Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

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

Department of Analysis, Algebra and Geometry Report AM-R8912 August

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models for Infectious Diseases in Heterogeneous Populations

O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands
&
Institute of Theoretical Biology, Leiden University
Kaiserstraat 63, 2311 GP Leiden, The Netherlands

The expected number of secondary cases produced by a *typical* infected individual during its entire period of infectiousness is mathematically defined as the dominant eigenvalue of a positive linear operator. It is shown that in certain special cases one can easily compute or estimate this eigenvalue. Several examples involving various structuring variables like age, sexual disposition and activity are presented.

Key Words & Phrases: epidemic models, heterogeneous populations, basic reproductive number, invasion
1980 Mathematics Subject Classification: 92A15.

1. Introduction

Suppose we want to know whether or not a contagious disease can "invade" into a population which is in a steady (at the time scale of disease transmission) demographic state with all individuals susceptible. To decide about this question we first of all *linearize,* i.e. we ignore the fact that the density of susceptibles decreases due to the infection process. It has become common practice in the analysis of the simplest models to consider next the associated *generation process* and to define the *basic reproduction ratio* (or reproductive number but not, as it is quite often wrongly called, "rate") R_0 as the *expected number of secondary cases* produced by a *typical* infected individual during its entire period of infectiousness. The famous *threshold criterion* then states:

the disease can invade if $R_0>1$, whereas it cannot if $R_0<1$.

It is the aim of this note to demonstrate how these ideas extend to less simple (though probably still highly oversimplified) models involving *heterogeneity* in the population and to explain the meaning of "typical" in the "definition" of R_0 above. Subsequently we shall deal with the actual computation of R_0 in certain special cases, in particular the so-called "proportionate mixing" case.

2. The definition

Let the individuals be characterized by a variable ξ, which we shall call the h-state variable (h for heterogeneity). Let $S(\xi)$ denote the density function describing the *steady* demographic state in the absence of the disease. Define $A(\tau,\xi,\eta)$ to be the expected infectivity of an individual which was infected τ units of time ago, while having h-state η, towards a susceptible which has h-state ξ. The expected number of infections produced during its entire infective life by an individual which was itself infected while having h-state η is then given by

$$\int_{\Omega} S(\xi) \int_0^{\infty} A(\tau,\xi,\eta) d\tau d\xi$$

where Ω denotes the h-state space, i.e. the domain of definition of ξ. We may call this quantity the next generation factor of η.

Report AM-R8912
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

REMARK. In order to have a unified notation for various cases we write integrals to denote sums whenever Ω is discrete (completely or just with respect to some component of ξ). A precise mathematical justification involves a dominant measure and Radon-Nikodym derivatives.

Since the new cases arise, in general, with h-states different from η, these numbers do not tell us exactly what happens under iteration, i.e. in subsequent generations (although it is clear that the supremum with respect to η yields an upper estimate for R_0).

So we abandon the idea of introducing an infected individual with a particular well-defined h-state and start instead with a "distributed" individual described by a density ϕ. The next-generation operator $K(S)$ defined by

$$(K(S)\phi)(\xi) = S(\xi)\int_{\Omega}\int_{0}^{\infty} A(\tau,\xi,\eta)d\tau\phi(\eta)d\eta \tag{2.1}$$

tells us both how many secondary cases arise from ϕ and how they are distributed over the h-state space. Ignoring the task of writing down conditions on S and A which guarantee that $K(S)$ is a bounded operator on $L_1(\Omega)$, we note that the next generation factor of ϕ is simply the $L_1(\Omega)$-norm of $K(S)\phi$, i.e.

$$\int_{\Omega} S(\xi)\int_{\Omega}\int_{0}^{\infty} A(\tau,\xi,\eta)d\tau\phi(\eta)d\eta d\xi$$

(note that we do not have to write absolute value signs since the biological interpretation requires all functions to be positive.) If we take the supremum of the next generation factor over all ϕ with $\|\phi\|=1$ we obtain, by definition, the operator norm of $K(S)$. This yields an upper estimate for R_0 for the same reason as above: the distribution with respect to ξ is changed in the next generation and consequently the factor of ϕ does not predict accurately what happens under iteration.

As a concrete example consider a host-vector model. Taking $\Omega=\{1,2\}$ we find that $K(S)$ is represented by the matrix

$$\begin{bmatrix} 0 & a_{12}S_1 \\ a_{21}S_2 & 0 \end{bmatrix}$$

and the operator norm is max $\{a_{12}S_1, a_{21}S_2\}$. These two numbers correspond to vector $\rightarrow$ host and host $\rightarrow$ vector transmission, respectively. No matter which of the two is the larger one, in the next generation it is necessarily the other of the two numbers which is the relevant factor. A moment of reflection is sufficient to realize that in the present case the *average* factor is

$$\sqrt{a_{12}S_1a_{21}S_2}$$

How can we define such a quantity in general?

After m generations the magnitude of the infected population is (in the linear approximation) $K(S)^m\phi$ and consequently the per-generation growth factor is $\|K(S)^m\|^{1/m}$. We want to know what happens to the population in the long run, so we let $m\rightarrow\infty$. The so-called *spectral radius* (SCHAEFER, 1974) $r(K(S))$ is defined by

$$r(K(S)) = \inf_{m\geq 1}\|K(S)^m\|^{1/m} = \lim_{m\rightarrow\infty}\|K(S)^m\|^{1/m} \tag{2.2}$$

Starting from the zero $^{\text{th}}$-generation ϕ, the m-th generation $K(S)^m\phi$ converges to zero for $m\rightarrow\infty$ if $r(K(S))<1$ whereas it can be made arbitrarily large by a suitable choice of ϕ and m when $r(K(S))>1$. Moreover, the *positivity* guarantees that in the latter case there is not really a restriction on ϕ. Indeed, $K(S)$ is a positive operator (i.e. nonnegative functions are mapped onto nonnegative functions) and this fact guarantees that $r(K(S))$ is an eigenvalue (SCHAEFER, 1960, 1974), which we

shall call the *dominant eigenvalue* (since $|\lambda| \leqslant r(K(S))$ for all λ in the spectrum of $K(S)$) and denote by ρ_d. Under minor technical conditions on A and S (see Remark 4 below) one has in addition that

$$K(S)^m \phi \sim c(\phi)\rho_d^m \phi_d \quad \text{for } m \to \infty \tag{2.3}$$

where ϕ_d is the corresponding eigenvector (which is positive) and $c(\phi)$ a scalar which is positive whenever ϕ is nonnegative and not identically zero. So after a certain period of transient behaviour each generation is (in an approximation which improves as time proceeds) ρ_d times as big as the preceding one and distributed over h-state space as described by ϕ_d.

If we rephrase this as: "the *typical* number of secondary cases is ρ_d ", we are ready for the

DEFINITION. $R_0 = r(K(S)) = \rho_d =$ the dominant eigenvalue of $K(S)$.

With this definition the threshold criterion remains valid, as can be verified as follows. The threshold criterion relates the generation process to the development of the epidemic in real time, both in the linearized version. The linearized real time equation is

$$i(t,\xi) = S(\xi) \int_{\Omega} \int_{0}^{\infty} A(\tau,\xi,\eta) i(t-\tau,\eta) d\tau d\eta \tag{2.4}$$

where $i(t,\xi)$ is the rate at which susceptibles with h-state ξ are infected at time t. This equation has a solution of the form $i(t,\xi) = e^{\lambda t}\psi(\xi)$ if and only if ψ is an eigenvector of the operator K_λ defined by

$$(K_\lambda \phi)(\xi) = S(\xi) \int_{\Omega} \int_{0}^{\infty} A(\tau,\xi,\eta) e^{-\lambda\tau} d\tau \phi(\eta) d\eta \tag{2.5}$$

with eigenvalue one. Positivity arguments can be used to show that among the set of such λ with *largest real part* there is a real one, which we shall denote by λ_d (and the corresponding eigenvector by ψ_d). Monotonicity arguments (see e.g. HEIJMANS (1986) p. 194) then imply that

$$\lambda_d > 0 \Leftrightarrow R_0 > 1 \text{ and } \lambda_d < 0 \Leftrightarrow R_0 < 1$$

REMARKS

1) Whereas R_0 is a number, λ_d is a rate.

2) Note that λ_d and ψ_d describe the growth and the h-state distribution in the exponential phase of an epidemic, when the influence of the precise manner in which the epidemic started has died off and the influence of the nonlinearity is not yet perceptible.

3) Let now S describe the susceptible population in a steady endemic state. Then necessarily $r(K(S)) = 1$. See Example 4.3.

4) In order to guarantee that *any* introduction of infectivity in the population leads to an epidemic when $R_0 > 1$ we need to make an *irreducibility* hypothesis (SCHAEFER, 1974).

5) To obtain a complete model one has to specify the demographic processes, and in particular how per capita birth - and death rates are affected by the disease. If one makes the obvious assumption that the disease leads to a lower (or equal) birth rate and to a higher (or equal) death rate one can use the linearized problem to obtain upper estimates for the nonlinear problem. Thus one can prove, in general, *global* rather than local stability for $R_0 < 1$. Or, in other words, endemic states are impossible when $R_0 < 1$.

6) We have restricted our attention to the bilinear case. However, replacing $S(\xi)$ in the definition (2.1) of $K(S)$ by $h(S(\xi))$ or $S(\xi)/(1+\int_\Omega S(\eta)d\eta)$ or something similar does not make any essential difference. See Examples 4.1 and 4.2 below. Note that for the invasion problem one will always have an expression involving the (known) function S only. Of course things are different if one wants to characterize endemic states, like in remark 3 above.

3. Computational aspects: easy special cases

3.1. Proportionate mixing

To compute the dominant eigenvalue of a positive operator is, in general, not an easy task. However, there is one special case in which the task is trivial: when the operator has one-dimensional range. Biologically this corresponds to the situation in which the distribution (over the h-state space Ω) of the "offspring" (i.e. the ones who become infected) is *independent* of the state of the "parent" (i.e. the one who transmits the infection). In the epidemic literature this is called *"proportionate-mixing"* (Barbour, 1978).

Assume that

$$A(\tau,\xi,\eta) = a(\xi)B(\tau,\eta) \tag{3.1}$$

then

$$(K(S)\phi)(\xi) = S(\xi)a(\xi)\int_{\Omega}\int_{0}^{\infty} B(\tau,\eta)d\tau\phi(\eta)d\eta. \tag{3.2}$$

So there can be but one eigenvector: $S(\xi)a(\xi)$. Since

$$K(S)Sa = (\int_{\Omega}\int_{0}^{\infty} B(\tau,\eta)d\tau S(\eta)a(\eta)d\eta)Sa$$

we conclude that

$$R_0 = \rho_d = \int_{\Omega}\int_{0}^{\infty} B(\tau,\eta)d\tau S(\eta)a(\eta)d\eta \tag{3.3}$$

Remarks 1. Note that the crucial assumption is that

$$\int_{0}^{\infty} A(\tau,\xi,\eta)d\tau = a(\xi)b(\eta) \tag{3.4}$$

rather than that A itself can be decomposed as in (3.1). In other words, proportionate mixing need only to occur over the generations.

2. A convenient normalization is

$$\int_{\Omega} S(\xi)a(\xi)d\xi = 1. \tag{3.5}$$

Then Sa is the probability density function for h-state at infection while $b(\xi)$ is the total expected number of "offspring" of an individual which was infected while having h-state ξ. This interpretation yields once more that

$$R_0 = \int_{\Omega} b(\eta)S(\eta)a(\eta)d\eta \tag{3.6}$$

3. Proportionate mixing in a strict sense is the special case in which a and b differ only by a multiplicative constant.

3.2. Proportionate mixing with enhanced infection within the own group.

A second case in which it is easy to derive an explicit threshold criterion, even if we cannot calculate R_0 explicitly, occurs when individuals preferentially mix with their own kind and otherwise mix proportionally. If we moreover assume that the h-state stays constant over epidemiological time (but see Example 4.3) then $K(S)$ is of the form

$$(K(S)\phi)(\xi) = S(\xi)\{a(\xi)\int_{\Omega} b(\eta)\phi(\eta)d\eta + c(\xi)\phi(\xi)\} \tag{3.7}$$

where $c(\xi)S(\xi)$ is the number of first generation "offspring" produced "directly" in one's own group.

The eigenvalue problem $K(S)\phi=\rho\phi$ can be rewritten as

$$\frac{1}{\rho-c(\xi)S(\xi)}S(\xi)a(\xi)\int_{\Omega}b(\eta)\phi(\eta)d\eta = \phi(\xi) \tag{3.8}$$

Multiplying both sides by $b(\xi)$ and integrating over Ω we obtain the characteristic equation

$$\int_{\Omega}\frac{b(\xi)S(\xi)a(\xi)}{\rho-c(\xi)S(\xi)}\,d\xi = 1 \tag{3.9}$$

The left hand side defines a decreasing function of ρ which tends to zero for $\rho\to\infty$. The largest real root R_0 is larger than one if and only if either

$$\begin{aligned}&\text{(i)}\quad c(\xi)S(\xi) > 1 \qquad \text{for some } \xi\in\Omega,\\ &\text{or, otherwise,}\\ &\text{(ii)}\quad \int_{\Omega}\frac{b(\xi)S(\xi)a(\xi)}{1-c(\xi)S(\xi)}\,d\xi > 1.\end{aligned} \tag{3.10}$$

(Of course a more precise formulation of (i) is *ess* sup $c(\xi)S(\xi)>1$.) When (i) holds a single just infected individual with h-state ξ will already start a full blown epidemic among its likes. If, on the other hand, $c(\xi)S(\xi)<1$ for all $\xi\in\Omega$ any epidemic has to be kept going by the additional cross infections among different types. To understand (ii) we distinguish cross infections and direct infections within the own group and argue as follows. As before Sa is, with the normalization (3.5), the probability density function for h-state at *cross* infection. The expected total number of cases, including its own, produced by an individual of h-state ξ through chains of infectives which stay wholly among its likes is $(1-S(\xi)c(\xi))^{-1}$. Each of these produces an expected number of cross infections equal to $b(\xi)$. So by treating the "clan's" as a kind of individuals we are back to our old proportionate mixing problem and we find

$$\int_{\Omega}b(\xi)\frac{1}{1-S(\xi)c(\xi)}S(\xi)a(\xi)d\xi \tag{3.11}$$

as the expected offspring number at the clan level. An epidemic occurs if and only if this number exceeds one.

REMARK. One of us had derived the result (3.10) in the context of the geographical spread of plant diseases (think of foci within fields). Recently our attention for this special case was revived by preprints of V. Andreasen & F.B. Christiansen (in which they derive the same result in the context of a finite h-state space) and of S.P. Blythe and C. Castillo-Chavez.

3.3. Multigroup proportionate mixing

An obvious mathematical generalization of proportionate mixing is to assume that $K(S)$ has a finite dimensional range. In general, however, this does not make biological sense. Therefore we restrict our elaboration to a special example in this category which does allow a biological interpretation.

Let ξ be of the form (i,ξ_i), where i can take the values $1,2,\ldots,n$ and ξ_i takes values in Ω_i. So $\Omega=\bigcup_{i=1}^{n}\{i\}\times\Omega_i$. Assume that

$$\int_{0}^{\infty}A(\tau,(i,\xi_i),(j,\xi_j))d\tau = a_i(\xi_i)b_{ij}(\xi_j) \tag{3.12}$$

which one could call a *local* form of proportionate mixing since, with the normalization

$$\int_{\Omega}a_i(\xi_i)S((i,\xi_i))d\xi_i = 1 \tag{3.13}$$

the *conditional* (on the first component being i) probability density function for h-state at infection is independent of the h-state of the one who infects and given by $a_i(\cdot)S((i,\cdot))$. Then

$$(K(S)\phi)(i,\xi_i) = S(i,\xi_i)a_i(\xi_i)\sum_j \int_{\Omega_j} b_{ij}(\xi_j)\phi((j,\xi_j))d\xi_j \tag{3.14}$$

and we conclude that, in order to be an eigenvector, necessarily

$$\phi(i,\xi_i) = \sigma_i S(i,\xi_i)a_i(\xi_i) \tag{3.15}$$

Substituting (3.15) into (3.14) we deduce that in addition the vector σ should be an eigenvalue of the matrix M with entries

$$m_{ij} = \int_{\Omega_j} b_{ij}(\xi_j)S(j,\xi_j)a_j(\xi_j)d\xi_j \tag{3.16}$$

In particular R_0 is the dominant eigenvalue of the matrix M.

4. Examples

4.1. Discrete and static h-state

In this case the operator $K(S)$ is represented by a matrix. We shall first show how this matrix can be derived in the special case of the conventional S-E-I-R compartment models.

Let M be the diagonal matrix of the per capita standard death rates of the various types. After infection individuals enter the exposed class E. From there they make the transition to the infective class I at a rate described by the diagonal matrix Σ whereafter they are removed at a rate described by the diagonal matrix D. Finally, let $T(S)$ be the *transmission matrix,* i.e. the matrix such that (beware: I denotes the vector of infectives, not the identity matrix)

$$\dot{E} = T(S)I - MI - \Sigma I$$

We claim that

$$K(S) = T(S)\Sigma(\Sigma+M)^{-1}(D+M)^{-1} \tag{4.1}$$

The (easy) argument goes as follows. The fraction of the infected individuals which enters I (before dying) is the diagonal of $\Sigma(\Sigma+M)^{-1}$. The mean time of staying in I is the diagonal of $(D+M)^{-1}$. While in I the transmission is described by $T(S)$.

If $M=0$ (4.1) simplifies to

$$K(S) = T(S)D^{-1} \tag{4.2}$$

(In section 9 of the paper by Jacquez et al. (1988), a special case of this matrix is introduced with $T(S)$ written out in some more detail.). Note that, as to be expected, Σ is irrelevant for the computation of R_0 in case $M=0$ even though it may, of course, have substantial influence on the magnitude of λ_d.

In the proportionate mixing case the entries of $T(S)$ are of the form

$$a_i S_i b_j$$

and according to (3.6) R_0 equals the trace of the matrix $K(S)$. See Hethcote & Yorke (1984) for another derivation of this fact.

4.2 Sexually transmitted diseases

4.2.1. Heterosexual transmission only. Let the index 1 refer to females and the index 2 to males. For each sex we distinguish individuals according to some variable ξ_i which is static (the interpretation of ξ_1 may or may not be the same as the interpretation of ξ_2). Adopting the local proportionate mixing assumption and neglecting homosexual transmission we arrive at the matrix

$$M = \begin{bmatrix} 0 & m_{12} \\ m_{21} & 0 \end{bmatrix} \tag{4.3}$$

where

$$m_{12} = \int_{\Omega_2} b_{12}(\xi_2) S_2(\xi_2) a_2(\xi_2) d\xi_2$$

$$m_{21} = \int_{\Omega_1} b_{21}(\xi_1) S_1(\xi_1) a_1(\xi_1) d\xi_1 \tag{4.4}$$

We conclude that

$$R_0 = \sqrt{m_{12} m_{21}}. \tag{4.5}$$

(See HETHCOTE & YORKE (1984) for a "discrete" version of this result)

Distinguishing not only males and females, but on top of that hetero -, *bi* - and homosexuals one easily arrives at a six by six matrix whose spectral radius one has to compute to obtain R_0.

4.2.2. Sexual activity.

Frequently the variables ξ_i are used to describe sexual activity (in the sense of: propensity to make sexual contacts), and a_i and b_{ji} are taken to be proportional to ξ_i. In the context of the heterosexual transmission model above we would, more precisely, take

$$a_1(\xi_1) = \frac{\xi_1}{\int_{\Omega_2} \xi_2 S_2(\xi_2) d\xi_2} \tag{4.6}$$

$$b_{12}(\xi_2) = \beta_{12} \xi_2$$

(with formulas for a_2 and b_{21} obtained by interchanging 1's and 2's; one may argue that $\int_{\Omega_1} \xi_1 S_1(\xi_1) d\xi_1 = \int_{\Omega_2} \xi_2 S_2(\xi_2) d\xi_2$ is required if ξ is interpreted as the actual number of sexual contacts per unit of time). Thus one arrives at

$$m_{12} = \frac{\beta_{12} \int_{\Omega_2} \xi_2^2 S_2(\xi_2) d\xi_2}{\int_{\Omega_1} \xi_1 S_1(\xi_1) d\xi_1}$$

and

$$R_0 = \left[\beta_{12} \beta_{21} \frac{\int_{\Omega_1} \xi_1^2 S_1(\xi_1) d\xi_1}{\int_{\Omega_1} \xi_1 S_1(\xi_1) d\xi_1} \frac{\int_{\Omega_2} \xi_2^2 S_2(\xi_2) d\xi_2}{\int_{\Omega_2} \xi_2 S_2(\xi_2) d\xi_2} \right]^{1/2} \tag{4.7}$$

Recalling that

$$\frac{\int_{\Omega_i} \xi_i^2 S_i(\xi_i) d\xi_i}{\int_{\Omega_i} \xi_i S_i(\xi_i) d\xi_i} = \text{mean} + \frac{\text{variance}}{\text{mean}} \tag{4.8}$$

we realize that this result is analogous to a result of Dietz (1980) and identical to formula (5.7) in May & Anderson (1988).

4.2.3. Two is worse than one.
Stimulated by work of May & Jose (to appear), as reported in MAY & ANDERSON (1988), we now investigate how the presence of some sexually transmitted disease causing ulcers and the like may enhance the possibility of the successful invasion of another sexually transmitted disease like HIV.

Assume that disease d is in an endemic steady state. We want to calculate R_0 for a disease D, assuming that the susceptibility to D is, for individuals having d, v times as large as for individuals without d. What we have in mind is that encounter rates are totally random, but that the success ratio for disease transmission, given that contact takes place, is enlarged by a factor v. Then the proportions of 0 (:= free of d) and + (:= having d) individuals that will be infected by D in the linear initial phase of an epidemic are described by the vector

$$\begin{bmatrix} S_0 \\ vS_+ \end{bmatrix}$$

where S_0 and S_+ are the steady (with respect to d) state population sizes of 0 and + individuals. As we will show below, this vector indeed spans the range of the operator $K(S)$.

Let ζ denote the force of d-infection in the steady state and let γ denote the probability per unit of time that d is cured (whereupon susceptibility to d returns). The dynamics of d are completely described by these parameters ζ and γ. Let μ denote the natural death rate.
Since

$$\frac{dS_+}{dt} = \zeta S_0 - \gamma S_+ - \mu S_+$$

we deduce that in steady state $\frac{S_+}{S_0} = \frac{\zeta}{\gamma+\mu}$ or

$$S_0 = \frac{\gamma+\mu}{\gamma+\mu+\zeta} S\,, \quad S_+ = \frac{\zeta}{\gamma+\mu+\zeta} S$$

where S denotes the total population size.

Any individual undergoes, as long as it does not die, transitions between 0 and + according to the matrix of rates

$$T = \begin{bmatrix} -\zeta & \gamma \\ \zeta & -\gamma \end{bmatrix}$$

So this is an example of a discrete but dynamic h-state.

We describe the success ratio's for D transmission by the matrix

$$p \begin{bmatrix} 1 & w \\ v & vw \end{bmatrix}$$

Here p is the success ratio when both individuals involved in the contact are free of d and $w>1$ is the factor by which the success ratio is enlarged when the D infectious individual is suffering from d. Note that we have to satisfy the requirement $pwv \leqslant 1$.

We assume that encounters occur independently of the 0 + distinction and that the rate is given by σ/S (i.e., the number of contacts per unit of time is independent of the population size).

Finally, assume that D causes an extra death rate ρ.
Then

$$K(S) = \frac{\sigma p}{S} \begin{bmatrix} S_0 & wS_0 \\ vS_+ & wvS_+ \end{bmatrix} \int_0^\infty e^{(T-\mu-\rho)\tau} d\tau$$

With the notation

$$M := \int_0^\infty e^{(T-\mu-\rho)\tau} d\tau = (\mu+\rho-T)^{-1} =$$

$$= \frac{1}{(\mu+\rho)(\mu+\rho+\gamma+\zeta)} \begin{bmatrix} \mu+\rho+\gamma & \gamma \\ \zeta & \mu+\rho+\zeta \end{bmatrix}$$

we now can express R_0 as

$$R_0 = \frac{\sigma p}{\gamma+\mu+\zeta} \begin{bmatrix} 1 \\ w \end{bmatrix}^T M \begin{bmatrix} \gamma+\mu \\ v\zeta \end{bmatrix}$$

(where x^T denotes the transpose of a vector x) or, written out in detail,

$$R_0 = \frac{\sigma p\{(\gamma+\mu)(\mu+\rho+\gamma+\zeta w) + \zeta v w(\frac{\gamma}{w}+\mu+\rho+\zeta)\}}{(\gamma+\mu+\zeta)(\mu+\rho)(\mu+\rho+\gamma+\zeta)}$$

Note that the special case $w=v=1$ yields $R_0=\frac{\sigma p}{\mu+\rho}$ as to be expected (since in that case the 0 + distinction is totally irrelevant).

Next, let us consider the variant of this model in which the sexual activity level figures as another component of the h-state. We now take, for a change, the sexual activity level as a discrete variable. The possible h-states are then $(i, 0)$ and $(i, +)$ with $i=0,1,2,3,...$ Assuming

$$\frac{dS_{(i,+)}}{dt} = i\zeta S_{(i,0)} - \gamma S_{(i,+)} - \mu S_{(i,+)}$$

we find

$$S_{(i,0)} = \frac{\gamma+\mu}{\gamma+\mu+i\zeta} S_i$$

$$S_{(i,+)} = \frac{i\zeta}{\gamma+\mu+i\zeta} S_i$$

where S_i denotes the size of class i. The relevant proportions are therefore described by the countably many two vectors

$$ip_i \begin{bmatrix} \gamma+\mu \\ iv\zeta \end{bmatrix}$$

where $p_i := \frac{S_i}{S}$ with $S := \Sigma_i S_i$, the total population size.

The transitions of the h-state are governed by the matrices

$$T_i = \begin{bmatrix} -i\zeta & \gamma \\ i\zeta & -\gamma \end{bmatrix}$$

The encounter rate of an i individual with a j individual is assumed to be

$$\frac{\sigma ij}{\sum_k k\, S_k}$$

The operator

$$K(S) = \frac{\sigma p i}{\sum_k k S_k} \begin{bmatrix} S_{(i,0)} & wS_{(i,0)} \\ vS_{(i,+)} & wvS_{(i,+)} \end{bmatrix} \sum_j j \int_0^\infty e^{(T_j-\mu-\rho)\tau}\, d\tau$$

acts on a sequence of two vectors. Its range is spanned by the sequence

$$i \begin{bmatrix} S_{(i,0)} \\ vS_{(i,+)} \end{bmatrix} \sim \frac{ip_i}{\gamma+\mu+i\zeta} \begin{bmatrix} \gamma+\mu \\ i\zeta \end{bmatrix}$$

Consequently

$$R_0 = \frac{\sigma p}{\sum_k k S_k} \begin{bmatrix} 1 \\ w \end{bmatrix}^T \sum_j j^2 (\mu+\rho-T_j)^{-1} \begin{bmatrix} S_{(j,0)} \\ \nu S_{(j,+)} \end{bmatrix}$$

$$= \frac{\sigma p}{\sum_k k p_k} \sum_j p_j j^2 \frac{(\gamma+\mu)(\mu+\rho+\gamma+jw\zeta) + j\zeta\nu w(\frac{\gamma}{w}+\mu+\rho+j\zeta)}{(\mu+\rho)(\mu+\rho+\gamma+j\zeta)(\gamma+\mu+j\zeta)}$$

Formula (4.32) in MAY & ANDERSON (1988) is the analogue of this expression when one starts from (2.5) in this paper to find the initial growth rate λ_d.

4.3. Age dependence

We now turn our attention to a continuous dynamic h-state variable.

Let $\mathcal{F}(a)$ denote the survival probability as a function of age a in the absence of the disease. Then, at population dynamical equilibrium

$$S(a) = S(0)\mathcal{F}(a) \tag{4.9}$$

Let $\gamma(\tau,a,\alpha)$ be the average infectivity of an infected individual of age α and d-age τ towards a susceptible individual of age a. Then

$$A(\tau,a,\alpha) = \gamma(\tau,a,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)} \tag{4.10}$$

and

$$(K(S)\phi)(a) = S(0)\mathcal{F}(a)\int_0^\infty\int_0^\infty \gamma(\tau,a,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)}\phi(\alpha)d\alpha d\tau \tag{4.11}$$

4.3.1. Proportionate mixing.

Under the proportionate mixing assumption

$$\gamma(\tau,a,\alpha) = f(a)g(\tau,\alpha) \tag{4.12}$$

we find

$$R_0 = S(0)\int_0^\infty\int_0^\infty g(\tau,\alpha+\tau)\mathcal{F}(\alpha+\tau)f(\alpha)d\alpha d\tau \tag{4.13}$$

4.3.2. Endemic steady states.

Recalling Remark 3 at the end of section 2 we shall now consider an endemic steady state. Let

$$\lambda(a) = \text{age specific force of infection} \tag{4.14}$$

i.e. the age specific probability per unit of time of becoming infected. The survival function

$$\mathcal{F}_i(a) = e^{-\int_0^a \lambda(\alpha)d\alpha} \tag{4.15}$$

describes the probability of being susceptible for those who did not die. Hence

$$S(a) = S(0)\mathcal{F}(a)\mathcal{F}_i(a) \tag{4.16}$$

The age specific incidence rate is $\lambda(a)S(a)$ and consistency now requires that

$$\lambda(a) = \int_0^\infty\int_0^\infty A(\tau,a,\alpha)\lambda(\alpha)S(\alpha)d\alpha d\tau \tag{4.17}$$

$$= S(0) \int_0^\infty \int_0^\infty \gamma(\tau,a,\alpha+\tau)\mathfrak{K}(\alpha+\tau)\mathfrak{F}_i(\alpha)\lambda(\alpha)d\alpha d\tau$$

which can be considered as a nonlinear (recall (4.15)) integral equation for the (unknown) function λ. Note that linearization at the trivial solution $\lambda \equiv 0$ and the transformation $\phi \rightarrow \mathfrak{F}\lambda$ lead us back to the eigenvalue problem for $K(S)$, as to be expected. If we make the proportionate mixing assumption (4.12) we find that necessarily

$$\lambda(a) = Qf(a) \tag{4.18}$$

where the scalar Q has to satisfy

$$1 = S(0) \int_0^\infty \int_0^\infty g(\tau,\alpha+\tau)\mathfrak{K}(\alpha+\tau)e^{-Q\int_0^a f(\sigma)d\sigma} f(\alpha)d\alpha d\tau \tag{4.19}$$

4.3.3. Vaccination.

DIETZ & SCHENZLE (1985) consider the effect of vaccination and take

$$S(a) = S(0)\mathfrak{K}(a)\mathfrak{F}_v(a)\mathfrak{F}_i(a) \tag{4.20}$$

where $\mathfrak{F}_v(a)$ denotes the probability that an individual which did not die is immune due to vaccination. The analogue of (4.19) now is

$$1 = S(0) \int_0^\infty \int_0^\infty g(\tau,\alpha+\tau)\mathfrak{K}(\alpha+\tau)\mathfrak{F}_v(\alpha)e^{-Q\int_0^a f(\sigma)d\sigma} f(\alpha)d\alpha d\tau \tag{4.21}$$

which alternatively can be written as

$$1 = S(0) \int_0^\infty \int_0^\infty g(\tau,\theta)\mathfrak{K}(\theta)\mathfrak{F}_v(\theta-\tau)e^{-Q\int_0^{\theta-\tau} f(\sigma)d\sigma} f(\theta-\tau)d\tau d\theta \tag{4.22}$$

If we adopt the further assumption that

$$g(\tau,\alpha) = h(\alpha)k(\tau)\mathfrak{F}_r(\tau) \tag{4.23}$$

where k describes the infectivity as a function of d-age and $\mathfrak{F}_r$ the "removal" from the infected class, we finally arrive at

$$1 = S(0) \int_0^\infty h(\theta)\mathfrak{K}(\theta) \int_0^\theta k(\tau)\mathfrak{F}_r(\tau)\mathfrak{F}_v(\theta-\tau)e^{-Q\int_0^{\theta-\tau} f(\sigma)d\sigma} f(\theta-\tau)d\tau d\theta \tag{4.24}$$

which is, apart from the notation, identical to formula (3) in DIETZ & SCHENZLE (1985). These authors introduce yet two other simplifications:

(i) $h = f$, i.e. susceptibles and infectives have the same age dependence in activity level

(ii) the duration of the disease is short on the time scale of ageing

Then (4.24) can be approximated by

$$1 = CS(0) \int_0^\infty f^2(\theta)\mathfrak{K}(\theta)\mathfrak{F}_v(\theta)e^{-Q\int_0^\theta f(\sigma)d\sigma} d\theta \tag{4.25}$$

where C is a constant (describing the "magnitude" of the total infectivity). One can now use data about the endemic state to estimate f, Q and $\mathfrak{F}$ and subsequently calculate whether or not a given $\mathfrak{F}_v$ suffices to eradicate the disease. We refer once more to DIETZ & SCHENZLE (1985) for some more details.

4.3.4. Proportionate mixing with enhanced within age group infection.
To conclude this subsection we show how to compute the analogue of the threshold condition (3.10) (ii) in the case of age dependence (recall that in deriving (3.10) we assumed that the h-state is constant which it is not if we consider age). Assume that

$$\gamma(\tau,a,\alpha) = f(a)g(\tau,\alpha) + h(\tau,\alpha)\delta(a-\alpha) \tag{4.26}$$

where δ denotes Dirac's delta "function". Then

$$(K(S)\phi)(a) = S(a)\{f(a)\int_0^\infty\int_0^\infty g(\tau,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)}\phi(\alpha)d\alpha d\tau \tag{4.27}$$

$$+ \int_0^a h(\tau,a)\frac{\mathcal{F}(a)}{\mathcal{F}(a-\tau)}\phi(a-\tau)d\tau\}$$

We define an operator L by

$$(L\psi)(a) = \int_0^a h(a-\alpha,a)\frac{\mathcal{F}(a)}{\mathcal{F}(\alpha)}\psi(\alpha)d\alpha \tag{4.28}$$

and rewrite the eigenvalue problem $K(S)\phi=\rho\phi$ as

$$\theta(\phi)Sf + L\phi = \rho\phi \tag{4.29}$$

where θ is the $\mathbb{C}$-valued mapping defined by

$$\theta(\psi) = \int_0^\infty\int_0^\infty g(\tau,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)}\psi(\alpha)d\alpha d\tau \tag{4.30}$$

For ρ real and sufficiently large we can invert $\rho I-L$. In fact one has the series expansion

$$(\rho I-L)^{-1} = \sum_{n=0}^\infty \frac{L^{(n)}}{\rho^{n+1}} \tag{4.31}$$

Substituting $\phi=(\rho I-L)^{-1}\theta(\phi)Sf$ in the definition of θ we find the characteristic equation

$$1 = \int_0^\infty\int_0^\infty g(\tau,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)}((\rho I-L)^{-1}Sf)(\alpha)d\alpha d\tau \tag{4.32}$$

Assuming that (4.31) keeps converging up to $\rho=1$ (this is the analogue of the assumption $c(\xi)S(\xi)<1$ for all $\xi\in\Omega$ in section 3) we find that $R_0>1$ if and only if

$$\int_0^\infty\int_0^\infty g(\tau,\alpha+\tau)\frac{\mathcal{F}(\alpha+\tau)}{\mathcal{F}(\alpha)}\sum_{n=0}^\infty (L^{(n)}Sf)(\alpha)d\alpha d\tau > 1 \tag{4.33}$$

This condition allows an interpretation similar to the one of (3.11).

References

1. V. Andreasen, & F.B. Christiansen, *Persistence of an infections disease in a subdivided population,* preprint.
2. A.D. Barbour. (1978). *MacDonald's model and the transmission of bilharzia.* Trans. Roy. Soc. Trop. Med. Hyg. **72:** 6-15.
3. S.P. Blythe & C. Castillo-Chavez, *Like-with-like preference and sexual mixing models,* preprint.
4. K. Dietz (1980). *Models for vector-borne parasitic diseases.* In: C. Barigozzi (ed.), Vito Volterra Symposium on Mathematical Models in Biology. Springer Lect. notes in Biomath. **39:** 264-277.
5. K. Dietz & D. Schenzle (1985). *Proportionate mixing models for age-dependent infection transmission,* J. Math. Biol. **22:** 117-120.

6. H.J.A.M. HEIJMANS, (1986). *The dynamical behaviour of the age-size distribution of a cell population.* In: J.A.J. Metz & O. Diekmann, eds., *The Dynamics of Physiologically Structured Populations,* Springer Lect. Notes in Biomath. **68**: 185-202.
7. H.W. HETHCOTE & J.A. YORKE (1984). *Gonorrhea, Transmission Dynamics and Control,* Springer Lect. Notes in Biomath. **56**.
8 MAY, R.M. & R.M. ANDERSON, 1988: *The transmission dynamics of human immunodeficiency* virns (HIV). Phil. Trans. R. Soc. Lond. B **321**: 565-607.
9. J.A. JACQUEZ, C.P. SIMON, J. KOOPMAN, L. SATTENSPIEL & T. PERRY. (1988). *Modeling and analyzing HIV transmission: the effect of contact patterns.* Math. Biosc. **92**: 119-199.
10. H.H. SCHAEFER, (1960). *Some spectral properties of positive linear operators,* Pacific J. Math. **10**: 1009-1019.
10. H.H. SCHAEFER, (1974). *Banach Lattices and Positive Operators.* Springer, Berlin.