

A new method for estimating the effort required to control an infectious disease

M. G. Roberts^{1*} and J. A. P. Heesterbeek²

¹*Institute of Information and Mathematical Sciences, Massey University, Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand*

²*Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3584 CL Utrecht, The Netherlands*

We propose a new threshold quantity for the analysis of the epidemiology of infectious diseases. The quantity is similar in concept to the familiar basic reproduction ratio, R_0 , but it singles out particular host types instead of providing a criterion that is uniform for all host types. Using this methodology we are able to identify the long-term effects of disease-control strategies for particular subgroups of the population, to estimate the level of control necessary when targeting control effort at a subset of host types, and to identify host types that constitute a reservoir of infection. These insights cannot be obtained by using R_0 alone.

Keywords: basic reproduction ratio; infectious diseases; epidemiology; reservoirs of infection; disease control

1. INTRODUCTION

The most pervasive and useful concept in the mathematical epidemiology of infectious diseases is the basic reproduction ratio, commonly written R_0 (for recent examples see Keeling & Gilligan 2000; Ferguson *et al.* 2001; Gandon *et al.* 2001; Gani & Leach 2001; Lloyd & May 2001; Dye *et al.* 2002; Kao 2002). It is defined as the expected number of secondary cases that would arise from a typical primary case in a susceptible population. If $R_0 < 1$, an infection cannot persist in a population, hence R_0 is a useful indicator of the effort required to eliminate an infection (Anderson & May 1991; Diekmann & Heesterbeek 2000; Heesterbeek 2002). For example, an infection will be eliminated over time if a proportion of the population greater than $1 - 1/R_0$ is permanently protected from becoming infected, by being effectively vaccinated at birth or by some other means. This criterion is based on the assumption that the population is homogeneous and well mixed, but may be extended to populations with a more complex structure. If we identify n distinct host types, we construct an $n \times n$ next-generation matrix (K) (Diekmann *et al.* 1990). The word *generation* here refers to the infection not the host, and different *host types* may be physiologically identical, but might have been infected in different ways that influence the course of the infection, might be in different locations, or might have some other characteristic that differentiates them epidemiologically. The elements of K are similar in concept to R_0 ; for example, K_{ij} is the expected number of secondary cases in host type i that would arise from a typical primary case in host type j in a susceptible population. The value of R_0 is calculated as the spectral radius (dominant eigenvalue) of K (Diekmann *et al.* 1990; Diekmann & Heesterbeek 2000), written $R_0 = \rho(K)$. Hence, to calculate the control effort required to eliminate an infection, we construct the next-generation matrix for the population with control in

operation, K_c , as a function of the control effort, c , and then determine the control effort that ensures $\rho(K_c) < 1$. In the literature on epidemic modelling, a real-time approach is often taken instead of one based on successive infection generations. Starting from the system of differential equations near the infection-free steady state, a threshold quantity is derived from the Jacobian matrix determining the steady state's stability. Although a quantity derived in this way will have the same threshold behaviour as the dominant eigenvalue of the next-generation matrix, it does not have the same biological interpretation and can therefore not be called the basic reproduction ratio or denoted by R_0 . This distinction is important if statements based on the value of R_0 are made about the relative ease of controlling different infections, or when gauging the control effort needed to eliminate an infection.

The concept of R_0 is less useful when the control effort is to be targeted at particular host types, or if the infection cycle includes another type such as a vector, intermediate host or reservoir host. For example, R_0 is well defined for malaria via the next-generation matrix, but its calculation results in a uniform value over the human and mosquito populations. Taking humans and mosquitoes as host types 1 and 2, respectively, the next-generation matrix would then have $K_{11} = K_{22} = 0$, as humans cannot infect humans and mosquitoes cannot infect mosquitoes, K_{12} as the expected number of humans that a single infected mosquito infects, and K_{21} as the expected number of mosquitoes infected by a single infected human, both in a susceptible population. Hence $R_0 = (K_{21}K_{12})^{1/2}$. However, the expected number of secondarily infected humans that result from a single infected human is R_0^2 , as two generations are required to transmit an infection from human to human, the first being from human to mosquito and the second being from mosquito to human. If a vaccine were available that could be administered to humans at birth, a proportion greater than $1 - 1/R_0^2$ should be vaccinated to eliminate the infection; hence, application of the original threshold of $1 - 1/R_0$ would lead to an underesti-

* Author for correspondence (m.g.roberts@massey.ac.nz).

mate of the control effort required. Similar expressions are obtained for other systems where the life cycle is obligatory, but a more detailed analysis is required when the life cycle is not obligatory. Examples of the latter in host-vector systems include cases where there could be vertical transmission in the vector, as with dengue disease (Rosen *et al.* 1983), or horizontal transmission in the human, as with Chagas' disease (Prata 2001). In the first example, it could take any number of infection generations greater than one for an initial infection in humans to result in a secondary human infection, because the infection can *cycle* within the vector population. In the second example, secondary human cases caused by horizontal transmission may occur after a single infection generation. The situation may be even more complicated for infections where there are a variety of host types, for example bovine tuberculosis in cattle with a wildlife reservoir (Jackson 2002). The introduction of an infected bovine could initiate an endemic infection in the wildlife population, which could then continue to infect cattle indefinitely. Hence, a single infected bovine could result in a large number of secondary infections of cattle, all of which are the result of transmission within the wildlife reservoir. Although reducing R_0 below one will lead to elimination of the infection in all these cases, the magnitude of R_0 does not indicate the control effort required to achieve this. By concentrating on one particular host type it is possible, however, to derive a threshold quantity that does convey this information.

We present a method to estimate the control effort required to eliminate an infectious disease when control is applied to a specific subpopulation of hosts, but taking into account the fact that the infection will pass through other subpopulations (of the same species or another species, in the same or in another geographical area) before causing secondary cases in the subpopulation of interest. The formula for the threshold quantity and its motivation are presented in § 2; we then provide examples of its application to vector-transmitted infections and infectious diseases of wild animals. Appendix A contains formal mathematical derivations of the properties of interest.

2. THE THRESHOLD QUANTITY

Consider an infection that has multiple host types, and concentrate without loss of generality on host type 1. With the entire population susceptible, introduce a single infected host of type 1, and apply the matrix K to determine the expected numbers of infected hosts of all types in the next infection generation. The first element of the vector Ke , where e is the unit vector with first element equal to one and other elements equal to zero, gives the number of new cases of type 1 expected in the next generation. We are, however, interested not only in the number of infected hosts of type 1 in this generation, which might well be zero if we are dealing with a pure host-vector system, but also in the future infections of host type 1 that will result from the other infected host types. Because we wish to calculate the infections of type 1 arising from a single type 1 infected individual, we reset the number of infected hosts of type 1 to zero and apply K again to the infected hosts of types 2 to n . In each infection generation we count and 'remove' the infections of type 1, then apply

K again to the numbers of infected hosts of types 2 to n . The cumulative number of infected hosts of type 1 that result in this process, as a result of chains of infection that link one or more of host types 2 to n without another infected host of type 1 being allowed to *reproduce*, is denoted by T_1 . Note that $(Ke)_1$ is the result of this process after only one infection generation, hence $T_1 \geq (Ke)_1$. The explicit formula for T_1 is

$$T_1 = e'K(I - (I - P)K)^{-1}e, \tag{2.1}$$

where I is the $n \times n$ identity matrix, P is the projection matrix, defined by $P_{11} = 1$, $P_{ij} = 0$ when $i \neq 1$ or $j \neq 1$, and a prime signifies that the vector e is transposed.

We observe the following properties:

- (i) $T_1 > 1$ if and only if $R_0 > 1$;
- (ii) an infection will be eliminated over time if a proportion of the population of hosts of type 1 greater than $1 - 1/T_1$ is permanently vaccinated at birth; and
- (iii) a reservoir host other than type 1 exists if $\rho((I - P)K) \geq 1$.

Equation (2.1) and the properties ((i)–(iii)) are derived in Appendix A. The result can be generalized further, to consider the transmission of infection from a subset of ℓ out of a total of n host types, where we can take the ℓ types to be type 1, ..., ℓ , without loss of generality. The threshold quantity T_ℓ is then the spectral radius of an $\ell \times \ell$ matrix defined by

$$M_\ell = E_\ell K(I - (I - P_\ell)K)^{-1}E_\ell, \tag{2.2}$$

where E_ℓ and P_ℓ are $n \times \ell$ and $n \times n$ projection matrices, respectively, defined by $(E_\ell)_{ii} = (P_\ell)_{ii} = 1$ for $i = 1 \dots \ell$, $(E_\ell)_{ij} = (P_\ell)_{ij} = 0$ otherwise. Equation (2.2) is also derived in Appendix A, and provides the basis for determining the correct control effort needed to eradicate an infection while targeting control at a subset of ℓ host types.

3. VECTOR-TRANSMITTED INFECTIONS

Although the malaria example described in § 1 illustrates the contrast between R_0 and T_ℓ , some of the confusion that has arisen in the literature and the practical implications, the point may still seem rather academic. Here, we present two examples of vector-transmitted infections where it is not obligatory to follow the life cycle, and demonstrate the use of the threshold quantity T_ℓ in these cases.

(a) Chagas' disease

The protozoan parasite *Trypanosoma cruzi*, which is the cause of Chagas' disease, is transmitted by haematophagous arthropods, principally Triatominae (Marsden 1998), but can also be transmitted horizontally between humans by blood transfusion (Velasco-Hernandez 1994; Prata 2001). We may wish to determine first whether the infection can be maintained by vector transmission alone under a particular control procedure, and, if not, what proportion of blood transfusions should be screened to eradicate the infection. Let host types 1 and 2 be humans infected by blood transfusion and by the arthropod,

respectively, and host type 3 be the arthropod. The next-generation matrix has the form

$$\begin{pmatrix} K_{11} & K_{12} & 0 \\ 0 & 0 & K_{23} \\ K_{31} & K_{32} & 0 \end{pmatrix}, \tag{3.1}$$

and $R_0 = \rho(K)$ is a solution of the characteristic equation of K , which is a cubic. The argument leading to equation (2.1) implies that if $\rho((I - P)K) = (K_{23}K_{32})^{1/2} \geq 1$ then the infection can persist through host-vector transmission alone. If a control intervention were to reduce host-vector transmission so that $\rho((I - P)K) < 1$, then this mode of transmission would no longer be sustainable in isolation, but it would still be possible that $R_0 \geq 1$ (equivalently $T_1 \geq 1$) owing to the additional human-to-human transmission via blood transfusion. Although R_0 cannot conveniently be expressed explicitly, T_1 can. From equation (2.1)

$$T_1 = K_{11} + \frac{K_{12}K_{23}K_{31}}{1 - K_{23}K_{32}}, \tag{3.2}$$

and the infection will be eliminated if a proportion greater than $1 - 1/T_1$ of horizontally transmitted cases can be prevented, for example by screening blood transfusions for infection.

(b) Dengue disease

Dengue disease is caused by the four serotypes of an arbovirus transmitted by the mosquitoes *Aedes aegypti* and *Ae. albopictus*. Whereas the former is primarily an urban mosquito, the latter is more abundant in rural areas and also exhibits vertical transmission of the virus to a much greater extent (Rosen *et al.* 1983; Esteva & Vargas 1998). If a vaccine were available for dengue disease, we may wish to determine the relative merits of vaccinating the rural or urban populations, or whether one should vaccinate the entire human population. Let host types 1 and 2 be urban and rural humans, respectively, host type 3 be *Ae. aegypti* and host types 4 and 5 be *Ae. albopictus* infected from biting humans and from vertical transmission, respectively. For illustration we consider only one virus serotype and neglect vertical transmission in *Ae. aegypti*. The next-generation matrix K has the form

$$\begin{pmatrix} 0 & 0 & K_{13} & K_{14} & K_{15} \\ 0 & 0 & K_{23} & K_{24} & K_{25} \\ K_{31} & K_{32} & 0 & 0 & 0 \\ K_{41} & K_{42} & 0 & 0 & 0 \\ 0 & 0 & 0 & K_{54} & K_{55} \end{pmatrix}. \tag{3.3}$$

To calculate $R_0 = \rho(K)$, it is necessary to find the largest root of the fifth-order characteristic polynomial of K , and an analytic expression can not be conveniently obtained. To evaluate the probable effects of a vaccination policy, appropriate values for the components of K must be substituted in equation (3.3) and the eigenvalues evaluated numerically. Alternatively, by applying equation (2.2) with $\ell = 2$ we find that M_2 , the matrix that expresses transmission from human to human, is

$$M_2 = \begin{pmatrix} K_{13}K_{31} + \left(K_{14} + \frac{K_{15}K_{54}}{1 - K_{55}}\right)K_{41} & K_{13}K_{32} + \left(K_{14} + \frac{K_{15}K_{54}}{1 - K_{55}}\right)K_{42} \\ K_{23}K_{31} + \left(K_{24} + \frac{K_{25}K_{54}}{1 - K_{55}}\right)K_{41} & K_{23}K_{32} + \left(K_{24} + \frac{K_{25}K_{54}}{1 - K_{55}}\right)K_{42} \end{pmatrix}. \tag{3.4}$$

The components of M_2 are similar in concept to the components of K , for example $(M_2)_{11}$ is the expected number of secondarily infected urban humans that would arise from a single infected urban human in a fully susceptible population. The matrix M_2 has spectral radius

$$T_2 = \frac{(M_2)_{11} + (M_2)_{22} + \sqrt{((M_2)_{11} - (M_2)_{22})^2 + 4(M_2)_{21}(M_2)_{12}}}{2}. \tag{3.5}$$

Equation (3.5) provides an analytic criterion for determining critical vaccination levels when treating both urban and rural populations. The formula to determine the proportion of urban humans alone that should be vaccinated to eradicate the disease is obtained by applying equation (2.1) to either K or M_2 :

$$T_1 = (M_2)_{11} + \frac{(M_2)_{21}(M_2)_{12}}{1 - (M_2)_{22}} = K_{13}K_{31} + \left(K_{14} + \frac{K_{15}K_{54}}{1 - K_{55}}\right)K_{41} + \frac{\left(K_{23}K_{31} + \left(K_{24} + \frac{K_{25}K_{54}}{1 - K_{55}}\right)K_{41}\right)\left(K_{13}K_{32} + \left(K_{14} + \frac{K_{15}K_{54}}{1 - K_{55}}\right)K_{42}\right)}{1 - \left(K_{23}K_{32} + \left(K_{24} + \frac{K_{25}K_{54}}{1 - K_{55}}\right)K_{42}\right)}. \tag{3.6}$$

To eradicate the infection, a proportion greater than $1 - 1/T_1$ of urban humans must be maintained in an immune state.

The reason for the term $1/(1 - K_{55})$ appearing in equations (3.4) and (3.6) merits a comment. By definition K_{55} is the proportion of *Ae. albopictus* mosquitoes that transmit the infection vertically. An infected mosquito is expected to infect K_{25} rural humans (for example) in one infection generation and produce K_{55} vertically infected mosquitoes who infect $K_{55}K_{25}$ humans in the next infection generation, $K_{55}^2K_{25}$ in the next and so on. Taking the sum to infinity yields $K_{25}/(1 - K_{55})$.

4. INFECTIOUS DISEASES OF WILD ANIMALS

The analysis presented in this paper is not restricted to infectious diseases with proscribed, or even non-obligatory, life cycles. The methodology can be used to analyse the epidemiology of complex situations to determine which parts of the system are essential to maintain transmission, and to determine control strategies where only a subset of host types is targeted. We present two further examples, one where the spatial structure of a wild-animal population means that only animals in a restricted area can be targeted, while infection dynamics have to be considered on a metapopulation; and the other the more general problem of defining host types that make up reservoirs of infection.

(a) Metapopulation infection dynamics

The brush-tailed possum (*Trichosurus vulpecula*) is a reservoir for bovine tuberculosis in New Zealand (Cowan 1990). The dynamics of the infection may be described in a metapopulation residing on a number of habitat

patches, with transmission between patches being maintained by the migration, as they mature into adults, of juvenile possums that had previously been exposed to infection (Fulford *et al.* 2002). Hence, in a metapopulation of m patches we define host types $2i - 1$ and $2i$ to be the juvenile and adult possums on patch i , respectively, for $i = 1 \dots m$. The next-generation matrix K is of dimension $2m \times 2m$, and may be regarded as being constructed from m^2 2×2 matrices. The submatrices in off-diagonal positions have zeros in their right-hand columns, as possums exposed as adults do not migrate to other patches, and hence cannot transmit infection directly to juveniles or adults on other patches. For example, in a system of just two patches we would have

$$K = \begin{pmatrix} K_{11} & K_{12} & K_{13} & 0 \\ K_{21} & K_{22} & K_{23} & 0 \\ K_{31} & 0 & K_{33} & K_{34} \\ K_{41} & 0 & K_{43} & K_{44} \end{pmatrix}. \tag{4.1}$$

Suppose that patch 2 is an area of forest in which the possum population is being regularly culled, so that $\rho((I - P_2)K) < 1$ and the infection cannot maintain itself in the forest alone. Suppose also that the plan is to vaccinate possums against bovine tuberculosis on patch 1, which is an area adjacent to farmland. We target juvenile and adult possums on patch 1, i.e. host types 1 and 2, and therefore calculate $T_2 = \rho(M_2)$ where

$$M_2 = \begin{pmatrix} K_{11} + \frac{K_{13}K_{31}(1 - K_{44}) + K_{13}K_{34}K_{41}}{(1 - K_{33})(1 - K_{44}) - K_{34}K_{43}} & K_{12} \\ K_{21} + \frac{K_{23}K_{31}(1 - K_{44}) + K_{23}K_{34}K_{41}}{(1 - K_{33})(1 - K_{44}) - K_{34}K_{43}} & K_{22} \end{pmatrix}. \tag{4.2}$$

If equal proportions of juveniles and adults can be maintained in an immune state on patch 1, then that proportion must exceed $1 - 1/T_2$ in order to eradicate bovine tuberculosis from the metapopulation. If the proportions are unequal, they may be calculated from the explicit relationship presented as equation (3.5), using values for the components of M_2 from equation (4.2). The infection may be eradicated from the metapopulation by introducing control measures so that $T_2 < 1$, which in turn guarantees that $R_0 < 1$.

(b) Reservoirs of infection

The identification of reservoirs of infection is an important topic in the epidemiology of infectious diseases, especially for zoonoses and diseases of farmed animals (Haydon *et al.* 2002). For example, the possum is considered to be a reservoir of bovine tuberculosis for cattle, and if we had needed to quantify the transmission to cattle, we would have had to include additional host types in the analysis presented in § 4a. With ℓ host types that are cattle, and $2m$ host types that are possums, comprising juveniles and adults on m patches of habitat, we would have a next-generation matrix K of size $(\ell + 2m) \times (\ell + 2m)$. In the terminology of Haydon *et al.* (2002), the cattle would then become the target species and the possums the maintenance community, provided $\rho((I - P_\ell)K) \geq 1$.

In general, consider an infection that is transmitted between n host types. Identify ℓ host types as target hosts and label these 1 to ℓ . Any subset ω of the remaining host types is a reservoir of infection for the target hosts if $\rho(P_\omega K) \geq 1$, where $(P_\omega)_{ii} = 1$ if $i \in \omega$ and $(P_\omega)_{ij} = 0$ otherwise. The matrix $P_\omega K$ has the rows of K corresponding to the host types in ω , and zeros elsewhere. When ω contains k host types, the eigenvalues of $P_\omega K$ are zero ($n - k$ times) and the eigenvalues of the $k \times k$ matrix whose elements are the entries in K that correspond to members of the putative reservoir of infection. As an example, consider the analysis based on dengue disease presented in § 3b. To ascertain whether the rural life cycle is a reservoir of infection for urban dengue, we would determine whether the matrix

$$P_\omega K = \begin{pmatrix} 0 & K_{24} & K_{25} \\ K_{42} & 0 & 0 \\ 0 & K_{54} & K_{55} \end{pmatrix} \tag{4.3}$$

had an eigenvalue greater than or equal to one.

5. DISCUSSION

The basic reproduction ratio (R_0) remains the most useful concept in the mathematical epidemiology of infectious diseases. Its threshold property provides a criterion for policies to eradicate an infection from a population, and its magnitude provides an indication of the control effort required to achieve eradication. When R_0 is used for the latter purpose some extra care is required in its definition and calculation. The quantity T_1 has the same threshold property as R_0 , but focuses attention on one particular host type rather than averaging over all types. It therefore provides a direct measure of the control effort required to achieve eradication when that effort is directed at only one host type, and in many of the examples that we have presented it provides an explicit formula where the calculation of R_0 would require the numerical determination of a matrix eigenvalue.

The models for infectious diseases with life cycles illustrate this point convincingly, and it is here that the literature is inconsistent. A formula equivalent to T_1 was used as R_0 for models of malaria (Anderson & May 1991; Dye 1994), canine leishmaniasis (Dye 1994), tick-borne infections (Randolph *et al.* 2002) and African horse sickness (Lord *et al.* 1996). By contrast, $R_0 = \rho(K)$ has been used in relation to models for dengue disease (Feng & Velasco-Hernandez 1997; Soewono & Supriatna 2001) as has $R_0 = \rho(K^2)$ with a statement that the basic reproduction number was equal to $\sqrt[3]{(R_0)}$ (Esteva & Vargas 1998, 2000). By adopting the distinction between R_0 and T_1 it is to be hoped that this confusion may be removed.

The concepts of T_1 and its multidimensional analogue T_ℓ have application in the epidemiology of infectious diseases with wild-animal reservoirs. In fact, the analysis provides a rigorous definition of a reservoir and a method for determining which subset of hosts constitutes it. A reservoir may consist of one or many host types, or the hosts of a particular type in one location or group, or other possible combinations. The key characteristic of a reservoir of infection, as defined by the criterion $\rho((I - P_\ell)K) \geq 1$, is

that if an infectious host of one of the target types were introduced to the population then the host types that make up the reservoir would subsequently become infected; the infection would then be maintained indefinitely in the reservoir hosts and continue to be transmitted back to the target hosts. Hence, a single primary infection of a target host would result in an indefinite number of secondary cases in the target population owing to the perpetuation of transmission between the reservoir hosts.

The methodology presented in this paper provides a framework for the analysis of the epidemiology of infectious diseases where there are a number of host types. It could be applied to many other infectious agents, for example HIV with transmission within and between different risk categories of hosts, and foot and mouth disease with short-range transmission between neighbouring premises and longer-range transmission between locations. By calculating the quantity T_1 under different control scenarios, the long-term effects of control interventions may be estimated. In particular, if a proportion v of transmission to the target host can be eliminated, then $v > 1 - 1/T_1$ will ensure that the infection is eradicated.

This work was commenced while the authors collaborated on a workshop on epidemic models at the Institute of Technology Bandung, Indonesia, funded by KNAW, the Royal Dutch Academy of Arts and Sciences. At the time M.G.R. worked for AgResearch Ltd.

APPENDIX A

(a) Derivation of equation (2.1)

Starting with one infected host of type 1, the second generation of infected hosts consists of an expected $(Ke)_i$ hosts of type i . We trace the transmission paths that have not yet yielded a type 1 infection, that is those that arise from the infections represented by the vector $(I - P)Ke$. At the third generation we have a vector $K(I - P)Ke$, which represents an expected $PK(I - P)Ke$ infected hosts of type 1 and $(I - P)K(I - P)Ke$ infected hosts of types $2 \dots n$. In general, at the $(j + 1)$ th infection generation there are expected to be $e'K((I - P)K)^{j-1}e$ infected hosts of type 1 that arise as a result of transmission through the life cycle without involving a host of type 1 in an intermediate infection generation (note that the prime denotes transpose, and $e'P = e'$). Hence the number of secondarily infected hosts of type 1 that arise from a typical primary infected host of type 1 is $T_1 = e'Me$ where

$$M = K \sum_{j=0}^{\infty} ((I - P)K)^j. \tag{A 1}$$

If $\rho((I - P)K) < 1$ then the summation in equation (A 1) converges, leading to the desired result.

(b) Derivation of equation (2.2)

A matrix

$$M = P_\ell K \sum_{j=0}^{\infty} ((I - P_\ell)K)^j \tag{A 2}$$

is constructed in a manner similar to that used to derive equation (A 1), but we now start with any vector e_ℓ for which $(e_\ell)_i = 0$ for $i > \ell$, $(e_\ell)_i \geq 0$ for $i \leq \ell$ and $(e_\ell)_i \neq 0$ for some $i \leq \ell$. When $\rho((I - P_\ell)K) < 1$ equation (A 2) leads to

$$M = P_\ell K (I - (I - P_\ell)K)^{-1}. \tag{A 3}$$

The vector e_ℓ represents a system state with infected hosts of types $1 \dots \ell$ only; the vector Me_ℓ represents the expected number of second-generation infected hosts of types $1 \dots \ell$ that arise from these, and so on. Hence the sequence $M^j e_\ell$ converges to zero as $j \rightarrow \infty$ whenever $\rho(M) < 1$. The matrix M has non-negative entries in rows 1 to ℓ and zero entries elsewhere, and hence has rank ℓ . The matrix $M_\ell = E'_\ell M E_\ell$ is the leading $\ell \times \ell$ submatrix of M (note $E'_\ell P_\ell = E'_\ell$), hence M_ℓ and M have the same eigenvalues and $T_\ell = \rho(M) = \rho(M_\ell)$.

(c) Derivation of property (i)

We prove this for the general case presented as equation (2.1), as the result for T_1 follows trivially from that for T_ℓ . Assume $\rho((I - P_\ell)K) < 1$. Rearranging equation (A 3), $M(I - K) = (I - M)P_\ell K$. The matrix K has a positive eigenvector w corresponding to the eigenvalue $R_0 = \rho(K)$, hence $(1 - R_0)Mw = R_0(I - M)P_\ell w$. All elements Mw and $P_\ell w$ are non-negative, hence $1 - R_0$ and $1 - \rho(M)$ are either zero or have the same sign (Minc 1988), and the proof is complete.

(d) Derivation of property (ii)

The row vector $e'K$ has elements $(e'K)_j = K_{1j}$ for $j = 1 \dots n$, whereas $(I - P)K$ does not contain any terms K_{1j} . Recalling that K_{1j} is the expected number of infections of host type 1 arising from an infection of host type j , if these elements of K are linear in S_1 , the steady-state number of susceptibles of host type 1, then T_1 is linear in S_1 . If a proportion v of hosts of type 1 are vaccinated at birth, then T_1 becomes $(1 - v)T_1$, and to eliminate the infection we require $v > 1 - 1/T_1$.

(e) Derivation of property (iii)

If $\rho((I - P)K) \geq 1$ then clearly $R_0 \geq 1$, but also the projection of the next-generation matrix K onto a subset of host types not including type 1 has spectral radius greater than or equal to one, implying that one of these other host types is a reservoir host.

REFERENCES

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*. Oxford University Press.
 Cowan, P. E. 1990 Brushtail possum. In *The handbook of New Zealand mammals* (ed. C. M. King), pp. 68–98. Oxford University Press.
 Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Chichester, UK: Wiley.
 Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.
 Dye, C. 1994 Vector control. In *Parasitic and infectious diseases: epidemiology and ecology* (ed. M. E. Scott & G. Smith), pp. 159–174. San Diego, CA: Academic.
 Dye, C., Williams, B. G., Espinal, M. A. & Raviglione, M. C. 2002 Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science* **295**, 2042–2046.

- Esteva, L. & Vargas, C. 1998 Analysis of a dengue disease transmission model. *Math. Biosci.* **150**, 131–151.
- Esteva, L. & Vargas, C. 2000 Influence of vertical and mechanical transmission on the dynamics of dengue disease. *Math. Biosci.* **167**, 51–64.
- Feng, Z. & Velasco-Hernandez, J. X. 1997 Competitive exclusion in a vector–host model for the dengue fever. *J. Math. Biol.* **35**, 523–544.
- Ferguson, N., Donnelly, C. A. & Anderson, R. M. 2001 The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science* **292**, 1155–1160.
- Fulford, G. R., Roberts, M. G. & Heesterbeek, J. A. P. 2002 The metapopulation dynamics of an infectious disease: tuberculosis in possums. *Theor. Popul. Biol.* **61**, 15–29.
- Gandon, S., Mackinnon, M. J., Nee, S. & Read, A. F. 2001 Imperfect vaccines and the evolution of pathogen virulence. *Nature* **414**, 751–756.
- Gani, R. & Leach, S. 2001 Transmission potential of smallpox in contemporary populations. *Nature* **414**, 748–751.
- Haydon, D. T., Cleaveland, S., Taylor, L. H. & Laurenson, K. M. 2002 Identifying reservoirs of infection: a conceptual and practical challenge. *Emerging Infect. Dis.* **12**, 1468–1473.
- Heesterbeek, J. A. P. 2002 A brief history of R_0 and a recipe for its calculation. *Acta Biotheoretica* **50**, 189–204.
- Jackson, R. 2002 The role of wildlife in *Mycobacterium bovis* infection of livestock in New Zealand. *N. Z. Vet. J.* **50**(Suppl.), 49–52.
- Kao, R. R. 2002 The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. *Trends Microbiol.* **10**, 279–286.
- Keeling, M. J. & Gilligan, C. A. 2000 Metapopulation dynamics of bubonic plague. *Nature* **407**, 903–906.
- Lloyd, A. L. & May, R. M. 2001 How viruses spread among computers and people. *Science* **292**, 1316–1317.
- Lord, C. C., Woolhouse, M. E. J., Heesterbeek, J. A. P. & Mellor, P. S. 1996 Vector-borne diseases and the basic reproduction number: a case study of African horse sickness. *Med. Vet. Entomol.* **10**, 19–28.
- Marsden, P. D. 1998 American trypanosomiasis. In *Zoonoses: biology, clinical practice, and public health control* (ed. S. R. Palmer, E. J. L. Soulsby & D. I. H. Simpson), pp. 513–525. Oxford University Press.
- Minc, H. 1988 *Non-negative matrices*. New York: Wiley.
- Prata, A. 2001 Clinical and epidemiological aspects of Chagas disease. *Lancet Infect. Dis.* **1**, 92–100.
- Randolph, S. E. (and 11 others) 2002 The ecology of tick-borne infections in wildlife reservoirs. In *The ecology of wildlife diseases* (ed. P. Hudson, A. Rizzoli, B. T. Grenfell, J. A. P. Heesterbeek & A. P. Dobson), pp. 119–138. Oxford University Press.
- Rosen, L., Shroyer, D. A., Tesh, R. B., Freier, J. E. & Lien, J. C. 1983 Transovarial transmission of dengue viruses by mosquitoes: *Aedes albopictus* and *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* **32**, 1108–1119.
- Soewono, E. & Supriatna, A. K. 2001 A two-dimensional model for the transmission of dengue fever disease. *Bull. Malaysian Math. Sci. Soc.* **24**, 49–57.
- Velasco-Hernandez, J. X. 1994 A model for Chagas disease involving transmission by vectors and blood transfusion. *Theor. Popul. Biol.* **46**, 1–31.