JOURNAL OF THE ROYAL SOCIETY Interface

Perspectives on the basic reproductive ratio

J.M Heffernan, R.J Smith and L.M Wahl

J. R. Soc. Interface 2005 **2**, doi: 10.1098/rsif.2005.0042, published 22 September 2005

JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2005) **2**, 281–293
doi:10.1098/rsif.2005.0042
Published online 7 June 2005

REVIEW

Perspectives on the basic reproductive ratio

J. M. Heffernan[1], R. J. Smith[2] and L. M. Wahl[1,†]

[1]*The Department of Applied Mathematics, The University of Western Ontario, London, Ontario N6A 5B7, Canada*
[2]*Department of Mathematics and College of Veterinary Medicine, The University of Illinois at Urbana-Champaign, Urbana, IL 61802, USA*

The basic reproductive ratio, R_0, is defined as the expected number of secondary infections arising from a single individual during his or her entire infectious period, in a population of susceptibles. This concept is fundamental to the study of epidemiology and within-host pathogen dynamics. Most importantly, R_0 often serves as a threshold parameter that predicts whether an infection will spread. Related parameters which share this threshold behaviour, however, may or may not give the true value of R_0. In this paper we give a brief overview of common methods of formulating R_0 and surrogate threshold parameters from deterministic, non-structured models. We also review common means of estimating R_0 from epidemiological data. Finally, we survey the recent use of R_0 in assessing emerging diseases, such as severe acute respiratory syndrome and avian influenza, a number of recent livestock diseases, and vector-borne diseases malaria, dengue and West Nile virus.

Keywords: R_0; epidemiology; population dynamics; mathematical modelling

1. INTRODUCTION

The basic reproductive ratio, R_0, is a key concept in epidemiology, and is inarguably 'one of the foremost and most valuable ideas that mathematical thinking has brought to epidemic theory' (Heesterbeek & Dietz 1996). Originally developed for the study of demographics (Böckh 1886; Sharp & Lotka 1911; Dublin & Lotka 1925; Kuczynski 1928), it was independently studied for vector-borne diseases such as malaria (Ross 1911; MacDonald 1952) and directly transmitted human infections (Kermack & McKendrick 1927; Dietz 1975; Hethcote 1975). It is now widely used in the study of infectious disease, and more recently, in models of in-host population dynamics. Two excellent surveys of the tangled history of R_0 can be found in Dietz (1993) and Heesterbeek (2002). An excellent overview of the demographic history can be found in Smith & Keyfitz (1977).

As a general definition, R_0 is the expected number of secondary individuals produced by an individual in its lifetime. The interpretation of 'secondary', however, depends on context. In demographics and ecology, R_0 is taken to mean the lifetime reproductive success of a typical member of the species. In epidemiology, we take R_0 to mean the number of individuals infected by a single infected individual during his or her entire infectious period, in a population which is entirely susceptible. For in-host dynamics, R_0 gives the number of newly infected cells produced by one infected cell during its lifetime, assuming all other cells are susceptible.

From this definition, it is immediately clear that when $R_0<1$, each infected individual produces, on average, less than one new infected individual, and we therefore predict that the infection will be cleared from the population, or the microparasite will be cleared from the individual. If $R_0>1$, the pathogen is able to invade the susceptible population. This threshold behaviour is the most important and useful aspect of the R_0 concept. In an endemic infection, we can determine which control measures, and at what magnitude, would be most effective in reducing R_0 below one, providing important guidance for public health initiatives.

The magnitude of R_0 is also used to gauge the risk of an epidemic or pandemic in emerging infectious disease. For example, the estimation of R_0 was of critical importance in understanding the outbreak and potential danger from severe acute respiratory syndrome (SARS) (Choi & Pak 2003; Lipsitch *et al.* 2003; Lloyd-Smith *et al.* 2003; Riley *et al.* 2003). R_0 has been likewise used to characterize bovine spongiform encephalitis (BSE) (Woolhouse & Anderson 1997; Ferguson *et al.* 1999; de Koeijer *et al.* 2004), foot and mouth disease (FMD) (Ferguson *et al.* 2001; Matthews *et al.* 2003), novel strains of influenza (Mills *et al.* 2004; Stegeman *et al.* 2004) and West Nile virus (Wonham *et al.* 2004). The incidence and spread of dengue (Luz *et al.* 2003), malaria (Hagmann *et al.* 2003), Ebola

†Author for correspondence (lwahl@uwo.ca).

Received 3 November 2004
Accepted 29 March 2005

(Chowell *et al.* 2004*b*) and scrapie (Gravenor *et al.* 2004) have also been assessed using R_0 in recent literature. Topical issues such as the risks of indoor airborne infection (Rudnick & Milton 2003), bioterrorism (Kaplan *et al.* 2002; Longini *et al.* 2004), and computer viruses (Lloyd & May 2001) also rely on this important concept.

Ongoing theoretical work has extended R_0 for a range of complex models, including stochastic and finite systems (Nasell 1995), models with spatial structure (Mollison 1995*b*; Lloyd & May 1996; Keeling 1999) or age-structure (Anderson & May 1991; Diekmann & Heesterbeek 2000; Hyman & Li 2000), and macroparasite models (Anderson & May 1991; Diekmann & Heesterbeek 2000). We note, however, that the *practical* use of R_0 has been, for the most part, restricted to very simple deterministic systems. For comparison with this 'field' literature in epidemiology, we restrict our attention in the following sections to deterministic, unstructured microparasite models.

The purpose of this paper is to review the various methods currently in use for the derivation of R_0, highlighting the difference between R_0 and surrogate parameters with equivalent threshold behaviour. We then discuss methods commonly used to estimate R_0 from incidence data. Finally, we give an overview of the recent use of R_0 in assessing emerging and endemic disease. Our aim in this final section of the paper is to determine the usefulness of this endeavour: to what extent has estimating R_0 informed public health measures?

2. DERIVATIONS OF R_0 FROM A DETERMINISTIC MODEL

The derivation of R_0 from a non-spatial, deterministic model is fairly straightforward from first principles. The survival function method (§2.1) gives the 'gold standard' determination of R_0, and is applicable even when non-constant transmission probabilities between classes (i.e. non-exponential lifetime distributions) are assumed. For models which include multiple classes of infected individuals, the next generation operator is the natural extension of this approach (§2.2). However, we note that the definition of R_0 may have more than one possible interpretation in the multi-class system, as discussed below.

2.1. Survival function

The method we describe as the 'survival function' approach is, in essence, a first-principles definition of R_0, and thus has a rich history of use. The approach is described in detail in Heesterbeek & Dietz (1996), who also give an interesting historical overview.

Consider a large population and let $F(a)$ be the probability that a newly infected individual remains infectious for at least time a. This is called the survival probability. Also, let $b(a)$ denote the average number of newly infected individuals that an infectious individual will produce per unit time when infected for total time a. Then, R_0 is given by:

$$R_0 = \int_0^\infty b(a)F(a)\mathrm{d}a. \tag{2.1}$$

As this expression yields R_0 by definition, this approach will be appropriate for any model in which closed-form expressions can be given for the underlying survival probability, $F(a)$, and the infectivity as a function of time, $b(a)$. In particular, it is straightforward to handle situations in which infectivity depends on time, since infection, or other transmission probabilities between states, vary with time. Thus, this derivation of R_0 is not restricted to systems described by ordinary differential equations (ODEs).

This method can also be naturally extended to describe models in which a series of states are involved in the 'reproduction' of an infected individual. As an example of the latter technique, consider epidemic modelling of malaria. An infected human may pass the infection to a mosquito, which may in turn infect more humans. This complete cycle must be taken into account in our derivation of R_0, which we might expect to yield the total number of infected humans produced by one infected human. In general, if only two distinct infectious states are involved in such an *infection cycle*, $F(a)$ can be defined as the probability that an individual in state 1 at time zero produces an individual who is in state 2 until at least time a. Similarly, $b(a)$ is the average number of new individuals in state 1 produced by an individual who has been in state 2 for time a. In modelling malaria, $F(a)$ could be the probability that a human infected at time zero produces an infected mosquito which remains alive until at least time a. In more concrete terms, $F(a)$ would be the integral of the following product:

$$\begin{aligned} F(a) = \int_0^a & \text{prob(human infected at time 0} \\ & \quad \text{exists at time } t) \\ & \times \text{prob(human infected for tot. time } t \\ & \quad \text{infects mosquito)} \\ & \times \text{prob(infected mosquito lives to be} \\ & \quad \text{age } a - t)\ \mathrm{d}t \end{aligned} \tag{2.2}$$

while $b(a)$ would simply be the average number of humans newly infected by a mosquito which has been infected for time a. (Note that we could also take the infected mosquito as state 1, deriving an analogous expression which would yield the same value of R_0.)

Unfortunately, derivations such as equation (2.2) become increasingly cumbersome as this method is extended to infection cycles involving three or more states (Hethcote & Tudor 1980; Lloyd 2001*b*; Huang *et al.* 2003). In these situations, the next generation operator offers an elegant solution, as described in the following section.

2.2. Next generation method

A rich history in the literature addresses the derivation of R_0, or an equivalent threshold parameter, when more than one class of infectives is involved (Rushton &

Mautner 1955; Hethcote 1978; Nold 1980; Hethcote & Thieme 1985).

The next generation method, introduced by Diekmann *et al.* (1990), is a general method of deriving R_0 in such cases, encompassing any situation in which the population is divided into discrete, disjoint classes. The next generation operator can thus be used for models with underlying age structure or spatial structure, among other possibilities. For typical implementations, continuous variables within the population are approximated by a number of discrete classes. This approximation assumes that transmission probabilities between states are constant, or equivalently, that the distribution of residence times in each state is exponential.

The next generation operator is fully described in Diekmann & Heesterbeek (2000) and a number of salient cases are elucidated in van den Driessche & Watmough (2002). Recent examples of this method are given in Matthews *et al.* (1999), Porco & Blower (2000), Castillo-Chavez *et al.* (2002), Hill & Longini (2003) and Wonham *et al.* (2004).

In the next generation method, R_0 is defined as the spectral radius of the 'next generation operator'. The formation of the operator involves determining two compartments, infected and non-infected, from the model. In this section, we outline the steps needed to find the next generation operator in matrix notation (assuming only finitely many types), and then employ this method for a susceptible–exposed–infectious–recovered (SEIR) model and a model of malaria. (For a detailed explanation on the formation of the next generation operator when there are infinitely many types see pp. 95–96 of Diekmann & Heesterbeek (2000).)

Let us assume that there are n compartments of which m are infected. We define the vector $\bar{x} = x_i$, $i = 1, \ldots, n$, where x_i denotes the number or proportion of individuals in the ith compartment. Let $F_i(\bar{x})$ be the rate of appearance of new infections in compartment i and let $V_i(\bar{x}) = V_i^-(\bar{x}) - V_i^+(\bar{x})$, where V_i^+ is the rate of transfer of individuals into compartment i by all other means and V_i^- is the rate of transfer of individuals out of the i th compartment. The difference $F_i(\bar{x}) - V_i(\bar{x})$, gives the rate of change of x_i. Note that F_i should include only infections that are newly arising, but does not include terms which describe the transfer of infectious individuals from one infected compartment to another.

Assuming that F_i and V_i meet the conditions outlined by Diekmann *et al.* (1990) and van den Driessche & Watmough (2002), we can form the next generation matrix (operator) FV^{-1} from matrices of partial derivatives of F_i and V_i. Specifically,

$$F = \left[\frac{\partial F_i(x_0)}{\partial x_j}\right] \quad \text{and} \quad V = \left[\frac{\partial V_i(x_0)}{\partial x_j}\right],$$

where $i, j = 1, \ldots, m$ and where x_0 is the disease-free equilibrium. The entries of FV^{-1} give the rate at which infected individuals in x_j produce new infections in x_i, times the average length of time an individual spends in a single visit to compartment j. R_0 is given by the spectral radius (dominant eigenvalue) of the matrix FV^{-1}.

As an example, let us consider an SEIR model. Since we are concerned with the populations that spread the infection we only need to model the exposed, E, and infected, I, classes. Let us define the model dynamics using the following equations:

$$\left.\begin{aligned} \dot{E} &= \beta SI - (\mu + k)E, \\ \dot{I} &= kE - (\gamma + \mu)I. \end{aligned}\right\} \tag{2.3}$$

where μ is the per capita natural death rate, β is the efficacy of infection of susceptible individuals S, k is the rate at which a latent individual becomes infectious and γ is the per capita recovery rate. For this system

$$F = \begin{pmatrix} 0 & \beta\lambda/\mu \\ 0 & 0 \end{pmatrix}$$

(where λ is the birth rate of susceptibles) and

$$V = \begin{pmatrix} \mu + k & 0 \\ -k & \gamma + \mu \end{pmatrix},$$

and thus

$$R_{0,N} = \frac{k\beta\lambda}{(\mu + k)(\mu + \gamma)\mu}. \tag{2.4}$$

Note that this is also the value of R_0 determined by the survivor function method.

For the second example, we consider a model of malaria. Let us describe the rate of change of the infected human, H_I, and mosquito, M_I, populations by the following equations:

$$\left.\begin{aligned} \dot{H}_\mathrm{I} &= \beta_\mathrm{MH} M_\mathrm{I} H_\mathrm{S} - (\mu_\mathrm{H} + \alpha + \sigma)H_\mathrm{I}, \\ \dot{M}_\mathrm{I} &= \beta_\mathrm{HM} M_\mathrm{S} H_\mathrm{I} - \mu_\mathrm{M} M_\mathrm{I}. \end{aligned}\right\} \tag{2.5}$$

Infected humans are produced by the infection of susceptible humans, H_S, by an infected mosquito with efficacy β_MH. We assume that they die with natural death rate μ_H, die due to infection with rate σ and recover from the infection with rate α. Infected mosquitoes are produced when susceptible mosquitoes, M_S, bite infected humans. We assume that this process has efficacy β_HM and assume that infected mosquitoes can only leave the infected compartment by dying naturally with rate μ_M. For this system we find that

$$F = \begin{pmatrix} 0 & \beta_\mathrm{MH} H_\mathrm{S}(0) \\ \beta_\mathrm{HM} M_\mathrm{S}(0) & 0 \end{pmatrix},$$

and

$$V = \begin{pmatrix} \mu_\mathrm{H} + \alpha + \sigma & 0 \\ 0 & \mu_\mathrm{M} \end{pmatrix}.$$

Since V is non-singular we can determine V^{-1}. Thus,

$$R_{0,\mathrm{M}} = \sqrt{\frac{\beta_\mathrm{MH}\beta_\mathrm{HM} H_\mathrm{S}(0) M_\mathrm{S}(0)}{(\mu_\mathrm{H} + \alpha + \sigma)\mu_\mathrm{M}}}. \tag{2.6}$$

For comparison, we also compute the value of R_0 for this system using the survival method:

$$R_{0,\mathrm{S}} = \frac{\beta_{\mathrm{MH}}\beta_{\mathrm{HM}}H_{\mathrm{S}}(0)M_{\mathrm{S}}(0)}{(\mu_{\mathrm{H}}+\alpha+\sigma)\mu_{\mathrm{M}}} = (R_{0,\mathrm{M}})^2. \qquad (2.7)$$

The difference here is a matter of definition: the survival function gives the total number of infectives *in the same class* produced by a single infective of that class, while the next generation operator gives the mean number of new infectives per infective in any class, *per generation*. Values corresponding to the latter definition thus depend on the number of infective classes in the infection cycle. We note that the latter definition is widely accepted as standard in the biomathematics literature (e.g. Diekmann & Heesterbeek 2000), but the former definition has also been used extensively (Anderson & May 1991; Barbour & Kafetzaki 1993; Nowak & May 2000), and is still in standard use in epidemiology (Hagmann *et al.* 2003; Luz *et al.* 2003) and immunology (Huang *et al.* 2003).

3. DERIVATIONS OF THRESHOLD CRITERIA

As mentioned in §1, the most important feature of R_0 is that it reflects the stability of the disease-free equilibrium. When $R_0<1$, this equilibrium is stable and we predict that the pathogen will be cleared.

Surveying the recent literature, it quickly becomes apparent that a number of related quantities, all of which share this 'threshold' behaviour, are used as surrogates for R_0. For example, R_0^n $(n>0)$ will give an equivalent threshold, but does not give the number of secondary infections produced by a single infectious individual.

The methods outlined in the following section each derive, from a deterministic model, a quantity which shares this predictive threshold with R_0. For some models, these methods will, in fact, yield the true value of R_0, but this is by no means guaranteed. If a prediction of whether the pathogen will persist or be cleared is the only feature of interest, a threshold criterion is sufficient—however, these methods cannot be used to compare risks associated with different pathogens.

We outline three such threshold criteria below, giving examples where each is used in the literature.

3.1. Jacobian and stability conditions

A predictive threshold is often found through the study of the eigenvalues of the Jacobian at the disease-free equilibrium (for an overview see Diekmann & Heesterbeek 2000). This is a simple, widely used method for ODE systems. Using this method, a parameter is derived from the condition that all of the eigenvalues of the Jacobian have a negative real part. This can easily be done using the characteristic polynomial and the Routh–Hurwitz stability conditions.

The Jacobian method clearly allows us to derive a parameter that reflects the stability of the disease-free equilibrium. The parameter obtained in this way, however, may or may not reflect the biologically meaningful value of R_0. An example where the Jacobian method does not yield R_0 is described in detail in Diekmann & Heesterbeek (2000; exercise 5.43). Despite this caveat, the technique remains popular; recent uses of this criterion in the literature include Porco & Blower (1998); Murphy *et al.* (2002); Kawaguchi *et al.* (2004); Laxminarayan (2004) and Moghadas (2004). In Roberts & Heesterbeek (2003), it is suggested that if this threshold parameter does not have the same biological interpretation as the dominant eigenvalue of the next generation matrix, then it should not be called the basic reproductive ratio, nor denoted R_0.

3.2. Existence of the endemic equilibrium

Similarly, we can often derive a condition based on parameter values such that when the condition holds, the endemic equilibrium exists, whereas when the condition is false, only the disease-free equilibrium exists. Mathematically, we are referring to a *transcritical bifurcation*, and we know that the condition must switch from being false to true at parameter values which give $R_0=1$.

For example, consider the model of herpes simplex virus described in Blower *et al.* (1998). For simplicity, we can ignore drug resistance (i.e. $p_1=p_2=0$). This model then consists of three differential equations

$$\begin{aligned}
\frac{\mathrm{d}X}{\mathrm{d}t} &= \pi - Xc\beta_{\mathrm{S}}\frac{H_{\mathrm{S}}}{N} - X\mu,\\
\frac{\mathrm{d}Q_{\mathrm{S}}}{\mathrm{d}t} &= H_{\mathrm{S}}(\sigma+q) - Q_{\mathrm{S}}(\mu+r),\\
\frac{\mathrm{d}H_{\mathrm{S}}}{\mathrm{d}t} &= Xc\beta_{\mathrm{S}}\frac{H_{\mathrm{S}}}{N} - H_{\mathrm{S}}(\mu+\sigma+q) + rQ_{\mathrm{S}},
\end{aligned}$$

where X is the susceptible population, Q_{S} represents those infected with the virus in the non-infection latent state, H_{S} represents those infected with the virus in infectious state and $N=X+Q_{\mathrm{S}}+H_{\mathrm{S}}$. (Other letters are positive parameters.) At equilibrium,

$$\begin{aligned}
N &= \frac{\pi}{\mu},\\
X &= \frac{\pi}{\mu} - \frac{\mu+\sigma+q+r}{\mu+r}H_{\mathrm{S}},\\
Q_{\mathrm{S}} &= \frac{\sigma+q}{\mu+r}H_{\mathrm{S}}.
\end{aligned}$$

Thus, either $H_{\mathrm{S}}=0$ (the disease-free equilibrium) or

$$H_{\mathrm{S}} = \frac{\pi}{\mu}\left[\frac{\mu+r}{\mu+\sigma+q+r} - \frac{\mu}{c\beta_{\mathrm{S}}}\right]$$

(the endemic equilibrium). It follows that the endemic equilibrium only exists when

$$R_{0,\mathrm{E}} \equiv c\beta_{\mathrm{S}}\left(\frac{r+\mu}{\mu(r+\mu+\sigma+q)}\right) > 1,$$

and does not exist if the reverse inequality holds.[1]

Outbreaks of infectious periods are brief, but continue over the course of the patients' lifetime, with the virus quiescent at other times. This makes calculating R_0 from other methods quite complicated.

[1]Note that this derivation of $R_{0,E}$ differs from that in Blower *et al.* (1998) due to a missing σ in eqn. (7), p. 678 of that manuscript.

3.3. Constant term of the characteristic equation

For more complex models, the characteristic equation may be of the form

$$\lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_1\lambda + p_0 = 0,$$

with $p_1, p_2, \ldots, p_{n-1} > 0$. In this special case, $n-1$ roots of the polynomial have negative real part. When $p_0 = 0$, the nth root, or largest eigenvalue, is zero, when $p_0 > 0$, all eigenvalues are negative, whereas when $p_0 < 0$, the largest eigenvalue has positive real part. Thus, the stability is determined solely by the sign of the constant term of the characteristic equation.

For example, consider the multi-strain tuberculosis model described in Blower & Chou (2004). In eqn (6) in their appendix, their characteristic polynomial is

$$(\lambda + \mu_0)\prod_{i=1}^{N}[\lambda^2 + B_i\lambda + C_i] = 0,$$

where

$$B_i = \mu_i^L + \mu_i^T - \beta_i^T S^* + \nu_i + c_i + k_{i,i+1},$$

$$C_i = (c_i + \mu_i^T - \beta_i^T S^* + k_{i,i+1})(\mu_i^L + \nu_i) - \beta_i^L \nu_i S^*,$$

and all parameters non-negative. Note that B_i has the property that $B_i > 0$ when $C_i = 0$. For each strain i, the equation for $C_i = 0$ is rearranged to produce

$$R_0(i) = S^* \frac{(\beta_i^T + \beta_i^L)\nu_i + \beta_i^T \mu_i^L}{(\nu_i + \mu_i^L)(c_i + k_{i,i+1} + \mu_i^T)}.$$

Each $R_0(i)$ value has the property that $R_0(i) = 1$ when $C_i = 0$, $R_0(i) < 1$ when $C_i > 0$ and $R_0(i) > 1$ when $C_i < 0$.

Calculating an $R_0(i)$ for each strain using the methods from previous sections is extremely difficult, as is calculating a formula for the endemic equilibrium. However, the Jacobian matrix at the disease-free equilibrium is relatively tractable, so an $R_0(i)$ for each strain can be calculated from the constant term. This method generally allows for the calculation of threshold criteria when other methods fail.

4. ESTIMATIONS FROM EPIDEMIOLOGICAL DATA

The previous sections addressed methods of formulating R_0 in terms of the parameters of some deterministic model. In order to estimate the value of R_0 from incidence data, however, we require numerical estimates of a number of these parameters. Typically, death rates and recovery rates are readily estimated; in contrast, the contact or transmission rate is difficult to determine from direct measures. For this reason, R_0 is rarely estimated using formulae such as equations (2.6) and (2.7) above. We outline a number of alternative approaches for estimating R_0 from available data in §§4.1–4.4. These approaches typically involve simplifying assumptions to reduce the number of unknown parameters. For more complete overviews of these techniques, we refer the reader to Mollison (1995*a*), Diekmann & Heesterbeek (2000) and Hethcote (2000).

4.1. Susceptibles at endemic equilibrium

This method assumes that an endemic equilibrium is attained and uses the prevalence of the infection at this equilibrium to estimate R_0. Following Mollison's (1995*a*) derivation, we consider a single infected individual and note that the number of successful contacts (in which the infection is passed on) for that individual should be given by $R_0\pi_s$, where π_s is the probability that a given contact is with a susceptible. At equilibrium, the average number of new infections per infected individual must be exactly one, allowing us to write $R_0 = 1/\pi_s$. Under the assumption of homogenous mixing, the unknown probability, π_s, can be estimated as the fraction of the host population that is susceptible at the endemic equilibrium. This yields an extremely simple estimate of the basic reproductive ratio, which has been used extensively (see Anderson & May (1991) for review).

An interesting point here is that R_0 reflects not only the behaviour of the system at the uninfected equilibrium (which is apparent by definition), but may also, under certain assumptions, reflect important features of the endemic equilibrium. Similar to other ODE methods, we must first assume that the host population is homogenous, that is, all hosts have intrinsically similar epidemiological properties, independent of age, genetic make-up, geography, and so on. We also assume mass-action transmission, specifically, that the number of contacts per infective is independent of the number of infectives. The accuracy of this estimate will clearly depend on the degree to which these assumptions hold; if infectivity or mortality vary with age, for example, the approximation suffers.

Mathematically, this method may seem unrealistic at first glance, as $R_0 < 1$ would imply that the fraction of susceptibles is greater than one. This is because there is a transcritical bifurcation at $R_0 = 1$ and the number of susceptibles of the 'endemic' equilibrium is actually negative. During this portion of the bifurcation diagram, the uninfected equilibrium is stable, and hence the initial condition ensures that negative individuals cannot be reached. Practically, this means that when $R_0 < 1$ we would never find a population at the endemic equilibrium, and could not apply this method. (Note that when the assumption of mass-action transmission is relaxed, a backward bifurcation may occur at $R_0 = 1$, and diseases with $R_0 < 1$ may persist (Dushoff 1996; Dushoff *et al.* 1998).)

Recent examples of this method include Heesterbeek (2003) and Ferguson *et al.* (2001).

4.2. Average age at infection

A related approach, also based on the endemic equilibrium, is that R_0 can be estimated as L/A, where L is the mean lifetime and A is the mean age of acquiring the disease (Dietz 1975). A derivation for this simple relation is also provided by Mollison (1995*a*) and Hethcote (2000); for further discussion, see Anderson & May (1991) and Brauer (2002). In brief, we must assume that all individuals are born susceptible, that after acquiring the disease they are no longer susceptible, that the population is at the endemic equilibrium

(i.e. $R_0>1$) and that homogenous mixing, particularly among age groups, occurs. While this strong set of assumptions might never be fully realized in a practical setting, the usefulness of this approach is clear since both L and A are readily measured. This method was recently used to estimate R_0 for endemic canine pathogens (Laurenson *et al.* 1998).

4.3. The final size equation

While the previous two methods estimate R_0 from the endemic equilibrium, the final size equation is applicable to closed populations only, where the infection leads either to immunity or death. In this situation, the number of susceptibles can only decrease and the final fraction of susceptibles, $s(\infty)$, can be used to estimate R_0:

$$R_0 = \frac{\ln s(\infty)}{s(\infty) - 1}.$$

This was first recognized by Kermack & McKendrick (1927); for a detailed derivation and discussion, see Diekmann & Heesterbeek (2000), Hethcote (2000) and Brauer (2002). This estimate holds when the disease itself does not interfere with the contact process, or when contact intensity is proportional to population density.

4.4. Calculation from the intrinsic growth rate

Finally, R_0 may be determined from the intrinsic growth rate of the infected population. This growth rate, often denoted r_0, is the rate at which the total number of infectives, I, grows in a susceptible population, such that $\mathrm{d}I/\mathrm{d}t = r_0 I$. We note that this is an *implicit* definition of r_0, and thus from a modelling perspective using r_0 is seldom elegant.

Using incidence data, however, r_0 can often be approximately measured from the growth rate of the infected class, and R_0 can subsequently be estimated from r_0. There are several possible problems with this approach: firstly, stochastic fluctuations in the early stages of the epidemic can obscure the measure of r_0 (see Heffernan & Wahl in press); secondly, reporting inaccuracies are very likely to bias the incidence data. Finally, even when r_0 can be measured with some confidence, the relationship between R_0 and r_0 is highly model dependent.

In the simplest possible models, when infectivity is constant throughout the infectious period, R_0 can be estimated as $1+r_0L$, where L is the expected duration of the infectious period. (The 'one' is necessary in this expression because R_0 reflects the total number of new infections, whereas the overall growth rate r_0 includes the death of the founding individual.) For more complex models, the relation between r_0 and R_0 can be derived by expressing both in terms of the model parameters, exploiting that fact that the spectral radius of the Jacobian, evaluated at the disease-free equilibrium, gives r_0. (This is apparent from the definition of r_0.) We also note that r_0 itself can be used as a threshold parameter, since $R_0<1$ implies $r_0<0$. Thus, the condition $r_0<0$ is actually equivalent to the 'Jacobian' method described in §3.1.

As an example, consider Nowak *et al.* (1997) and Lloyd (2001*a*), who studied the within-host dynamics of viral disease. From standard models of viral dynamics, they find that the relationship between R_0 and r_0 is

$$R_0 = 1 + \frac{r_0(r_0 + a + u)}{au}, \tag{4.1}$$

where a is the death rate of the infected cells and u is the clearance rate of the virions. If $r_0 + a \ll u$ then the relation approaches

$$R_0 = 1 + \frac{r_0}{a}. \tag{4.2}$$

Since $1/a$ is the expected lifetime of an infected cell, this expression is consistent with our previous approximation of R_0.

This method proves useful since r_0 can be readily estimated from viral load data, for in-host models, or from incidence data in epidemiology. A number of recent studies have used this approach, including Pybus *et al.* (2001) and Lipsitch *et al.* (2003).

5. RECENT USE OF R_0 IN THE EPIDEMIOLOGY OF MICROPARASITES

5.1. SARS and influenza

5.1.1. SARS. The emergence of SARS underscored the need for careful epidemiological modelling, in order to better understand and contain such novel pathogens. A number of models were developed to study SARS and to compute R_0 for outbreaks in Hong Kong, Singapore and Canada.

Lipsitch *et al.* (2003) estimated R_0 for the outbreaks in Canada and Singapore, including the effects of superspreaders (infected individuals who directly infect a large number of people). The exponential growth rate of the cumulative number of cases in the epidemic was taken as an estimate for r_0. R_0 was then estimated by computing the largest eigenvalue of a linearized SEIR model (assuming no depletion of susceptibles), and expressing this spectral radius as a function of R_0, the ratio of the infectious period to the serial interval, f, and the length of the serial interval, L. This technique yielded the following equation for R_0:

$$R_0 = 1 + r_0L + f(1-f)(r_0L)^2.$$

R_0 were approximately 2.2–3.6 for serial intervals of 8–12 days. The serial intervals were estimated from the data, but at the time were not well defined for SARS. A strength of this approach is that the various parameters of the SEIR model 'collapse', such that epidemiological estimates of only three parameters are necessary: r_0, f and L. Although the usual problems of underreporting before an epidemic, overreporting during an epidemic and stochasticity are unavoidable in estimates of r_0, Lipsitch *et al.* conducted thorough sensitivity analyses, concluding that R_0 will still have a relatively low value. This suggests that the spread of SARS can be contained when proper control protocols are put into place.

Lipsitch *et al.* then extended the SEIR model to explore the effects of isolation of symptomatic cases and quarantine of asymptomatic contacts on the spread of the disease. They found that to reduce R_0 from approximately 3 to 1, isolation and quarantine must reduce total infectiousness by at least two-thirds. Further analysis of these control policies enabled Lipsitch *et al.* to conclude that quarantine would impose a large burden on the population if SARS was allowed to spread over a long period with an $R_0 > 1$ in a susceptible population. Individuals could be quarantined multiple times over the course of the infection or for very long periods of time. These conclusions offer useful guidance for public health initiatives, but as several parameters of this model are unknown, Lipsitch *et al.* were unable to give concrete estimates for the levels of quarantine and isolation necessary to decrease the value of R_0 below one.

Chowell *et al.* (2003) developed a system of ODEs to describe the spread of SARS in the three geographical populations mentioned above. Their model includes two classes of susceptibility, low risk and high risk, and also includes two types of infected individuals, symptomatic and asymptomatic, which differ in their rate of diagnosis and mode of transmission. The main goal of this study was not to determine R_0, but to estimate the diagnostic rate and isolation effectiveness for the three separate regions, with an emphasis on the Toronto outbreak. These two parameter values were estimated by first determining the exponential growth rates from SARS incidence data in all three regions and fitting the model to the data assuming that all of the other model parameters were roughly constant between regions. In a brief section the parameter estimates were used to calculate R_0 using the next generation approach. R_0 was 1.2 for Hong Kong, approximately 1.2 for Toronto and 1.1 for Singapore. A weakness of this model is that R_0 depends on estimating many (approximately 10) model parameters. These estimates of R_0 are comparable to those estimated by Lipsitch *et al.* (2003) when the latter group assumed the serial interval to be small, around 4 days. However, the serial interval in this study was taken to be between 7 and 10 days. This disparity was not discussed in detail.

Using the same model, Chowell *et al.* (2004*a*) conducted sensitivity analyses for R_0, quantifying the effects of changes in the model parameters. They found that the transmission rate and the relative infectiousness after isolation have the largest effect on R_0. They also found that it is unlikely that the implementation of a single control measure will reduce R_0 below one. The practical conclusion of this work is that control measures that affect the diagnostic rate, relative infectiousness after isolation and the per capita transmission rate should be implemented.

In another study (Riley *et al.* 2003), R_0 was determined by fitting a stochastic mathematical model to incidence data for SARS. Riley *et al.* developed a stochastic, compartmental metapopulation model capturing both spatial variability and the growth dynamics at the early stages of the epidemic. Using data from the Hong Kong epidemic, Riley *et al.* determined probability distributions for transitions between the model compartments of susceptible, latent, infectious, hospitalized, recovered and deceased individuals. R_0 was calculated using multiple realizations of the model to be approximately 3.

Riley *et al.* also found that the SARS control measures were effective and, most importantly, concluded that the Hong Kong epidemic was under control by early April. This conclusion was made by determining R_0 when control measures were implemented. An advantage of this approach is that multiple realizations of the model can generate predicted case incidence time-series, quantifying any reduction in the transmission rates after control measures are in place. However, this complex model relies heavily on the quality of the data. Another drawback of this model is that the effects of superspreaders were not included.

Lloyd-Smith *et al.* (2003) developed a stochastic model of a SARS outbreak in a community and its hospital. The goal of this model was to evaluate contact precautions, quarantine and isolation as containment procedures while assuming a particular value of R_0. Using a value of $R_0 \approx 3$ for the Hong Kong and Singapore outbreaks they found that isolation alone could control the spread of SARS if it met very stringent requirements. However, they concluded that the control measures that were most successful were limiting contact between people in hospitals and decreasing the number of contacts between people inside and outside of the hospital.

Summarizing the results above, we can conclude that the estimated value of R_0 for SARS is relatively low, suggesting that the epidemic can be controlled. We can also conclude that the control policies studied are most effective when used in combination. These conclusions are reassuring and give direction to public health initiatives. These results should be viewed with some caution, however, as the data used in these studies are limited, the models are complex, and aspects of the virulence and persistence of SARS that might affect public health initiatives have not yet been addressed.

5.1.2. 1918 Pandemic influenza. Mills *et al.* (2004) used mortality data to estimate R_0 for the 1918 influenza pandemic in 45 cities in the USA. Interestingly, this approach relied on none of the mathematical techniques described in previous sections; instead, the number of susceptibles, incident infections and infectious hosts were estimated using a discrete time simulation. Using a case fatality proportion of 2%, the total number of deaths was estimated and this was compared with 'excess' mortality data, that is, the number of deaths in 1918 above the median for 1910–1916. A value of R_0 was determined which minimized the sum of squared differences between the simulated and observed data. The median estimate for R_0 was 2.9.

It is interesting to note that in this study, one of the most careful and recent investigations of R_0 in the literature, the authors relied on a very simple simulation and least-squares fitting, rather than any more sophisticated mathematical approaches. The advantage of the simulation is that the many assumptions which must be made are explicit, and their effects

can be examined individually, as these authors have done in extensive supplementary material. In all cases, the sensitivity analyses predicted that the overall conclusion of the work—that R_0 was approximately 3–4—was robust.

As noted by the authors, various possible sources of downward bias, including heterogenous mixing, intervention measures, and the depletion of susceptibles, are ignored in this approach. To correct for this, for each city, the two weeks in which the growth rate of mortality data was highest were also fit separately; this increased the median estimate of R_0 to 3.9. It seems likely, however, that any heterogenous mixing and intervention measures were in place during these two weeks of rapid epidemic growth as well, since these weeks were not always the first weeks of the epidemic. Thus, this 'extreme' estimate of R_0 is only the most extreme value that can be observed from the data, under the same assumptions regarding lack of control measures and homogenous mixing. The extent to which any control measures were in place and their mitigation of R_0 was not addressed.

The aim of the study was to evaluate the risk of an impending pandemic from a novel strain of influenza. The results suggest that control of such a pandemic will be possible, given the 'modest' reproductive number of the 1918 strain. From a statistical point of view, however, R_0 for the 1918 pandemic was a single observation of an extreme value, and it is very difficult to predict the magnitude of a single future extreme value drawn from the same distribution. Thus, the conclusions only hold under the assumption that a future influenza strain will be 'similarly' infectious. Nonetheless, it is important to have demonstrated that even for the worst influenza pandemic in recent history, R_0 was probably not large relative to other diseases.

5.1.3. Avian influenza. Stegeman *et al.* (2004) quantified between-flock transmission characteristics of high-pathogenicity avian influenza, a virus in the Netherlands that led to the culling of 30 million birds in 2003. R_0 was calculated as the product of the infectious period at flock level and the transmission rate at flock level; however, neither parameter was measured directly. Instead, the infectious period was estimated as the period between the moment of detection and the moment of culling, plus 4 days. The transmission probability of the stochastic SEIR model was estimated by means of a generalized linear model. An estimate of the variance of R_0 was used to calculate the confidence interval for the period of infection and the transmission probability. A variety of potential control measures were evaluated.

The results of this study estimated that R_0 reached as high as 6.5 in some regions and was decreased to 1.2 after the outbreak. Although R_0 still exceeded one, between-flock transmission nevertheless decreased significantly after the outbreak. This discrepancy between the calculated value of R_0 and the ultimate course of the epidemic suggested that control measures designed to reduce the transmission rate were inadequate. It was instead hypothesized that containment of the epidemic was probably owing to the reduction in the number of susceptible flocks caused by culling rather than the reduction of the transmission rate by other control measures. From these observations, it was suggested that effective control in the future could be achieved only by depopulation of the whole affected area.

5.2. Livestock disease

5.2.1. Bovine spongiform encephalopathy (BSE). Bovine spongiform encephalopathy affects populations of cattle and other livestock and may pose a threat to human health. A number of models of BSE have been analysed; these models include key transmission routes and evaluate the efficacy of various control policies.

Ferguson *et al.* (1999) developed a model to describe the spread of BSE. The goal of this paper was to demonstrate how different assumptions regarding the infectivity of BSE affect R_0. Two models of infectivity that represent epidemiological extremes were considered: the first assumes that infectivity rises exponentially with a growth coefficient of two per year throughout the incubation period of BSE; the second assumes that infectivity is constant during this time. Using the next generation approach, Ferguson *et al.* estimated that $R_0 \approx 10$–12 for the first case and that $R_0 \approx 2$–2.5 for the second. These values were determined using a back calculation model (see Gail & Rosenberg 1992) to estimate the force of infection of BSE in Great Britain between 1980 and 1996. The transmission coefficient of BSE was estimated using a model for infectivity as a function of incubation stage.

Ferguson *et al.* also determined the effect that the 1988 ban on MBM (recycling of animals into ruminant-based meat and bone meal) had on R_0. They found that, for both cases of infectivity, R_0 was reduced to a value of approximately 0.15. This result has important implications as it shows that the spread of BSE can be controlled for the extreme cases of infectivity, implying that this will be true for all intermediate models. These estimates of R_0 also suggest that BSE will not become endemic in the UK. A drawback of this model is that it assumes that underreporting of BSE cases does not exist after 1988. This assumption can result in a lower value of R_0. Also, the effects of clustering were not modelled; instead, homogenous mixing was assumed. However, Ferguson *et al.* concluded that this would have only a minor effect on the conclusions of the study.

In a more recent study by de Koeijer *et al.* (2004), R_0 was calculated for BSE assuming five different transmission routes: horizontal, vertical, diagonal (the disease can be spread to other animals close by during a birth), feed-based transmission and infectious material in the environment (use of MBM as fertilizer). Separating the infected population into two classes of infected individuals, those that are infected from birth and those that become infected by all other routes, de Koeijer *et al.* determined the expected number of new infections during the whole infectious period for both classes.

These expressions were then used to formulate the next generation matrix to determine R_0. Using parameter estimates from BSE data from the United Kingdom and the Netherlands, values for R_0 were determined for separate outbreaks in 1986, 1991, 1995 and 1998. The estimated values of R_0 were approximately 14 and 0.7 in 1986 for the United Kingdom and the Netherlands, respectively, whereas R_0 values were far less than unity in later years when control measures were in effect.

This study also attempted to quantify the impact of the control policies in use. They found that there are three major control measures: a feed ban on MBM to cattle, optimization of the rendering process (how cattle feed is made, temperature, etc.) and removing and incinerating any materials that increase the risk of contracting BSE. They also found that, in order to reduce R_0 to a value less than unity, at least two of the three control measures should be applied. However, the authors stated that even when all three control measures are in place, infection routes other than via feed will remain difficult to control, and therefore, R_0 cannot be reduced to zero. This is not a serious concern, as they find that R_0 is only 0.06 when transmission via feed has been eliminated. In this study, then, the primary use of R_0 was as a measure of the efficacy of control measures, with the goal of predicting control measures that reduce R_0 to below unity. A drawback of this model is that calculating R_0 relied on estimating many model parameters using BSE data and procedures that have high uncertainty. This resulted in a very wide confidence interval around R_0. The effects of clustering were also ignored.

5.2.2. Scrapie. Matthews *et al.* (1999) developed a model of scrapie transmission within a single flock of sheep. The model includes both horizontal and vertical transmission, as well as genetic variation in susceptibility. R_0 was calculated through the next generation operator.

Using parameters for a single, well-studied flock of Chevriot sheep, an estimate of 3.9 was obtained for R_0 in a natural outbreak of scrapie between 1970 and 1982. We note, however, that the detailed parameters needed for this estimate, including the initial frequencies of the susceptible and resistant alleles, are not likely to be routinely available.

The real importance of this study, however, is in the accompanying sensitivity analyses. R_0 is found to vary little with the vertical transmission rate, but is sensitive to the horizontal transmission rate. Thus, measures reducing the latter are recommended. Similarly, slaughter of preclinically infected animals is able to reduce R_0 by over 90%. This paper thus encourages using early diagnostic tests as effective control measures. Finally, this model allows genetic control measures to be evaluated, and predicts that inbreeding may increase R_0 if the susceptibility allele is recessive. Although the precise value of R_0 may be impossible to determine in a given flock, this study demonstrates the use of R_0 as an important predictor of the efficacy of control measures.

In a more recent study by Gravenor *et al.* (2004), the estimated flock-to-flock value of R_0 for scrapie in Cyprus was between 1.4 and 1.8. This model uses a four-compartment ODE system, and evaluates R_0 using the survival function. The model is then fitted to weekly incidence data to estimate three unknown parameters.

This study also investigates the impact of interventions, estimating both the epidemiological impact and the cost of each intervention. The usefulness of each control measure, however, is gauged not by changes in R_0, but by estimating the total number of farms affected by the epidemic. The estimate of R_0 in this paper is thus somewhat peripheral to the main conclusions of the work.

5.2.3. Foot and mouth disease. Determining the magnitude of R_0 for FMD has also proved important, guiding policies for culling and vaccination, the two major control measures implemented for FMD.

Ferguson *et al.* (2001) determined R_0 for FMD by considering contact tracing data and the number of susceptibles at equilibrium. They found that $R_0 \approx 4.5$ and that is reduced to approximately 1.6 when control measures were implemented. Also, by developing a model of differential equations to describe FMD dynamics and fitting this model to R_0 values over time, they were able to conclude that slaughtering on all farms within 24 h of case reporting (without necessarily waiting for laboratory confirmation) can significantly slow the epidemic. However, they found that even these improvements in slaughter times did not reduce R_0 below one. They concluded that it is necessary to consider other interventions, especially those capable of rapidly controlling infections established in multiple regions.

Ring culling and vaccination were also explored using the model. Ferguson *et al.* concluded that both are highly effective strategies if implemented rigorously, but that this may be very costly. The high initial value of R_0 estimated in this study confirmed that FMD is highly transmissible, and estimates of R_0 were essential in determining which control measures might be effective against this pathogen.

Matthews *et al.* (2003) extended previous models of FMD by defining an optimal control policy. This policy included removing newly discovered infected holdings and the pre-emptive removal of holdings deemed to be at enhanced risk of infection. Matthews *et al.* employed a simple SIR model to determine the magnitude of the effect of different control policies on a chosen value of R_0. They found, not surprisingly, that the level of control required to minimize the number of animals removed increases with R_0. They also found that non-zero levels of control can optimize the outcome of the epidemic even when $R_0 < 1$. In this case, the impact of the control measure was assessed using the fraction of animals removed.

Extending their model to a metapopulation, Matthews *et al.* concluded that a greater level of

control is needed in this case, but most importantly, they found that to minimize losses to livestock populations, R_0 should be only *sufficiently* reduced; there is a tradeoff between the amount by which R_0 can be reduced and the fraction of animals removed. The key points which emerge are that total losses are not highly sensitive to small variations in the control effort around the optimal values, and that losses increase only gradually as control effort increases beyond the optimal value. They concluded that some leeway is acceptable in practice, but that over-control is generally safer than under-control when trying to avoid large losses to the population. Similar arguments were also applied for variation in R_0; that is, over-control should be implemented if there is any uncertainty or variability in the value of R_0.

5.3. Vector-borne disease

5.3.1. Dengue. Luz *et al.* (2003) used R_0 to evaluate the risk of dengue fever outbreaks in Rio de Janeiro, and to assess possible control measures. R_0 was calculated from the survival function, assuming two spatial compartments with high and low vector density, respectively. Latin hypercube sampling of probability density functions was used to explore the effects of uncertain parameter values.

The goal of this paper was not so much to calculate an accurate value of R_0, but to assess which of the many unknown parameter values are most important to the model. Luz *et al.* concluded that field estimates of mosquito mortality and the incubation period of dengue in mosquitoes are of critical importance. We note that although dengue is a vector-borne disease and multiple classes of infectives are defined in the model, the definition of R_0 used here is the number of infected humans per infected human, not the square root of this value as would be obtained by the next generation operator.

5.3.2. Malaria. Although quantifying the incidence and spread of malaria has an extremely rich history (Garrett-Jones 1964), work in characterizing R_0 for malaria is ongoing. A recent paper investigates the incidence of malaria on an island in the Gulf of Guinea with a population of 6000 (Hagmann *et al.* 2003).

The paper estimates the 'vectorial capacity' of malaria (Garrett-Jones 1964), that is, the rate at which future human infections arise from a currently infective human host. This capacity C is estimated using maximum likelihood fits to observed age-prevalence data, and R_0 is predicted as a function of C. We note once again that the value of R_0 thus obtained corresponds to the definition provided in §2.1, not that of §2.2. The paper also reports detailed incidence data, stratified by age, sex, residence of patient and grade of malarial infection. Finally, a detailed survey on the use of mosquito nets, dwelling types, etc., was conducted; fully 17% of the population participated in the survey.

The low value of R_0 obtained in this study (1.6) was used to justify the overall conclusion of the work that malaria can probably be eliminated from the island through simple control measures. However, calculating R_0 was otherwise incidental; arguably the most important findings in this study were obtained through the detailed surveying and reporting of incidence and demographic data.

5.3.3. West Nile virus. Wonham *et al.* (2004) derived a system of ODEs to describe the behaviour of West Nile virus. Their model consisted of susceptible, infectious, recovered and dead birds, and larval, susceptible, exposed and infectious mosquitoes. The next generation method was used to calculate R_0 from this model in order to evaluate the ability of the virus to invade the system. The calculated value of R_0 was then interpreted biologically as the square root of the product of (i) the disease R_0 from mosquitoes to birds and (ii) the R_0 from birds to mosquitoes. Each of these R_0 values was further analysed as a product of disease transmission and infectious lifespan in case (i) and the product of the transmission probability, the number of initially susceptible mosquitoes per bird that survive the exposed period and the bird's infectious lifespan in case (ii). R_0 was then used to establish a threshold mosquito level, above which the virus will invade a constant population of susceptible mosquitoes.

The R_0 value derived was then used to evaluate public health policy markers. Two such policies were evaluated: mosquito control and bird control. It was demonstrated that a small increase in mosquito mortality can lead to a disproportionately large increase in the outbreak threshold. More surprisingly, however, R_0 was used to show that reducing crow densities would have the opposite effect and actually enhance disease transmission (unless extremely low densities limited mosquito biting rates). Thus, R_0 was used to show that reducing the initial mosquito population below the calculated threshold would have prevented the West Nile outbreak for New York in 2000. Conversely, bird control would have had the opposite effect.

6. DISCUSSION

Our review of the practical use of R_0 has focused, largely, on literature from a 2 year period, 2003 and 2004. The number of papers included here—and our review was by no means exhaustive—testifies to the current relevance of this important concept.

The methods used to calculate R_0 from incidence data vary. Model fitting using standard optimization techniques is often used to estimate parameters, which are then used to determine R_0 by either the survival function or next generation methods. Estimating the initial growth rate, r_0, has also been widely used. For multiple classes of infectives (e.g. vector-borne disease), we find examples both where R_0 is defined per generation, and examples where it is defined per infection cycle (see §2.2). Owing to the usual limitations in using real data, we note that models typically used 'in the field' are simple, deterministic and non-structured (but see Ferguson *et al.* 1999; Lloyd-Smith

et al. 2003; Matthews *et al.* 2003; and Riley *et al.* 2003 for counter examples).

The basic reproductive ratio for emerging or endemic pathogens described above has been estimated for two main purposes. First, R_0 is estimated in order to gauge the relative risk associated with a pathogen. These estimates are then used to compare the transmissibility of the disease to other well-known (and better understood) pathogens. Unfortunately, some time is needed to accrue sufficient incidence data for these estimates of R_0, and typically, R_0 is only quantified after the epidemic has run its course, or is at least well established. The degree to which R_0 for one emerging infectious agent might be predictive of future novel pathogens is questionable (Mills *et al.* 2004). Furthermore, a numerical estimate of R_0 for a specific disease does not, in and of itself, inform public health measures. These values are instead used to justify severe or costly control measures (e.g. FMD; Ferguson *et al.* 2001; Matthews *et al.* 2003), or less severe, more sustainable measures (e.g. malaria on Principe; Hagmann *et al.* 2003).

Evaluating these control measures reveals the second, and more important, use of R_0 in the recent literature. In most of the studies outlined above, R_0 is evaluated both before and after a putative control measure is applied, with the aim of determining which measures, at what magnitudes and in what combinations, are able to reduce R_0 to a value less than one. The results of these efforts have clearly offered useful practical guidelines: in some cases the results are counter-intuitive (e.g. West Nile virus: Wonham *et al.* 2004), in many cases they are sobering.

Although R_0 offers a simple, universal measure of control efficacy, it is important to note that using R_0 for this purpose ignores other important issues, such as the timing of secondary infections, or the negative impact of control measures on the population. For example, it is possible that some patterns of quarantine may be roughly equivalent in their effect on R_0, but may have different effects on the growth rate of the epidemic. Matthews *et al.* (2003) discuss the trade-off between reducing R_0 and culling as few animals as possible; Lipsitch *et al.* (2003) discuss similar trade-offs between reducing R_0 and burdening the population with excessive quarantine. These studies suggest that R_0 may not always be the best overall measure of control efficacy. In contrast, the total mortality or morbidity, the total number of affected farms and other such measures may offer more practical indicators of control success, and can be balanced against the associated costs (e.g. Gravenor *et al.* 2004). We argue that R_0 may be somewhat overused in evaluating control measures, presumably because it is more readily calculated than these alternative indicators, and is widely recognized and understood.

For host–pathogen interactions, R_0 stresses the role of the pathogen. An alternative, more host-centred characterization has been suggested by Bowers (2001). Nicknamed the basic depression ratio, D_0 measures the degree to which the infected host population is depressed below its uninfected equilibrium level. Consideration of both R_0 and D_0 allows modelling of the complex trade-offs in the evolution of host–pathogen interactions.

When control is targeted at specific subgroups, R_0 is not a good indicator of the required control effort, and the type-reproduction number, T, has been suggested instead (Roberts & Heesterbeek 2003; Heesterbeek & Roberts in press). This quantity is equivalent to R_0 in homogeneous populations, but in heterogeneous populations it singles out the control effort required to achieve eradication when control is targeted towards a particular host type (or subset of types), rather than the population as a whole, assuming the other types cannot sustain an epidemic by themselves. In many cases, T is easier to formulate than R_0 and both share the same threshold behaviour.

This work followed from a mini-symposium on the same topic at the Society for Mathematical Biology Annual Meeting, Ann Arbor, Michigan, 2004. We are indebted to Hans Heesterbeek and an anonymous referee for a number of insightful suggestions; we also thank Sally Blower, Erin Bodine, Romulus Breban, Elissa Schwartz, Raffaello Vardavas and David Wilson for valuable discussions. We are also grateful to the Natural Sciences and Engineering Research Council of Canada and to the Ontario Ministry of Science, Technology and Industry for their support.

REFERENCES

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press.

Barbour, A. D. & Kafetzaki, M. 1993 A host–parasite model yielding heterogeneous parasite loads. *J. Math. Biol.* **31**, 157–176.

Blower, S. M. & Chou, T. 2004 Modeling the emergence of the "Hot Zones": tuberculosis and the amplification dynamics of drug resistance. *Nat. Med.* **10**, 1111–1116.

Blower, S. M., Porco, T. C. & Darby, G. 1998 Predicting and preventing the emergence of antiviral drug resistance in HSV-2. *Nat. Med.* **4**, 673–678.

Böckh, R. 1886 *Statistisches Fahrbuch der Stadt Berlin, Zwölfter Jahrgang. Statistik des Jahres 1884,* pp. 30–31. Berlin: P. Stankiewicz.

Bowers, R. G. 2001 The basic depression ratio of the host: the evolution of host resistance to microparasites. *Proc. R. Soc. B* **268**, 243–250. (doi:10.1098/rspb.2000.1360.)

Brauer, F. 2002 Basic ideas of mathematical epidemiology. In *Mathematical approaches for emerging and reemerging infection diseases: an introduction* (ed. C. Castillo-Chavez, P. van den Driessche, D. Kirschner & A.-A. Yakubu). The IMA Volumes in Mathematics and its Applications, vol. 125, pp. 31–65. New York: Springer.

Castillo-Chavez, C., Feng, Z. & Huang, W. 2002 On the computation of R_0 and its role in global stability. In *Mathematical approaches for emerging and reemerging infection diseases: an introduction* (ed. C. Castillo-Chavez, P. van den Driessche, D. Kirschner & A.-A. Yakubu). The IMA Volumes in Mathematics and its Applications, vol. 125, pp. 229–250. New York: Springer.

Choi, B. C. K. & Pak, A. W. P. 2003 A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths. *J. Epidemiol. Commun.* **57**, 831–835.

Chowell, G., Genimore, P. W., Castillo-Garsow, M. A. & Castill-Chavez, C. 2003 SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. *J. Theor. Biol.* **224**, 1–8.

Chowell, G., Castillo-Chavez, C., Fenimore, P. W., Kribs-Zaleta, C. M., Arriola, L. & Hymnan, J. M. 2004*a* Model parameters and outbreak control for SARS. *Emerg. Infect. Dis.* **10**, 1258–1263.

Chowell, G., Hengartner, N. W., Castillo-Chavez, C., Fenimore, P. W. & Hyman, J. M. 2004*b* The basic reproductive ratio number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *J. Theor. Biol.* **229**, 119–126.

de Koeijer, A., Heesterbeek, H., Schreuder, B., Oberthür, R., Wilesmith, J., van Roermund, H. & de Jong, M. 2004 Quantifying BSE control by calculating the basic reproduction ratio R_0 for the infection among cattle. *J. Math. Biol.* **48**, 1–22.

Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation.* New York: Wiley.

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. *J. Math. Biol.* **35**, 503–522.

Dietz, K. 1975 Transmission and control of arbovirus diseases. In *Epidemiology* (ed. D. Ludwig & K. L. Cooke), pp. 104–121. Philadelphia: Society for Industrial and Applied Mathematics.

Dietz, K. 1993 The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* **2**, 23–41.

Dublin, L. I. & Lotka, A. J. 1925 On the true rate of natural increase of a population. *J. Am. Stat. Assoc.* **20**, 305–339.

Dushoff, J. 1996 Incorporating immunological ideas in epidemiological models. *J. Theor. Biol.* **180**, 181–187.

Dushoff, J., Huang, W. & Castillo-Chavez, C. 1998 Backwards bifurcations and catastrophe in simple models of fatal diseases. *J. Math. Biol.* **36**, 227–248.

Ferguson, N. M., Donnelly, C. A., Woolhouse, M. E. J. & Anderson, R. M. 1999 Estimation of the basic reproduction number of BSE: the intensity of transmission in British cattle. *Proc. R. Soc. B* **266**, 23–32. (doi:10.1098/rspb.1999.0599.)

Ferguson, N. M., 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.

Gail, M. H. & Rosenberg, P. S. 1992 Perspective on using backcalculation to estimate HIV prevalence and projects AIDS incidence. In *AIDS epidemiology: methodological issues* (ed. N. P. Jewell, K. Dietz & V. T. Farewell), pp. 1–38. Boston: Birkhauser.

Garrett-Jones, C. 1964 Prognosis for interruption of malaria transmission through assessment of the mosquito's vectorial capacity. *Nature* **4964**, 1173–1175.

Gravenor, M. B., Papasozomenos, P., McLean, A. R. & Neophytou, G. 2004 A scrapie epidemic in Cyprus. *Epidemiol. Infect.* **132**, 751–760.

Hagmann, R., Charlwood, J. D., Gil, V., Conceicao, F., do Rosario, V. & Smith, T. A. 2003 Malaria and its possible control on the island of Principe. *Malar. J.* **2** published online 18 June 2003.

Heesterbeek, J. A. P. 2002 A brief history of R_0 and a recipe for its calculation. *Acta Biotheret.* **50**, 189–204.

Heesterbeek, J. A. P. & Dietz, K. 1996 The concept of R_0 in epidemic theory. *Stat. Neerl.* **50**, 89–110.

Heesterbeek, J. A. P., Roberts, M. G. In press. The type-reproduction number T in models for infectious disease control. *Math. Biosci.*

Heffernan, J. M., Wahl, L. M. In press. Monte Carlo estimates of natural variation in HIV infection. *J. Theor. Biol.* (doi:10.1016/jojtbi/2005.03.002.)

Hethcote, H. W. 1975 Mathematical models for the spread of infectious diseases. In *Epidemiology* (ed. D. Ludwig & K. L. Cooke), pp. 122–131. Philadelphia: Society for Industrial and Applied Mathematics.

Hethcote, H. W. 1978 An immunization model for a heterogeneous population. *Theor. Popul. Biol.* **14**, 338–349.

Hethcote, H. W. 2000 The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653.

Hethcote, H. W. & Thieme, H. R. 1985 Stability of the endemic equilibrium in epidemic models with subpopulations. *Math. Biosci.* **75**, 205–227.

Hethcote, H. W. & Tudor, D. W. 1980 Integral equation models for endemic infectious diseases. *J. Math. Biol.* **9**, 37–47.

Hill, A. N. & Longini Jr, I. M. 2003 The critical vaccination fraction for heterogeneous epidemic models. *Math. Biosci.* **181**, 85–106.

Huang, Y., Rosenkranz, S. L. & Wu, H. 2003 Modeling HIV dynamics and antiviral response with consideration of time-varying drug exposures, adherence and phenotypic sensitivity. *Math. Biosci.* **184**, 165–186.

Hyman, J. M. & Li, J. 2000 An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations. *Math. Biosci.* **167**, 65–86.

Kaplan, E. H., Craft, D. L. & Wein, L. M. 2002 Emergency response to a smallpox attack: the case for mass vaccination. *Proc. Natl Acad. Sci. USA* **99**, 10 935–10 940.

Kawaguchi, I., Sasaki, A. & Mogi, M. 2004 Combining zooprophylaxis and insecticide spraying: a malaria-control strategy limiting the development of insecticide resistance in vector mosquitoes. *Proc. R. Soc. B* **271**, 301–309. (doi:10.1098/rspb.2003.2575.)

Keeling, M. J. 1999 The effects of local spatial structure on epidemiological invasion. *Proc. R. Soc. B* **266**, 859–867. (doi:10.1098/rspb.1999.0716.)

Kermack, W. O. & McKendrick, A. G. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc.* **115**, 700–721.

Kuczynski, R. R. 1928 *The balance of births and deaths, vol. 1.* New York: Macmillan.

Laurenson, K., Silleo-Zubiri, C., Thompson, H., Shiferaw, F., Thirgood, S. & Malcolm, J. 1998 Disease as a threat to endangered species: Ethiopian wolves, domestic dogs and canine pathogens. *Anim. Conserv.* **1**, 273–280.

Laxminarayan, R. 2004 Act now or later? Economics of malaria resistance. *Am. J. Trop. Med. Hyg.* **71**(Suppl. 2), 187–195.

Lipsitch, M. *et al.* 2003 Transmission dynamics and control of severe acute respiratory syndrome. *Science* **300**, 1966–1970.

Lloyd, A. L. 2001*a* The dependence of viral parameter estimates on the assumed viral life cycle: limitations of studies of viral load data. *Proc. R. Soc. B* **268**, 847–854. (doi:10.1098/rspb.2000.1592.)

Lloyd, A. L. 2001*b* Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proc. R. Soc. B* **268**, 985–993. (doi:10.1098/rspb.2001.1599.)

Lloyd, A. L. & May, R. M. 1996 Spatial heterogeneity in epidemic models. *J. Theor. Biol.* **179**, 1–11.

Lloyd, A. L. & May, R. M. 2001 How viruses spread among computers and people. *Science* **292**, 1316–1317.

Lloyd-Smith, J. O., Galvani, A. P. & Getz, W. M. 2003 Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. *Proc. R. Soc. B* **270**, 1979–1989. (doi:10.1098/rspb.2003.2481.)

Longini Jr, I. M., Halloran, M. E., Nizam, A. & Yang, Y. 2004 Containing pandemic influenza with antiviral agents. *Am. J. Epidemiol.* **159**, 623–633.

Luz, P. M., Codeco, C. T., Massad, E. & Struchiner, C. J. 2003 Uncertainties regarding dengue modeling in Rio de Janeiro, Brazil. *Mem. Inst. Oswaldo Cruz.* **98**, 871–878.

MacDonald, G. 1952 The analysis of equilibrium in malaria. *Trop. Dis. Bull.* **49**, 813–829.

Matthews, L., Woolhouse, M. E. J. & Hunter, N. 1999 The basic reproduction number for scrapie. *Proc. R. Soc. B* **266**, 1085–1090. (doi:10.1098/rspb.1999.0747.)

Matthews, L., Haydon, D. T., Shaw, D. J., Chase-Topping, M. E., Keeling, M. J. & Woolhouse, M. E. J. 2003 Neighbourhood control policies and the spread of infectious diseases. *Proc. R. Soc. B* **270**, 1659–1666. (doi:10.1098/rspb.2003.2429.)

Mills, C. E., Robins, J. M. & Lipsitch, M. 2004 Transmissibility of 1918 pandemic influenza. *Nature* **432**, 904–906.

Moghadas, S. M. 2004 Modelling the effect of imperfect vaccines on disease epidemiology. *Discrete Contin. Dynam. Syst. B* **4**, 999–1012.

Mollison, D. 1995*a* The structure of epidemic models. In *Epidemic models: their structure and relation to data* (ed. D. Mollison), pp. 17–33. Cambridge: Cambridge University Press.

Mollison, D. (ed.) 1995*b* Epidemic models: their structure and relation to data. Cambridge: Cambridge University Press.

Murphy, B. M., Singer, B. H., Anderson, S. & Kirschner, D. 2002 Comparing epidemic tuberculosis in demographically distinct heterogeneous populations. *Math. Biosci.* **180**, 161–185.

Nasell, I. 1995 The threshold concept in stochastic epidemic and endemic models. In *Epidemic models: their structure and relations to data* (ed. D. Mollison), pp. 71–83. Cambridge: Cambridge University Press.

Nold, A. 1980 Heterogeneity in disease-transmission modelling. *Math. Biosci.* **52**, 227–240.

Nowak, M. A. & May, R. M. 2000 *Virus dynamics*. Oxford: Oxford University Press.

Nowak, M. A. *et al.* 1997 Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *J. Virol.* **71**, 7518–7525.

Porco, T. C. & Blower, S. M. 1998 Designing HIV vaccination policies: subtypes and cross-immunity. *Interfaces* **28**, 167–190.

Porco, T. C. & Blower, S. M. 2000 HIV vaccines: the effect of the mode of action on the coexistence of HIV subtypes. *Math. Popul. Stud.* **8**, 205–229.

Pybus, O. G., Charleston, M. A., Gupta, S., Rambaut, A., Holmes, E. C. & Harvey, P. H. 2001 The epidemic behaviour of the Hepatitis C virus. *Science* **292**, 2323–2325.

Riley, S. *et al.* 2003 Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* **300**, 1961–1966.

Roberts, M. G. & Heesterbeek, J. A. P. 2003 A new method for estimating the effort required to control and infectious disease. *Proc. R. Soc. B* **270**, 1359–1364. (doi:10.1098/rspb.2003.2339.)

Ross, R. 1911 *The prevention of malaria*. London: John Murray.

Rudnick, S. N. & Milton, D. K. 2003 Risk of indoor airborne infection transmission estimated from carbon dioxide concentration. *Indoor Air* **13**, 237–245.

Rushton, S. & Mautner, A. J. 1955 The deterministic model of a simple epidemic for more than one community. *Biometrika* **42**, 126–132.

Sharp, F. R. & Lotka, A. J. 1911 A problem in age distribution. *Phil. Mag.* **6**, 435–438.

Smith, D. & Keyfitz, N. 1977 *Mathematical demography: selected papers. Biomathematics*, vol. 6. Berlin: Springer.

Stegeman, A., Bouma, A., Elbers, A. R., de Jong, M. C., Nodelijk, G., de Klerk, F., Koch, G. & van Boven, M. 2004 Avian influenza A virus (H7N7) epidemic in The Netherlands in 2003: course of the epidemic and effectiveness of control measures. *J. Infect. Dis.* **190**, 2088–2095.

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

Wonham, M. J., de-Camino-Beek, T. & Lewis, M. A. 2004 An epidemiological model for West Nile virus: invasion analysis and control applications. *Proc. R. Soc. B* **271**, 501–507. (doi:10.1098/rspb.2003.2608.)

Woolhouse, M. E. J. & Anderson, R. M. 1997 Understanding the epidemiology of BSE. *Trends Microbiol.* **5**, 421–424.