due to demographics (births) and infection. Our equations, therefore, combine multiple processes to establish a more realistic model of changing host behaviour.

We now return to the issue of group switching and the details surrounding the switching terms τ_{ij} and ϕ_{ij}, ultimately replacing these constants with functions of x, y, and v. Individuals do not always adhere to beneficial measures such as taking medications (Osterberg and Blaschke, 2005), exercise regimes (Robison and Rogers, 1994), or dietary restrictions (Patton, 2011), so we include group switching in our model to account for these types of effects. Following Pharaon and Bauch (2018), we use the replicator dynamic (Taylor and Jonker, 1978) to model the total rate at which individuals move from one group to another. We assume that the decision to switch groups is driven by utility, as discussed in Matessi and Di Pasquale (1996). Specifically, an individual's utility will be determined by their risk of infection less the utility cost of taking prophylactic measures. For an i-susceptible, the risk of infection it faces will be quantified by the force of infection,

$$\sum_j -\frac{\beta_{ij}I_j}{N} = -(\beta_{i0}(1-y) + \beta_{i1}y)v.$$

Consequently, the utility of a susceptible individual adopting prophylactic measures is

$$-(\beta_{10}(1-y) + \beta_{11}y)v - \chi$$

where χ is the utility cost to engaging in prophylaxis. Infective individuals have no risk of infection, so they only pay the utility cost of the prophylactic measure should they choose to engage in it. Thus, $-\chi$ represents the utility of an infective individual engaging in prophylaxis. A susceptible (resp. infective) individual will then choose to switch into a given group when the utility of an individual in that group exceeds the utility of the average susceptible (resp. infective) in the population. This replicator-dynamics model of switching gives us

$$-\tau_{01}x + \tau_{10}(1-x) = kx(1-x)[y(\beta_{01} - \beta_{11}) + (1-y)(\beta_{00} - \beta_{10})]v - k\chi x \tag{3}$$

and

$$-\phi_{01}y + \phi_{10}(1-y) = -k\chi y \tag{4}$$

where k is a constant that reflects the rate at which individuals change their behaviour based on the behaviour of others in the population (see Appendix A for a more detailed derivation of equations 3 and 4). We note that our model of switching makes the simplification (made in Pharaon and Bauch (2018)) of assuming individuals have up-to-date information about the global state of the population. Our final model of disease dynamics can now be stated as,

$$\frac{du}{dt} = 1 - [y(\beta_{01}(1-x) + \beta_{11}x) + (1-y)(\beta_{00}(1-x) + \beta_{10}x)]uv - u \tag{5a}$$

$$\frac{dv}{dt} = [y(\beta_{01}(1-x) + \beta_{11}x) + (1-y)(\beta_{00}(1-x) + \beta_{10}x)]uv - (1+\gamma)v \tag{5b}$$

$$\frac{dx}{dt} = \frac{-x}{u} + (k+1)x(1-x)[y(\beta_{01} - \beta_{11}) + (1-y)(\beta_{00} - \beta_{10})]v - k\chi x \tag{5c}$$

$$\frac{dy}{dt} = -[y(\beta_{01}(1-x)y - \beta_{11}x(1-y)) + (1-y)(\beta_{00}(1-x)y - \beta_{10}x(1-y))]u - k\chi y. \tag{5d}$$

A sample trajectory of this system is shown in figure 1, panel A.

Note that, if no individual in the population takes prophylactic measures, then $x = y = 0$ and we recover the standard endemic SIR model (Appendix B). When some fraction of the population takes prophylactic measures, however, standard predictions of the SIR model may or may not hold. The linear stability analysis presented in Appendix B shows that the endemic equilibrium without individuals engaging in prophylaxis remains stable as long as

$$\chi > \mathcal{R}_0 - \left(\frac{k+1}{k}\right)\left(\frac{\beta_{10}}{\beta_{00}}(\mathcal{R}_0 - 1) + 1\right) \triangleq \chi_c, \tag{6}$$

where $\mathcal{R}_0 = \beta_{00}/(1+\gamma)$ is the basic reproductive number (Heffernan et al., 2005) of the system in the absence of prophylactic measures. Below this threshold, the cost to taking prophylactic measures is low enough that the system moves towards a new endemic equilibrium $(\hat{u}, \hat{v}, \hat{x}, \hat{y})$ at which some non-zero fractions of susceptible and infective individuals adopt prophylactic measures (figure 1, panel B). It is this new endemic equilibrium that will frame the evolutionary model we pursue in the next section.

2.2. Evolutionary dynamics

We consider the evolution of the level of pathogen exploitation of its host, denoted $\xi > 0$. Exploitation affects disease transmission, with a greater ξ value corresponding to a greater β_{ij}. To reflect this, we now write $\beta_{ij}(\xi)$ where

$$\beta_{ij}(\xi) = \frac{\beta_{\max}(1-\epsilon)^{i+j}\xi}{\kappa + \xi}. \tag{7}$$

In words, we are treating β_{ij} as an increasing function of ξ that saturates at a value of $\beta_{\max}(1-\epsilon)^{i+j}$, where ϵ is the probability that the prophylactic measures prevent disease transmission. We assume that prophylactic measures taken by individuals fail independently, so $(1-\varepsilon)^2$ gives the probability that the disease is transmitted between two individuals engaging in prophylaxis. The rate at which transmission saturates is controlled by $\kappa > 0$, with larger values of this constant corresponding to a reduced rate of saturation.

Exploitation also affects recovery. To reflect this assumption, we write $\gamma(\xi)$, where $\gamma(\xi) = c\xi$ for a constant c with units of inverse time (the exploitation level ξ is dimensionless). Without loss of generality, we take $c = 1$. Here, increased exploitation acts to reduce the expected duration $1/(1+\gamma(\xi))$ of an infection. This penalty of larger ξ, then, trades off against the transmissibility benefits described above. Previous authors have either assumed (or shown) that such trade-offs exist, though they are often mediated by disease-related mortality (Anderson and May, 1982; Day, 2002; Ewald, 1983) or viral load (Fraser et al., 2014). Here, we follow Úbeda and Jansen (2016) and Alizon (2008) by assuming the trade-off faced by the pathogen involves recovery. For example,

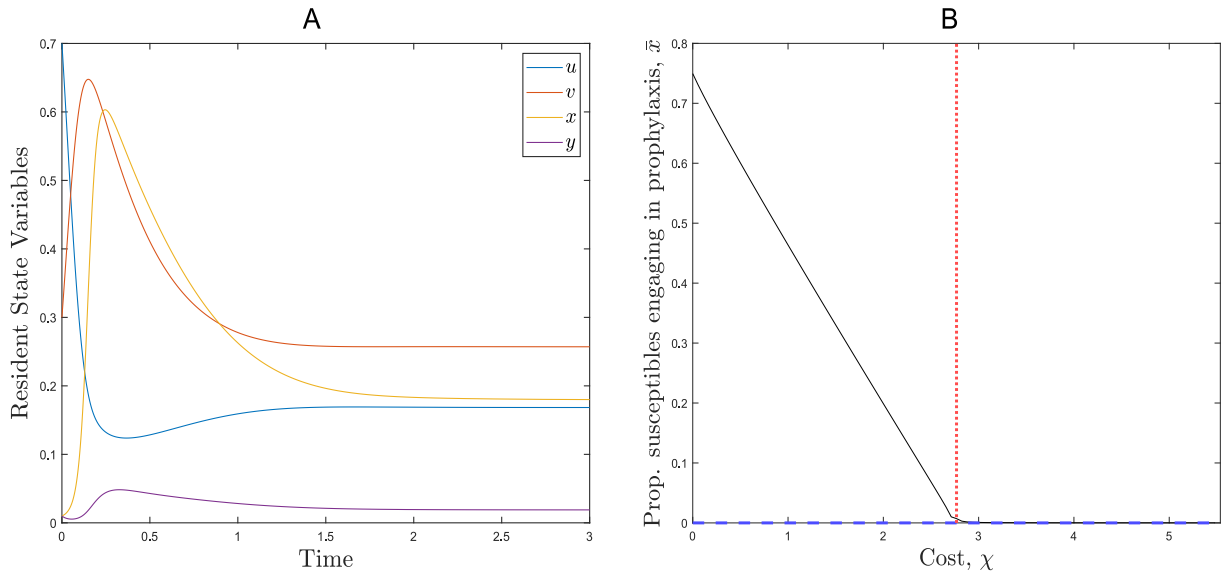

Fig. 1. Panel A shows a sample trajectory for the resident system (5). As described in the main text, time has been rescaled so that one time unit is equivalent to the average lifetime of an individual in the population. Panel B shows a bifurcation plot indicating the cost threshold (dotted red line) at which the endemic equilibrium without prophylaxis (dashed blue line) and the endemic equilibrium with prophylaxis (solid black line) undergo an exchange of stability.

through increased viral load making it more likely that the pathogen is detected by the host's immune system, or increased exploitation leading to more antigens presented on the surfaces of target cells.

We use an adaptive dynamics approach to model the evolution of pathogen exploitation under the primary influence of natural selection (Dercole and Rinaldi (2008); Dieckmann and Law, 1996; Metz et al., 1992; and see Day and Burns, 2003; Úbeda and Jansen, 2016; Hurford et al., 2010 for examples specifically related to virulence evolution). We introduce a rare mutant pathogen with exploitation trait ξ_m into a resident pathogen population with exploitation trait ξ. It is assumed that the resident system has reached equilibrium prior to introducing the mutant, that there is no co-infection, and that the prophylactic measures are equally effective at preventing transmission of both strains.

Let v_m denote the fraction of individuals in the population infected with the mutant strain. While the mutant is rare, its dynamics are well approximated by

$$\frac{dv_m}{dt} = [y_m(\beta_{01}(\xi_m)(1-\bar{x}) + \beta_{11}(\xi_m)\bar{x}) + (1-y_m)(\beta_{00}(\xi_m)(1-\bar{x}) + \beta_{10}(\xi_m)\bar{x})]\bar{u}\,v_m - (1+\gamma(\xi_m))v_m, \tag{8a}$$

where overbars denote equilibrium values of the respective variables. As we do with the resident population, we can also track the proportion of individuals infected with the mutant strain engaging in prophylaxis. Denoting this proportion by y_m, we can describe its dynamics by

$$\frac{dy_m}{dt} = -[y_m(\beta_{01}(\xi_m)(1-\bar{x})y_m - \beta_{11}(\xi_m)\bar{x}(1-y_m)) + (1-y_m)(\beta_{00}(\xi_m)(1-\bar{x})y_m - \beta_{10}(\xi_m)\bar{x}(1-y_m))]\bar{u} - k\chi y_m. \tag{8b}$$

If the mutant strain becomes common and the mutant-free equilibrium becomes unstable, we say that the mutant has successfully invaded the resident population. Provided the system is sufficiently close to an evolutionarily steady state, a mutant who successfully invades will become the new resident (Dercole and Rinaldi, 2008). Since v_m does not appear in (8b), we can first solve for the equilibrium value $\bar{y}_m$ of y_m, substitute that value into (8a), and study (8a) alone. If the right-hand side of (8a) is positive (resp. negative), the mutant invades (resp. is eliminated) because it is favoured (resp. disfavoured) by natural selection. The sign of the right-hand side of (8a) is the same as the sign of the difference between

$$W(\xi_m, \xi) = \frac{\beta_{\text{avg}}(\xi_m, \xi)\,\bar{u}(\xi)}{1+\gamma(\xi_m)} \tag{9}$$

and unity, where

$$\beta_{\text{avg}}(\xi_m, \xi) = [\,1-\bar{x}(\xi) \quad \bar{x}(\xi)\,]\begin{bmatrix}\beta_{00}(\xi_m) & \beta_{01}(\xi_m)\\ \beta_{10}(\xi_m) & \beta_{11}(\xi_m)\end{bmatrix}\begin{bmatrix}1-\bar{y}_m(\xi_m)\\ \bar{y}_m(\xi_m)\end{bmatrix}$$

represents an average transmission rate taking into account the different groups of susceptible and infective individuals. W then has a clear biological interpretation, made in previous work (Day and Burns, 2003), in terms of the basic reproductive number of the mutant strain. In particular, an infection with the mutant strain lasts an average of $1/(1+\gamma)$ time units and an average of $\beta_{\text{avg}}\bar{u}$new infections are created during this time. If this quantity is larger than one (resp. smaller than one), then the mutant population will grow (resp. shrink). Writing W as we have done in Eq. (9) also highlights the transmission-recovery trade-off described above, captured through the β_{avg} and γ terms. Through the $\bar{y}_m$ terms, the pathogen is also able to influence whether a new infection occurs in an individual engaging or not engaging in prophylaxis. We can, therefore, use W as an invasion fitness function in the adaptive dynamics analysis, even though the derivation of the function proceeded in a non-standard way.

Following the discussion above, the direction of evolution of ξ that is favoured by natural selection is given by the sign of $\partial W/\partial\xi_m|_{\xi_m=\xi}$. Consequently, the selective process is at equilibrium whenever $\xi = \bar{\xi}$ where $\bar{\xi}$ satisfies

$$\frac{\partial W}{\partial \xi_m}\Big|_{\xi_m=\xi=\bar{\xi}} = 0. \tag{10}$$

An equilibrium value $\bar{\xi}$, i.e., one that satisfies condition (10), may or may not be stable. If the equilibrium value $\bar{\xi}$ attracts nearby resident populations, then

$$\frac{d}{d\xi}\left[\frac{\partial W}{\partial \xi_m}\Big|_{\xi_m=\xi}\right]_{\xi=\bar{\xi}} \leqslant 0 \tag{11}$$

and we say that $\bar{\xi}$ is convergence stable (Christiansen, 1991). If the equilibrium $\bar{\xi}$ resists invasion from nearby mutants, then

$$\frac{\partial^2 W}{\partial \xi_m^2}\Big|_{\xi_m=\xi=\bar{\xi}} \leqslant 0 \tag{12}$$

and we say that $\bar{\xi}$ is evolutionarily stable (sensu Maynard Smith (1982)). A sufficient condition for either type of stability is obtained by replacing the weak inequality with a strict one. It is the strict sufficient versions that we use here. When $\bar{\xi}$ is both convergence stable and evolutionarily stable, we say it is a continuously stable strategy (CSS) (Eshel, 1983).

3. Results

3.1. Simple cases

There are two special cases that can be analyzed with relative ease. The first special case assumes the cost of prophylaxis exceeds the threshold χ_c. In this case, no one in a population supporting the resident endemic disease is adopting prophylactic measures and so $\bar{x} = \bar{y} = 0$. The equilibrium value of y_m can be shown to be $\bar{y}_m = 0$ (see Appendix C) and so the fitness function simplifies to

$$W(\xi_m, \xi) = \frac{\beta_{00}(\xi_m)\bar{u}(\xi)}{1+\gamma(\xi_m)} = \frac{\mathcal{R}_0(\xi_m)}{\mathcal{R}_0(\xi)}. \tag{13}$$

The mutant strain is then able to invade (resp. is eliminated) if $W(\xi_m, \xi)$ exceeds (resp. is less than) unity; equivalently, if $\mathcal{R}_0(\xi_m)$ exceeds or is less than $\mathcal{R}_0(\xi)$. More importantly, the CSS level of exploitation, $\bar{\xi}$, will maximize $\mathcal{R}_0(\xi)$ and so $\bar{\xi} = \sqrt{\kappa}$. This will serve as a benchmark against which more general results will be compared.

The second special case assumes that prophylactic measures are cost-free, i.e., $\chi = 0$. Since $\chi < \chi_c$, the system moves towards an endemic equilibrium where some non-zero proportions of susceptible and infective individuals are engaging in prophylaxis. While the replicator dynamics predict that all individuals will begin adopting prophylactic measures in the absence of cost, the terms noted in equations (2c) and (2d) relating to demographics and disease dynamics counteract this effect. This will result in the evolution of intermediate levels of the proportions of susceptible and infective individuals engaging in prophylaxis. Under the assumption of zero cost, the distribution of the mutant strain, captured by $\bar{y}_m$, does not depend on the mutant exploitation level and so the pathogen is not able to influence whether new infections occur in individuals engaging or not engaging in prophylaxis. In this case, the selection gradient simplifies and results in a CSS of $\bar{\xi} = \sqrt{\kappa}$ (see Appendix C). This is the same result as the benchmark established in our first special case, despite the presence of individuals taking prophylactic measures. Absence of cost, it seems, decouples the pathogen's evolution from the evolution of host behaviour due to the pathogen no longer being able to influence the relative proportions of infections in individuals engaged or not engaged in prophylaxis.

3.2. Evolution near the critical cost

In general, the model cannot be explored analytically. However, there are certain analytical results we can derive for cost values other than those discussed in the previous section. In particular, we can show that for any cost value, there is a unique stable equilibrium value of y_m on the interval $[0, 1]$. Moreover, we can show that the derivative of this equilibrium value with respect to the mutant exploitation ξ_m is always positive, and that this complicates the relationship between pathogen exploitation and transmissibility (see Appendix D for a derivation of these results). This leads to a change in the evolutionarily stable level of pathogen exploitation away from that which would be expected in the absence of prophylaxis.

If we are near the critical cost χ_c we can derive quasi-analytic results to predict the direction of this change in the CSS value of ξ. When the cost χ is slightly below its critical threshold χ_c, we can approximate the CSS exploitation level as $\bar{\xi} \approx \sqrt{\kappa} + \sigma(\chi - \chi_c)$ where $\chi - \chi_c < 0$ and σ is a constant such that

$$\sigma \propto -\frac{\partial}{\partial \bar{x}}\left[\frac{\partial W}{\partial \xi_m}\Big|_{\xi_m=\xi}\right]_{\substack{\chi=\chi_c \\ \xi=\sqrt{\kappa}}}. \tag{14}$$

If σ is positive (resp. negative), then $\bar{\xi}$ is below (resp. above) the benchmark value of $\sqrt{\kappa}$. As Eq. (14) shows, whether we are above or below this benchmark depends on how small changes in cost lead to small changes in the proportion of susceptible individuals engaged in prophylaxis which, in turn, lead to changes in the selection gradient acting on exploitation (see Appendix E for a derivation of Eq. (14)).

We can show with a quasi-analytic approach that Eq. (14) is always positive. Our evidence relies on first choosing feasible values of our parameters. In particular, we need $\mathcal{R}_0 > 1$. If $\mathcal{R}_0 > 1$, then we need also to choose the probability $\epsilon > \frac{1}{k+1}\frac{\mathcal{R}_0}{\mathcal{R}_0-1}$, thus ensuring that $\chi_c > 0$. To ensure that $\epsilon < 1$, we then need to choose $k > \frac{1}{\mathcal{R}_0-1}$.

Using feasible parameters and working to zeroth order in $\chi - \chi_c$, we use the computer algebra software (CAS) Maple (version 2019.1) to investigate the sign of σ as described in Eq. (14). We find that the requirement that $\mathcal{R}_0 > 1$ necessarily restricts our choices of maximal transmissibility, $\beta_{\max}$, to values that lie above the curve traced out by $\left(1+\sqrt{\kappa}\right)^2$ (Fig. 2, panel A). The CAS shows that nullclines of the partial derivative in (14) never exceed the $\left(1+\sqrt{\kappa}\right)^2$ curve for the wide range of feasible parameters we investigated (Fig. 2, panel B). Thus, the sign of σ does not change provided the $\mathcal{R}_0 > 1$ restriction is met. Moreover, test points show that the sign of σ itself is positive when feasible model parameters are chosen. Based on CAS investigations described in Appendix F, then, we conclude that, just below the critical cost, selection acts to reduce the CSS level of host exploitation exhibited by the pathogen.

3.3. Evolution for arbitrary cost

Our results can be extended numerically for costs that are possibly much smaller than the critical value, χ_c, using a Matlab (version R2019a) procedure detailed in Appendix G. We build the procedure around the observation that locally asymptotically stable equilibrium solutions to $d\xi/dt = (\partial W/\partial \xi_m)|_{\xi_m=\xi}$ are also convergence-stable evolutionary equilibria as defined by conditions (10) and (11), respectively. As a result, numerical iteration of this differential equation can be used to find candidate CSS strategies. The evolutionary stability of candidate CSS strategies can be confirmed with a centred finite-difference approximation of (12). Since the error is on the order of the square of the distance between ξ values used in the approximation, we consider any value within this error to satisfy the ESS condition (12).

The results of our numerical procedure confirm that the benchmark CSS level of $\bar{\xi} = \sqrt{\kappa}$ is obtained when $\chi = 0$ and $\chi = \chi_c$. Second, numerical results confirm the reduction in the CSS level of pathogen exploitation for χ slightly smaller than χ_c. Third, and most important, numerical results indicate that the CSS exploitation level $\bar{\xi}$ changes in a simple way as cost is reduced from its

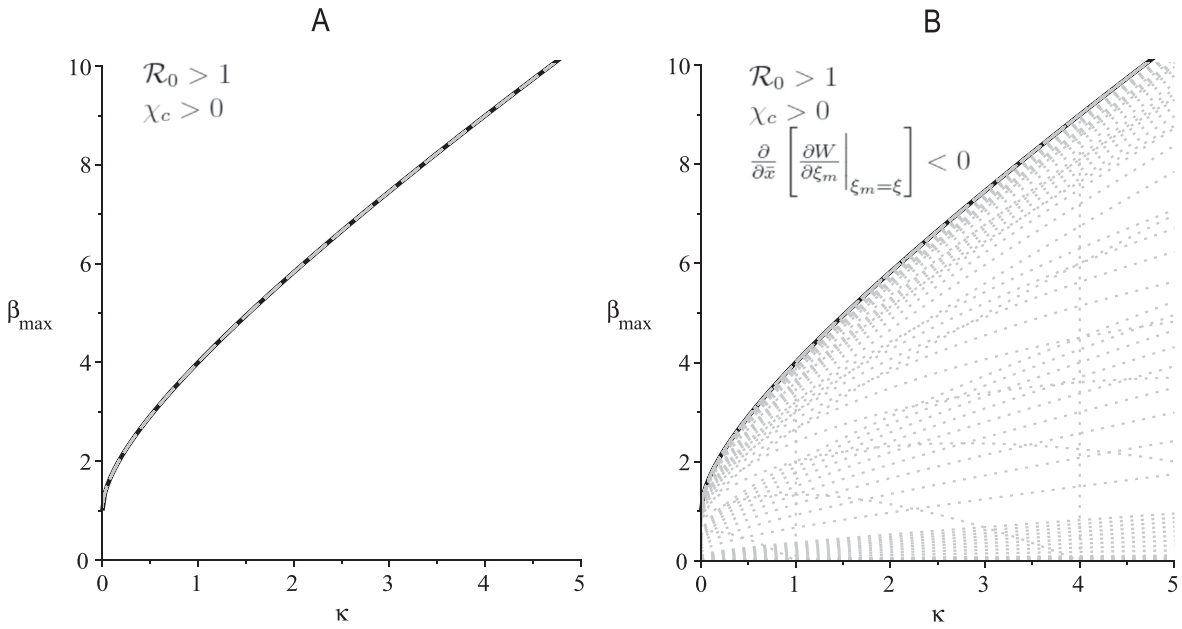

Fig. 2. Panel A shows the curve $\chi_c = 0$ (dashed grey) overlaying the curve $\mathcal{R}_0 = 1$ (solid black), and the region of parameter space where $\chi_c > 0$ and $\mathcal{R}_0 > 1$. Dotted grey curves represent the roots of the partial derivative on the right-hand side of Eq. (14) and panel B shows that these all occur on or below the black and dashed grey curves (note that vertical dotted grey lines are an artifact of jump discontinuities). Choosing parameter values in the region above the curves in panel A results in the partial derivative on the right-hand side of equation (14) being negative, as noted in panel B. The implication is that the parameter σ is positive and so pathogen exploitation will decrease relative to the benchmark level of $\sqrt{\kappa}$ close to the cost threshold χ_c.

critical value to its natural lower limit at zero (Fig. 3). In particular, as cost is reduced $\bar{\xi}$ decreases monotonically from the benchmark value until it reaches a minimum. Once at the minimum, the direction of selection changes and $\bar{\xi}$ increases monotonically, ultimately returning to the benchmark when cost disappears.

This pattern holds for a wide range of parameter values chosen to satisfy the conditions described in the previous section. Specifically, we investigate four different values of $\kappa : \kappa = 0.1,\ \kappa = 1, \kappa = 10$, and $\kappa = 100$. For each of these values, we choose five values of β_{max} above the $\left(1+\sqrt{\kappa}\right)^2$ threshold: $\beta_{max} = \left(1+\sqrt{\kappa}\right)^2 + 5$, $\beta_{max} = \left(1+\sqrt{\kappa}\right)^2 + 10$, $\beta_{max} = \left(1+\sqrt{\kappa}\right)^2 + 50$, $\beta_{max} = \left(1+\sqrt{\kappa}\right)^2 + 100$, and $\beta_{max} = \left(1+\sqrt{\kappa}\right)^2 + 500$. We then choose five values of k ($k = 1/(\mathcal{R}_0 - 1) + 5, k = 1/(\mathcal{R}_0 - 1) + 10, k = 1/(\mathcal{R}_0 - 1) + 50$, $k = 1/(\mathcal{R}_0 - 1) + 100$, and $k = 1/(\mathcal{R}_0 - 1) + 500$) and five values of ϵ spread evenly between the threshold $\mathcal{R}_0/((k+1)(\mathcal{R}_0 - 1))$ and 1, for a total of 500 combinations of parameter values.

Although the decline in the CSS value of exploitation shown in Fig. 3 appears modest, recall that one time unit is equivalent to the average lifetime of an individual in the population. This means that the change in the duration of infection as $\bar{\xi}$ changes is on the order of years. For example, if the average lifespan of an individual in the population is 79 years, then a decrease from $\bar{\xi} = 1$ to $\bar{\xi} = 0.86$ (as seen in Fig. 3, panel B) corresponds to an increase in the duration of infection of approximately three years.

4. Discussion

We study the impact of measures taken by hosts to limit disease transmission. Here, the willingness among hosts to engage in these prophylactic behaviours responds to changing utility costs and benefits. We focus on long-term evolution of a pathogen, defined by successive mutations until an equilibrium state is reached (Matessi and Di Pasquale, 1996), alongside the rapid evolution of host behaviour. We find that when prophylactic behaviour among hosts occurs, pathogen host exploitation is always lower than it is in the absence of prophylaxis. Moreover, we find that stable exploitation is lowest for an intermediate frequency of prophylactic behaviour among hosts (indirectly, intermediate cost of prophylaxis).

This study contributes to the growing body of work that shows host behaviour, in general, influences pathogen evolution. Much of this work has considered vaccination behaviour, in particular, and has described both beneficial and detrimental evolutionary outcomes. In the case of human papillomavirus (HPV), for example, theoretical work predicted HPV vaccination will select for higher levels of virulence (Murall et al., 2015). By contrast, empirical evidence suggests that vaccination can actually limit the ecological opportunity open to certain HPV types (Poolman et al., 2008). In keeping with the mixed nature of results, Gandon et al. (2001); Gandon et al., 2003 find that the direction of selection acting on pathogen virulence depends on the mechanism by which vaccination works.

More closely related to the current study are the conclusions of Pharaon and Bauch (2018). They show that host prophylactic behaviour in response to an endemic disease can allow for the invasion of a pathogen strain that is more virulent than the resident, and that the conditions for such a result are an increased perceived severity for the more virulent strain and more effective prophylactic measures against the less virulent strain. While our model considers sublethal disease effects and does not explicitly include virulence, the positive relationship between pathogen exploitation and virulence allows us to predict that such an outcome cannot occur in our model. This is a result of the fact that our benchmark result, established in the absence of prophylactic behaviour, is our worst case scenario. For example, consider a mutant pathogen with an exploitation level $\xi > \sqrt{\kappa}$. When the cost is above its critical value so that no one is engaging in prophylaxis, then this mutant cannot invade a resident population at the CSS $\bar{\xi} = \sqrt{\kappa}$. If we then decrease the cost below its critical value so that individuals begin to take prophylactic measures, the CSS exploitation level decreases away from $\bar{\xi} = \sqrt{\kappa}$ and so the mutant is still unable to invade the resident population.

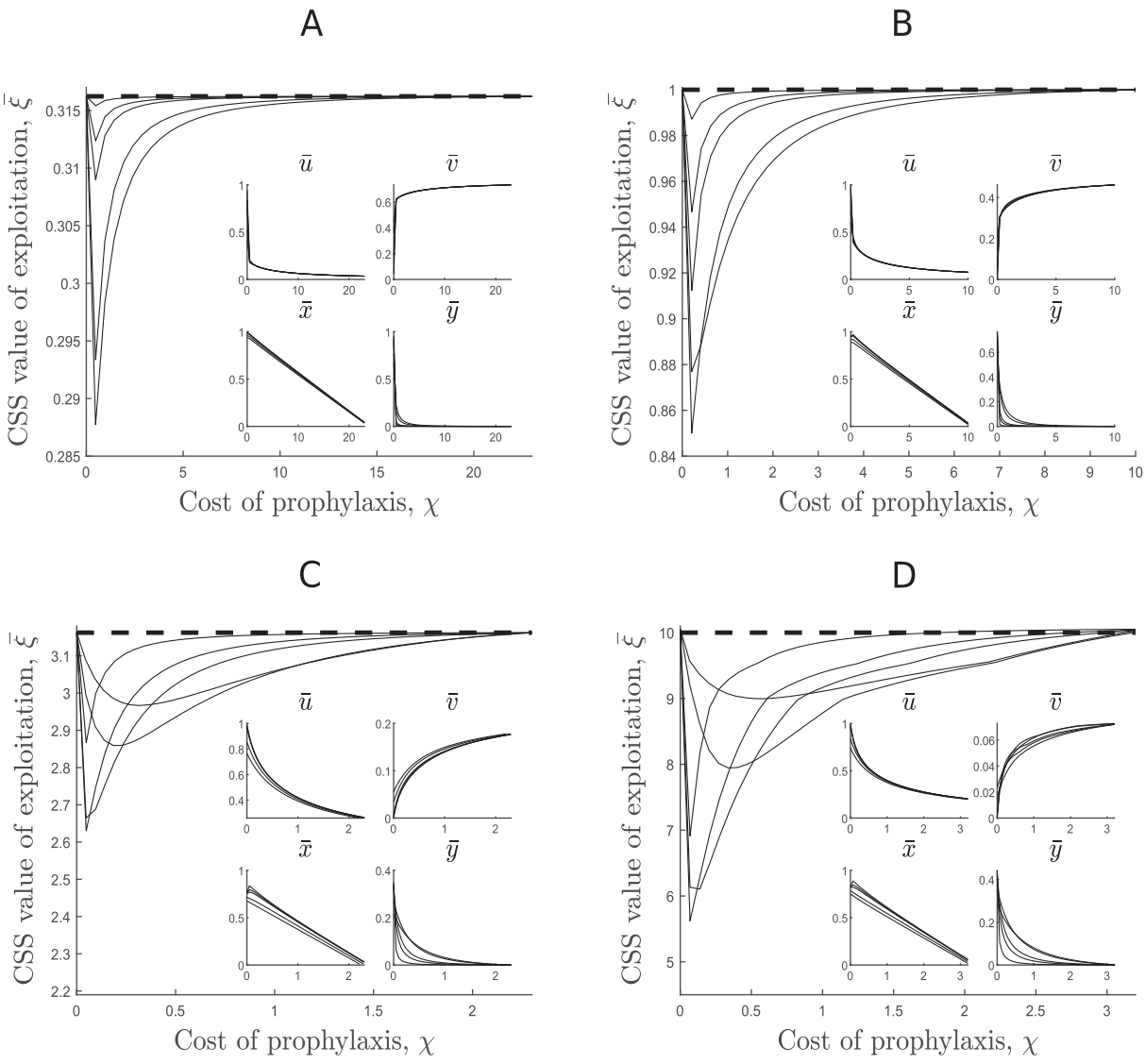

Fig. 3. Plots of the CSS level of pathogen exploitation found using the numerical procedure described in the main text. Panel A contains sample results for $\kappa = 0.1$, panel B for $\kappa = 1$, panel C for $\kappa = 10$, and panel D for $\kappa = 100$. Inset figures show the equilibrium values of the epidemiological variables u, v, x, and y. The cost values presented are all below the critical cost threshold χ_c. In all cases, the CSS level of exploitation is lower than the benchmark value $\bar{\xi} = \sqrt{\kappa}$, plotted as a dashed black line, observed in the absence of prophylaxis.

The discrepancy between our predictions and those of Pharaon and Bauch (2018) is due to the differences in how we model the dynamics of host behaviour. In addition to looking at actual risk instead of perceived risk (as in Pharaon and Bauch (2018)), the evolution of host behaviour in our work is not governed solely by the replicator dynamics. Our model consists of additional terms to represent the effects of births and infection on the proportion of individuals engaging in prophylactic behaviour. Moreover, we track the proportion of infective individuals taking prophylactic measures, which adds an extra layer of complexity to the relationship between pathogen exploitation and transmissibility (Appendix D). These key differences are missing from previous work (Pharaon and Bauch, 2018) and lead us to the conclusion that, for all feasible sets of parameter values, we expect a decrease in pathogen exploitation away from the benchmark value found in the absence of prophylactic behaviour. To obtain a more direct comparison to previous work, future iterations of our model should explicitly include virulence and explore the subsequent predictions on pathogen evolution.

To get some intuition into the decrease in exploitation that we find in this paper, we need to understand two things. In the absence of prophylaxis, standard theory (e.g., Day and Burns (2003); Úbeda and Jansen, 2016; Alizon, 2008) predicts that increased pathogen exploitation results in a decrease in the duration of infection and an increase in transmission, leading to a balance between these two competing effects at evolutionary equilibrium. In our model, increasing exploitation still reduces the duration of infection, but it affects transmission in a more complicated way. This more complicated effect on tranmissibility disrupts the balance that would be achieved in standard models. As we show in Appendix D, there is a benefit to increased exploitation through a direct increase in transmission, but also a marginal cost through increased exposure to the host behavioural immune system. In fact, the marginal cost is indirect as it results in fewer mutant infections in hosts not engaging in prophylaxis (i.e., smaller $1 - \bar{y}_m$). Ultimately, this added cost tips the scales in favour of lower pathogen exploitation.

We have already pointed to specific differences between our work and similar work of others (Pharaon and Bauch, 2018). While those differences will undoubtedly affect the model predictions, the intuition developed above suggests a more concrete explanation for the contrast is possible. In particular, Pharaon and Bauch (2018) do not expose mutant infectives to host prophylactic behaviour in a way that differs from resident infectives. Granted, Pharaon and Bauch (2018) do allow for the efficacy of prophylaxis to differ between resident and mutant strains, but the host landscape looks the same from both resident and mutant perspectives in that model. In our model, mutant infections in individuals engaging in prophylaxis happen in different proportions than resident infections (i.e., $\bar{y}_m \neq \bar{y}$ in general). Simply put, the host landscape in our model differs meaningfully between resident and mutant infectives.

Arguably, our main result is reminiscent of other pathogens that control host behaviour for their own gain. While our model pathogens are not directly controlling hosts like the pathogen *Ophiocordyceps unilateralis* does with the ant *Camponotus leonardi* (Hughes et al., 2011) (or *Schistocephalus solidus* with the stickleback fish *Gasterosteus aculeatus* (Øverli et al., 2001), or *Leucochloridium paradoxum* with the snail *Succinea putris* (Wesołowska and Wesołowska, 2014)), one might speculate that ours indirectly manipulate the hosts' economic agency. This suggests that there may be some cryptic parasite manipulation to further investigate.

As with any modelling endeavour, we have made some simplifying assumptions to reduce the mathematical complexity of our model. For example, we have assumed that individuals instantly update their behaviour when receiving new information about the progression of the disease. In reality, there is a time delay between receiving information and deciding to modify behaviour. Previous work has studied the effects of including a delay in the form of waning immunity either independently (Hethcote et al., 1989) or together with a delay in the form of a latent period following infection (Cooke and van den Driessche, 1996), and found that this can create periodicity in the model solutions. Other work has also investigated the effects of these delays on effective vaccination strategies (Gao et al., 2006). Future work could extend our model to include the lag in information gain and behaviour modification, with the expectation that this would cause oscillations in the predicted exploitation level.

One could also relax the assumption that individuals sample from all other individuals in the population when deciding whether or not to take prophylactic measures. Epidemic models have previously been extended to include spatial structure through the use of networks, with different types of networks providing qualitatively different predictions (Pastor-Satorras and Vespignani, 2001). Others have studied the interaction between network models and host heterogeneities in the case of STIs (Newman, 2002), and the effects of adaptive networks where individuals may build and sever connections during the progression of a disease (Gross et al., 2006). Networks could be incorporated here to explore the decoupling of social interactions related to disease transmission and those related to information transmission. Based on previous work (Wild et al., 2009), it is expected that this decoupling could lead to a lower level of pathogen exploitation.

It is tempting to use our evolutionary predictions to inform public–health policy. While pathogen exploitation does reach a minimum value for an intermediate level of cost of prophylaxis, our results (Fig. 3) also show that lowering the cost even farther below this level leads to fewer infections even if those infections are from a more exploitative pathogen strain. Moreover, the pathogen's exploitation is always below the level predicted in the absence of individuals engaging in prophylactic behaviour. There is a balance, then, between the level of cost that is optimal for minimizing prevalence and the level optimal for minimizing pathogen exploitation. It is important to recognize that efforts to minimize the cost of prophylaxis will result in a pathogen strain that is more exploitative of its host than it otherwise might be. More broadly, our work suggests that conversations about disease management and infection should be more inclusive towards the effects of human behaviour.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Evan Mitchell: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing. **Geoff Wild:** Writing - review & editing, Supervision, Funding acquisition.

Acknowledgements

This work was supported by an NSERC–CRSNG Discovery Grant (RGPIN: 2019-06626). We also thank anonymous reviewers for their helpful and insightful comments, which helped to strengthen the manuscript.

Appendix A. Details of the replicator dynamics

Here, we present the details of the derivation of equations (3) and (4). The replicator dynamic tells us that an individual's decision to start or stop engaging in prophylaxis is proportional to the utility difference between the focal susceptible (resp. infective) individual and the average susceptible (resp. infective) individual in the population, taking into account the utility cost that the focal individual must pay to engage in the prophylactic behaviour. Specifically, the replicator dynamic equation allows us to replace the τ_{ij} terms in our model with the following:

$$-\tau_{01}x + \tau_{10}(1-x) = kx\big((U_{\text{foc}} - \chi) - U_{\text{avg}}\big), \tag{15}$$

where U_{foc} represents the utility of the focal susceptible individual, U_{avg} represents the utility of the average susceptible individual in the population, and k is a constant that reflects the rate at which individuals change their behaviour based on the behaviour of others. As described in the main text, we measure utility by the force of infection so that the benefit gained by an individual choosing to engage in prophylaxis is a reduced force of infection relative to the average individual. The utility of a focal susceptible individual engaging in prophylaxis, then, is given by

$$U_{\text{foc}} = -(\beta_{10}(1-y) + \beta_{11}y)v, \tag{16}$$

while the utility of the average susceptible individual in the population is given by

$$U_{\text{avg}} = -[x(\beta_{10}(1-y) + \beta_{11}y) + (1-x)(\beta_{00}(1-y) + \beta_{01}y)]v. \tag{17}$$

Substituting these into the replicator dynamic equation gives

$$-\tau_{01}x + \tau_{10}(1-x) = kx(1-x)[y(\beta_{01} - \beta_{11}) + (1-y)(\beta_{00} - \beta_{10})]v - k\chi x, \tag{18}$$

which is equation (3) in the main text.

We can derive equation (4) in a similar way. Since there is no force of infection acting on an infective individual, the utility of both a focal infective individual engaging in prophylaxis and an average infective individual in the population is zero. However, an infective individual who decides to engage in prophylaxis must still pay the cost of that behaviour. In this case, then, the replicator dynamic equation gives

$$-\phi_{01}y + \phi_{10}(1-y) = -k\chi y. \tag{19}$$

Appendix B. Linear stability analysis

The full resident system of our model (5) in the main text is built on the standard two-dimensional SIR model. This standard two-dimensional model has a Jacobian matrix of

$$J = \begin{bmatrix} -\beta_{00}v - 1 & -\beta_{00}u \\ \beta_{00}v & \beta_{00}u - 1 - \gamma \end{bmatrix} \tag{20}$$

which, when evaluated at the disease-free equilibrium (DFE) $(\bar{u}, \bar{v}) = (1, 0)$, has eigenvalues $\lambda_1 = -1$ and $\lambda_2 = \beta_{00} - 1 - \gamma$. The DFE is stable when both eigenvalues are negative, which leads us to the condition that $\mathcal{R}_0 = \beta_{00}/(1+\gamma) < 1$ for stability. When $\mathcal{R}_0$ exceeds this threshold, the DFE becomes unstable and the standard two-dimensional system moves towards an endemic equilibrium $(\bar{u}, \bar{v}) = (1/\mathcal{R}_0, (1 - 1/\mathcal{R}_0)/(1+\gamma))$. To determine the region in which this equilibrium remains stable after we incorporate the host behavioural dynamics, we need to consider the Jacobian matrix of our four-dimensional system (5) evaluated at $(\bar{u}, \bar{v}, \bar{x}, \bar{y}) = (1/\mathcal{R}_0, (1 - 1/\mathcal{R}_0)/(1+\gamma), 0, 0)$:

$$\begin{bmatrix} -\beta_{00}\bar{v} - 1 & -1 - \gamma & * & * \\ \beta_{00}\bar{v} & 0 & * & * \\ 0 & 0 & k(\chi_c - \chi) & 0 \\ 0 & 0 & \frac{\beta_{10}(1+\gamma)}{\beta_{00}} & -1 - \gamma - k\chi \end{bmatrix}, \tag{21}$$

where $\chi_c = \mathcal{R}_0 - \left(\frac{k+1}{k}\right)\left(\frac{\beta_{10}}{\beta_{00}}(\mathcal{R}_0 - 1) + 1\right)$ and asterisks denote entries that are possibly non-zero. Since this matrix is block upper triangular, the eigenvalues are given by the eigenvalues of the 2×2 matrices on the diagonal. The 2×2 matrix in the upper left is the Jacobian matrix of the standard two-dimensional SIR model evaluated at the endemic equilibrium $(\bar{u}, \bar{v}, \bar{x}, \bar{y})$, which we know has negative eigenvalues whenever $\mathcal{R}_0 > 1$. The 2×2 matrix in the bottom right is lower triangular, so its eigenvalues are the entries on the main diagonal. The second of these entries is always negative, while the first is negative as long as $\chi > \chi_c$. This defines a critical cost threshold where for $\chi > \chi_c$ the endemic equilibrium $(\bar{u}, \bar{v}, \bar{x}, \bar{y})$ is stable, while for $\chi < \chi_c$ our system tends towards an endemic equilibrium that contains individuals engaging in prophylaxis in some non-zero quantities.

Appendix C. Evolutionary analysis of simple cases

Two simple cases, where evolutionary analysis is possible analytically, were discussed in Section 3.1 of the main text. Here, we outline the details of that analysis.

The first special case is when the cost of prophylaxis exceeds the threshold χ_c. In this case, our linear stability analysis shows that the stable endemic equilibrium is $(\bar{u}, \bar{v}, \bar{x}, \bar{y}) = (1/\mathcal{R}_0, (1 - 1/\mathcal{R}_0)/(1+\gamma), 0, 0)$, where $\mathcal{R}_0 = \beta_{00}/(1+\gamma)$. We can then solve equation (8b) to find two possible equilibria for $y_m : \bar{y}_m = 0$ or $\bar{y}_m = 1/\epsilon + (k\chi/(\epsilon(1+\xi_m))$. Since $\epsilon \leqslant 1$, the second of these values is always larger than one and so is not biologically sensible, as y_m is defined to be a proportion. Checking the sign of $(d/dy_m)(dy_m/dt)$ at the first of these equilibrium values shows that $\bar{y}_m = 0$ is the stable equilibrium. From this, the fitness function in equation (9) reduces to

$$W(\xi_m, \xi) = \frac{\beta_{00}(\xi_m)\bar{u}(\xi)}{1 + \gamma(\xi_m)} = \frac{\mathcal{R}_0(\xi_m)}{\mathcal{R}_0(\xi)}, \tag{22}$$

and so to find the CSS we need to solve the equation

$$\frac{\partial W}{\partial \xi_m}|_{\xi_m = \xi = \bar{\xi}} = \frac{d\mathcal{R}_0(\xi_m)}{d\xi_m}|_{\xi_m = \bar{\xi}} = 0. \tag{23}$$

This results in a CSS value of $\bar{\xi} = \sqrt{\kappa}$.

The second special case is when the prophylactic measures are cost-free, i.e., $\chi = 0$. As noted in the main text, the system moves towards an endemic equilibrium where some non-zero proportions of susceptible and infective individuals are engaging in prophylaxis. While we do not have exact analytic expressions for these equilibrium values, we can still solve for $\bar{y}_m$ analytically. Under the assumption of zero cost, solving equation (8b) gives two possible equilibrium values for $y_m : \bar{y}_m = 1/\epsilon$ and $\bar{y}_m = (\epsilon\bar{x} - \bar{x})/(\epsilon\bar{x} - 1)$. Since $\epsilon \leqslant 1$ and $\bar{x} \leqslant 1$, the first of these is always larger than one and the second is always on the interval $[0, 1]$. Substituting the second of these values into $(d/dy_m)(dy_m/dt)$ shows that $\bar{y}_m = (\epsilon\bar{x} - \bar{x})/(\epsilon\bar{x} - 1)$ is the stable equilibrium. Notably, this equilibrium value is independent of the mutant exploitation ξ_m and so the fitness function has the form

$$W(\xi_m, \xi) = \frac{\xi_m}{(\kappa + \xi_m)(1 + \gamma(\xi_m))} B(\xi)\bar{u}(\xi), \tag{24}$$

where $B(\xi)$ represents what remains from $\beta_{\text{avg}}(\xi_m, \xi)$ after factoring out the terms involving ξ_m. To find the CSS value, we then solve

$$\frac{\partial W}{\partial \xi_m}|_{\xi_m = \xi = \bar{\xi}} = \frac{\kappa - \bar{\xi}^2}{(\kappa + \bar{\xi})^2(1 + \bar{\xi})^2} B(\bar{\xi})\bar{u}(\bar{\xi}) = 0. \tag{25}$$

This results in a CSS value of $\bar{\xi} = \sqrt{\kappa}$.

Appendix D. Evolutionary analysis in the case of arbitrary cost

While a full evolutionary analysis of the model in the case of an arbitrary cost value is possible only numerically, we expand here on some analytical details of this case discussed in the main text. In particular, we show that there is a unique stable equilibrium value of y_m on the interval $[0, 1]$, and that the derivative $d\bar{y}_m/d\xi_m$ evaluated at $\xi_m = \xi$ is always positive and discuss the implications of this on the fitness function.

In order to numerically perform the evolutionary invasion analysis, we first need to know that there is a unique equilibrium value of y_m on the interval of $[0, 1]$ and that this value is stable. To find the equilibrium values of y_m, we need to solve equation (8b). The right-hand side of this equation is a quadratic polynomial in y_m, so we know there are two possible equilibrium values for y_m. Evaluating (8b) at $y_m = 0$ gives

$$\frac{dy_m}{dt} = \frac{\beta_{\max}(1 - \epsilon)\xi_m\bar{x}\bar{u}}{\kappa + \xi_m}, \tag{26}$$

which is positive since $\epsilon \in [0, 1]$. If we then evaluate (8b) at $y_m = 1$, we find that

$$\frac{dy_m}{dt} = -\frac{\bar{u}(1 - \epsilon)(1 - \bar{x})\beta_{\max}\xi_m + k\chi\xi_m + k\chi\kappa}{\kappa + \xi_m}, \tag{27}$$

which is negative since $\epsilon \in [0, 1]$ and $\bar{x} \in [0, 1]$. By the Intermediate Value Theorem and the fact that the right-hand side of (8b) is quadratic in y_m, this then shows that there is a unique equilibrium value of y_m in the interval $[0, 1]$. Moreover, since the derivative dy_m/dt is positive at $y_m = 0$ and negative at $y_m = 1$, this equilibrium value is stable.

With this knowledge, we are able to solve (8b) and get an explicit expression for the equilibrium value $\bar{y}_m$. To understand how this value interacts with the fitness function (9), we need to understand how $\bar{y}_m$ changes with ξ_m. Differentiating with respect to ξ_m, we are able to get an explicit expression for $d\bar{y}_m/d\xi_m$ and evaluate when $\xi_m = \xi$. Doing so results in an expression of the form

$$\frac{d\bar{y}_m}{d\xi_m}|_{\xi_m = \xi} = \frac{a - \sqrt{a^2 - b}}{c}, \tag{28}$$

where $a = k\chi(\kappa + \xi) + \beta_{\max}(1 - \epsilon^2\bar{x})\bar{u}, b = 4\beta_{\max}^2\epsilon(1 - \epsilon)(1 - \epsilon\bar{x})\xi^2\bar{u}^2\bar{x}$, and $c = 2\beta_{\max}\epsilon(1 - \epsilon\bar{x})\xi^2\bar{u}\sqrt{a^2 - b}/k\chi\kappa$. The fact that $\epsilon \in [0, 1]$ and

$\bar{x}\in[0,1]$ allows us to conclude that a,b, and c are all positive. Some simplification of the radicand also shows that $a^2-b\geqslant 0$, so we are guaranteed that equation (28) is real-valued. This proves that $\sqrt{a^2-b}\leqslant a$ and, thus, $(d\bar{y}_m/d\xi_m)|_{\xi_m=\xi}\geqslant 0$.

To see how this affects the evolution of pathogen exploitation, we need to pull apart the fitness function in equation (9). Finding a potential CSS value of ξ involves first solving $(\partial W/\partial\xi_m)|_{\xi_m=\xi=\bar{\xi}}=0$. Using equation (9), this gives

$$0=\frac{\partial W}{\partial\xi_m}|_{\xi_m=\xi=\bar{\xi}}=\frac{\left(\beta'_{\text{avg}}(1+\gamma)-\beta_{\text{avg}}\gamma'\right)\bar{u}}{(1+\gamma)^2}, \tag{29}$$

where primes denote derivatives with respect to ξ_m evaluated when $\xi_m=\xi=\bar{\xi}$. This further reduces to solving the equality

$$\frac{\beta'_{\text{avg}}}{\beta_{\text{avg}}}=\frac{\gamma'}{1+\gamma}=\frac{(1+\gamma)'}{1+\gamma}. \tag{30}$$

In the absence of individuals engaging in prophylaxis, $\beta_{\text{avg}}=\beta_{00}$ and so solving for the CSS value $\bar{\xi}$ amounts to balancing the standard trade-off between transmission and recovery. However, when individuals take prophylactic measures, β_{avg} becomes more complicated. To understand how β_{avg} responds to changes in exploitation in this case, we take a closer look at β'_{avg}. Using the definition of β_{avg} given in the main text, we have that

$$\beta'_{\text{avg}}=\begin{bmatrix}1-\bar{x} & \bar{x}\end{bmatrix}\left(\underbrace{\begin{bmatrix}\beta'_{00} & \beta'_{01}\\ \beta'_{10} & \beta'_{11}\end{bmatrix}\begin{bmatrix}1-\bar{y}_m\\ \bar{y}_m\end{bmatrix}}_{\text{I}}-\underbrace{\bar{y}'_m\begin{bmatrix}\beta_{00}-\beta_{01}\\ \beta_{10}-\beta_{11}\end{bmatrix}}_{\text{II}}\right). \tag{31}$$

It follows from the fact that $\beta_{00}\geqslant\beta_{01}=\beta_{10}\geqslant\beta_{11}$ and the above proof that $\bar{y}'_m\geqslant 0$ that terms I and II in equation (31) are non-negative. Accounting for individuals engaging in prophylaxis—in particular, infective individuals engaging in prophylaxis—reduces β'_{avg} below the level we would expect in the absence of those measures. If the pathogen increases its exploitation, it leads to an increase in the proportion of infective individuals engaging in prophylactic measures due to the fact that $\bar{y}'_m\geqslant 0$. This increases the likelihood that the rate of transmission between a susceptible individual and an infective individual will be one of β_{01} or β_{11}, instead of β_{00} or β_{10}. Since β_{01} and β_{11} are always smaller than β_{00} and β_{10}, this leads to a reduction in the rate of change of the average transmission rate in the population. This influence on the rate of change of transmission intertwines with the more standard trade-off between transmission and recovery, and reduces the evolutionarily stable level of pathogen exploitation below what we would expect in the absence of prophylaxis.

Appendix E. Perturbation analysis

Letting $\mathbf{r}=(u,v,x,y)$ and $s=\xi$, we can express the four equations governing the epidemiological dynamics of our system as the vector-valued function $\mathbf{F}(\mathbf{r},s;\chi)$ and the equation describing the evolutionary dynamics of pathogen exploitation as the scalar-valued function $G(\mathbf{r},s;\chi)$. We know that below the critical cost threshold χ_c the endemic equilibrium $(\hat{\mathbf{r}},\hat{s})=\left(\hat{u},\hat{v},\hat{x},\hat{y},\hat{\xi}\right)$ is stable, while above this threshold our system tends towards the endemic equilibrium $(\bar{\mathbf{r}},\bar{s})=(\bar{u},\bar{v},0,0,\bar{\xi})$ where individuals engaging in prophylaxis are absent. The critical cost level represents a bifurcation point where these two equilibria coincide and undergo an exchange of stability. To study how our system reacts as we decrease the cost away from this threshold, we introduce a perturbation parameter $\delta=\chi-\chi_c$ and take a first-order approximation to our new equilibrium point $(\hat{\mathbf{r}},\hat{s})=(\bar{u}+\rho_1\delta,\bar{v}+\rho_2\delta,\rho_3\delta,\rho_4\delta,\bar{\xi}+\sigma\delta)$. Knowing that this equilibrium point must satisfy $\mathbf{F}(\hat{\mathbf{r}},\hat{s};\chi)=0$ and $G(\hat{\mathbf{r}},\hat{s};\chi)=0$, our goal is to find expressions for the perturbation coefficients $\rho_1,\rho_2,\rho_3,\rho_4$, and σ.

If we treat $(\hat{\mathbf{r}},\hat{s})$ as a function of χ, we can make a first-order Taylor series approximation centred around χ_c:

$$\mathbf{F}(\bar{\mathbf{r}},\bar{s};\chi_c)+\delta\left[D_{\hat{\mathbf{r}}}\mathbf{F}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{\mathbf{r}}+\mathbf{F}_{\hat{s}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{s}+\mathbf{F}_{\chi}(\bar{\mathbf{r}},\bar{s};\chi_c)\right]=0 \tag{32a}$$
$$G(\bar{\mathbf{r}},\bar{s};\chi_c)+\delta\left[G_{\hat{\mathbf{r}}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{\mathbf{r}}+G_{\hat{s}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{s}+G_{\chi}(\bar{\mathbf{r}},\bar{s};\chi_c)\right]=0, \tag{32b}$$

where subscripts denote partial derivatives and $d\hat{\mathbf{r}}=(\rho_1,\rho_2,\rho_3,\rho_4)$ and $d\hat{s}=\sigma$ are the derivatives with respect to χ of $\hat{\mathbf{r}}$ and $\hat{s}$, respectively. We know that $(\bar{\mathbf{r}},\bar{s})$ is an equilibrium point, so the first term in Eqs. (32a) and (32b) evaluates to zero. We also observe that every term in $\mathbf{F}$ and G involving χ is multiplied by at least one of x or y, and so the partial derivatives with respect to χ vanish when we evaluate at $(\bar{\mathbf{r}},\bar{s})$. This simplifies (32) to:

$$D_{\hat{\mathbf{r}}}\mathbf{F}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{\mathbf{r}}+\mathbf{F}_{\hat{s}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{s}=0 \tag{33a}$$
$$G_{\hat{\mathbf{r}}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{\mathbf{r}}+G_{\hat{s}}(\bar{\mathbf{r}},\bar{s};\chi_c)d\hat{s}=0. \tag{33b}$$

We can write (33) more succinctly as $J\begin{bmatrix}d\hat{\mathbf{r}}\\ d\hat{s}\end{bmatrix}=0$ where the matrix J has the following structure:

$$J=\begin{bmatrix}* & * & * & * & *\\ * & * & * & * & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & * & * & 0\\ 0 & 0 & \frac{\partial}{\partial\bar{x}}\left[\frac{\partial W}{\partial\xi_m}|_{\xi_m=\xi}\right]_{\chi=\chi_c,\,\xi=\sqrt{\kappa}} & 0 & \frac{\partial}{\partial\bar{\xi}}\left[\frac{\partial W}{\partial\xi_m}|_{\xi_m=\xi}\right]_{\chi=\chi_c,\,\xi=\sqrt{\kappa}}\end{bmatrix}, \tag{34}$$

with asterisks denoting entries that are possibly non-zero. Since J is a block triangular matrix, the eigenvalues are given by the eigenvalues of the matrices on the main diagonal. The 2×2 matrix in the upper left is the Jacobian matrix arising from the linearization of the standard SIR model around the endemic equilibrium $(\bar{\mathbf{r}},\bar{s})$. Since the lower-right 3×3 block is lower triangular and has a zero entry on its main diagonal, we can see that zero is an eigenvalue of J. This allows us to interpret $\begin{bmatrix}d\hat{\mathbf{r}}\\ d\hat{s}\end{bmatrix}$ as the eigenvector of J associated with the zero eigenvalue, and so shows that there is a non-trivial solution for our perturbation coefficients.

While an analytic expression for this eigenvector can be found, it is unwieldy. Of more interest is the sign of the perturbation coefficient σ, as this tells us in which direction $\bar{\xi}$ moves as we decrease the cost below its critical value. The third row of (34) tells us that $\bar{x}$ is a free variable, and the last row tells us that there is a simple relationship between this free variable and $\bar{\xi}$. In particular, if we consider finding the eigenvector $\begin{bmatrix}d\hat{\mathbf{r}}\\ d\hat{s}\end{bmatrix}$ by solving the expression $J\begin{bmatrix}d\hat{\mathbf{r}}\\ d\hat{s}\end{bmatrix}=0$, then the last row of (34) tells us that

$$\sigma=-\frac{\frac{\partial}{\partial\bar{x}}\left[\frac{\partial W}{\partial\xi_m}|_{\xi_m=\xi}\right]_{\chi=\chi_c,\,\xi=\sqrt{\kappa}}}{\frac{\partial}{\partial\bar{\xi}}\left[\frac{\partial W}{\partial\xi_m}|_{\xi_m=\xi}\right]_{\chi=\chi_c,\,\xi=\sqrt{\kappa}}}\rho_3. \tag{35}$$

We know that the denominator of (35) is always negative since $\bar{\xi}=\sqrt{\kappa}$ is convergence stable (see (11)). Furthermore, the proportion of susceptible individuals engaging in prophylaxis increases as the cost is decreased below its critical value and $\delta=\chi-\chi_c<0$ below this cost threshold, so we must have that $\rho_3<0$. It follows, then, that the sign of σ is controlled only by the numerator of (35) and so we arrive at Eq. (14) in the main text.

Appendix F. Maple code

Here, we present the Maple code, described in Section 3.2 of the main text, used to check the sign of the perturbation coefficient σ for a range of parameter values near the critical cost threshold. Using $\delta = \chi - \chi_c$ as our perturbation parameter, we first define the differential equation for y_m and solve for the equilibrium value:

```
restart:
with(LinearAlgebra): with(VectorCalculus): with(plots):
dymdt := -(y[m]*(beta[max]*(1 - epsilon)*xi[m]*(1 - x)*
y[m]/(kappa + xi[m]) - beta[max]*(1 - epsilon)^2*xi[m]*
x*(1 - y[m])/(kappa + xi[m])) + (1 - y[m])*(beta[max]*
xi[m]*(1 - x)*y[m]/(kappa + xi[m]) - beta[max]*
(1 - epsilon)*xi[m]*x*(1 - y[m])/(kappa + xi[m])))*u
- k*(chi[c] + delta)*y[m]:
```

Solving this equation returns two possible equilibrium values:

```
ym := solve(dymdt, y[m]):
simplify(subs(delta = chi - chi[c], ym[1])):
simplify(subs(delta = chi - chi[c], ym[2])):
```

These equilibrium values are of the form $\left(a \pm \sqrt{b}\right)/c$. Knowing that $0 \leqslant \epsilon \leqslant 1$ and $0 \leqslant \bar{x} \leqslant 1$ allows us to conclude that $a \leqslant 0, b \geqslant 0$, and $c \leqslant 0$. Furthermore, some algebra shows that $b \leqslant a^2$ and so $\sqrt{b} \leqslant |a|$. This allows us to conclude that $0 \leqslant \mathrm{ym}[1] \leqslant \mathrm{ym}[2]$. We now check the derivative of the differential equation to find the stable equilibrium:

```
simplify(subs(y[m] = ym[1], diff(dymdt, y[m]))):
simplify(subs(y[m] = ym[2], diff(dymdt, y[m]))):
```

Since the first of these quantities is negative and the second is positive, this shows that the first root ym[1] is the stable equilibrium. The proof in Appendix D shows that there is a unique stable equilibrium on the interval $[0, 1]$, and so we are guaranteed that ym[1] is on $[0, 1]$. Thus, we define this as the equilibrium value of y_m:

```
ym := ym[1]:
```

We now define the differential equations for u, v, x, and y, as well as the partial derivative of the fitness function with respect to the mutant pathogen exploitation level ξ_m:

```
dudt := 1 - (beta[max]*xi*(1 - x)*(1 - y)/(kappa + xi)
+ beta[max]*(1 - epsilon)*xi*(1 - x)*y/(kappa + xi)
+ beta[max]*(1 - epsilon)*xi*x*(1 - y)/(kappa + xi)
+ beta[max]*(1 - epsilon)^2*xi*x*y/(kappa + xi))*u*v - u:
dvdt := (beta[max]*xi*(1 - x)*(1 - y)/(kappa + xi)
+ beta[max]*(1 - epsilon)*xi*(1 - x)*y/(kappa + xi)
+ beta[max]*(1 - epsilon)*xi*x*(1 - y)/(kappa + xi)
+ beta[max]*(1 - epsilon)^2*xi*x*y/(kappa + xi))*u*v - (1 + xi)*v:
dxdt := -x/u + (k + 1)*x*(1 - x)*(y*(beta[max]*(1 - epsilon)*
xi/(kappa + xi) - beta[max]*(1 - epsilon)^2*xi/(kappa + xi))
+ (1 - y)*(beta[max]*xi/(kappa + xi) - beta[max]*(1 - epsilon)*
xi/(kappa + xi)))*v - k*(chi[c] + delta)*x:
dydt := -(beta[max]*xi*(1 - x)*y*(1 - y)/(kappa + xi)
+ beta[max]*(1 - epsilon)*xi*(1 - x)*y^2/(kappa + xi)
- beta[max]*(1 - epsilon)*xi*x*(1 - y)^2/(kappa + xi)
- beta[max]*(1 - epsilon)^2*xi*x*y*(1 - y)/(kappa + xi))*u
- k*(chi[c] + delta)*y:
W := (beta[max]*xi[m]*(1 - x)*(1 - ym)/(kappa + xi[m]) + beta[max]*
(1 - epsilon)*xi[m]*(1 - x)*ym/(kappa + xi[m]) + beta[max]*
(1 - epsilon)*xi[m]*x*(1 - ym)/(kappa + xi[m]) + beta[max]*
(1 - epsilon)^2*xi[m]*x*ym/(kappa + xi[m]))*u/(1 + xi[m]):
dWdxi := subs(xi[m] = xi, diff(W, xi[m])):
```

Using this, we put together the matrix J described in Appendix E:

```
J1 := Jacobian([dudt, dvdt, dxdt, dydt], [u, v, x, y] =
[(1 + sqrt(kappa))*(kappa + sqrt(kappa))/(beta[max]*sqrt(kappa)),
(1 - (1 + sqrt(kappa))*(kappa + sqrt(kappa))/(beta[max]*
sqrt(kappa)))/(1 + sqrt(kappa)), 0, 0]):
J2 := <0, 0, subs(u = (1 + sqrt(kappa))*
(kappa + sqrt(kappa))/(beta[max]*sqrt(kappa)),
x = 0, delta = 0, diff(dWdxi, x)), 0>^%T:
J3 := subs(u = (1 + sqrt(kappa))*(kappa + sqrt(kappa))/
(beta[max]*sqrt(kappa)), v = (1 - (1 + sqrt(kappa))*(kappa
+ sqrt(kappa))/(beta[max]*sqrt(kappa)))/(1 + sqrt(kappa)),
x = 0, y = 0, delta = 0, Jacobian([dudt, dvdt, dxdt, dydt, dWdxi],
[xi] = [sqrt(kappa)])):
J4 := <J1, J2>:
J5 := <J4 | J3>:
J := subs(chi[c] = beta[max]*sqrt(kappa)/((1 + sqrt(kappa))*(kappa
+ sqrt(kappa))) - (k + 1)*((1 - epsilon)*(beta[max]*sqrt(kappa)/((1
+ sqrt(kappa))*(kappa + sqrt(kappa))) - 1) + 1)/k, xi = sqrt(kappa),
delta = 0, J5):
```

We then extract the entry of J corresponding to the partial derivative in (14). We also define the critical cost threshold χ_c:

```
sigmanum := simplify(J[5, 3]):
costthresh := beta[max]*sqrt(kappa)/((1 + sqrt(kappa))*(kappa
+ sqrt(kappa))) - (k + 1)*((1 - epsilon)*(beta[max]*
sqrt(kappa)/((1 + sqrt(kappa))*(kappa + sqrt(kappa))) - 1) + 1)/k:
```

If we choose $\beta_{\max}$ and κ values so that $\mathcal{R}_0 > 1$, we get the following curve (where above the curve $\mathcal{R}_0 > 1$ and below the curve $\mathcal{R}_0 < 1$):

```
R0plot := plot((1 + sqrt(kappa))^2, kappa = 0 .. 5,
color = BLACK, thickness = 3, view = [0 .. 5, 0 .. 10],
labels = [kappa, beta[max]]):
```

If we also choose ϵ values so that $\chi_c > 0$ and k values so that $\epsilon < 1$, we can generate a series of $\beta_{\max}$-κ curves and plot them together with the previous curve for $\mathcal{R}_0$:

```
expressions := [seq(seq(subs([epsilon = beta[max]/(((1
+ sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2) + 10^n
+ 1)*(beta[max] - (1 + sqrt(kappa))^2)) + m*(((1
+ sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2) + 10^n)
- (1 + sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2))/
(10*((1 + sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2)
+ 10^n + 1)), k = (1 + sqrt(kappa))^2/(beta[max]
- (1 + sqrt(kappa))^2) + 10^n], costthresh), m = 1 .. 9),
n = -2 .. 2)]:
for i to 45 do
    p[i] := plot({solve(expressions[i], beta[max])},
    kappa = 0 .. 5, color = GREY, linestyle = DASH):
end do:
display({R0plot, seq(p[i], i = 1 .. 45)},
view = [0 .. 5, 0 .. 10], labels = [kappa, beta[max]]):
```

This generates the plot in panel A of Fig. 2 and suggests that all of the $\chi_c = 0$ curves overlap with the $\mathcal{R}_0 = 1$ curve, meaning that the region in which $\mathcal{R}_0 > 1$ coincides with the region in which $\chi_c > 0$. We can confirm this by looking at the difference between these two curves; running the following line of code will show that this difference is always exactly zero:

```
for i to 45 do
    simplify((1 + sqrt(kappa))^2 - solve(expressions[i],
    beta[max])[2]):
end do:
```

If we use the ϵ and k values chosen above, we can also look at the value of the partial derivative in Eq. (14). This gives a second set of $\beta_{\max}$-κ expressions that we can plot over top of the expressions for χ_c and $\mathcal{R}_0$ above:

```
expressions2 := [seq(seq(subs([epsilon = beta[max]/(((1
+ sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2) + 10^n
+ 1)*(beta[max] - (1 + sqrt(kappa))^2)) + m*(((1
+ sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2) + 10^n)
- (1 + sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2))/(10*
((1 + sqrt(kappa))^2/(beta[max] - (1 + sqrt(kappa))^2)
+ 10^n + 1)), k = (1 + sqrt(kappa))^2/(beta[max] - (1
+ sqrt(kappa))^2) + 10^n], sigmanum), m = 1 .. 9),
n = -2 .. 2)]:
for i to 45 do
    p2[i] := plot({solve(expressions2[i], beta[max])},
    kappa = 0 .. 5, color = GREY, linestyle = DOT):
end do:
display({R0plot, seq(p[i], i = 1 .. 45), seq(p2[i], i = 1 .. 45)},
view = [0 .. 5, 0 .. 10], labels = [kappa, beta[max]]):
```

This produces the plot shown in panel B of Fig. 2 and shows that all of these curves that represent the roots of the partial derivative in (14) lay below the curve for $\mathcal{R}_0 = 1$. So for all sensible sets of parameter values (i.e., all parameter values that satisfy $\chi_c > 0$ and $\mathcal{R}_0 > 1$), (14) has the same sign. We can take a test point in this region of parameter space to show that this sign is always positive, meaning that in a neighbourhood below the critical cost value the CSS pathogen exploitation level will decrease below its benchmark value.

Appendix G. Matlab code

Here, we present the Matlab code, described in Section 3.3 of the main text, used to numerically find the CSS pathogen exploitation level. We start by defining functions for the transmission rates and the recovery rate:

Listing 1:Matlab function `b00.m` that computes the transmission rate $\beta_{00}(\xi)$.

```
function trans00 = b00(xi, bmax, kappa)
% A function to compute beta00 transmission rate
trans00 = bmax*xi/(kappa + xi);
```

Listing 2: Matlab function `b01.m` that computes the transmission rate $\beta_{01}(\xi)$.

```
function trans01 = b01(xi, epsil, bmax, kappa)
% A function to compute beta01 transmission rate
trans01 = bmax*(1 - epsil)*xi/(kappa + xi);
```

Listing 3: Matlab function `b10.m` that computes the transmission rate $\beta_{10}(\xi)$.

```
function trans10 = b10(xi, epsil, bmax, kappa)
% A function to compute beta10 transmission rate
trans10 = bmax*(1 - epsil)*xi/(kappa + xi);
```

Listing 4: Matlab function `b11.m` that computes the transmission rate $\beta_{11}(\xi)$.

```
function trans11 = b11(xi, epsil, bmax, kappa)
% A function to compute beta11 transmission rate
trans11 = bmax*(1 - epsil)^2*xi/(kappa + xi);
```

Listing 5: Matlab function `g.m` that computes the recovery rate $\gamma(\xi)$.

```
function recov = g(xi)
% A function to compute gamma recovery rate
recov = xi;
```

We also need functions to define the resident system and the partial derivative with respect to ξ_m of the fitness function:

Listing 6: Matlab function `resident.m` that computes the system of ODEs for the u, v, x, y system.

```
function dres = resident(res, xi, k, epsil, bmax, ...
    cost, kappa)
% A function to define the resident (u, v, x, y) ODE
% system; res is a vector of four entries representing
% u, v, x, and y; other inputs are as detailed in the
% model in the main paper.
dres = zeros(1, 4);
dres(1) = 1 - (b11(xi, epsil, bmax, kappa)*res(3)*...
    res(4) + b01(xi, epsil, bmax, kappa)*...
    (1 - res(3))*res(4) + b10(xi, epsil, ...
    bmax, kappa)*res(3)*(1 - res(4)) + ...
    b00(xi, bmax, kappa)*(1 - res(3))*...
    (1 - res(4)))*res(1)*res(2) - res(1);
dres(2) = (b11(xi, epsil, bmax, kappa)*res(3)*...
    res(4) + b01(xi, epsil, bmax, kappa)*...
    (1 - res(3))*res(4) + b10(xi, epsil, ...
    bmax, kappa)*res(3)*(1 - res(4)) +...
    b00(xi, bmax, kappa)*(1 - res(3))*...
    (1 - res(4)))*res(1)*res(2) - ...
    (1 + g(xi))*res(2);
dres(3) = -res(3)/res(1) + (k + 1)*res(3)*...
    (1 - res(3))*(res(4)*...
    (b01(xi, epsil, bmax, kappa) - ...
    b11(xi, epsil, bmax, kappa)) + (1 - res(4))*...
    (b00(xi, bmax, kappa) - ...
    b10(xi, epsil, bmax, kappa)))*res(2) ...
    - k*res(3)*cost;
dres(4) = (b11(xi, epsil, bmax, kappa)*res(3)*res(4)*...
    (1 - res(4)) - b01(xi, epsil, bmax, kappa)*...
    (1 - res(3))*(res(4))^2 + ...
    b10(xi, epsil, bmax, kappa)*res(3)*...
    (1 - res(4))^2 - b00(xi, bmax, kappa)*...
    (1 - res(3))*res(4)*(1 - res(4)))*res(1) - ...
    k*res(4)*cost;
```

Listing 7: Matlab function `dFitness.m` that computes the partial derivative with respect to ξ_m of the fitness function.

```
function dWdxi = dFitness(xi, ubar, xbar, k, epsil, bmax, ...
    cost, kappa)
% A function to compute the partial derivative with respect to
% xi_m of the fitness function; ubar and xbar are the
% equilibrium values of u and x, respectively.
dWdxi = (1/(2*sqrt(k^2*(kappa + xi)^2*cost^2 - 2*k*ubar*xi*bmax*...
    (kappa + xi)*(epsil^2*xbar - 1)*cost + ubar^2*xi^2*bmax^2*...
    (epsil^2*xbar - 2*xbar*epsil + 1)^2)*(1 + xi)^2*(kappa ...
    + xi)^2))*(-(((ubar*(epsil^2*xbar - 2*xbar*epsil + 1)*bmax ...
    - k*cost)*xi^2 - 2*cost*k*kappa*xi - (ubar*(epsil^2*xbar ...
    - 2*xbar*epsil + 1)*bmax + k*kappa*cost)*kappa)*...
    sqrt(k^2*(kappa + xi)^2*cost^2 - 2*k*ubar*xi*bmax*(kappa ...
    + xi)*(epsil^2*xbar - 1)*cost + ubar^2*xi^2*bmax^2*...
    (epsil^2*xbar - 2*xbar*epsil + 1)^2) + (ubar^2*bmax^2*...
    (epsil^2*xbar - 2*xbar*epsil + 1)^2 - 2*k*ubar*cost*...
    (epsil^2*xbar - 1)*bmax + cost^2*k^2)*xi^3 + 3*k*cost*...
    ((-epsil^2*xbar + 1)*ubar*bmax + k*cost)*kappa*xi^2 ...
    + 3*(-(1/3)*ubar^2*bmax^2*(epsil^2*xbar - 2*xbar*epsil ...
    + 1)^2 - (1/3)*k*ubar*cost*(kappa - 1)*(epsil^2*xbar ...
    - 1)*bmax + cost^2*k^2*kappa)*kappa*xi + k*(ubar*bmax*...
    (epsil^2*xbar - 1) + k*kappa*cost)*cost*kappa^2));
```

Finally, we use all of these to define a function that numerically approximates the CSS level of pathogen exploitation:
Listing 8: Matlab function `findCSS.m` that approximates the CSS level of pathogen exploitation.

```
function css = findCSS(xires, tol, step, maxit, k, ...
    epsil, bmax, cost, kappa)
% A function to find a candidate CSS exploitation level;
% input xires should be a vector of the lower and upper
% initial resident exploitation levels; step is the step
% size used for choosing a new exploitation level and
% should be bigger than tol, the tolerance level for
% ending the program; maxit is the maximum number of
% iterations to compute before ending the program; other
% parameters are as detailed in the model in
% the main paper.
r0 = bmax/(1 + sqrt(kappa))^2;
usir = 1/r0;
vsir = (1/(1 + sqrt(kappa)))*(1 - 1/r0);
for i = 1:2
    % find equilibrium using Euler's method
    x0 = [usir, vsir, 0.01, 0.02]; % initialize
    dxidt = dFitness(xires(i), x0(1), x0(3), k, ...
        epsil, bmax, cost, kappa);
    itn = 1;
    if abs(dxidt) < tol
        flagxi = 0;
    else
        flagxi = 1;
    end
    while flagxi == 1
        if itn <= maxit % check to make sure maxit hasn't
        % been exceeded
```

```
flagx0 = 1; % flag to indicate resident system is
% away from equilibrium
while flagx0 == 1
    dx0dt = resident(x0, xires(i), k, ...
            epsil, bmax, cost, kappa);
    if max(abs(dx0dt)) < 1e-03
        flagx0 = 0;
    end
    x0 = x0 + (1e-05)*dx0dt;
    for j = 1:4
        if x0(j) < 0
            x0(j) = 0;
        end
    end
    if x0(1) + x0(2) > 1
        x0(1) = 0.5*(1 + x0(1) - x0(2));
        x0(2) = 1 - x0(1);
        x0(1) = x0(1) - 1e-05;
        x0(2) = x0(2) - 1e-05;
    end
    for j = 3:4
        if x0(j) > 1
            x0(j) = 1 - 1e-05;
        end
    end
end
dxidt = dFitness(xires(i), x0(1), x0(3), ...
    k, epsil, bmax, cost, kappa);
if abs(dxidt) < tol
```

```
                flagxi = 0;
            end
            xires(i) = xires(i) + step*dxidt;
            itn = itn + 1; % increase iteration counter
        else
            % print message if maximum number of iterations
            % has been exceeded
            out = sprintf('%f ', xires);
            fprintf(['Maximum number of iterations ', ...
                'exceeded; increase maxit and try ', ...
                'again. Most recent xi values: %s.\n'], ...
                out); % print message to console
            fprintf('Number of iterations: %d.\n', itn);
            return
        end
    end
end
% return average of both trajectories and take steps on either side
% to approximate second derivative
css = [mean(xires) + step, mean(xires), mean(xires) - step];
% compute equilibrium values of u, v, x, and y using Euler's method
x0 = [[usir, vsir, 0.01, 0.02]; [usir, vsir, 0.01, 0.02]; ...
    [usir, vsir, 0.01, 0.02]]; % initialize
for i = 1:3
    flagx0 = 1; % flag to indicate resident system is
    % away from equilibrium
    while flagx0 == 1
        dx0dt = resident(x0(i, :), css(i), k, epsil, ...
            bmax, cost, kappa);
```

```
        if max(abs(dx0dt)) < 1e-03
            flagx0 = 0;
        end
        x0(i, :) = x0(i, :) + (1e-05)*dx0dt;
        for j = 1:4
            if x0(i, j) < 0
                x0(i, j) = 0;
            end
        end
        if x0(i, 1) + x0(i, 2) > 1
            x0(i, 1) = 0.5*(1 + x0(i, 1) - x0(i, 2));
            x0(i, 2) = 1 - x0(i, 1);
            x0(i, 1) = x0(i, 1) - 1e-05;
            x0(i, 2) = x0(i, 2) - 1e-05;
        end
        for j = 3:4
            if x0(i, j) > 1
                x0(i, j) = 1 - 1e-05;
            end
        end
    end
end
% evaluate payoff function at these points
dWcssplus = dFitness(css(1), x0(1, 1), x0(1, 3), k, epsil, ...
    bmax, cost, kappa);
dWcssminus = dFitness(css(3), x0(3, 1), x0(3, 3), k, epsil, ...
    bmax, cost, kappa);
% use these values to evaluate the ESS condition
essCondition = (dWcssplus - dWcssminus)/(2*step);
```

```
% save the equilibrium values and CSS value, along with an
% indicator of whether the ESS condition is satisfied
if essCondition < (step^2)*10
    css = [x0(2, :), css(2), 1, essCondition];
else
    css = [x0(2, :), css(2), 0, essCondition];
end
out = sprintf('%f ', css);
fprintf(['Equilibrium values of u, v, x, y at the ', ...
    'CSS, the CSS level of xi, and whether the ESS condition ', ...
    'is satisfied (with value of second derivative): %s.\n'], ...
    out);
```

References

Alizon, S., 2008. Transmission-recovery trade-offs to study parasite evolution. Am. Nat. 172, E113–E121. https://doi.org/10.1086/589892.

Alizon, S., Michalakis, Y., 2015. Adaptive virulence evolution: the good old fitness-based approach. Trends Ecol. Evol. 30, 248–254. https://doi.org/10.1016/j.tree.2015.02.009.

Anderson, R.M., May, R.M., 1982. Coevolution of hosts and parasites. Parasitology 85, 411–426. https://doi.org/10.1017/S0031182000055360.

Anderson, R.M., Heesterbeek, H., Klinkenberg, D., Hollingsworth, T.D., 2020. How will country-based mitigation measures influence the course of the covid-19 epidemic? Lancet 395, P931–P934. https://doi.org/10.1016/S0140-6736(20)30567-5.

Antonovics, J., Iwasa, Y., Hassell, M.P., 1995. A generalized model of parasitoid, venereal, and vector-based transmission processes. Am. Nat. 145, 661–675. https://doi.org/10.1086/285761.

Bauch, C.T., Earn, D.J.D., 2004. Vaccination and the theory of games. Proc. Natl. Acad. Sci. USA 101, 13391–13394. https://doi.org/10.1073/pnas.0403823101.

Bauch, C.T., Galvani, A.P., 2013. Social factors in epidemiology. Science 342, 47–49. https://doi.org/10.1126/science.1244492.

Bremermann, H.J., Pickering, J., 1983. A game-theoretical model of parasite virulence. J. Theor. Biol. 100, 411–426. https://doi.org/10.1016/0022-5193(83)90438-1.

Britton, N.F., 2003. Essential Mathematical Biology. Springer. ISBN: 9781852335366.

Christiansen, F.B., 1991. On conditions for evolutionary stability for a continuously varying character. Am. Nat. 138, 37–50. https://doi.org/10.1086/285203.

Cooke, K.L., van den Driessche, P., 1996. Analysis of an SEIRS epidemic model with two delays. J. Math. Biol. 35, 240–260. https://doi.org/10.1007/s002850050051.

Cressler, C.E., McLeod, D.V., Rozins, C., van den Hoogen, J., Day, T., 2016. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. Parasitology 143, 915–930. https://doi.org/10.1017/S003118201500092X.

Day, T., 2002. On the evolution of virulence and the relationship between various measures of mortality. Proc. R. Soc. B 269, 1317–1323. https://doi.org/10.1098/rspb.2002.2021.

Day, T., Burns, J.G., 2003. A consideration of patterns of virulence arising from host-parasite coevolution. Evolution 57, 671–676. https://doi.org/10.1111/j.0014-3820.2003.tb01558.x.

Dercole, F., Rinaldi, S., 2008. Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications. Princeton University Press. ISBN: 9780691120065.

Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. J. Math. Biol. 34, 579–612. https://doi.org/10.1007/BF02409751.

Eshel, I., 1983. Evolutionary and continuous stability. J. Theor. Biol. 103, 99–111. https://doi.org/10.1016/0022-5193(83)90201-1.

Ewald, P.W., 1983. Host-parasite relations, vectors, and the evolution of disease severity. Annu. Rev. Ecol. Evol. Syst. 14, 465–485. https://doi.org/10.1146/annurev.es.14.110183.002341.

Frank, S.A., 1992. A kin selection model for the evolution of virulence. Proc. R. Soc. B 250, 195–197. https://doi.org/10.1098/rspb.1992.0149.

Fraser, C., Lythgoe, K., Leventhal, G.E., Shirreff, G., Hollingsworth, T.D., Alizon, S., Bonhoeffer, S., 2014. Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. Science 343, 1243727. https://doi.org/10.1126/science.1243727.

Gandon, S., 2004. Evolution of multihost parasites. Evolution 58, 455–469. https://doi.org/10.1111/j.0014-3820.2004.tb01669.x.

Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2001. Imperfect vaccines and the evolution of pathogen virulence. Nature 414, 751–756. https://doi.org/10.1038/414751a.

Gandon, S., Mackinnon, M., Nee, S., Read, A., 2003. Imperfect vaccination: some epidemiological and evolutionary consequences. Proc. R. Soc. B 270, 1129–1136. https://doi.org/10.1098/rspb.2003.2370.

Gao, S., Chen, L., Nieto, J.J., Torres, A., 2006. Analysis of a delayed epidemic model with pulse vaccination and saturation incidence. Vaccine 24, 6037–6045. https://doi.org/10.1016/j.vaccine.2006.05.018.

Gross, T., D'Lima, C.J.D., Blasius, B., 2006. Epidemic dynamics on an adaptive network. Phys. Rev. Lett. 96,. https://doi.org/10.1103/PhysRevLett.96.208701 208701.

Heffernan, J.M., Smith, R.J., Wahl, L.M., 2005. Perspectives on the basic reproductive ratio. J.R. Soc. Interface 2, 281–293. https://doi.org/10.1098/rsif.2005.0042.

Hethcote, H.W., Lewis, M.A., van den Driessche, P., 1989. An epidemiological model with a delay and a nonlinear incidence rate. J. Math. Biol. 27, 49–64. https://doi.org/10.1007/BF00276080.

Hughes, D.P., Andersen, S.B., Hywel-Jones, N.L., Himaman, W., Billen, J., Boomsma, J.J., 2011. Behavioral mechanisms and morphological symptoms of zombie ants dying from fungal infection. BMC Ecol. 11. https://doi.org/10.1186/1472-6785-11-13.

Hurford, A., Cownden, D., Day, T., 2010. Next-generation tools for evolutionary invasion analyses. J.R. Soc. Interface 7, 561–571. https://doi.org/10.1098/rsif.2009.0448.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proc. R. Soc. A 115, 700–721. https://doi.org/10.1098/rspa.1927.0118.

Knolle, H., 1989. Host density and the evolution of parasite virulence. J. Theor. Biol. 136, 199–207. https://doi.org/10.1016/S0022-5193(89)80226-7.

Matessi, C., Di Pasquale, C., 1996. Long-term evolution of multilocus traits. J. Math. Biol. 34, 613–653. https://doi.org/10.1007/BF02409752.

Maynard Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press. ISBN:0521288843.

McCallum, H., Barlow, N., Hone, J., 2001. How should pathogen transmission be modelled? Trends Ecol. Evol. 16, 295–300. https://doi.org/10.1016/S0169-5347(01)02144-9.

Metz, J.A.J., Nisbet, R.M., Geritz, S.A.H., 1992. How should we define 'fitness' for general ecological scenarios? Trends Ecol. Evol. 7, 198–202. https://doi.org/10.1016/0169-5347(92)90073-K.

Murall, C.L., Bauch, C.T., Day, T., 2015. Could the human papillomavirus vaccines drive virulence evolution? Proc. R. Soc. B 282. https://doi.org/10.1098/rspb.2014.1069.

Newman, M.E.J., 2002. Spread of epidemic disease on networks. Phys. Rev. E 66,. https://doi.org/10.1103/PhysRevE.66.016128 016128.

Osterberg, L., Blaschke, T., 2005. Adherence to medication. N. Engl. J. Med. 353, 487–497. https://doi.org/10.1056/NEJMra050100.

Øverli, Ø., Páll, M., Borg, B., Jobling, M., Winberg, S., 2001. Effects of Schistocephalus solidus infection on brain monoaminergic activity in female three-spined sticklebacks Gasterosteus aculeatus. Proc. R. Soc. B 268, 1411–1415. https://doi.org/10.1098/rspb.2001.1668.

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic dynamics and endemic states in complex networks. Phys. Rev. E 63,. https://doi.org/10.1103/PhysRevE.63.066117 066117.

Patton, S.R., 2011. Adherence to diet in youth with type 1 diabetes. J. Am. Diet. Assoc. 111, 550–555. https://doi.org/10.1016/j.jada.2011.01.016.

Pharaon, J., Bauch, C.T., 2018. The influence of social behaviour on competition between virulent pathogen strains. J. Theor. Biol. 455, 47–53. https://doi.org/10.1016/j.jtbi.2018.06.028.

Poolman, E.M., Elbasha, E.H., Galvani, A.P., 2008. Vaccination and the evolutionary ecology of human papillomavirus. Vaccine 26, C25–C30. https://doi.org/10.1016/j.vaccine.2008.04.010.

Reluga, T.C., 2005. Simple models of antibiotic cycling. Math. Med. Biol. 22, 187–208. https://doi.org/10.1093/imammb/dqi002.

Robison, J., Rogers, M.A., 1994. Adherence to exercise programmes. Sports Med. 17, 39–52. https://doi.org/10.2165/00007256-199417010-00004.

Ross, R., 1916. An application of the theory of probabilities to the study of a priori pathometry.—Part I. Proc. R. Soc. A 92, 204–230. https://doi.org/10.1098/rspa.1916.0007.

Schaller, M., 2011. The behavioural immune system and the psychology of human sociality. Philos. Trans. R. Soc. B 366, 3418–3426. https://doi.org/10.1098/rstb.2011.0029.

Taylor, P.D., Jonker, L.B., 1978. Evolutionary stable strategies and game dynamics. Math. Biosci. 40, 145–156. https://doi.org/10.1016/0025-5564(78)90077-9.

Úbeda, F., Jansen, V.A.A., 2016. The evolution of sex-specific virulence in infectious diseases. Nat. Commun. 7. https://doi.org/10.1038/ncomms13849.

van Baalen, M., 1998. Coevolution of recovery ability and virulence. Proc. R. Soc. B 265, 317–325. https://doi.org/10.1098/rspb.1998.0298.

Wesołowska, W., Wesołowska, T., 2014. Do Leucochloridium sporocysts manipulate the behaviour of their snail hosts? J. Zool. 292, 151–155. https://doi.org/10.1111/jzo.12094.

Wild, G., Gardner, A., West, S.A., 2009. Adaptation and the evolution of parasite virulence in a connected world. Nature 459, 983–986. https://doi.org/10.1038/nature08071.