The exclusion problem in seasonally forced epidemiological systems

J. V. Greenman*
Computing Science and Mathematics
University of Stirling
Stirling, FK9 4LA, Scotland, UK
tel: +44(0)1786-467421; fax: +44(0)1786-464551
email: j.v.greenman@stir.ac.uk
and

B. Adams
Department of Mathematical Sciences
University of Bath
Bath, BA2 7AY, UK
tel: +44(0)1225-386320; fax: +44(0)1225-386492
email: b.adams@bath.ac.uk

Keywords: phase control, optimal scheduling, basic reproduction number R_0, reservoir of infection, resonance

Running Title: Exclusion in seasonally forced systems

* Corresponding author

First submitted 12 February 2014

Revised: 19 August 2014
Revised: 10 October 2014

Abstract

The pathogen exclusion problem is the problem of finding control measures that will exclude a pathogen from an ecological system or, if the system is already disease-free, maintain it in that state. To solve this problem we work within a holistic control theory framework which is consistent with conventional theory for simple systems (where there is no external forcing and constant controls) and seamlessly generalises to complex systems that are subject to multiple component seasonal forcing and targeted variable controls. We develop, customise and integrate a range of numerical and algebraic procedures that provide a coherent methodology powerful enough to solve the exclusion problem in the general case. An important aspect of our solution procedure is its two-stage structure which reveals the epidemiological consequences of the controls used for exclusion. This information augments technical and economic considerations in the design of an acceptable exclusion strategy. Our methodology is used in two examples to show how time-varying controls can exploit the interference and reinforcement created by the external and internal lag structure and encourage the system to 'take over' some of the exclusion effort. On-off control switching, resonant amplification, optimality and controllability are important issues that emerge in the discussion.

1. Introduction

It is well known that variability in the environment can have a significant impact on the dynamic behaviour of epidemiological systems whether they involve humans or animals or both (Chesson, 1982; Grassly and Fraser, 2006). Of particular interest is periodic variation, especially seasonality, but there are other examples where the period is greater than a year (e.g. El Nino (Koelle et al., 2005) and African rain patterns (Wichmann et al., 2003)) or less than a year (e.g. marine life subject to tidal or light intensity cycles (Rinaldi et al., 1993)). In simple epidemiological systems environmental forcing acts primarily through infection transmission. This is the case in childhood diseases such as measles (Dietz, 1976) where the seasonal variation is caused by the term structure of the school year. In other cases several forcing components are in play. A study of conjunctivitis in house finches (Hosseini et al., 2004) found that infection is transmitted in the autumn/winter when there is population aggregation but breeding takes place in the summer when there is dispersal. These two seasonal effects are 'out of phase'. A third example, also highlighting the importance of lags, is that of a managed game-bird population subject to two forms of variable external forcing: A seasonally transmitted disease and 'harvesting' restricted to particular times of the year. The choice of lag between transmission and harvesting determines whether harvesting reduces or increases the impact of the disease (Choisy and Rohani, 2006).

Another important factor that influences how external forcing affects a system is the number and nature of the different infected host types (Diekmann et al., 1990, 2010). Unforced systems with more than one host type have received a lot of attention in the literature, for example Anderson and May (1981, 1986) and many others. Of particular interest recently has been the spread of bovine TB between badgers and livestock (Cox et al. 2005, Lintott et al. 2013) and the dominance of the grey over the red squirrel population because of reinforcement between direct competition and apparent competition mediated by a parapox virus (Tompkins et al. 2003). Much less work has been carried out on how external forcing affects transmission between species and there remains much more to do in this area (Brassil, 2006). However, seasonality in host-vector systems has received some attention. Bacaer and Guernaoui (2006) analysed a seasonal model for *leishmaniasis* in Chichaoua, Morocco while Wang and Zhao (2008) studied a simple seasonal model for dengue.

The specific problem studied in this paper is how to exclude a pathogen from an epidemiological system or how to maintain that exclusion if the system is already disease-free. In the absence of forcing and with constant controls the exclusion problem can be solved explicitly for a standard model in terms of its resident asymptotic state even when there are multiple host types (Diekmann et al. 1990, 2010). With forcing present the exclusion problem is much more difficult to solve (Heesterbeek and Roberts, 1995; Bacaer and Guernaoui, 2006; Wang and Zhao 2008). Most applications have been limited to the simplest cases with forcing only on infection transmission and no structure in the resident subsystem. Further advances in solution methods are necessary to study the new opportunities for bringing about exclusion that are created when these limitations are removed.

Our primary objective in this paper therefore is to contribute to the development of a sufficiently powerful coherent and insightful methodology to solve the exclusion problem for the general case of complex epidemiological systems that have a structured resident subsystem (with predation or competitive forces in play for example), multiple infected host types and subject to variable controls and multiple

seasonal forcing components. Since the exclusion problem is a control problem, involving intervention with a set of control measures, we work within a control theory framework to find the levels of these controls that bring about exclusion. Our methodology involves the following three main components: An approximation procedure that replaces nonlinearities by explicitly solvable linear equations (Greenman and Pasour, 2012); monodromy theory on which to base the numerical calculations (Hale, 1969); optimal control theory of use in exploring the impact of variable controls (Lenhart and Workman, 2007). Integrating these different procedures creates an efficient 'fit for purpose' exclusion methodology and, in so doing, divides the exclusion process into two distinct stages that provide insight into the epidemiology of exclusion and connectivity with other approaches to be found in the literature.

The paper is set out as follows. In section 2 there is a general discussion on how to solve the exclusion problem for a special control u that will later provide the link to all other controls of interest. In section 3 the exclusion procedure is applied to invasion systems with one infected state and in section 4 to systems with 2 or more such states. It is in sections 3.4 and 4.4 where it is shown how to extend the theory to handle a general set of pre-emptive controls. Examples illustrate what difference forcing can make to the exclusion dynamics and what mechanisms are activated during forcing to explain the changes. This involves comparing systems with single or multiple host types, forcing with single or multiple components and controls that are constant or variable.

2. The pathogen exclusion problem from a control theory perspective

In section 2.1 we introduce the rare invader approximation that simplifies the solution of the exclusion problem by dividing it into two stages. Further we introduce the special control u that removes a proportion of the newly infecteds and define the 'effort' required to remove the pathogen using this control. In section 2.2 we describe how to apply the zero invader growth condition for exclusion by relating this condition to the eigenvalues of the monodromy matrix. We discuss the relationship between exclusion effort and the basic reproduction number R_0 when control u is constant and highlight the strengths of the control approach.

2.1 The Rare Invader Approximation

Consider the controlled epidemiological system modeled by the equations:

$$\frac{d\underline{x}}{dt} = \underline{f}(\underline{x}, \underline{y}, \underline{u}, t) \tag{1a}$$

$$\frac{d\underline{y}}{dt} = \underline{g}(\underline{x}, \underline{y}, \underline{u}, t) \tag{1b}$$

where $\underline{x}$ is the vector of uninfected (resident) populations (for example the susceptibles and the immune), $\underline{y}$ the vector of the infected (invader) populations and $\underline{u}$ the vector of controls. Equations (1a) are the 'resident equations' and (1b) the 'invasion equations'. This system is subject to periodic environmental forcing as indicated by the explicit time dependence t of functions $\underline{f}$, $\underline{g}$. We are particularly interested in the case that control $\underline{u}$ is variable in time but first we consider the simpler case where it is constant.

The solution of the (pathogen) exclusion problem (to exclude or prevent invasion of a pathogen) is simplified by using the Rare Invader Approximation (RIA)

which assumes that the number of infecteds is so small that they can be ignored in the resident equations (1a) and so small that the invasion equations (1b) can be linearised about the disease free equilibrium. This approximation holds in the early stage of an invasion or in the final stage of exclusion. The RIA defines a two-stage solution procedure: First solve the decoupled resident equations: $d\underline{x}/dt = \underline{f}(\underline{x}, \underline{0}, \underline{u}, t)$ for $\underline{x}$. Then solve the linear invasion equations after substitution of the asymptotic resident solution, $\underline{x}_\infty$, i.e. solve: $d\underline{y}/dt = G(\underline{x}_\infty, \underline{0}, \underline{u}, t)\underline{y}$ where $G(\underline{x}, \underline{y}, \underline{u}, t) = \partial \underline{g}/\partial \underline{y}$ is the matrix of derivatives of vector function $\underline{g}$. Matrix G is the Jacobian for the invasion subsystem and will be labeled more simply as J:

$$\frac{d\underline{y}}{dt} = J.\underline{y} \tag{2}$$

The problem is to find controls $\underline{u}$ where the asymptotic growth rate of the infected populations in (2) is zero. These solutions define the 'pathogen threshold' separating solutions where the growth rate is negative (i.e. the pathogen is excluded) and positive (i.e. the pathogen invades). On this threshold the RIA becomes exact and so its use in solving the exclusion problem is appropriate.

In epidemiological models with compartmental structure and with the controls inactive, matrix J in (2) can be written as $J = F - V$ where F is the transmission matrix specifying the number of newly infecteds (per infected individual) for each host type and V is the transition matrix that specifies the flow rates between compartments and with the external world. The term 'host type' identifies the state a host enters at the point of infection and hence the different ways in which infection can occur (Diekmann and Heesterbeek (2000); Hartemink et al. 2008). We will also use the term 'infected state' to identify the states a host can be in throughout its infected lifetime. The host types identify a subset of the infected states. For example, for the SEIR model there is one host type (the latent state E), but two infected states (the latent state E and the infectious state I). For the SISI model, describing the transmission of disease without latency within and between two host species, there are two host types corresponding to the I states and these are also the infected states. This model is discussed in section 4.

Now consider the effect of the special control, constant in time and acting alone, that allows only the proportion u of the newly infecteds for each host type to survive, the remainder being removed (by culling or quarantine for example). Then the invasion matrix equation (2) becomes:

$$\frac{d\underline{y}}{dt} = (uF - V)\underline{y}\,. \tag{3}$$

Control value u satisfies the constraints: $0 < u \leq 1$ with $u = 1$ indicating an inactive control. To exclude the pathogen we need to reduce u sufficiently but common experience suggests that the smaller u is the disproportionately greater the effort required to reduce it further. To model this feature in a simple but realistic way we will take the effort to be inversely proportional to u so that the effort required to exclude the pathogen (i.e. the 'exclusion effort') is the inverse of the value of u that will move the system to the pathogen threshold. We will denote the 'exclusion effort' with constant control u by E_0 when environmental forcing is present and $\bar{E}_0$ when not.

2.2 Solving the invasion equation when control u remains constant in time

If there is no variation in the environment, controls or the decoupled resident asymptotic state then standard linear algebraic methods can be used to solve the matrix invasion equation (3) explicitly (in terms of the asymptotic resident population levels) and the exclusion problem (to find the exclusion value $\bar{\bar{E}}_0$ for control u) by imposing the condition of zero asymptotic growth for the invaders. This condition requires that 0 is the leading eigenvalue of the now constant matrix J. A necessary condition for this to be the case is that the determinant (det) of J is 0. We can pursue this argument further by rewriting equation (3) as:

$$\frac{d\underline{\underline{y}}}{dt} = (FV^{-1} - u^{-1}I)(uV)\underline{\underline{y}}$$

with V non-singular (i.e. $\det(V) \neq 0$) and I the identity matrix. A condition for zero growth is that $\det(FV^{-1} - u^{-1}I) = 0$, i.e. u^{-1} is an eigenvalue of FV^{-1}. So if exclusion is to happen in the long term $u^{-1} = \bar{\bar{E}}_0$ must be the dominant eigenvalue of matrix FV^{-1}, the so-called Next Generation Matrix. This eigenvalue is in fact equal to the basic reproduction number R_0, i.e. $\bar{\bar{E}}_0 = R_0$ (Diekmann et al., 2010; Hartemink et al., 2008). Conventionally R_0 is expressed in epidemiological terms as the expected number of secondary infections arising from a 'typical' infected individual in an otherwise naïve population. (For a more precise definition see Diekmann et al., 2010). In control terms, R_0 is equal to the effort required for pathogen exclusion using the constant control u.

If the system is subject to periodic environmental forcing then in general there will no longer be an explicit solution to equation (3) because even though it remains linear the elements of J will be time-dependent (Hale, 1969). (An exception is when there is just one infected population. Then (3) can be solved directly by integration.) We have two choices: Solve (3) numerically or use an approximate method to obtain an analytic solution. In the numerical method, solving (3) and then imposing the zero growth condition is equivalent to constructing the linear operator that shifts the system forward in time through a complete cycle. If the control is chosen so that the dominant eigenvalue of the matrix defining this operator, the so-called monodromy matrix, is equal to 1 then the system will return to its original state after completing a cycle if started in the corresponding eigenstate or will eventually reach this state otherwise. In this eigenstate there will be no growth or decay in the infected populations over a cycle. The eigenvalues of the monodromy matrix are called Floquet multipliers and the theory on which it is based is called Floquet theory. The algorithm for constructing the monodromy matrix is presented in Section 4**.**

The second method for solving the exclusion problem is algebraic, yielding an approximate formula for the invader growth rate (Greenman and Pasour, 2012) which can be used to explore the general properties of the threshold dynamics of the system. In this second method the invading populations are written as power series in the strengths of the environmental variations. This enables the equations (3) to be transformed into an unbounded set of linear differential equations that can be solved explicitly in sequence. For medium strength forcing a good enough approximation can usually be obtained to second order in forcing strength, yielding a Quadratic Approximation (QA) for the average growth rate (over a cycle), with the linear term zero. It is important to note that this approximation can also be used for solving the resident equations when nonlinearities are present and when the resident subsystem is asymptotically unstable.

With forcing, the exclusion effort E_0 associated with constant control u is equal to the 'generalised' basic reproduction number introduced by Bacaer and Gueranoui (2006). It is necessary to generalise R_0 because, as originally defined, R_0 would depend on the time when secondary infections were initiated by infected individuals. The generalised R_0 is equal to the limit of the ratio of new infections in two successive generations, as the number of elapsed generations becomes large (Bacaer and Ait Dads, 2011), thus steering around this time- dependency problem. The interpretation in control terms does not change. It remains clear, direct and operationally useful.

One of the most important advantages of the control theory approach to pathogen exclusion over conventional theory is the seamless transition that can be made from constant to variable controls. Such a transition is essential when there is environmental forcing since variable controls can often take advantage of the fluctuations generated by the forcing to reduce the exclusion effort, for example by reducing the size of a susceptible population when infection transmission is at its highest. Of particular interest is the least effort control strategy which can be found by harnessing the power of optimal control theory. However this particular optimisation problem is highly singular and requires a special algorithm involving the monodromy matrix to solve it.

3. Solving the pathogen exclusion problem for a model with one infected state

To keep the analysis simple, attention in this section is restricted to environmentally forced systems with just one host type and no latency. In this case the single invasion equation (3) can be solved explicitly and a formula for exclusion effort E_0, obtained by QA (section 3.1). This formula can be used to identify the factors that contribute to the distortion of the pathogen threshold (lying between the exclusion and invasion regions) and the effect of lags between forcing components on exclusion effort (section 3.2). The least effort when control u is varying is found as the exact solution to an optimal control problem (section 3.3). The final step is establishing the relationship between the special control u and a general pre-emptive control using the two-stage structure of the algorithm to solve the exclusion problem.

3.1 Constant control in a variable environment

When infection transmission is density-dependent (McCallum et al., 2001) the single host invasion equation (3) becomes:

$$\frac{dI}{dt} = u\beta SI - d_0 I = (uF - V)I \tag{4}$$

where I is the infectious and S the susceptible population, β the infection transmission coefficient varying periodically (with period p) due to environmental forcing and d_0 the exit rate from the infectious state. Comparison with (3) gives $F = \beta S$, $V = d_0$ and $y = I$. Suppose the susceptible population S is known from the solution to the resident equation(s) in RIA. Then we can solve (4) exactly to find the average growth rate ξ_{ave} of the infected population I over a cycle, by first dividing by $I(t)$, then integrating over a cycle. Precisely:

$$\xi_{ave} = \frac{1}{p}\ln\left(\frac{I(p)}{I(0)}\right) = \frac{1}{p}\int \frac{1}{I}\frac{dI}{dt}dt = \frac{1}{p}\int u\beta S dt - d_0 \tag{5}$$

When u is constant it can be factored out of the integral and the condition $\xi_{\text{ave}} = 0$ trivially solved for u (i.e. $u = (pd_0/\int \beta S \mathrm{d}t)$). This is the level of control u that will bring about pathogen exclusion. If, further, S is constant in RIA then $u = (\beta_0 S/d_0)^{-1}$ for

exclusion where β_0 is the average infection transmission coefficient, unchanged from the unforced value if β is held at β_0. However if resident populations are forced as well as β then both S and β will vary in time and the exclusion level of (constant) control u will change due to the forcing. This is because the integral $\int \beta S dt$ in (5) will depend on the 'covariance' between the S and β oscillations. Formula (5) holds whatever the structure of the resident subsystem provided there is only one susceptible population S. For example in the SIR model there are two residents, S and R (= recovered). Similarly in a predator (P) - prey (S) system with immune predator and susceptible prey there are also two residents, S and P.

To examine in more detail what can happen when there is covariance between S and β consider the familiar SIS model with just one resident (S) and one invader (I):

$$\frac{dS}{dt} = aH - r\frac{H^2}{K} - bS + (\gamma I - \beta SI) \tag{6a}$$

$$\frac{dI}{dt} = u\beta SI - d_0 I \tag{6b}$$

where $H = S + I$ and K is the carrying capacity. Note that the density-dependence indicated by K works through fertility rather than mortality. Parameters $r = a - b$ and $d_0 = b + \alpha + \gamma$ where a is the initial (per capita) birth rate, b is the natural mortality rate, α is the virulence and γ is the recovery rate. In RIA the term in brackets in the resident equation (6a) is deleted and H is replaced by S (since $I = 0$). The invasion equation (6b) is already linear in I. Consider the case where there are variations in β and K with time profiles:

$$\beta = \beta_0(1 + \delta_1 \cos(\omega t - \phi)) \text{ and } K = K_0(1 + \delta_2 \cos(\omega t)) \tag{7}$$

where β_0 and K_0 are the average values, δ_1 and δ_2 are the forcing strengths, p is the period of the cycle, $\omega = 2\pi/p$ and ϕ is the (phase) lag between these two forcing components. The analysis is not straightforward because the resident equation (6a) cannot be solved exactly in scenario (7) even in RIA. An approximate formula for the control level u that achieves exclusion is obtained using a quadratic approximation (QA) in forcing strengths δ_1, δ_2. The end result is that the exclusion effort E_0 (equal to u^{-1} at exclusion) is given by:

$$E_0 = A_0 - \delta_2^2 A_2 + \delta_1 \delta_2 A_1 \cos(\phi - \psi) \tag{8}$$

where A_0 is the exclusion effort for no forcing (i.e. $A_0 = \bar{E}_0$), ψ the internal lag (with S taking time to adjust to changes in K) and A_1, A_2 are two non-negative constants dependent on the unforced model parameters and the forcing period. For details of the calculation see Appendix A.

3.2 Interference and reinforcement

Equation (8) is a general formula for the pathogen threshold of the SIS model for scenario (7) in (ϕ, E_0) space. For the parameter values given as set 1 in Table B1 (Appendix B) the threshold has the shape (bb) shown in Fig 1. In the region above the threshold the pathogen is excluded but below the threshold it can invade. To understand the mechanisms that create the change in position and shape of the threshold under environmental forcing, suppose that initially the forcing is inactive. Then the threshold lies along the line (aa) (Fig 1) at effort level $\bar{E}_0$. Now activate the forcing on K, but not on β. Then the threshold 'shifts' from line (aa) to line (cc), due to the average value of the now oscillating S population being lowered. This 'shift'

corresponds to the second term on the right hand side of (8). If the forcing on β is also activated then there is a nonlinear distortion (bb) of the threshold about (cc) caused by the 'covariance' between the oscillations in β and K, generating the dependence of E_0 on the lag ϕ. This distortion corresponds to the third term in (8). Of special interest is the minimum point A on threshold (bb) where least effort is required to exclude the pathogen. At this value of ϕ the oscillations of β and S are exactly 'out of phase'. β is at its highest when S is at its lowest (Fig 2) thereby reducing the impact of high infection transmission, the average number of newly infecteds and hence E_0. At the highest point B of threshold (bb) there is 'reinforcement' (rather than the 'interference' at point A) with β and S exactly 'in phase' rising and falling together, increasing the effectiveness of infection transmission and hence E_0. The shape of the forced threshold shown in Fig 1 will hold for all choices of model parameters since (8) is a general formula (for medium strength forcing). However there are differences in the detail. For example in Fig 1 forcing will reduce the exclusion effort required whatever the lag. This is not always the case. For larger δ_1 but smaller δ_2 (e.g. $\delta_1 = 0.95$, $\delta_2 = 0.5$) it is possible for the point B to lie above the (unforced) line (aa) in which case exclusion effort will be increased with forcing. It is also possible for point A to lie below the line $E_0 = 1$ (e.g. $\delta_1 = 0.5$, $\delta_2 = 0.85$). But $E_0 < 1$ means that the pathogen has already been excluded without the need for intervention. ($E_0 < 1$ implies $u > 1$, i.e. newly infecteds have to be added to the system to ensure that the zero growth condition $\xi_{ave} = 0$ is satisfied.) So in this case forcing brings about exclusion without intervention. (Note that formula (8) provides a good approximation for relatively high values of the forcing strengths δ_i (~ 0.6) because the effective forcing strength in (8) is quadratic in the δ_i.)

The analysis also applies to other scenarios, for example when birth rate a or mortality b rather than K is forced. But then $A_2 = 0$ and so there is no shift. Similar shapes and behaviour reappear in more complex structured models as we will see.

3.3 Variable control: optimality

Previous work (e.g. Choisy and Rohani, 2006) has shown that exclusion effort can often be reduced by using controls that vary in time. How much reduction is achievable is found from the minimum exclusion effort (E^*) given by the solution of the optimal control problem (9) below. This allows us to compare minimum effort E^* (for a varying control) with effort E_0 (for a constant control) and hence with R_0 (the basic reproduction number).

For a model with invasion equation (3) the least effort required to achieve pathogen exclusion is found by solving the optimal control problem for control $u = u(t)$ that ensures zero growth (9c) and minimises average effort (9a) over a cycle, i.e.

$$\text{minimise: } \frac{1}{p}\int \frac{1}{u(t)}\,dt \tag{9a}$$

$$\text{subject to: } \frac{dy}{dt} = \big(u(t)F(t) - V(t)\big).\underline{y} \tag{9b}$$

$$\text{and: } \underline{y}(0) = \underline{y}(p) \tag{9c}$$

$$\text{and: } 0 < u(t) \le 1 \quad \text{over a cycle} \quad 0 \le t \le p. \tag{9d}$$

This optimal control problem is formulated for a general (compartmental) epidemiological system. It can be reduced to a set of algebraic and differential equations using Hamiltonian theory (Lenhart and Workman 2007) but in this case the standard numerical algorithm to solve these equations is highly unstable. So in section 4.3 we introduce a novel algorithm that overcomes this difficulty. However,

when there is just one infected state an exact solution can be found. Precisely, the optimal control $u(t)$ and minimum effort E^* are given by:

$$\frac{1}{u(t)} = A\sqrt{F(t)} \text{ where } A = \frac{\int \sqrt{F(w)}\, dw}{\int V(w)\, dw} \tag{10a}$$

$$E^* = \frac{1}{p} \frac{\left(\int \sqrt{F(w)}\, dw\right)^2}{\int V(w)\, dw} \tag{10b}$$

where the integrals are evaluated over a cycle and provided the constraints (9d) are satisfied. (For details see Appendix C.) To interpret these results it is useful to introduce the notion of (instantaneous) effort $E(t)$ being exerted at moment t. Then $E(t) = 1/u(t)$ in (10a) and $E^* = (1/p)\int E(t)\, dt$ in (10b). So the distribution of effort $E(t)$ over time that minimises exclusion effort is proportional to $\sqrt{F(t)}$ following the variation in $F(t)$ but modulated by the square root damping. The fact that $E(t)$ and $F(t)$ are in phase, rising or falling together, means that the survival factor $u(t)$ is out of phase with the newly infecteds, i.e the smallest proportion that survive is when the number of newly infected, $F(t)$, is at its greatest. The net effect is that the total number of surviving newly infecteds falls and as a result so does the exclusion effort.

Problem (9) is in fact an optimal scheduling problem. E^* is the least effort required in scheduling the removal of newly infecteds to achieve pathogen exclusion. E^* therefore maintains a direct link between control and second generation infecteds. This optimal scheduling formulation can be advantageously compared with ways of handling external forcing that find alternative solutions to the time-dependency problem in second generation infection (Williams and Dye, 1997; Omori and Adams, 2011).

As an example consider again the SIS model (6) in scenario (7) with $F(t) = \beta(t)S(t)$, $V(t) = d_0$ and parameter values given in Table B1 (set 1). When $\phi = 0.33\pi$ (at point B in Fig 1), least effort under variable control is $E^*=1.53$ (calculated from (10b)), less than the effort under constant control $E_0 = 1.67$. This compares with $E^* = 1.396$ and $E_0 = 1.403$ (at point A) when $\phi = 1.34\pi$. Fig 3 shows how the optimal control effort varies over a cycle. Path d is for lag $\phi = 0.33\pi$ (point B in Fig 1) and path b for lag 1.34π (point A in Fig 1). Also superimposed are the optimal control paths for $\phi = \pi$ (path a) and $3\pi/2$ (path c) to show how path d evolves into path b and conversely. As indicated in (10a) these paths are synchronised with the variations in the newly infecteds $F = \beta S$. With reinforcement (path d) there are large amplitude fluctuations about a high average level (because S and β are in phase) while with interference (path b) the amplitude and average level are much lower (because S and β are out of phase). In fact for interference, constant control gives a good approximation for path b and for E^* (as we have seen from 10b).

These calculations illustrate the general inequalities $E^* \le E_0$ (by optimality) and $1 < E^*$ that hold when exclusion takes place under forcing. The inequality $1 < E^*$ follows from the constraint: $u(t) \le 1$ (9d) at each point of a cycle. ($u(t) > 1$ would mean 'adding' rather than removing newly infecteds). However for exclusion, $u(t)$ has to be less than 1 for at least part of the cycle and inactive ($u(t) = 1$) for the remaining part otherwise the inequality would not be strict. In these inactive intervals, between switching the control off and on again, the system 'takes over'. If $u(t) = 1$ for all t (with $E^* = 1$) then there is control inactivity throughout the cycle, i.e. the pathogen is already at threshold. An example of switching behaviour is seen on

path d (Fig 3). Control u is switched off at $t = 2.8$ and on again at $t = 3.8$. Switching happens when forcing is strong enough that the large fluctuations for path d breach the constraint. The switching that occurs is a 'clever' solution found by the optimality algorithm that efficiently reschedules removal of newly infecteds in response to the constraint.

The inequality $E^* \leq E_0$ can be rewritten in terms of the basic reproduction number as $E^* \leq R_0$ (since $E_0 = R_0$ (Bacaer and Gueranoui, 2006)). So in control terms R_0 will in general overestimate the effort required to exclude the pathogen when a system is forced because of the assumption that the control is constant in time.

3.4 Generalisation: Pathogen exclusion using pre-emptive controls

At first sight it might appear that the special control u used to remove newly infecteds would be of little interest since in the great majority of situations it would be difficult to implement. However we now show that u plays a much more significant role as a key component in the analysis of other more practical exclusion controls. In fact we can use the methods that have been developed in this section and that will be generalised in the next to help solve the exclusion problem for a general pre-emptive control P (acting only on the resident populations) such as culling or vaccination for example.

The decoupling of the resident subsystem in RIA means that the analysis of the impact of P can be divided into two stages. Applying control P leads first to changes in the resident populations and then to changes in the invasion populations through the links that exist between these two subsystems. It is in this second stage that we reintroduce the previous control u that will mimic the action of control P to find the invader growth rate and hence the exclusion effort.

To illustrate this process consider again the single invasion equation (4) with just one link to the resident subsystem. Equation (4) will take the form (11a) if u is 'turned off' (i.e. $u = 1$) and the susceptible population S becomes a function of P and t. If instead P is 'turned off' (with $P = 0$ say) then (4) will take the form (11b):

$$\frac{dI}{dt} = \beta(t)S(P,t)I - d_0 I \tag{11a}$$

$$\frac{dI}{dt} = u\beta(t)S(0,t)I - d_0 I \tag{11b}$$

The equivalence relationship between the values of the two controls u, P that bring about exclusion when acting on their own is read off from (11) as:

$$u = S(P, t)/S(0, t) = u(P, t) \tag{12}$$

i.e. u will depend on P and (in general) t. This 'indirect' control, u, will be varying (but not necessarily optimal) if there is periodic environmental forcing on the residents or the 'direct' control P is varied or both. The level of control P that gives exclusion can be found by integrating (11a) and imposing the condition $\xi_{ave} = 0$, i.e. by solving the integral equation: $\int \beta(t)S(P, t)dt = pd_0$. This can be done by iterating on P using bisection or the secant method (Lenhart and Workman, 2007). Control u then follows from (12). This indirect control u has an important role to play in giving insight into the epidemiological impact of the direct control P. This impact can be measured by the (average) exclusion effort $E_1 = p^{-1}\int u^{-1}dt$ applicable whether or not variable control u is optimal. (Subscript '1' in E_1 indicates variable control.)

Conversely, one can reverse the calculation to find the direct control P that corresponds to an indirect control u. For example, it would be of interest to find the control P that corresponds to the least effort control (10a). However, in relating direct

and indirect controls, controllability issues can arise. There may not be a direct control that corresponds to least epidemiological impact.

As an example of exclusion under direct control consider again the SIS model (6) with culling control P applied to the resident population and modeled in (6a) by the addition of the term (-cPS) where c measures culling 'effectiveness'. For definiteness suppose that both the birth rate a and transmission rate β are environmentally forced with β lagging a by phase ϕ_1 while (direct) control P is varied (sinusoidally) with lag ϕ_2 relative to the birth rate. Lag ϕ_1 is pre-determined by the environment while lag ϕ_2 is a control variable. For a given choice of ϕ_2 we can find the value of P_0 (the average value of P) that brings about exclusion by solving (11a). Of particular interest is the value of ϕ_2 that gives the lowest P_0 for a given ϕ_1. How P_0 depends on these lags is shown in the 'direct control' threshold of Fig 4a based on the parameters specified in Table B1 (set 2). The epidemiological consequences of these control choices are shown in Fig 4b where the (indirect) exclusion effort E_1 (for the equivalent second stage control u) is plotted against the two lags to form the 'indirect control' threshold.

Of special interest is the strong peak in the indirect effort E_1 (Fig 4b) when lag ϕ_1 is close in value to the internal lag ψ generated by the external forcing (i.e. when $\phi_1 \sim \pi/2$) and when a and P are close to being out of phase (i.e. $\phi_2 \sim \pi$). Approaching this peak the proportionate reduction in the newly infected population increases sharply. At the peak the population is, in effect, being reduced by two thirds compared to one half when $\phi_2 \sim 0$ (and $\phi_1 \sim \pi/2$). However, interference between control and environmental variations means that this increase in E_1 is achieved with just a modest increase (11%) in the direct control P_0 away from its point of least value when $\phi_2 \sim 0$ (Fig 4a). Intuitively we would want to reduce u (i.e. survival of the newly infected) by as much as possible. This is discouraged by the high effort in doing this by indirect control. But this 'resonance' effect allows this increased effort to be substantially reduced by using a direct control.

This resonant (phase) amplification is the result of applying maximum forcing ($\delta_1 = 1$). It does not occur for medium strength forcing where QA would be valid. For example for forcing at half strength ($\delta_1 = 0.5$) the percentage changes in direct and indirect efforts that would lower the survival rates at peak position are comparable. The amplification emerges as exploration of the model at higher strengths is carried out using numerical methods. Sinusoidal forcing at maximum strength is an important case to consider since it can be taken as an approximate model for the many seasonal activities, such as breeding, hibernation or harvesting, that only occur for a restricted part of the 'year'. The ability to explore this part of control space is therefore essential.

Note also that there is a range of values of lag ϕ_1 where the controls are inactive ($1.15\pi \le \phi_1 \le 1.58\pi$) indicating that the system is capable on its own of keeping the pathogen excluded. This inactive interval could reflect, for example, a latitudinal gradient or a time gradient created by climate change. Travelling along such a gradient would eventually require switching the control back on again.

4. Exclusion with multiple infected states

In this section we go back to equation (3) and follow the development of the control theory approach taken in section 3 but with two or more infected states and hence a Jacobian for the invasion equation (3) of size 2×2 or greater. Since (3) can no longer be solved explicitly we will discuss two alternative approaches: A review of how to

construct the monodromy matrix and a summary of how to develop and customise the quadratic approximation method. For clarity of exposition we limit discussion of these methods to systems with 2 infected states since they are straightforward to generalise to higher dimensions. We discuss the properties of the special control u in (3) when constant, variable and optimal and its use in solving the exclusion problem for pre-emptive controls.

4.1 The monodromy matrix for 2 (or more) infected states and constant control u

The following three-step algorithm generates the monodromy matrix and hence the solution of the exclusion problem when there are two infected states (and at least 1 host type) and control u is constant (Hale 1969): [1] Solve equations (3) over one cycle (with period p) starting with initial condition $\underline{y}(0) = (1, 0)^T$. Let $\underline{y}(p) = (v_1, w_1)^T$ be the solution at the end of that cycle. (T indicates transpose.) [2] Repeat this calculation but with $\underline{y}(0) = (0, 1)^T$ and let $\underline{y}(p) = (v_2, w_2)^T$. The monodromy matrix $M = \begin{bmatrix} v_1 & v_2 \\ w_1 & w_2 \end{bmatrix}$. [3] Adjust control u so that the dominant eigenvalue of matrix M equals 1. This is the value of u that gives zero growth over a cycle and hence exclusion.

As an example consider a compartmental epidemiological model with two host species subject to density-dependent seasonally transmitted infection with no latency. It will have a Jacobian of the form:

$$J = \begin{bmatrix} uS_1\beta_{11} - d_1 & uS_1\beta_{12} \\ uS_2\beta_{21} & uS_2\beta_{22} - d_2 \end{bmatrix} = u\begin{bmatrix} S_1\beta_{11} & S_1\beta_{12} \\ S_2\beta_{21} & S_2\beta_{22} \end{bmatrix} - \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}. \qquad (13)$$

The constant control variable u defines the common survival factor for each cohort of newly infecteds. Parameters β_{ij} are the inter- and intra-transmission coefficients (from host j to host i) and d_1, d_2 the exit rates from the infectious states. There is forcing on the β_{ij} according to the time profiles:

$$\beta_{ij} = \beta_{ij0}(1 + \delta_i \cos(\omega t + \phi_{ij})) \qquad (i, j = 1, 2) \qquad (14)$$

where β_{ij0} are the average values, δ_i the forcing strengths and ϕ_{ij} the lags with $\phi_{1j} = 0$, $\phi_{2j} = \phi \geq 0$. This models seasonal variation in the susceptibility of the two hosts to the disease, with maximum susceptibility of the two hosts occurring at different times of the year if lag $\phi > 0$.

S_1, S_2 are the susceptible host populations determined by the resident dynamics decoupled in RIA. In the simplest case with no forcing and no direct interaction the two hosts will be at their carrying capacities (i.e. $S_1 = K_1$, $S_2 = K_2$). But one or both host populations can be below capacity if, for example, they are prey to an immune predator or one of them is in fact a predator as well as a host. When the residents are subject to forcing S_1 or S_2 or both will be time-dependent.

As a numerical example take the parameters given in Table B2 (set 1) with forcing at strength $\delta_1 = \delta_2 = 0.95$ and with variation in susceptibility of the two host species exactly out of phase (i.e. $\phi \sim \pi$ in (14)). For the residents suppose there is no forcing and no direct interaction and hence that $S_1 = K_1$, $S_2 = K_2$. Applying the algorithm above, the value of the control u leading to zero growth over a cycle turns out to be $u = 0.663$, in which case the monodromy matrix $M = \begin{bmatrix} 0.765 & 0.881 \\ 0.203 & 0.238 \end{bmatrix}$ has dominant eigenvalue 1. When there is no environmental variation ($\delta_1 = \delta_2 = 0$ in (14)) $u = 0.738$, so a further reduction in the newly infected population of 8% is

required due to the forcing. In terms of exclusion effort $E_0 = 1.509$ (with forcing) and $\bar{E}_0 = 1.355$ (without). In contrast, if the calculation is repeated for $\phi = 0$ (i.e. the host susceptibilities are exactly in phase) then forcing results in almost no change in the exclusion control level or effort. These results are shown in the plot of control u against lag ϕ in Fig 5a which shows that the imposition of environmental variation necessitates a further reduction in control u to achieve exclusion whatever the lag, i.e. the forced threshold in terms of u lies entirely below the unforced threshold. (Equivalently the plot of exclusion effort E_0 against lag ϕ lies entirely above the line of unforced effort $\bar{E}_0$.)

4.2 The Quadratic Approximation for 2 (or more) infected states

It is not always the case that more exclusion effort is needed when there is forcing. This can be deduced from the second method for finding the exclusion value of u. This method uses a quadratic approximation (QA) in the forcing strengths δ_1, δ_2 to generate an approximate formula for the growth rate (Greenman and Pasour 2012). The first step in setting up this procedure is to separate out the growth rate dynamics in order to study the zero growth condition more directly. For a two host model with infectious states I_1, I_2, and Jacobian $J = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ we can rewrite matrix equation (2) as the equations:

$$\frac{dz}{dt} = C + (D - A)z - Bz^2 \qquad (15a)$$

$$\xi = \frac{1}{X}\frac{dX}{dt} = A + Bz \qquad (15b)$$

where $X = I_1$, $z = I_2/I_1$, $\underline{y} = (I_1, I_2)^T$ and ξ is the growth rate. For J given by (13) with residents at their carrying capacities, $A = uK_1\beta_{11} - d_1$, $B = uK_1\beta_{12}$, $C = uK_2\beta_{21}$, $D = uK_2\beta_{22} - d_2$. In scenario (14) elements A, B, C, D will depend on t, δ_1, δ_2, ϕ and control u.

Equation (15a) is decoupled and to be solved first, with its solution then substituted in (15b) to determine ξ. Although (15a) is not linear it can be transformed into an infinite set of linear equations that can be exactly solved in sequence. This can be done by writing z as a series in powers of the forcing strengths $\delta \equiv \delta_1 = \delta_2$ supposed equal for simplicity, i.e. $z = z_0 + \delta z_1 + \delta^2 z_2 + \ldots$, and then collecting together terms of the same power in (15a). From (15b) we can then find the average growth rate (ξ_{ave}) by integration over a cycle and hence the level of the control u and effort E_0 needed to bring about exclusion when the zero growth condition is imposed.

The algebraic analysis is made easier if the simpler question is asked as to whether forcing increases or decreases the exclusion control value. This can be answered if the growth rate in (15b) is evaluated for control value $u^{(0)}$ that brings about exclusion when there is no forcing. Then the forced average growth rate ξ_{ave} is given by the following power series in δ:

ξ_{ave} = (unforced growth rate) + δ(linear term) + δ^2(quadratic term) + …. (16).

The first of these terms is zero as is the linear term (because the variations average to zero) while the quadratic term equals N/Δ where numerator N and denominator Δ are given by:

$$N = \rho_1\rho_2(\Gamma_1\cos(\phi) + \Gamma_2\sin(\phi) - \Gamma_0) \qquad (17a)$$

$$\Delta = 2(\rho_1 + \rho_2)[(\rho_1 + \rho_2)^2 + \omega^2]. \qquad (17b)$$

Here $\rho_1 = u^{(0)}K_1\beta_{110} - d_1$, $\rho_2 = u^{(0)}K_2\beta_{220} - d_2$, $\Gamma_1 = (\rho_1 + \rho_2)(d_1 + d_2) + 2d_1d_2$, $\Gamma_2 = (d_1 - d_2)\omega, \Gamma_0 = (d_1 - d_2)^2 + \Gamma_1$. (For details of the calculation see Appendix A.) On the unforced threshold, ρ_1 and ρ_2 are both negative since at exclusion neither host species can support the pathogen. So at $\phi = 0$ the quadratic term is positive (if $d_1 \neq d_2$) and hence control u needs to decrease further to achieve exclusion. Whether this is the case for all lag values depends on whether or not equation $N = 0$ has a solution for ϕ. For the parameters in Table B2 (set 1) there is no solution but for the parameters in Table B2 (set 2) there is. So for this second parameter set there are lag values ϕ where control u needs to be reduced further (ξ_{ave} positive) and other values where u has been reduced too much (ξ_{ave} negative) (Fig 5b).

The threshold distortions of Fig 5 produced by forcing on this multiple host model have a structure similar to those shown in Fig 1. In both cases the forced threshold is obtained from the unforced threshold by a shift (cf Γ_0 in (17a)) followed by a sinusoidal-like distortion. The lag positioning of the stationary points of the distortion is determined by the internal lag of the system (cf Γ_1/Γ_2 in (17a)). Whether exclusion effort is increased or decreased by the forcing depends on the relative magnitudes of the shift and distortion amplitude. It would not be surprising if these features were to be present in most models with sinusoidal forcing.

4.3 The optimal scheduling problem for multiple infected states

For invasion systems with multiple infected states such as (13) it is rarely possible to solve the optimal scheduling problem (9) exactly. Instead we introduce a new iterative algorithm based on the Hamiltonian method of solution (Lenhart and Workman, 2007) described in Appendix C. For (9) this means solving the following set of equations:

$$(a)\quad u^{-1} = s\sqrt{\underline{\lambda}(t).F.\underline{y}(t)} \qquad (b)\quad \frac{dy}{dt} = J.\underline{y} \qquad (c)\quad -\frac{d\underline{\lambda}}{dt} = \underline{\lambda}.J \tag{18}$$

subject to end point conditions on (b). Here $J = uF - V$, $\underline{\lambda} = (\lambda_1, \lambda_2)$ is the vector of adjoint (dual) variables and s is a 'free' parameter that plays a key role in the calculation. The algorithm iterates on the single control $u = u(t)$ which will be a function of t at each stage of the iteration except for the first where we take as the (default) starting value the exclusion level of u when assumed constant in time. We will still be able to use the monodromy matrix method for (numerically) solving equations (b), (c) even though u is now variable. At each stage of the iteration the algorithm proceeds as follows: [1] Find the monodromy matrix for system (18b) using the current approximation for u. Adjust s so that the dominant eigenvalue is 1. The variation $\underline{y}(t)$ over a cycle is then calculated from (18b) by taking as initial vector an eigenvector for this eigenvalue 1 of the monodromy matrix. [2] Find the variation $\underline{\lambda}(t)$ over a cycle in a similar fashion, but simplified by the knowledge that the 'adjoint' monodromy matrix for (18c) will necessarily have eigenvalue 1 for the previously calculated value of s. [3] Finally construct the next approximation for u from (18a) where s is left free and where $\underline{y}(t)$, $\underline{\lambda}(t)$ in (18a) are the solutions of (18b), (18c) just calculated.

The need for a special algorithm to solve this particular optimal control problem arises because the standard solution method with end point conditions on the state variables ($\underline{y}$) is highly unstable. This is understandable given that we are searching for the knife-edge threshold between exponential growth and exponential decay. To get around this instability we use the property that if one of the monodromy matrices for (18b), (18c) has eigenvalue 1 then so has the other matrix.

This follows from the fact that the inner product: $\underline{\lambda}.\underline{y}$ has the same value at the end of successive cycles, a result that can be proved from equations (18b), (18c). That parameter 's' is 'free' reflects the fact that the solutions of the linear equations (18b), (18c) with periodic boundary solutions are defined up to a multiplicative constant. The values of these constants are chosen so that the parameter 's' in (18a) renders the dominant eigenvalue of the monodromy matrix of (18b) equal to 1.

As an example of this algorithm consider again Jacobian (13) in scenario (14) with parameter values given in Table B2 (set 1). Our iterative algorithm finds optimal solutions for given choices of lag ϕ, starting the iteration with the constant solutions for those choices. Of particular interest are the lag values: $\phi = 0, \pi$ (see (14) and Fig 5a). The optimal control paths corresponding to these two values are shown in Fig 6. For curve 'a' ($\phi = 0$) we find that ($\bar{E}_0$, E_0, E^*) = (1.355, 1.356, 1.236) where $\bar{E}_0$; E_0; E^* is the constant unforced; constant forced; optimal forced exclusion effort. So effort can be reduced to well below the constant control level by varying the control. It therefore looks to be worthwhile using varying controls in this case. However for control curve 'b' ($\phi = \pi$) we find that ($\bar{E}_0$, E_0, E^*) = (1.355, 1.509, 1.506) so there is little advantage in varying the control with forcing present. This is reminiscent of the results obtained for model (6) in Fig 3 but the mechanism is very different, not least because of the very strong forcing. For $\phi \sim 0$ the two hosts go through the infection cycle in synchrony with little net infection transfer. In effect they act as single hosts with optimal control in phase with the newly infecteds, hence the similarity with path d in Fig 3, in particular the switching behaviour. At $\phi \sim \pi$, with one host on the up part of the infection cycle and the other on the down part, there is significant transfer to sustain the infection at a high level with little fluctuation. So the constant control solution is a good approximation but the required exclusion effort is much higher.

4.4 Pathogen exclusion using pre-emptive controls

As a generalisation of the discussion in section 3.4, consider a resident subsystem in RIA with pre-emptive vector control $\underline{P}$ and links to the invasion subsystem involving the vector of susceptible populations $\underline{S} = \underline{S}(\underline{P}, t)$. The invader growth rate can be found from the monodromy matrix and the exclusion levels of $\underline{P}$ from the zero growth condition. We can separate out the equivalent indirect controls for the second stage of the exclusion process by defining: $u_i = S_i(\underline{P}, t)/S_i(\underline{0}, t)$ where S_i denotes the host i susceptible population. In the multiple host case therefore we will be using a vector $\underline{u}$ of indirect controls, its components targeting specific hosts. This makes sense. It is likely to be highly inefficient to use the same level of control for hosts which have substantially different dynamical characteristics.

As an example consider the 2 host SISI model with resident equations in RIA given by:

$$dS_i/dt = r_i S_i - (r_i/K_i)S_i^2 - c_i P_i S_i \quad (i = 1, 2), \qquad (19)$$

where $\underline{P} = (P_1, P_2)^T$. Its Jacobian J_1 is obtained from (13) with $u = 1$ and is subject to scenario (14). If there is no forcing on the residents and the (culling) controls P_i are constant in time then (asymptotically) $S_i = K_i(1 - (c_i P_i)/r_i)$. The solution of the exclusion problem can now be obtained by inserting these values for S_i in J_1 and then solving (2) subject to the condition $\xi_{ave} = 0$. From the exclusion values for P_i one can find the indirect controls u_i as $u_i = S_i/K_i = (1 - (c_i P_i)/r_i)$. In terms of targeted indirect controls u_1, u_2 the Jacobian for the model now reads as:

$$J=\begin{bmatrix} u_1K_1\beta_{11}-d_1 & u_1K_1\beta_{12} \\ u_2K_2\beta_{21} & u_2K_2\beta_{22}-d_2 \end{bmatrix}=\begin{bmatrix} u_1K_1\beta_{11} & u_1K_1\beta_{12} \\ u_2K_2\beta_{21} & u_2K_2\beta_{22} \end{bmatrix}-\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}. \qquad (20)$$

Having assembled all the necessary pieces we can now explore the exclusion dynamics of the SISI model in both direct and indirect control spaces, using the parameter values to be found in TableB2 (set 3). The 'control map' of Fig 7 shows the pathogen exclusion threshold in (u_1, u_2) space for different values of lag ϕ (see (14)). (The unforced threshold (for clarity not shown) lies close to but above the threshold (a) for $\phi = 0$.) Q_0 is the point where the controls are both inactive and lies in the exclusion region for each threshold. On threshold (a) only u_2 is used at point Q_1; at Q_2, on the diagonal line $u_1 = u_2$, the control has the same action on both hosts; Q_3 is inaccessible because it would require infinite exclusion effort to reach $u_1 = 0$. Similarly for points on the other thresholds, namely (b), (c) for $\phi = \pi/2, \pi$. Fig 7 illustrates the fact that for this model, with constant controls and forcing only on infection transmission, exclusion effort increases whatever the lag.

Intuitively it makes sense to focus control effort on the reservoirs of infection, in this case at point Q_1 with $u_1 = 0$ and $u_2 \neq 0$ (i.e. host 2 is a reservoir of infection but not host 1). However a closer look at Fig 7 suggests that this might not always be the best strategy. At Q_1 it would be necessary to reduce the newly infected by over 60% (for threshold (a)) which might not be achievable. However at point Q_4 where there is removal of 15% of the host 1 newly infected, the required percentage reduction in host 2 newly infected to achieve exclusion is reduced to about 35%. For comparison, at Q_2 both controls u_1, u_2 are equal to 0.75, i.e. there is exclusion when the newly infected of both hosts are reduced by 25%. The strategy of control only on the reservoir of infection is cast in even more doubt if lag $\phi = \pi$ for then about 75% of host 2 newly infected have to be removed (threshold (c)). The reason for the high level of newly infected reduction required at Q_1 is the fact that host 1 is not far away from also being a reservoir of infection. In fact if (β_{110}) is increased to 1.1 (with both hosts then being reservoirs of infection) the threshold is shifted to position (d) in Fig 7 with both controls u_1, u_2 having to be used since the shifted end points are inaccessible, both requiring an infinite amount of effort to reach. Finally, note that the thresholds corresponding to (a), (b), (c) for direct controls P_i can be constructed from Fig 7 using the relations: $c_iP_i = (r_i/c_i)(1 - u_i)$.

These observations are relevant to pathogen mediated competition between livestock and wildlife, e.g. bovine TB in badgers. The badger (host 2) is undoubtedly a reservoir of infection in the UK and therefore has to be controlled for exclusion to happen. The situation with livestock is less clear, but it is not far from being a reservoir of infection if it is not already (Cox et al., 2005).

An alternative exclusion strategy for the example of Fig 7 (with $\beta_{110} = 0.93$) is to control only the (single) reservoir of infection (i.e with $u_1 = 1$, $P_1 = 0$) but now with a time varying control P_2 to see whether exclusion effort can be reduced. In Fig 8 we show the contour map of the threshold surface in which the exclusion value of P_{20} (the average of variable P_2) is plotted against lag $\phi_1 = \phi$ (the lag between hosts) and lag ϕ_2 between P_2 and host 1. The map shows a saddle point S with 'mountains' M on either side and 'valleys' V in between. It is along the valley (V – S – V) that one looks for least P_{20} for given ϕ_1. Comparison of the data for Figs 7 and 8 shows that one can always find a value of ϕ_2 where P_{20} (Fig 8) is lower than constant P_2 (Fig 7) whatever ϕ_1 when $P_1 = 0$. So varying P_2 can reduce (average) direct effort, often substantially. We can continue the analysis by finding out how close is the optimal

indirect control u_2* to the indirect control u_2 generated by P_2 when its average value is at a minimum (Fig 8). Conversely we can find the shape of the direct control P_2 equivalent to u_2* since the sinusoidal shape we have chosen may not be the most appropriate.

This example illustrates the importance of working in indirect as well as direct control space. The indirect space, defined by the set of universal controls (u_i), deals with epidemiological issues, such as reservoirs of infection. Since the structures of the pathogen threshold in direct and indirect spaces are closely related knowledge about one of these structures provides information about the other. As a result we can compare in greater depth possible exclusion strategies against both epidemiological and economic criteria.

5. Discussion

Our primary objective has been to solve the pathogen exclusion problem for eco-epidemiological model systems with a high level of complexity. This includes systems with multiple (infected) host types and subject to multi-component (periodic) environmental forcing and a general set of direct (pre-emptive) controls. The natural way to approach this problem is as a control problem since it involves intervention in the system with a set of control measures to achieve a given objective. Our adoption of an explicit holistic control framework opens up access to an extensive body of knowledge (from a range of scientific disciplines) based on essential concepts such as 'effort', 'optimisation' and 'controllability' and it allows a seamless development of the model system when new features such as constraints on controls or system parameters are to be added. In short, this flexible framework provides an ideal 'laboratory' for the exploration of exclusion dynamics and the design of suitable control strategies that take into account all relevant factors.

Examples in the literature (e.g. Choisy and Rohani, 2006; Packer et al., 2005; Omori and Adams, 2011) suggest that the complexity created by a system's internal structure or by a varying environment (or both) opens up a range of new opportunities for bringing about exclusion, for example by exploiting interference between fluctuating controls and system components or by encouraging the system to 'take over' some of the effort required in exclusion. What are the specific mechanisms driving such behaviour and what other complex interactions exist that might be of advantage in excluding the pathogen? To be able to answer such questions it is essential to have efficient numerical and analytic solution methods that provide insight into the dynamics of exclusion.

The obvious way of solving the exclusion problem is to use exhaustive simulation, by searching over a grid to find out when there is invasion and when there is not. But this is highly inefficient. A major gain in efficiency is achieved by using the Rare Invader Approximation to create a two-stage process, first solving the disease-free resident equations and then using the monodromy matrix to find the asymptotic invader growth rate and hence the direct controls that will bring about exclusion. One can be more ambitious by using a polynomial approximation in the forcing strengths to generate approximate formulae for the equations of the pathogen thresholds and hence test for the generality of the results obtained by numerical means within and between different families of models.

An important additional advantage of the two-stage structure of our numerical and analytic procedures is that we can use it to explore in more detail the epidemiological aspects of exclusion. Precisely we can construct special (indirect) controls targeting the different types of the newly infecteds in the invasion equations

which have the same effect on the invaders as the original direct (pre-emptive) controls. Taken together the direct and indirect controls enable epidemiological considerations as well as the practical, technical and economic issues associated with implementing the direct controls to be taken into account when designing an exclusion strategy.

These special controls, that act on the newly infected, also allow us to connect and compare our control theory approach with the conventional approach based on the basic reproduction number R_0. Precisely, for both forced and unforced systems, if the special control is the same on all types of newly infecteds and is constant in time the exclusion effort is equal in value to R_0. In this restricted sense the control theory approach is consistent with conventional theory. The properties of R_0 can be translated into control terms and conversely. However the control theory approach has distinct advantages in forced systems. The transition to forced systems is seamless, an operationally meaningful link with second generation infecteds is retained and the theory is readily generalised by relaxing the constraints that the impact of the special control is the same for all infected host types and is constant in time. With variability in time we can find the least effort special control as a solution of the optimal scheduling problem. Finally, to bring us full circle, the overarching control framework allows us to link the usually unrealisable special (indirect) controls with the (direct) controls that are realisable (Heffernan et al., 2005).

The examples of pathogen thresholds distorted under external forcing have shown suggest that there are three basic 'building blocks' involved: A shift due to changes in average population levels; a sinusoidal-like distortion due to covariance between forcing and population variations; a repositioning of this distortion due to internal lags in the system in response to the forcing. The relative 'sizes' of these effects, together with the strength of the forcing, determine the overall impact of the forcing on exclusion effort. Individually these effects suggest various ways of reducing the effort. For example one may take advantage of interference in the covariance effect when the lag is a control variable; offset reinforcement by using a variable control; use the shift effect when there is no seasonality by also varying the control.

These basic exclusion strategies can be hidden in the complexity of the systems responses to the forcing and they may also generate more subtle ways to exclude the pathogen. Of interest is 'persuading' the system to 'take over' some of the exclusion effort while the control remains inactive (Figs 3, 4, 6). Switching a control on and off is familiar in harvesting (Bairagi et al., 2008), in game-bird management (Choisy and Rohani, 2006) and is also relevant to pathogen exclusion. We have also looked at 'extreme' forcing which can lead to period resonance (Greenman and Pasour, 2012) and phase resonance (Fig 4) (Zambrano et al., 2008).

There is much more work to be carried out in studying pathogen and other types of exclusion when there is external forcing and internal structure. For example, we are currently investigating a model for dengue in which control measures aimed at reducing exposure to the mosquito population are targeted at different age groups in the host population. Further development and application of the theory and techniques we have introduced in this paper could lead to practical improvements in epidemic control across a range of human and wildlife ecological systems.

Acknowledgements
The University of Stirling, through the good offices of Leslie Smith, is to be thanked for providing the funding, direct and indirect, that made this research possible.

Appendix

A. The Quadratic Approximation method

Formula (8): To solve the resident equation (6a) in RIA substitute the QA for S, namely: $S \approx K_0 + \delta_2 x_1 + (\delta_2)^2 x_2$ using the binomial approximation $(1 + \delta_2 a)^{-1} \approx 1 - \delta_2 a + (\delta_2)^2 a^2$ in QA to invert K. Collect terms in powers of δ_2 to obtain linear equations of the form:

$dx_1/dt = \theta x_1 + C_1\cos(\omega t)$; $dx_2/dt = \theta x_2 + C_2 x_1 \cos(\omega t) + C_3(1 + \cos(2\omega t)) + C_4 x_1^2$

where θ, C_i are constants with values $\theta = -r$, $C_1 = rK_0$, $C_2 = 2r$, $C_3 = -rK_0/2$, $C_4 = -r/K_0$. Solve these linear equations in sequence to obtain S asymptotically as:

$$S \approx K_0 + \delta_2 C_5 \cos(\omega t - \psi) + (\delta_2)^2 C_6 + (\delta_2)^2(\text{cosine terms}) \qquad \text{(A1)}$$

where $C_5 = (C_1/\sqrt{(\theta^2 + \omega^2)})$, $C_6 = -C_3\omega^2/(\theta(\theta^2 + \omega^2))$, $\psi = \tan^{-1}(-\omega/\theta)$, $\omega = 2\pi/p$. From (5) with $\xi_{ave} = 0$, exclusion effort $E_0 = (\int \beta S dt)/(pd_0)$. Evaluate E_0 by substituting S from (A1) and β from (7) to obtain (8). (Note that the last term in (A1) averages out to zero on integration or generates terms of third or higher order.)

Parameter	a	b	K	β	α	γ	p	δ_1	δ_2	δ_3	c
Set 1	1.3	0.5	4.5	0.4	0.5	0.0	5	0.5	0.6		
Set 2	1.3	0.7	4.5	0.4	0.5	0.0	5	1.0	1.0	1.0	1.0

Formula 17: In this example we use the QA method to solve linear invasion equations. To solve (15a) we use the QA: $z \approx z_0 + \delta z_1 + \delta^2 z_2$ and decompose the elements of J into unforced and forced parts by writing:

$$A = A_0 + \delta A_1;\ B = B_0 + \delta B_1;\ C = C_0 + \delta C_1,\ D = D_0 + \delta D_1 \qquad \text{(A2)}$$

If there is no forcing ($\delta = 0$) then $z = z_0$ with $z_0 = -A_0/B_0 = -C_0/D_0$ obtained from (15a) and $\det J = 0$ when $u = u^{(0)}$, the unforced exclusion value of u. With forcing, substitute z and (A2) in (15a) and collect together powers of δ to obtain:

$dz_1/dt = \theta z_1 + C_1\cos(\omega t) + C_2 \cos(\omega t + \phi)$

$dz_2/dt = \theta z_2 + C_3 z_1 \cos(\omega t) + C_4 z_1 \cos(\omega t + \phi) + C_5 (z_1)^2$

where $\theta = A_0 + D_0$ (when $u = u^{(0)}$) and C_i are constants (different from those for formula (8)). Solve these linear equations in sequence. Substitute for z in (15b) and integrate over a cycle. Symbolic manipulation software helps in carrying out these tasks.

B. Parameter values for models (6), (13)

Table B1: Parameter values for the SIS model (6)

(Note: periodic forcing can always be made seasonal by appropriately scaling the unit of time.)

Table B2: Parameter values for the SISI model defined by (13) and (19).

Parameter	K_1	K_2	d_1	d_2	β_{110}	β_{220}	β_{120}	β_{210}	δ	p
Set 1	1.5	1.0	1.50	1.35	0.93	1.50	0.47	0.30	0.95	6
Set 2	1.5	1.0	1.50	0.50	0.93	1.50	0.47	0.30	0.75	5

Set 3 is formed from set 1 by adding parameter values: $r_2 = 0.8$, $c_2 = 0.4$.

C. Optimal control problems

Consider the general optimal control problem: Minimise $\int f(\underline{y}, \underline{u}, t)dt$ subject to $d\underline{y}/dt = \underline{g}(\underline{y}, \underline{u}, t)$ and $\underline{y}(t_1) = \underline{a}$, $\underline{y}(t_2) = \underline{b}$. The integral in the objective function is taken over the control interval: $t_1 \leq t \leq t_2$ and $\underline{u}$ denotes the controls. To solve this problem construct the Hamiltonian H:

$$H = f(\underline{y}, \underline{u}, t) + \underline{\lambda}.\underline{g}(\underline{y}, \underline{u}, t)$$

with $\underline{\lambda}$ the vector of adjoint (dual) variables. A necessary condition for optimality is:

$$\{\text{(i) } \partial H/\partial \underline{u} = \underline{0} \quad \text{(ii) } d\underline{y}/dt = \partial H/\partial \underline{\lambda} \quad \text{(iii) } -d\underline{\lambda}/dt = \partial H/\partial \underline{y}\} \qquad \text{(C1).}$$

For problem (9) $t_2 = t_1 + p$, $f(\underline{y}, u, t) = p^{-1}u^{-1}$, $\underline{g}(\underline{y}, u, t) = (uF - V).\underline{y}$ so: $H = p^{-1}u^{-1} + \underline{\lambda}.(uF - V).\underline{y}$ and:

$$\{\text{(i) } u^{-1} = \sqrt{(\underline{\lambda}.F.\underline{y})} \text{ (ii) } d\underline{y}/dt = (uF - V).\underline{y} \text{ (iii) } -d\underline{\lambda}/dt = \underline{\lambda}.(uF - V)\} \qquad \text{(C2)}$$

with $p = 1$ without loss of generality. The exact single host solution (10) for model (6) can be obtained from equations (C2), with the vectors now scalars. From (C2ii) and (C2iii) $\lambda y = \lambda(0)y(0) = B$ (a constant). From (C2i) $u^{-1} = \sqrt{\lambda y F)} = \sqrt{(BF)}$ so $uF = \sqrt{F}/\sqrt{B}$. Solving (C2ii) gives $\ln(y(p)/y(0)) = (\int \sqrt{(F)}dt)/\sqrt{(B)} - \int V\mathrm{d}t = 0$ using the end conditions on (C2ii). This gives an equation for $\sqrt{(B)}$ and

$$u^{-1}(t) = A\sqrt{(F(t))} \text{ where } A = (\int \sqrt{(F(u))}\mathrm{d}u)/\int V(u)\mathrm{d}u) \qquad \text{(C3).}$$

To handle boundary conditions such as $0 < u(t) \leq 1$ see for example Bryson and Ho (1975) or Lenhart and Workman (2007).

References

Anderson, R. M., May, R. M., 1981. The population dynamics of microparasites and their invertebrate hosts. Phil. Trans. R. Soc. Lond. B 291, 451-524.

Anderson, R. M., May, R.M., 1986. The invasion, persistence and spread of infectious diseases within animal and plant communities. Phil. Trans. R. Soc. Lond. B 314: 533-573.

Bacaer, N., Guernaoui, S., 2006. The epidemic threshold of vector-borne diseases with seasonality. J. Math. Biol. 53, 421-436.

Bacaer, N., Aits Dads, E. H., 2011. Genealogy with seasonality, the basic reproduction number and the influenza pandemic, J. Math. Biol. 62, 741-762.

Bairagi, N., Chaudhuri, S., Chattopadhyay, J., 2009. Harvesting as a disease control measure in an eco-epidemiological system – A theoretical study. Mathematical Biosciences 217, 134-144.

Brassil, C. E., 2006. Can environmental variation generate positive indirect effects in a model of shared predation? Am. Nat. 167, 43-54.

Bryson, A. E., Ho, Y-C., 1975. Applied Optimal Control. New York, Halstead Press.

Chesson, P., 1982. The stabilising effect of a random environment. J. Math. Biol. 15, 1-36.

Choisy, M., Rohani, P., 2006. Harvesting can increase severity of wildlife disease epidemics. Proc. R. Soc. B 273, 2025-2034.

Cox, D. R., Donnelly, C. A., Bourne, F. J., Gettinby, G., McInerney, J. P., Morrison, W. I., Woodroffe, R., 2005. Simple model for tuberculosis in cattle and badgers. PNAS 102, 17588-17593.

Diekmann, O., Heesterbeek, J. A. P., Metz, J. A. J., 1990. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. J. Math. Biol. 28, 365-382.

Diekmann, O., Heesterbeek, J. A. P., 2000. Mathematical Epidemiology of Infectious Diseases. John Wiley & Son, Ltd., Chichester, UK.

Diekmann, O., Heesterbeek, J. A. P., Roberts, M. G., 2010. The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface 7, 873-885.

Dietz, K., 1976. The incidence of infectious diseases under the influence of seasonal fluctuations. Lect. Notes Biomath. 11, 1-15.

Grassly, N. C., Fraser, C., 2006. Seasonal infectious disease epidemiology. Proc. R. Soc. B 273, 2541-2550.

Greenman, J. V., Pasour, V. B., 2012. Threshold dynamics for periodically forced ecological systems: The control of population invasion and exclusion. J. Theor. Biol. 295, 154-167.

Hale, J. K., 1969. Ordinary Differential Equations. Wiley-Interscience, New York.

Hartemink, N. A., Randolph, S. E., Davis, S. A., Heesterbeek, J. A. P., 2008. The basic reproduction number for complex disease systems: defining R_0 for tick-borne infections. Am. Nat. 171: 743-754.

Heesterbeek, J. A. P., Roberts, M. G., 1995. Threshold quantities for helminth infections. J. Math. Biol. 33, 415-434.

Heffernan, J. M., Smith, R. J., Wahl, L. M., 2005. Perspectives on the basic reproduction ratio. J. R. Soc. Interface 2, 281-293.

Hosseini, P. R., Dhondt, A. A., Dobson, A., 2004. Seasonality and wildlife disease: how seasonal birth, aggregation and variation in immunity affect dynamics of *Mycoplasma gallisepticum* in house finches. Proc. R. Soc. B 271, 2569-2577.

Koelle, K., Rodo, X., Pascual, M., Yunus, Md., Mostafa, G., 2005. Refractory periods and climate forcing in cholera dynamics. Nature 436, 696-700.

Lenhart, S., Workman, J. T., 2007. Optimal Control applied to biological problems. London: Chapman and Hall (CRC).

Lintott, R.A., Norman, R. A., Hoyle, A. S., 2013. The impact of increased dispersal in response to disease control in patchy environments. Journal of Theoretical Biology 323, 57-68.

McCallum, H., N. Barlow, and J. Hone. 2001. How should pathogen transmission be modelled? Trends in Ecology and Evolution 16, 295-300.

Omori, R., Adams, B., 2011. Disrupting seasonality to control disease outbreaks: The case of koi herpes virus. J. Theor. Biol. 271, 159-165.

Packer, C., Holt, R.D, Hudson, P.J., Lafferty, K.D., and Dobson, A.P. 2003. Keeping the Herds healthy and alert: implications of predator control for infectious disease. Ecology Letters. 6: 797-802.

Rinaldi, S., Muratori, S., Kuznetsov, Y., 1993. Multiple attractors, catastrophes and chaos in seasonally perturbed predator-prey communities. Bull. Math. Biol. 55, 15-35.

Tompkins, D. M., White, A. R., Boots, M., 2003. Ecological replacement of native red squirrels by invasive greys driven by disease. Ecol. Lett. 6, 189-196.

Wang, W., Zhao, X-Q., 2008. Threshold Dynamics for Compartmental Epidemic Models in Periodic Environments. J. Dyn. Diff. Equat. 20, 699-717.

Williams, B. G., Dye, C., 1997. Infectious disease persistence when transmission varies seasonally. Math. Biosci. 145, 77- 88.

Wichmann, M. C., Johst, K., Moloney, K. A., Wissel, C., Jeltsch, F., 2003. Extinction risk in periodically fluctuating environments. Ecolo. Modell. 167, 221-231.

Zambrano, S., Seoane, J. M., Marino, I. P., Sanjuan, M. A. F., Euzzor, S., Meucci, R., Arecchi, F. T., 2008. Phase control of excitable systems. New J. Phys. 10, 073030.

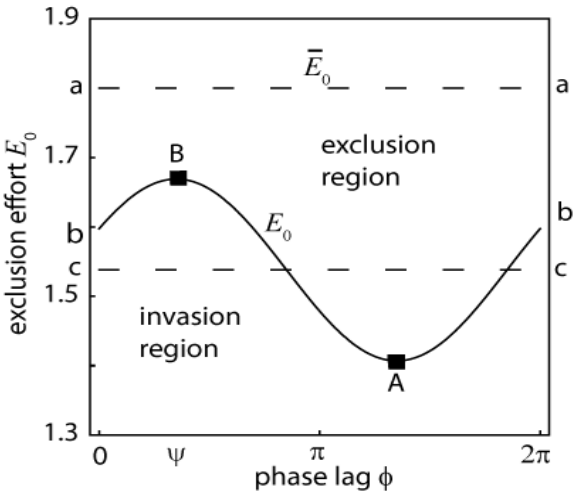

Figure 1: Pathogen exclusion thresholds as a function of phase lag ϕ for SIS model (6) with forcing on β, K (7). The threshold is (bb) when the system is forced (exclusion effort E_0) and (aa) when unforced ($\bar{E}_0$). Forcing strengths (δ_1, δ_2) have values: (aa) (0, 0); (bb) (0.5, 0.6); (cc) (0, 0.6). Other parameter values are given in Table B1 (set 1) in Appendix B.

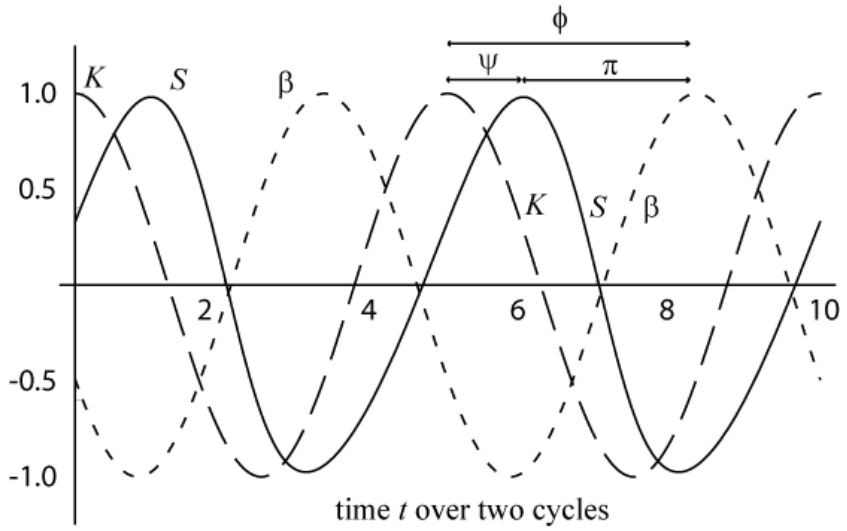

Figure 2: Time series showing the relative positioning of β, S, K for ϕ at point A in Fig 1 (S and β 'out of phase'). For clarity, averages have been made zero and amplitudes made equal. (Multiply ϕ by $p/2\pi$ to get the lag when the period is p.)

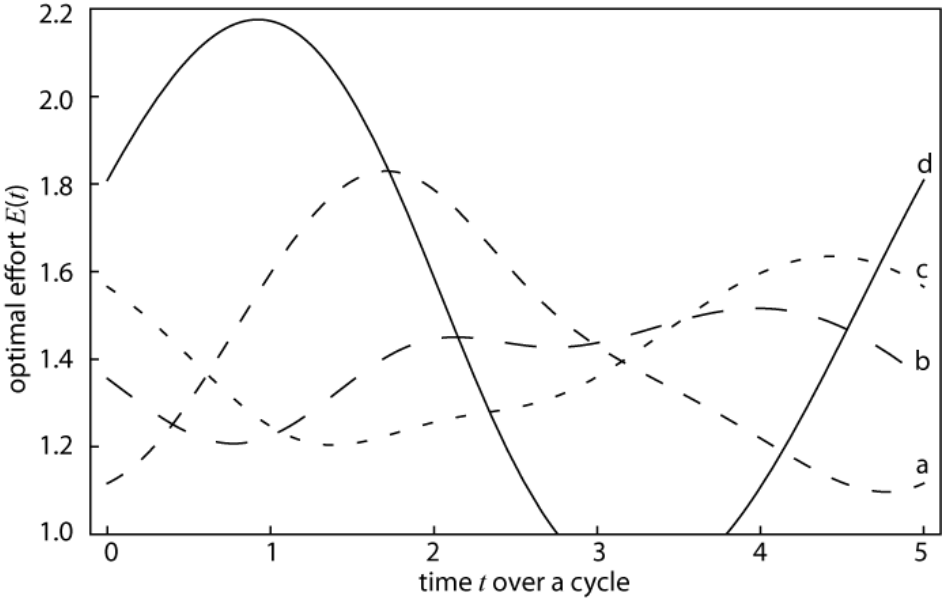

Figure 3: Variation of optimal effort $E(t)$ over a single cycle for SIS model (6) in scenario (7). Paths: (a, b, c, d) correspond to lag ϕ equal to (π, 1.34π (point A in Fig 1), $3\pi/2$, 0.33π (point B in Fig 1)). Path d exhibits off-on optimal control switching, Parameter values are given in Table B1 (set 1).

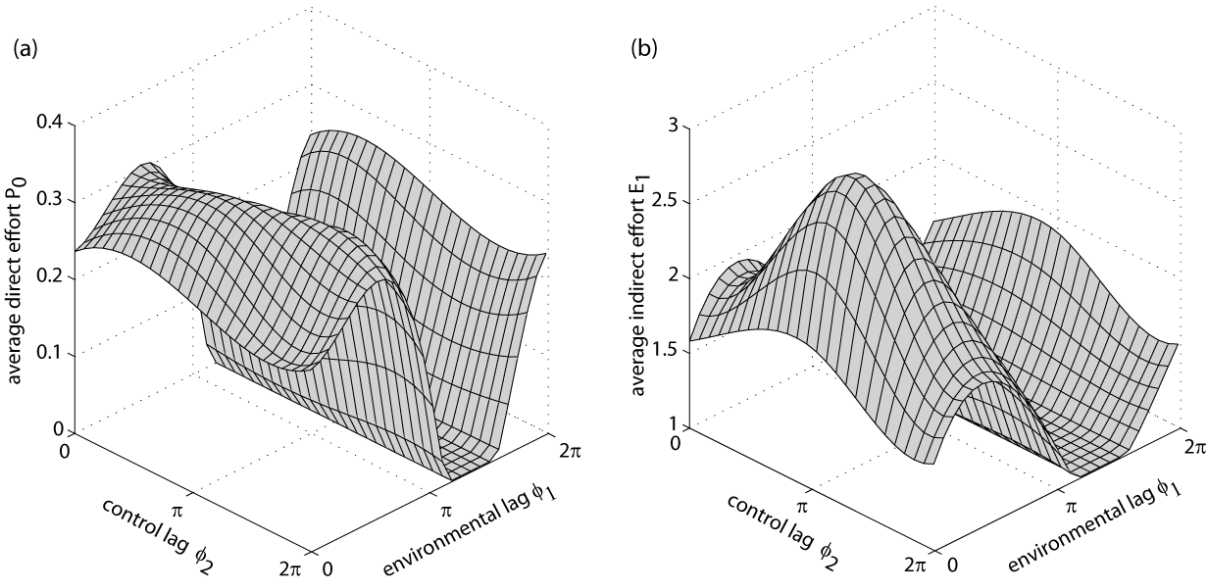

Figure 4: The SIS model with periodic forcing on birth rate and infection transmission and with varying culling control P, to illustrate the 'amplification effect'. Figures show the threshold surfaces in terms of control and seasonal lags ϕ_2, ϕ_1, in (a) for direct effort P_0 and in (b) for indirect effort E_1.

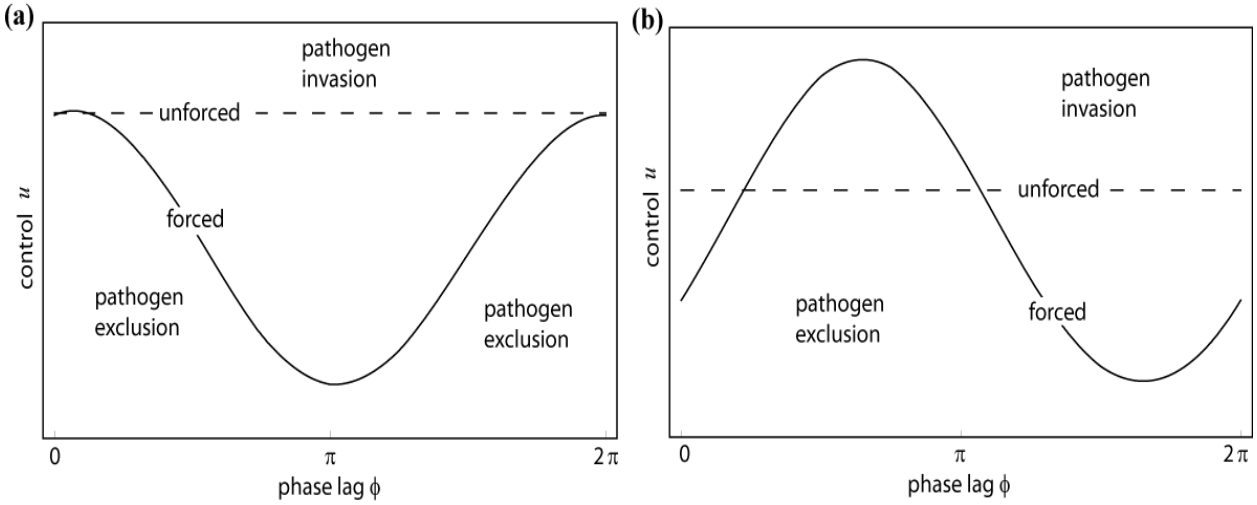

Figure 5: The level of control u required to achieve pathogen exclusion in the SISI model with Jacobian (13) is shown as a function of lag ϕ (see (14)). Parameters given in Table B2 (set 1) for (a) and Table B2 (set 2) for (b).

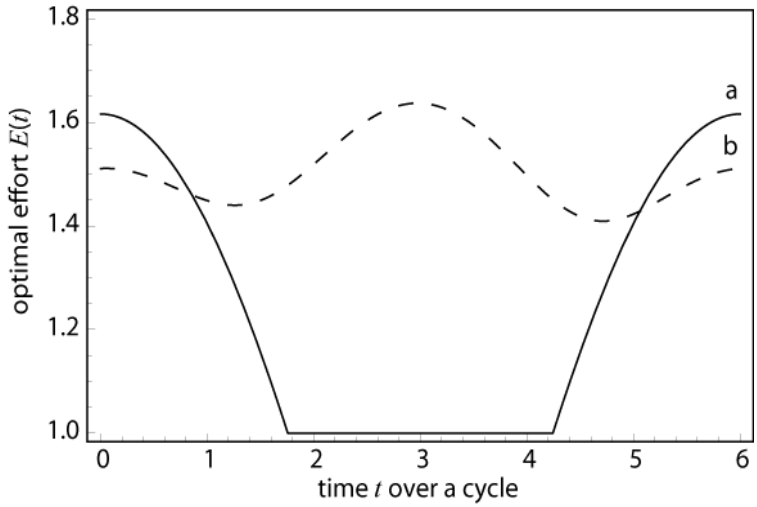

Figure 6: Variation of optimal control $E(t)$ over a single cycle for SISI model with Jacobian (13). Paths: (a, b) correspond to lags ϕ: (0, π) in Fig 5a. The control is switched off and on again on path a. Parameter values are given in Table B2 (set 1).

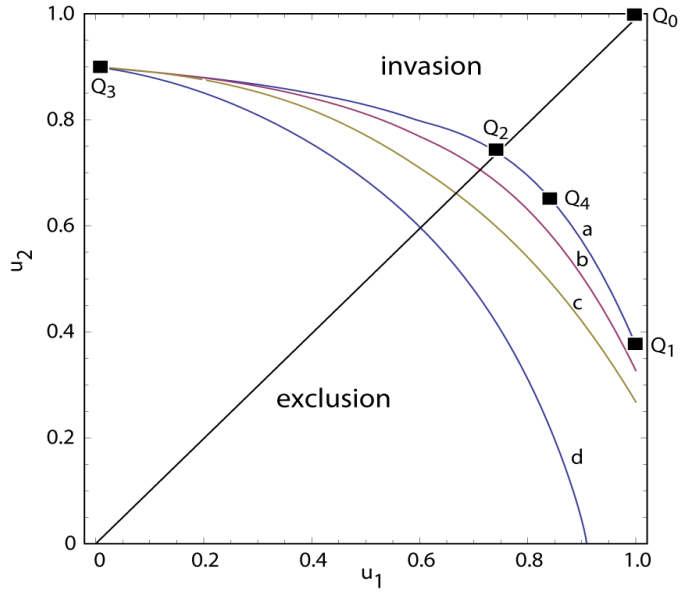

Figure 7: SISI model with constant direct (culling) controls on the residents. Shown are the thresholds in indirect control space (u_1, u_2). Threshold *a*; *b*; *c* corresponds to ϕ value 0; $\pi/2$; π. Host 2 is the only reservoir of infection. For threshold *d* ($\beta_{110} = 1.1$, $\phi = \pi$) both hosts are reservoirs. Parameters given in Table B2 (set 3).

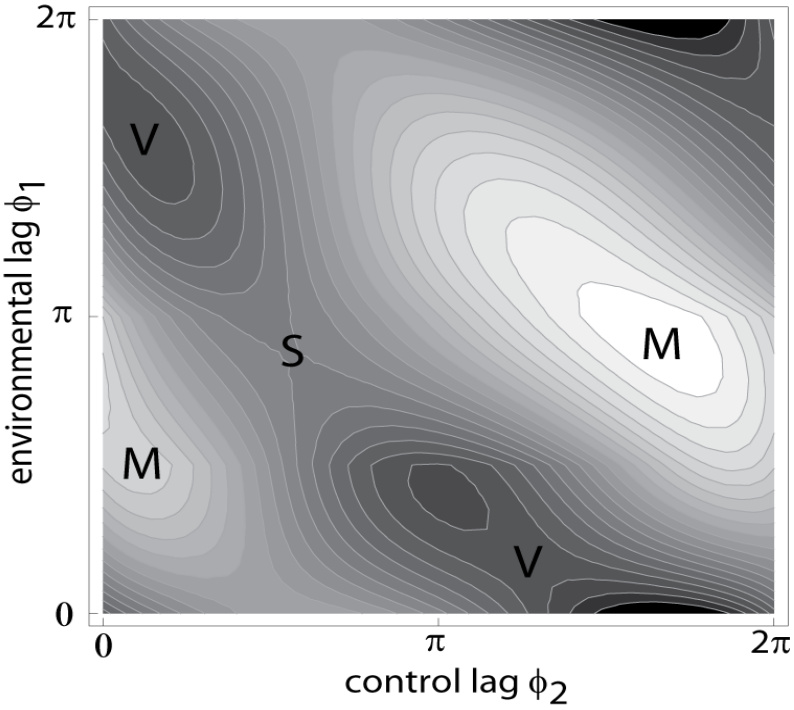

Figure 8: Contour map of the direct control threshold surface when there is one reservoir and one direct now-varying control P_2. The map relates the average value P_{20} of P_2 against control and seasonal lags ϕ_2, ϕ_1. Comparison with Fig 7 indicates when there is advantage in varying P_2.