



The type-reproduction number T in models for infectious disease control

J.A.P. Heesterbeek ^{a,*}, M.G. Roberts ^b

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

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

Received 31 March 2004; received in revised form 21 September 2004; accepted 28 October 2004

Available online 10 March 2006

Abstract

A ubiquitous quantity in epidemic modelling is the basic reproduction number R_0 . This became so popular in the 1990s that ‘All you need know is R_0 !’ became a familiar catch-phrase. The value of R_0 defines, among other things, the control effort needed to eliminate the infection from a homogeneous host population, but can be misleading when applied to a heterogeneous population for the same purpose. We have defined the type-reproduction number T for an infectious disease, and shown that this not only has the required threshold behaviour, but also correctly determines the critical control effort for heterogeneous populations. The two quantities coincide for homogeneous populations. In this paper we further develop the new threshold quantity as an indicator of control effort required in a system where multiple types of individuals are recognised when control targets a specific type.

© 2007 Published by Elsevier Inc.

1. Introduction

In Roberts and Heesterbeek [1], we introduced a new threshold quantity for use in the analysis of models for infectious disease control. It is defined for situations where a discrete number of

* Corresponding author.

E-mail address: j.a.p.heesterbeek@vet.uu.nl (J.A.P. Heesterbeek).

types of individuals are distinguished by some epidemiologically significant characteristic, for example species, age group, or route by which they were infected. For homogeneous populations the quantity coincides with the well-known basic reproduction number R_0 ; in heterogeneous situations it is related to R_0 , but singles out the control effort needed when control is targeted at a particular host type (or a subset of types) rather than at the population as a whole. We refer to the quantity as T when a single type is targeted. We now propose to call this quantity the *type-reproduction number*.

In our previous paper, we introduced T and showed that the relation: $T < 1 \iff R_0 < 1$ holds. We then applied the new concept to one class of possible control interventions, to be defined below as *S-control*. With S-control we mean that control acts primarily on reducing the availability of susceptibles of the target type, for example by vaccination, or the pre-emptive culling of hosts in the control of animal infections. We dealt rigorously with a number of properties of T in that context, notably that in the simplest case of a perfect vaccine, an epidemic is prevented in the vaccinated population if we vaccinate at least a fraction $1 - \frac{1}{T}$ of the susceptibles of the target type. This ‘control relation’ is well known for R_0 in a homogeneous population, but we will see that no such simple explicit relation in term of R_0 exists for structured populations, even in the situation with only two types. In the present paper we extend the theory in that we also deal with another class of control measures, to be defined below as *I-control*. With I-control we mean that control acts primarily on the infectives of the target type, with measures aimed at reducing infectious output or the length of the infectious period being prime examples. We see that the definition of T leads directly to the quantification of the control effort needed.

In Section 2 we introduce the type-reproduction number in the simplest case of a population with two types of host individuals, and explain the differences between this quantity and R_0 . We introduce and analyse the S-control situation. In Section 3 we develop the theory for the situation with n types of host individuals and derive ‘control relations’ for special cases of I-control. We do not give proofs for control relations of the S-control case, since these are already contained in Roberts and Heesterbeek [1]. In Section 4 we derive the control relation for the I-control case in detail. We conclude with a discussion of the possible usefulness, some caveats, and a primary focus for future research that is needed in order for the type-reproduction number to become a useful quantity in practice.

2. The simplest case of two types

Suppose we have two types of individuals, and interpret these for the purpose of illustration as host (type 1) and vector (type 2). We will always number the types in such a way that control is aimed at type 1. The next-generation matrix K (introduced in [2], for a full treatment see [3]) gives the next generation of infectious individuals distributed over the different types, as a matrix acting on the present infection generation. The element k_{ij} of K may be interpreted as the expected number of cases of type i produced by one infected individual of type j . Matrix K is therefore positive by definition, and we will assume throughout this paper that it is also irreducible. The basic reproduction number R_0 is the dominant eigenvalue of K .

First assume that hosts do not infect hosts and that vectors do not infect vectors. We then have

$$K = \begin{pmatrix} 0 & k_{12} \\ k_{21} & 0 \end{pmatrix}, \quad R_0 = \sqrt{k_{12}k_{21}}.$$

Suppose we concentrate control effort on the host. For example we vaccinate a fraction v of susceptible hosts with a perfect vaccine and k_{12} is linear in the number of susceptibles. We can then calculate the reproduction number in the population where this control measure is active (R_v) as the dominant eigenvalue of the matrix

$$K_v = \begin{pmatrix} 0 & (1-v)k_{12} \\ k_{21} & 0 \end{pmatrix}, \quad R_v = \sqrt{(1-v)k_{12}k_{21}} = R_0\sqrt{1-v}.$$

We then derive immediately that invasion of the infectious agent will be unsuccessful, i.e. $R_v < 1$ and a large epidemic will be prevented, if and only if

$$v > 1 - \frac{1}{R_0^2}.$$

We are interested in seeing what happens when an infection chain can cycle in one type before going back to the other type. For example, if the agent could be transmitted vertically in the vector, we would have $k_{22} > 0$. This is assuming that vertically infected vectors are, epidemiologically speaking, otherwise not different from horizontally infected vectors. The reason for making this assumption is for expositional purposes only, when vertically infected vectors behave differently (e.g. by having a longer infectious period, a lower life expectancy, et cetera), it makes more sense to introduce them as a third type. We would then obtain a 3×3 -matrix, see Roberts and Heesterbeek [1] for further details. The theory still holds, but as an example it is unnecessarily complicated.

If we repeat the argument above with $k_{22} > 0$ we find

$$K = \begin{pmatrix} 0 & k_{12} \\ k_{21} & k_{22} \end{pmatrix}, \quad R_0 = \frac{1}{2}(k_{22} + \sqrt{k_{22}^2 + 4k_{12}k_{21}}).$$

and, for the control effort targeted at the hosts as before

$$K_v = \begin{pmatrix} 0 & (1-v)k_{12} \\ k_{21} & k_{22} \end{pmatrix}, \quad R_v = \frac{1}{2}(k_{22} + \sqrt{k_{22}^2 + 4(1-v)k_{12}k_{21}}).$$

A large epidemic in the population will be prevented, i.e. $R_v < 1$, if

$$v > 1 - \frac{1 - k_{22}}{k_{12}k_{21}}.$$

We then see that the simple relation with R_0 that we had before no longer exists and that R_0 cannot be used to estimate the control effort needed. Furthermore, when the number of types becomes larger than two, the next-generation matrix generally does not lead to convenient explicit expressions for R_0 . Note that when $k_{22} > 1$, i.e. when the vector population alone is capable of sustaining the agent, then no amount of control effort devoted to the host population will prevent invasion.

3. The type-reproduction number T

When we have n types of epidemiologically distinct host types, we define the type-reproduction number T as

$$T = e'K(I - (I - P)K)^{-1}e$$

where I is the identity matrix, e is the first unit vector and P is the projection matrix on type 1 (i.e. $p_{11} = 1$, and $p_{ij} = 0$ for all other entries). We refer to Roberts and Heesterbeek [1] for the derivation of the main property

$$T < 1 \iff R_0 < 1$$

provided K is irreducible. A few remarks are in order:

- The interpretation of T is as follows: it is the expected number of cases in individuals of type 1, caused by one infected individual of type 1 in a completely susceptible population, either directly or through chains of infection passing through any sequence of the other types. The picture in Fig. 1 illustrates this.
- Generalisations to other subsets of more than one type at which the same control method is targeted are possible, see Roberts and Heesterbeek [1].
- The derivation of the explicit expression is given in [1]. In brief: Ke is the second generation vector of infecteds if we start with one infective of type 1 in generation 1. Of these we trace further any type except type 1. The third generation is then given by $K(I - P)Ke$, leading to $e'K(I - P)Ke$ expected cases of type 1 in that generation. If we do this for infinitely many generations we obtain a series that, when it converges, converges to the expression given above. The requirement for convergence is that the spectral radius of $(I - P)K$ is less than 1. In words this means that T is well defined if the types $2, \dots, n$, at which the control effort is not aimed, cannot sustain an epidemic by themselves. Note that if these types could sustain the agent, then no amount of control effort aimed at type 1 alone would be able to prevent an epidemic.

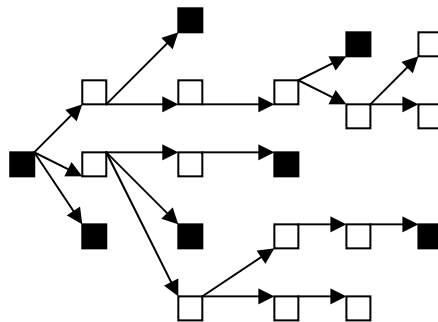


Fig. 1. Schematic definition of the type-reproduction number T . Infected individuals of type 1 are indicated by black squares, infecteds of type 2 by white squares. Shown are the first generations of the epidemic tree emanating from one individual of type 1, where the branches from any new infections of type 1 that arise are cut. For T we count all black squares arising in this way, over all generations.

Up to now we have concentrated on S-control: reducing the fraction of susceptibles of type 1 that are available for transmission from infecteds of all types in the population. Examples are vaccination, isolation, preemptive culling, prophylaxis. S-control reduces the number of cases in type 1 coming from all source types. For S-control with control effort v , we replace the elements of the first row of the next generation matrix k_{1j} with $(1 - v)k_{1j}$ for all $j = 1, \dots, n$ and $0 \leq v \leq 1$. In general one could say that S-control leads to a next-generation matrix K_v , and a corresponding type-reproduction number. Direct calculation gives that this type-reproduction number is given by $T_v = (1 - v)T$. The effort devoted to type-1 individuals alone is sufficient to prevent an epidemic in the whole population when v is such that $(1 - v)T < 1$, and a large epidemic is therefore prevented by S-control of type 1 only when

$$v > 1 - \frac{1}{T}.$$

We recover in general the simple well-known formula, which in terms of R_0 only holds for homogeneous populations. In fact for the simple example in Section 2 we have $T = k_{12}k_{21}/(1 - k_{22})$.

There are different control options, however, which have other effects on the entries of the next-generation matrix. In what we call I-control, control effort is aimed at the infecteds of target type 1. This class of control measure could increase the rate at which infected individuals of type 1 are removed (either to a non-infectious state or removed completely from the system), or could prevent them from transmitting infection. These methods therefore reduce the potential of type 1 infecteds to produce cases of all types. In an analogous way to our analysis of S-control, we can say that with control effort w , we replace k_{i1} with $(1 - w)k_{i1}$ for all $i = 1, \dots, n$ and $0 \leq w \leq 1$. The two next generation matrices are equal when $w = 0$, and total control is achieved when $w = 1$.

4. The case of I-control

In the simple host-vector example of Section 2 we can illustrate I-control by discussing control efforts aimed at the vector population. Vectors are now type 1, hosts are designated by type 2. Assume infected vectors have a removal rate μ_1 and produce β_{11} vertically infected offspring per unit of time, and β_{21} infected hosts per unit of time. We can regard

$$(K_w)_{i1} = \frac{(1 - w)\beta_{i1}}{\mu_1} \quad i = 1, 2, \quad 0 \leq w \leq 1.$$

Therefore

$$K_w = \begin{pmatrix} (1 - w)k_{11} & k_{12} \\ (1 - w)k_{21} & 0 \end{pmatrix}.$$

The factor $(1 - w)$ could represent, for example, an increase in the removal rate of infectives (i.e. a shortening of the infectious period). We then have in mind that

$$(K_w)_{i1} = \frac{\beta_{i1}}{\mu_1 + q} \quad i = 1, 2,$$

leading to the choice $w = q/(\mu_1 + q)$. For other examples, the control measure represented here might be a quarantine policy that prevents a proportion w of contacts by infectives. Calculation of T for the corresponding non-controlled matrix K leads to $T = k_{12}k_{21} + k_{11}$, and the type-reproduction number corresponding to K_w is $T_w(1 - w)T$. I-control effort aimed at type 1 alone is therefore successful if

$$w > 1 - \frac{1}{T}.$$

For n types we obtain the same result. This can be seen as follows. The $n \times n$ matrix K_w and the type-reproduction number is given by

$$T_w = e'K_wA_w^{-1}e$$

where $A_w = I - (I - P)K_w$ and hence

$$A_w = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -(1-w)k_{21} & 1-k_{22} & -k_{23} & -k_{2n} \\ \vdots & -k_{32} & \ddots & -k_{n-1,n} \\ -(1-w)k_{n1} & -k_{n2} & \cdots & 1-k_{nn} \end{pmatrix}.$$

It is therefore clear that the determinant $\Delta(A_w)$ is independent of w . Since we are only interested in element (1, 1) of $K_wA_w^{-1}$, we only need to know the first column of A_w^{-1} . The first element of A_w^{-1} is 1, and therefore the first term of $(K_wA_w^{-1})_{11}$ is $(1-w)k_{11}$. The only elements of A_w that depend on w are the elements of the first column, apart from the first element, and all dependence is through the multiplicative factor $(1-w)$. Therefore, all elements of the first column of A_w^{-1} will have $(1-w)$ as a factor, and consequently, all further terms of $(K_wA_w^{-1})_{11}$ will have $(1-w)$ as a factor. Together with the first term, and the fact that the first row of K_w does not depend on w apart from the first term, we finally have that

$$T_w = (K_wA_w^{-1})_{11} = (1-w)T$$

where T is defined as before, based on the corresponding matrix K in the absence of control.

It is clear that there are also control methods that would have effect on both the susceptibles and the infecteds of type 1. In fact, vector control is likely not to distinguish between shortening the life of infected or susceptible individuals. Control aimed at increasing the removal or death rate of vector individuals, will also lower the demographic steady state of the susceptible vector population. In terms of the above example, it will therefore also influence the element k_{12} , since this element describes the expected number of new cases among susceptible vectors, caused by one infective host. One can easily check that looking at I-control of the vector, while taking the multiplicative effect on the susceptible vectors into account, can be incorporated in the above arguments. The simple relations between the controlled and uncontrolled type-reproduction numbers are lost, however, even in the simplest case of the vector-host example above. One can check, for example, that the case

$$K_{v,w} = \begin{pmatrix} (1-w)k_{11} & (1-v)k_{12} \\ (1-w)k_{21} & 0 \end{pmatrix}.$$

has type-reproduction number $T_{v,w} = (1 - w)(k_{11} + (l - v)k_{12}k_{21})$, where one should remember that now v and w are linked, and one cannot choose both freely. For example, when we have a fixed vector birth rate b , leading to a demographic steady state b/μ_1 in the uncontrolled population, we will have a steady state $(1 - w)b/\mu_1$ in the case of additive vector control. We are then in the above situation with $v = w$, leading to an equation that is in this simple explicit case already quadratic in the control efforts w .

5. Discussion

We have shown that R_0 is not the right quantity to look at if one wishes to obtain insight into the control effort needed when targeting selected types of individuals in a heterogeneous population. In the situation where we target one host type for control, the type-reproduction number T is more closely related to the actual control effort required. T is the cumulative number of new cases of type 1, per case of type 1, from all chains of infection with no intermediate cases of type 1. In a homogeneous system $R_0 = T$, but in a heterogeneous system the two quantities only share their threshold behaviour at $R_0 = T = 1$. We have seen that in the simplest setting of control, we obtain explicit and simple relations between T and the control effort required, for two broad classes of control effects, and we obtain the same relations for any finite discrete number of types in the population.

When judging the effect of control measures one can of course also study the dominant eigenvalue R_v or R_w of the next-generation matrices K_v and K_w , respectively. The disadvantage of that approach, however, is that for situations where more than two types are recognised, one cannot get an explicit expression for the dominant eigenvalue (save for some special cases). We have shown that the type-reproduction number T serves the same purpose, but that it is defined by an explicit relation. Moreover, for simple control strategies, we find simple explicit relations between T and the control effort required, which cannot be obtained in terms of R_0 when focussing on the dominant eigenvalues.

We have focussed in this paper on simple relations where the effect of control effort v and w on the generation factors k_{ij} can be written as a linear multiplicative factor. This is not a prerequisite and other functional dependencies of the k_{ij} on control effort can be imagined for less idealised control measures.

One use of the calculation of T could be to repeat the analysis for all host types, allowing all types to become the target type for control and thus obtaining control efforts for all these cases separately. This would allow comparison, for example with additional information on cost, of the likely control targets and their relative efficacy.

We are able to look at n types of individual, but several important types of heterogeneity need a continuous type space. While it is perfectly clear how the calculation of R_0 generalises to that situation (see [2,3]), it is less immediate that the same holds for T (Barbara Boldin, unpublished results). One needs to formulate the problem in a measure-theoretical sense; since we would no longer be able to fix one target type for control, we would have to fix a set of positive measure.

Calculation of T could also be of benefit in the theory of matrix population models ([5]). Here one could calculate T targeting conservation effort at various stages in the life cycle, or age classes, to determine where the effort would have the most impact on the population growth rate.

What is lacking at the moment is insight into ways in which a quantity like T could be estimated from data. In the history of the basic reproduction number, it was certainly this aspect that brought to R_0 much wider applicability (see [4], for a historical account). We should investigate whether T can be linked to directly measurable population or demographic parameters to obtain estimates. It should be noted that these relationships exist for R_0 , but most rely on some assumption of homogeneity (e.g. the inverse of the endemic fraction of susceptibles, the ratio of life expectancy and average age at infection). For T we have, by definition, to deal with a heterogeneous population, which makes the problem much harder.

In the case of T we have the idea to study realisations of epidemic trees from various data sets to see whether estimates from these sources are possible and whether statistical techniques can be developed to say something of their accuracy. Such data sets are becoming available in various places where contact tracing is essential in targeting control, such as the 2001 FMD outbreak in the United Kingdom (see [6]) and the world-wide SARS outbreak.

References

- [1] M.G. Roberts, J.A.P. Heesterbeek, A new method to estimate the effort required to control an infectious disease, *Proceedings of the Royal Society London, Series B* 270 (2003) 1359.
- [2] O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, *Journal of Mathematical Biology* 28 (1990) 365.
- [3] O. Diekmann, J.A.P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, John Wiley & Sons, Chichester, 2000.
- [4] J.A.P. Heesterbeek, A brief history of R_0 and a recipe for its calculation, *Acta Biotheoretica* 50 (2002) 189.
- [5] H. Caswell, *Matrix Population Models: Construction, Analysis and Interpretation*, second ed., Sinauer Associates, Sunderland, 2001.
- [6] D.T. Haydon, M. Chase-Topping, D.J. Shaw, L. Matthews, J.K. Friar, J. Wilesmith, M.E.J. Woolhouse, The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak, *Proceedings of the Royal Society London, Series B* 270 (2003) 121.