Chapter 5
Models with Heterogeneous Mixing

5.1 A Vaccination Model

To cope with annual seasonal influenza epidemics there is a program of vaccination before the "flu" season begins. Each year a vaccine is produced aimed at protecting against the three influenza strains considered most dangerous for the coming season. We formulate a model to add vaccination to the simple SIR model under the assumption that vaccination reduces susceptibility (the probability of infection if a contact with an infected member of the population is made).

We consider a population of total size N and assume that a fraction γ of this population is vaccinated prior to a disease outbreak. Thus we have unvaccinated and vaccinated sub-populations of sizes $N_U = (1-\gamma)N$ and $N_V = \gamma N$, respectively. We assume that vaccinated members have susceptibility to infection reduced by a factor σ, $0 \le \sigma \le 1$, with $\sigma = 0$ describing a perfectly effective vaccine and $\sigma = 1$ describing a vaccine that has no effect. We assume also that vaccinated individuals who are infected have infectivity reduced by a factor δ and both vaccinated and unvaccinated individuals have a recovery rate α. The number of contacts in unit time per member is a_U for unvaccinated individuals and a_V for vaccinated individuals. These may be equal.

In this chapter we will study models in which there is more than one susceptible or infective compartment, and it is convenient to formulate such models using the number of contacts in unit time instead of the number of contacts multiplied by the total population size. Thus, for example, instead of writing the simple epidemic model as

$$S' = \beta SI, \quad I' = \beta SI - \alpha I,$$

F. Brauer et al., *Mathematical Models in Epidemiology*, Texts in Applied Mathematics 69, https://doi.org/10.1007/978-1-4939-9828-9_5

we would write it as

$$S' = -aS\frac{I}{N}, \quad I' = aS\frac{I}{N} - \alpha I.$$

We let S_U, S_V, I_U, I_V denote the number of unvaccinated susceptibles, the number of vaccinated susceptibles, the number of unvaccinated infectives, and the number of vaccinated infectives, respectively. The resulting model is

$$\begin{aligned} S_U' &= -a_U S_U \left[\frac{I_U}{N_U} + \delta\frac{I_V}{N_V}\right] \\ S_V' &= -\sigma a_V S_V \left[\frac{I_U}{N_U} + \delta\frac{I_V}{N_V}\right] \\ I_U' &= a_U S_U \left[\frac{I_U}{N_U} + \delta\frac{I_V}{N_V}\right] - \alpha I_U \\ I_V' &= \sigma a_V S_V \left[\frac{I_U}{N_U} + \delta\frac{I_V}{N_V}\right] - \alpha I_V. \end{aligned} \tag{5.1}$$

The initial conditions prescribe $S_U(0), S_V(0), I_U(0), I_V(0)$, with

$$S_U(0) + I_U(0) = N_U, \quad S_V(0) + I_V(0) = N_V.$$

Since the infection now is beginning in a population which is not fully susceptible, we speak of the *c*ontrol reproduction number $\mathscr{R}_c$ rather than the basic reproduction number. However, as we will soon see, calculation of the control reproduction number will require a more general definition and a considerable amount of technical computation. The computation method is applicable to both basic and control reproduction numbers. We will use the term reproduction number to denote either a basic reproduction number or a control reproduction number. We are able to obtain final size relations without knowledge of the reproduction number but these final size relations do contain information about the reproduction number, and more.

Since S_U and S_V are decreasing non-negative functions they have limits $S_U(\infty)$ and $S_V(\infty)$, respectively, as $t \to \infty$. The sum of the equations for S_U and I_U in (5.1) is

$$(S_U + I_U)' = -\alpha I_U,$$

from which we conclude, just as in the analysis of the simple SIR model in Sect. 2.4, that $I_U(t) \to 0$ as $t \to \infty$, and that

$$\alpha \int_0^\infty I_U(t)dt = N_U - S_U(\infty). \tag{5.2}$$

Similarly, using the sum of the equations for S_V and I_V, we see that $I_V(t) \to 0$ as $t \to \infty$, and that

$$\alpha \int_0^\infty I_V(t)dt = N_V - S_V(\infty). \tag{5.3}$$

Integration of the equation for S_U in (5.1) and use of (5.2), (5.3) gives

$$\begin{aligned} \log \frac{S_U(0)}{S_U(\infty)} &= a_U \left[\int_0^\infty I_U(t)dt + \delta \int_0^\infty I_V(t)dt \right] \\ &= \frac{a_U}{\alpha} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \frac{\delta a_U}{\alpha} \left[1 - \frac{S_V(\infty)}{N_V} \right]. \end{aligned} \tag{5.4}$$

A similar calculation using the equation for S_V gives

$$\log \frac{S_V(0)}{S_V(\infty)} = \frac{\sigma a_V}{\alpha} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \frac{\delta \sigma a_V}{\alpha} \left[1 - \frac{S_V(\infty)}{N_V} \right]. \tag{5.5}$$

This pair of Eqs. (5.4) and (5.5) are the final size relations. They make it possible to calculate $S_U(\infty)$, $S_V(\infty)$ if the parameters of the model are known.

It is convenient to define the matrix

$$\mathscr{R} = \begin{bmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} \\ \mathscr{R}_{21} & \mathscr{R}_{22} \end{bmatrix} = \begin{bmatrix} \frac{a_U}{\alpha} & \frac{\delta a_U}{\alpha} \\ \frac{\sigma a_V}{\alpha_U} & \frac{\delta \sigma a_V}{\alpha_V} \end{bmatrix}.$$

The element $\mathscr{R}_{ij}$ can be interpreted as the average number of susceptibles of group i infected by an infective of type j over its infectious period.

Then the final size relations (5.4), (5.5) may be written as

$$\begin{aligned} \log \frac{S_U(0)}{S_U(\infty)} &= \mathscr{R}_{11} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \mathscr{R}_{12} \left[1 - \frac{S_V(\infty)}{N_V} \right] \\ \log \frac{S_V(0)}{S_V(\infty)} &= \mathscr{R}_{21} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \mathscr{R}_{22} \left[1 - \frac{S_V(\infty)}{N_V} \right]. \end{aligned} \tag{5.6}$$

The matrix $\mathscr{R}$ is closely related to the reproduction number. In the next section we describe a general method for calculating reproduction numbers that will involve a matrix which is similar to this matrix.

5.2 The Next Generation Matrix and the Basic Reproduction Number

Up to this point, we have calculated reproduction numbers by following the secondary cases caused by a single infective introduced into a population. However, if there are sub-populations with different susceptibility to infection, as in the vaccination model introduced in Sect. 5.1, it is necessary to follow the secondary infections in the sub-populations separately, and this approach will not yield the reproduction number. It is necessary to give a more general approach to the meaning of the reproduction number, and this is done through the *next generation matrix* [18, 19, 45]. The underlying idea is that we must calculate the matrix whose (i, j) entry is the number of secondary infections caused in compartment i by an infected individual in compartment j. We will follow the development in [45, 46] for ordinary differential equation models, even though the approach of [18, 19] is more general.

In a compartmental disease transmission model, we sort individuals into compartments based on a single, discrete state variable. A compartment is called a *disease compartment* if the individuals therein are infected. Note that this use of the term disease is broader than the clinical definition and includes stages of infection such as exposed stages in which infected individuals are not necessarily infective. Suppose there are n disease compartments and m non-disease compartments, and let $x \in R^n$ and $y \in R^m$ be the sub-populations in each of these compartments. Further, we denote by $\mathscr{F}_i$ the rate at which secondary infections increase the i^{th} disease compartment and by $\mathscr{V}_i$ the rate at which disease progression, death, and recovery decrease the i^{th} compartment, that is, $\mathscr{V}_i$ is the net outflow from the i th compartment due to disease progression, death, and recovery, with inflow from other compartments yielding a negative contribution. The compartmental model can then be written in the form

$$\begin{aligned} x_i' &= \mathscr{F}_i(x, y) - \mathscr{V}_i(x, y), \quad i = 1, \ldots, n, \\ y_j' &= g_j(x, y), \quad j = 1, \ldots, m. \end{aligned} \tag{5.7}$$

Note that the decomposition of the dynamics into $\mathscr{F}$ and $\mathscr{V}$ and the designation of compartments as infected or uninfected may not be unique; different decompositions correspond to different epidemiological interpretations of the model. The definitions of $\mathscr{F}$ and $\mathscr{V}$ used here differ slightly from those in [45].

The derivation of the basic reproduction number is based on the linearization of the ODE model about a disease-free equilibrium. We make the following assumptions:

- $\mathscr{F}_i(0, y) = 0$ and $\mathscr{V}_i(0, y) = 0$ for all $y \geq 0$ and $i = 1, \ldots, n$.
- The disease-free system $y' = g(0, y)$ has a unique equilibrium that is asymptotically stable, that is, all solutions with initial conditions of the form $(0, y)$ approach a point $(0, y_o)$ as $t \to \infty$. We refer to this point as the disease-free equilibrium.

The first assumption says that all new infections are secondary infections arising from infected hosts; there is no immigration of individuals into the disease compartments. It ensures that the disease-free set, which consists of all points of the form $(0, y)$, is invariant. That is, any solution with no infected individuals at some point in time will be free of infection for all time. The second assumption ensures that the disease-free equilibrium is also an equilibrium of the full system.

Next, we further assume that

- $\mathscr{F}_i(x, y) \geq 0$ for all non-negative x and y and $i = 1, \ldots, n$.
- $\mathscr{V}_i(x, y) \leq 0$ whenever $x_i = 0, i = 1, \ldots, n$.
- $\sum_{i=1}^{n} \mathscr{V}_i(x, y) \geq 0$ for all non-negative x and y.

The reasons for these assumptions are that the function $\mathscr{F}$ represents new infections and cannot be negative, each component, $\mathscr{V}_i$, represents a net outflow from compartment i and must be negative (inflow only) whenever the compartment is empty, and the sum $\sum_{i=1}^{n} \mathscr{V}_i(x, y)$ represents the total outflow from all infected compartments. Terms in the model leading to increases in $\sum_{i=1}^{n} x_i$ are assumed to represent secondary infections and therefore belong in $\mathscr{F}$.

Suppose that a single infected person is introduced into a population originally free of disease. The initial ability of the disease to spread through the population is determined by an examination of the linearization of (5.7) about the disease-free equilibrium $(0, y_o)$. It is easy to see that the assumption $\mathscr{F}_i(0, y) = 0$, $\mathscr{V}_i(0, y) = 0$ implies

$$\frac{\partial \mathscr{F}_i}{\partial y_j}(0, y_o) = \frac{\partial \mathscr{V}_i}{\partial y_j}(0, y_o) = 0$$

for every pair (i, j). This implies that the linearized equations for the disease compartments, x, are decoupled from the remaining equations and can be written as

$$x' = (F - V)x, \tag{5.8}$$

where F and V are the $n \times n$ matrices with entries

$$F = \frac{\partial \mathscr{F}_i}{\partial x_j}(0, y_o) \qquad \text{and} \qquad V = \frac{\partial \mathscr{V}_i}{\partial x_j}(0, y_o).$$

Because of the assumption that the disease-free system $y' = g(0, y)$ has a unique asymptotically stable equilibrium, the linear stability of the system (5.7) is completely determined by the linear stability of the matrix $(F - V)$ in (5.8).

The number of secondary infections produced by a single infected individual can be expressed as the product of the expected duration of the infective period and the rate at which secondary infections occur. For the general model with n disease compartments, these are computed for each compartment for a hypothetical index case. The expected time the index case spends in each compartment is given by the integral $\int_0^\infty \phi(t, x_0)\, dt$, where $\phi(t, x_0)$ is the solution of (5.8) with $F = 0$

(no secondary infections) and non-negative initial conditions, x_0, representing an infected index case:

$$x' = -Vx, \qquad x(0) = x_0. \tag{5.9}$$

In effect, this solution shows the path of the index case through the disease compartments from the initial exposure through death or recovery with the i^{th} component of $\phi(t, x_0)$ interpreted as the probability that the index case (introduced at time $t = 0$) is in disease state i at time t. The solution of (5.9) is $\phi(t, x_0) = e^{-Vt}x_0$, where the exponential of a matrix is defined by the Taylor series

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots.$$

This series converges for all t and $\int_0^\infty \phi(t, x_0)\, dt = V^{-1}x_0$ (see, for example, [29]). The (i, j) entry of the matrix V^{-1} can be interpreted as the expected time an individual initially introduced into disease compartment j spends in disease compartment i.

The (i, j) entry of the matrix F is the rate at which secondary infections are produced in compartment i by an index case in compartment j. Hence, the expected number of secondary infections produced by the index case is given by

$$\int_0^\infty Fe^{-Vt}x_0\, dt = FV^{-1}x_0.$$

Following Diekmann and Heesterbeek [18], the matrix $K_L = FV^{-1}$ is referred to as the next generation matrix with large domain for the system at the disease-free equilibrium. The (i, j) entry of K is the expected number of secondary infections in compartment i produced by individuals initially in compartment j, assuming, of course, that the environment seen by the individual remains homogeneous for the duration of its infection.

Shortly, we will describe some results from matrix theory which imply that the matrix, $K_L = FV^{-1}$ is non-negative and therefore has a non-negative eigenvalue, $\mathscr{R}_0 = \rho(FV^{-1})$, such that there are no other eigenvalues of K with modulus greater than $\mathscr{R}_0$ and there is a non-negative eigenvector ω associated with $\mathscr{R}_0$ [7, Theorem 1.3.2]. This eigenvector is in a sense the distribution of infected individuals that produces the greatest number, $\mathscr{R}_0$, of secondary infections per generation. Thus, $\mathscr{R}_0$ and the associated eigenvector ω suitably define a "typical" infective and the basic reproduction number can be rigorously defined as the spectral radius of the matrix, K_L. The spectral radius of a matrix K_L, denoted by $\rho(K_L)$, is the maximum of the moduli of the eigenvalues of K_L. If K_L is irreducible, then $\mathscr{R}_0$ is a simple eigenvalue of K_L and is strictly larger in modulus than all other eigenvalues of K_L. However, if K_L is reducible, which is often the case for diseases with

multiple strains, then K_L may have several positive real eigenvectors corresponding to reproduction numbers for each competing strain of the disease.

We have interpreted the reproduction number for a disease as the number of secondary infections produced by an infected individual in a population of susceptible individuals. If the reproduction number $\mathscr{R}_0 = \rho(FV^{-1})$ is consistent with the differential equation model, then it should follow that the disease-free equilibrium is asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.

This is shown through a sequence of lemmas.

The *s*pectral bound (or abscissa) of a matrix A is the maximum real part of all eigenvalues of A. If each entry of a matrix T is non-negative we write $T \geq 0$ and refer to T as a non-negative matrix. A matrix of the form $A = sI - B$, with $B \geq 0$, is said to have the Z sign pattern. These are matrices whose off-diagonal entries are negative or zero. If in addition, $s \geq \rho(B)$, then A is called an M-matrix. Note that in this section, I denotes an identity matrix, not a population of infectious individuals. The following lemma is a standard result from [7].

Lemma 5.1 *If A has the Z sign pattern, then $A^{-1} \geq 0$ if and only if A is a non-singular M-matrix.*

The assumptions we have made imply that each entry of F is non-negative and that the off-diagonal entries of V are negative or zero. Thus V has the Z sign pattern. Also, the column sums of V are positive or zero, which, together with the Z sign pattern, implies that V is a (possibly singular) M-matrix [7, condition M_{35} of Theorem 6.2.3]. In what follows, it is assumed that V is non-singular. In this case, $V^{-1} \geq 0$, by Lemma 5.1. Hence, $K_L = FV^{-1}$ is also non-negative.

Lemma 5.2 *If F is non-negative and V is a non-singular M-matrix, then $\mathscr{R}_0 = \rho(FV^{-1}) < 1$ if and only if all eigenvalues of $(F - V)$ have negative real parts.*

Proof Suppose $F \geq 0$ and V is a non-singular M-matrix. By the proof of Lemma 5.1, $V^{-1} \geq 0$. Thus, $(I - FV^{-1})$ has the Z sign pattern, and by Lemma 5.1, $(I - FV^{-1})^{-1} \geq 0$ if and only if $\rho(FV^{-1}) < 1$. From the equalities $(V - F)^{-1} = V^{-1}(I - FV^{-1})^{-1}$ and $V(V - F)^{-1} = I + F(V - F)^{-1}$, it follows that $(V-F)^{-1} \geq 0$ if and only if $(I-FV^{-1})^{-1} \geq 0$. Finally, $(V-F)$ has the Z sign pattern, so by Lemma 5.1, $(V - F)^{-1} \geq 0$ if and only if $(V - F)$ is a non-singular M-matrix. Since the eigenvalues of a non-singular M-matrix all have positive real parts, this completes the proof. □

Theorem 5.1 *Consider the disease transmission model given by (5.7). The disease-free equilibrium of (5.7) is locally asymptotically stable if $\mathscr{R}_0 < 1$, but unstable if $\mathscr{R}_0 > 1$.*

Proof Let F and V be defined as above, and let J_{21} and J_{22} be the matrices of partial derivatives of g with respect to x and y evaluated at the disease-free equilibrium. The Jacobian matrix for the linearization of the system about the disease-free equilibrium has the block structure

$$J = \begin{bmatrix} F - V & 0 \\ J_{21} & J_{22} \end{bmatrix}.$$

The disease-free equilibrium is locally asymptotically stable if the eigenvalues of the Jacobian matrix all have negative real parts. Since the eigenvalues of J are those of $(F - V)$ and J_{22}, and the latter all have negative real parts by assumption, the disease-free equilibrium is locally asymptotically stable if all eigenvalues of $(F-V)$ have negative real parts. By the assumptions on $\mathscr{F}$ and $\mathscr{V}$, F is non-negative and V is a non-singular M-matrix. Hence, by Lemma 5.2 all eigenvalues of $(F - V)$ have negative real parts if and only if $\rho(FV^{-1}) < 1$. It follows that the disease-free equilibrium is locally asymptotically if $\mathscr{R}_0 = \rho(FV^{-1}) < 1$.

Instability for $\mathscr{R}_0 > 1$ can be established by a continuity argument. If $\mathscr{R}_0 \leq 1$, then for any $\epsilon > 0$, $((1 + \epsilon)I - FV^{-1})$ is a non-singular M-matrix and, by Lemma 5.1, $((1 + \epsilon)I - FV^{-1})^{-1} \geq 0$. By the proof of Lemma 5.2, all eigenvalues of $((1 + \epsilon)V - F)$ have positive real parts. Since $\epsilon > 0$ is arbitrary, and eigenvalues are continuous functions of the entries of the matrix, it follows that all eigenvalues of $(V - F)$ have non-negative real parts. To reverse the argument, suppose all the eigenvalues of $(V - F)$ have non-negative real parts. For any positive ϵ, $(V + \epsilon I - F)$ is a non-singular M-matrix, and by the proof of Lemma 5.2, $\rho(F(V + \epsilon I)^{-1}) < 1$. Again, since $\epsilon > 0$ is arbitrary, it follows that $\rho(FV^{-1}) \leq 1$. Thus, $(F - V)$ has at least one eigenvalue with positive real part if and only if $\rho(FV^{-1}) > 1$, and the disease-free equilibrium is unstable whenever $\mathscr{R}_0 > 1$. □

These results validate the extension of the definition of the reproduction number to more general situations. In the vaccination model (5.1) of Sect. 5.1 we calculated a pair of final size relations which contained the elements of a matrix K. This matrix is precisely the next generation matrix with large domain $K_L = FV^{-1}$ that has been introduced in this section.

Example 1 Consider the $SEIR$ model with infectivity in the exposed stage,

$$\begin{aligned} S' &= -\frac{a}{N} S(I + \varepsilon E) \\ E' &= \frac{a}{N} S(I + \varepsilon E) - \kappa E \\ I' &= \kappa E - \alpha I \\ R' &= \alpha I. \end{aligned} \tag{5.10}$$

Here the disease states are E and I,

$$\mathscr{F} = \begin{bmatrix} \varepsilon E a + I a \\ 0 \end{bmatrix}$$

and

$$F = \begin{bmatrix} \varepsilon a & a \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \kappa & 0 \\ -\kappa & \alpha \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} 1/\kappa & 0 \\ 1/\alpha & 1/\alpha \end{bmatrix}.$$

Then we may calculate

$$K_L = FV^{-1} = \begin{bmatrix} \dfrac{\varepsilon a}{\kappa} + \dfrac{a}{\alpha} & \dfrac{a}{\alpha} \\ 0 & 0 \end{bmatrix}.$$

It is clear that since $\mathscr{R}_0$ is equal to the trace of FV^{-1},

$$\mathscr{R}_0 = \frac{\varepsilon a}{\kappa} + \frac{a}{\alpha},$$

the element in the first row and first column of FV^{-1}. If all new infections are in a single compartment, as is the case here, the basic reproduction number is the trace of the matrix FV^{-1}.

In general, it is possible to reduce the size of the next generation matrix with large domain to the number of *state at infection* [18]. The states at infection are those disease states in which there can be new infections. Suppose that there are n disease states and k states at infection with $k < n$. Then we may define an auxiliary $n \times k$ matrix P in which each column corresponds to a state at infection and has 1 in the corresponding row and 0 elsewhere. Then the next generation matrix is the $k \times k$ matrix

$$K = P^T K_L P.$$

It is easy to show, using the fact that $PP^T K_L = K_L$, that the $n \times n$ matrix K_L and the $k \times k$ matrix K have the same non-zero eigenvalues and therefore the same spectral radius. Construction of the next generation matrix which has lower dimension than the next generation matrix with large domain may simplify the calculation of the basic reproduction number.

In Example 1 above, the only disease state at infection is E, the matrix P is

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and the next generation matrix K is the 1×1 matrix

$$K = \left[\frac{\varepsilon a}{\kappa} + \frac{a}{\alpha}\right].$$

5.2.1 Some More Complicated Examples

The next generation approach is very general and can be applied to models with heterogeneous mixing and control measures applied differently in different groups.

Example 2 Consider the vaccination model (5.1) of Sect. 5.1. The disease states are I_U and I_V. Then

$$\mathscr{F} = \begin{bmatrix} a_U(I_U + \delta I_V) \\ \sigma a_V(I_U + \delta I_V \end{bmatrix}$$

and

$$F = \begin{bmatrix} a_U \frac{N_U}{N} & \delta a_U \frac{N_U}{N} \\ \sigma a_V \frac{N_V}{N} & \sigma\delta a_V \frac{N_V}{N} \end{bmatrix}, \quad V = \begin{bmatrix} \alpha_U & 0 \\ 0 & \alpha_V \end{bmatrix}.$$

It is easy to see that the next generation matrix with large domain is the matrix K calculated in Sect. 5.1. Since each disease state is a disease state at infection, the next generation matrix is K, the same as the next generation matrix with large domain. As in Example 1, the determinant of K is zero and K has rank 1. Thus the control reproduction number is the trace of K,

$$\mathscr{R}_c = \frac{a_U}{\alpha_U}\frac{N_U}{N} + \delta\sigma\frac{a_V}{\alpha_V}\frac{N_U}{N}.$$

5.3 Heterogeneous Mixing

In disease transmission models not all members of the population make contacts at the same rate. In sexually transmitted diseases there is often a "core" group of very active members who are responsible for most of the disease cases, and control measures aimed at this core group have been very effective in control [27]. In epidemics there are often "super-spreaders", who make many contacts and are instrumental in spreading disease and in general some members of the population make more contacts than others. Recently there has been a move to complicated network models for simulating epidemics [23, 24, 32–34, 38]. These assume knowledge of the mixing patterns of groups of members of the population and

make predictions based on simulations of a stochastic model. A basic description of network models may be found in [44]. While network models can give very detailed predictions, they have some serious disadvantages. For a detailed network model, simulations take long enough to make it difficult to examine a significant range of parameter values, and it is difficult to estimate the sensitivity with respect to parameters of the model. The theoretical analysis of network models is a very active and rapidly developing field [38–40].

However, it is possible to consider models more realistic than simple compartmental models but simpler to analyze than detailed network models. To model heterogeneity in mixing we may assume that the population is divided into subgroups with different activity levels. We will analyze an SIR model in which there are two groups with different contact rates. The approach extends easily to models with more compartments, such as exposed periods or a sequence of infective stages and also to models with an arbitrary number of activity levels. In this way, we may hope to give models intermediate between the too simple compartmental models and the too complicated network models.

In this section, we describe the formulation of models for two groups with different activity levels and give the main results for the simplest compartmental epidemic models. The analysis of models of the same type with more complicated compartmental structure is given in [11] and the analysis of models with more groups is given in [12]. There is no problem, other than technical calculation difficulties, in extending everything in this section to an arbitrary number of subpopulations.

Consider two sub-populations of constant sizes N_1, N_2, respectively, each divided into susceptibles, infectives, and removed members with subscripts to identify the sub-population. In this section, we will assume that the number of contacts per member in unit time is a constant. Suppose that each group i member makes a_i contacts sufficient to transmit infection in unit time, and that the fraction of contacts made by a member of group i that is with a member of group j is p_{ij}, $(i, j = 1, 2)$. Then

$$p_{i1} + p_{i2} = 1, \quad i = 1, 2.$$

We assume that all contacts between a susceptible and an infective transmit infection to the susceptible. Suppose the mean infective period in group i is $1/\alpha_i$. We assume that there are no disease deaths, so that the population size of each group is constant.

A two-group SIR epidemic model is

$$\begin{aligned} S_i' &= -a_i S_i \left[p_{i1} \frac{I_1}{N_1} + p_{i2} \frac{I_2}{N_2} \right] \\ I_i' &= a_i S_i \left[p_{i1} \frac{I_1}{N_1} + p_{i2} \frac{I_2}{N_2} \right] - \alpha_i I_i, \quad i = 1, 2. \end{aligned} \tag{5.11}$$

The initial conditions are

$$S_i(0) + I_i(0) = N_i, \quad i = 1, 2.$$

The two-group model includes two possibilities. It may describe a population with groups differing by activity levels and possibly by vulnerability to infection. For an epidemic model, in which we assume the time scale is short enough that members do not age over the course of the epidemic, the grouping could be by age. However, for a longer term disease transmission model with age-dependent transmission it would be necessary to take into account the fact that the ages of members of the population change over the course of the disease and use a different type of model, to be studied in Chap. 13.

The two-group model (5.11) assumes the same infectivities and susceptibilities in each group. A more general model would be

$$\begin{aligned} S_i' &= -\sigma_i a_i S_i \left[\delta_1 p_{i1} \frac{I_1}{N_1} + \delta_2 p_{i2} \frac{I_2}{N_2} \right] \\ I_i' &= \sigma_i a_i S_i \left[\delta_1 p_{i1} \frac{I_1}{N_1} + \delta_2 p_{i2} \frac{I_2}{N_2} \right] - \alpha_i I_i, \quad i = 1, 2. \end{aligned} \tag{5.12}$$

This is just the model (5.11) with the addition of susceptibility factors σ_1, σ_2 for susceptibles in the two groups and infectivity factors δ_1, δ_2 for infectives in the two groups. As before a_1, a_2 are effective contact rates, and this model adds transmission probabilities to (5.11).

It is not possible to calculate the reproduction number for the two-group model (5.11) directly by counting secondary infections. It is necessary to use the next generation matrix approach of [45] described in Section 5.2 and calculate the reproduction number as the largest eigenvalue of the matrix FV^{-1}, where

$$F = \begin{bmatrix} p_{11}a_1 & p_{12}a_1 \dfrac{N_1}{N_2} \\ p_{21}a_2 \dfrac{N_2}{N_1} & p_{22}a_2 \end{bmatrix}, \quad V = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}.$$

Then

$$FV^{-1} = \begin{bmatrix} \dfrac{p_{11}a_1}{\alpha_1} & \dfrac{p_{12}a_1}{\alpha_2} \dfrac{N_1}{N_2} \\ \dfrac{p_{21}a_2}{\alpha_1} \dfrac{N_2}{N_1} & \dfrac{p_{22}a_2}{\alpha_2} \end{bmatrix}.$$

The eigenvalues of the matrix FV^{-1} are the roots of the quadratic equation

$$\lambda^2 - \left(\frac{p_{11}a_1}{\alpha_1} + \frac{p_{22}a_2}{\alpha_2} \right) \lambda + (p_{11}p_{22} - p_{12}p_{21}) \frac{a_1 a_2}{\alpha_1 \alpha_2} = 0. \tag{5.13}$$

The basic reproduction number $\mathscr{R}_0$ is the larger of these two eigenvalues,

$$\mathscr{R}_0 = \frac{\frac{p_{11}a_1}{\alpha_1} + \frac{p_{22}a_2}{\alpha_2} + \sqrt{\left(\frac{p_{11}a_1}{\alpha_1} - \frac{p_{22}a_2}{\alpha_2}\right)^2 + 4\frac{p_{12}p_{21}a_1a_2}{\alpha_1\alpha_2}}}{2}.$$

In order to obtain a more useful expression for $\mathscr{R}_0$, it is necessary to make some assumptions about the nature of the mixing between the two groups. The mixing is determined by the two quantities p_{12}, p_{21} since $p_{11} = 1 - p_{12}$ and $p_{22} = 1 - p_{21}$.

There has been much study of mixing patterns, see, for example, [8, 9, 13]. One possibility is proportionate mixing, that is, that the number of contacts between groups is proportional to the relative activity levels. In other words, mixing is random but constrained by the activity levels [42]. Under the assumption of proportionate mixing,

$$p_{ij} = \frac{a_j N_j}{a_1 N_1 + a_2 N_2},$$

and we may write

$$p_{11} = p_{21} = p_1, \quad p_{12} = p_{22} = p_2,$$

with $p_1 + p_2 = 1$. In particular,

$$p_{11}p_{22} - p_{12}p_{21} = 0,$$

and thus

$$\mathscr{R}_0 = a_1\frac{p_1}{\alpha_1} + a_2\frac{p_2}{\alpha_2}.$$

Another possibility is preferred mixing [42], in which a fraction π_i of each group mixes randomly with its own group and the remaining members mix proportionately. Thus, preferred mixing is given by

$$\begin{aligned} p_{11} &= \pi_1 + (1-\pi_1)p_1, & p_{12} &= (1-\pi_1)p_2 \\ p_{21} &= (1-\pi_2)p_1, & p_{22} &= \pi_2 + (1-\pi_2)p_2, \end{aligned} \tag{5.14}$$

with

$$p_i = \frac{(1-\pi_i)a_i N_i}{(1-\pi_1)a_1 N_1 + (1-\pi_2)a_2 N_2}, \quad i = 1, 2.$$

Proportionate mixing is the special case of preferred mixing with $\pi_1 = \pi_2 = 0$.

It is also possible to have like-with-like mixing, in which members of each group mix only with members of the same group. This is the special case of preferred mixing with $\pi_1 = \pi_2 = 1$. For like-with-like mixing,

$$p_{11} = p_{22} = 1, \quad p_{12} = p_{21} = 0.$$

Then the roots of (5.13) are a_1/α_1 and a_2/α_2, and the reproduction number is

$$\mathscr{R}_0 = \max\left\{\frac{a_1}{\alpha_1}, \frac{a_2}{\alpha_2}\right\}.$$

By calculating the partial derivatives of p_{ij} $(i, j = 1, 2)$ with respect to π_1 and π_2, we may show that p_{11} and p_{22} increase when either π_1 or π_2 is increased, while p_{12} and p_{21} decrease when either π_1 or π_2 is increased. From this, we may see from the general expression for $\mathscr{R}_0$ that increasing either of the preferences π_1, π_2 increases the basic reproduction number.

We may follow the analysis of the SIR model (5.11) to obtain the basic reproduction number for the SIR model (5.12) with susceptibility and infectivity reduction factors

$$\mathscr{R}_0 = \frac{\displaystyle\sum_{i=1}^{2} \sigma_i \delta_i \frac{p_{ii} a_i}{\alpha_i} + \sqrt{\left(\sigma_1 \delta_1 \frac{p_{11} a_1}{\alpha_1} - \sigma_2 \delta_2 \frac{p_{22} a_2}{\alpha_2}\right)^2 + 4\sigma_1 \delta_2 \sigma_2 \delta_1 \frac{p_{12} p_{21} a_1 a_2}{\alpha_1 \alpha_2}}}{2}.$$

In the special case of proportionate mixing, where $p_{11}p_{22} - p_{12}p_{21} = 0$, this reduces to

$$\mathscr{R}_0 = \sum_{i=1}^{2} \sigma_i \delta_i \frac{p_{ii} a_i}{\alpha_i}.$$

The vaccination model of Sect. 5.1 is an example of a two-group model of the form (5.12), with

$$\sigma_1 = \sigma_2 = \delta_1 = 1, \quad \delta_2 = \delta.$$

It is easy to show [11] that, just as for a one-group model [10],

$$S_1 \to S_1(\infty) > 0, \quad S_2 \to S_2(\infty) > 0,$$

as $t \to \infty$.

Example 1 Consider a two-group $SEIR$ model,

$$\begin{aligned} S_i' &= -a_i S_i \left[p_{ii} \tfrac{I_i}{N_i} + p_{ij} \tfrac{I_j}{N_j} \right] \\ E_i' &= a_i S_i \left[p_{ii} \tfrac{I_i}{N_i} + p_{ij} \tfrac{I_j}{N_j} \right] - \kappa_i E_i \\ I_i' &= \kappa_i E_i - \alpha_i I_i \\ R_i' &= \alpha_i I_i, \quad i, j = 1, 2, \ i \neq j. \end{aligned} \tag{5.15}$$

The disease states are E_i and I_i $(i = 1, 2)$.

Now

$$\mathscr{F} = \begin{bmatrix} a_1 p_{11} I_1 + a_1 p_{12} I_2 \dfrac{N_1}{N_2} \\ 0 \\ a_2 p_{21} I_1 \dfrac{N_2}{N_1} \\ 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 0 & a_1 p_{11} & 0 & a_1 p_{12} \dfrac{N_1}{N_2} \\ 0 & 0 & 0 & 0 \\ 0 & a_2 p_{21} \dfrac{N_2}{N_1} & 0 & a_2 p_{22} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$V = \begin{bmatrix} \kappa_1 & 0 & 0 & 0 \\ -\kappa_1 & \alpha_1 & 0 & 0 \\ 0 & 0\kappa_2 & 0 & \\ 0 & 0 & -\kappa_2 & \alpha_2 \end{bmatrix}, \qquad V^{-1} = \begin{bmatrix} \frac{1}{\kappa_1} & 0 & 0 & 0 \\ \frac{1}{\alpha_1} & \frac{1}{\alpha_1} & 0 & 0 \\ 0 & 0 & \frac{1}{\kappa_2} & 0 \\ 0 & 0 & \frac{1}{\alpha_2} & \frac{1}{\alpha_2} \end{bmatrix}.$$

Then we may calculate

$$K_L = FV^{-1} = \begin{bmatrix} a_1 \dfrac{p_{11}}{\alpha_1} & a_1 \dfrac{p_{11}}{\alpha_1} & a_1 \dfrac{p_{12}}{\alpha_1} \dfrac{N_1}{N_2} & a_1 \dfrac{p_{12}}{\alpha_1} \dfrac{N_1}{N_2} \\ 0 & 0 & 0 & 0 \\ a_2 \dfrac{p_{21}}{\alpha_1} \dfrac{N_2}{N_1} & a_2 \dfrac{p_{21}}{\alpha_1} \dfrac{N_2}{N_1} & a_2 \dfrac{p_{22}}{\alpha_2} & a_2 \dfrac{p_{22}}{\alpha_2} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this example, it is advantageous to construct the next generation matrix by reducing the next generation matrix with large domain K_L. In order to do this, we use the auxiliary matrix

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

to construct the next generation matrix

$$K = E^T K_L E = \begin{bmatrix} \frac{p_{11}a_1}{\alpha_1} & \frac{p_{12}a_1}{\alpha_2}\frac{N_1}{N_2} \\ \frac{p_{21}a_2}{\alpha_1}\frac{N_2}{N_1} & \frac{p_{22}a_2}{\alpha_2} \end{bmatrix}. \tag{5.16}$$

This is the same matrix as the next generation matrix obtained for the two-group SIR model (5.11) introduced in this section.

For a one-group epidemic model there is a final size relation that makes it possible to calculate the size of the epidemic from the reproduction number [10, 35]. There is a corresponding final size relation for the two-group model (5.1), established in much the same way, combining expressions for the integrals of $(S_1 + I_1)'$, $(S_2 + I_2)'$, $(S_1)'/S_1$, $(S_2)'/S_2$). This relation does not involve the reproduction number explicitly but still makes it possible to calculate the size of the epidemic from the model parameters.

The final size relation for the model (5.11) is the pair of equations

$$\log \frac{S_i(0)}{S_i(\infty)} = a_i \left[\frac{p_{ii}}{\alpha_i} \left(1 - \frac{S_i(\infty)}{N_i} \right) + \frac{p_{ij}}{\alpha_j} \left(1 - \frac{S_j(\infty)}{N_j} \right) \right], \quad i, j = 1, 2, \quad i \neq j. \tag{5.17}$$

Just as with the vaccination model (5.1), the final size relation may be written in terms of the matrix

$$\mathscr{R} = \begin{bmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} \\ \mathscr{R}_{21} & \mathscr{R}_{22} \end{bmatrix} = \begin{bmatrix} \frac{a_1 p_{11}}{\alpha_1} & \frac{a_1 p_{12}}{\alpha_2} \\ \frac{a_2 p_{21}}{\alpha_1} & \frac{a_2 p_{22}}{\alpha_2} \end{bmatrix}.$$

The matrix $\mathscr{R}$ is similar to the next generation matrix K (and therefore has the same eigenvalues), since

$$\mathscr{R} = T^{-1} K T,$$

where

$$T = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}.$$

The final size relation makes it possible to calculate $S_1(\infty)$ and $S_2(\infty)$ and thence the number of disease cases

$$[N_1 - S_1(\infty)] + [N_2 - S_2(\infty)].$$

The final size relation takes a simpler form if the mixing is proportionate. With proportionate mixing, since

$$p_{11} = p_{21} = p_1, \quad p_{12} = p_{22} = p_2,$$

(5.17) implies

$$a_2 \log \frac{S_1(0)}{S_1(\infty)} = a_1 \log \frac{S_2(0)}{S_2(\infty)},$$

and we may write the final size relations as

$$\begin{aligned} &\log \frac{S_1(0)}{S_1(\infty)} = \frac{a_1 p_1}{\alpha_1}\left[1 - \frac{S_1(\infty)}{N_1}\right] + \frac{a_1 p_2}{\alpha_2}\left[1 - \frac{S_2(\infty)}{N_2}\right], \\ &\left[\frac{S_1(\infty)}{S_1(0)}\right]^{a_2} = \left[\frac{S_2(\infty)}{S_2(0)}\right]^{a_1}. \end{aligned} \tag{5.18}$$

We recall that in the case of proportionate mixing

$$\mathscr{R}_0 = \frac{p_1 a_1}{\alpha_1} + \frac{p_2 a_2}{\alpha_2}.$$

The second equation of (5.18) implies that if $a_1 > a_2$, then

$$1 - \frac{S_1(\infty)}{S_1(0)} > 1 - \frac{S_2(\infty)}{S_2(0)},$$

that is, the attack ratio is greater in the more active group.

It is not difficult to show that the final size relations (5.18) give a unique set of final numbers of susceptibles in each group. The final size relation can also be obtained in a similar way for more complicated compartmental models [4–6].

The model can be extended easily to an arbitrary number of groups with different activity levels. It is also important to be able to describe models with more stages in the progression through compartments, and models in which there are differences between groups in susceptibility. For example, influenza has two characteristics not included in the model (5.11) that are of importance. There is a latent period between infection and the development of infectivity and influenza symptoms. Also, only a fraction of latent members develop symptoms, while the remainder go through an asymptomatic stage in which there is some infectivity. Another important aspect is treatment, which could be directed at either or both group, and could be used for making decisions on how to target groups for treatment. A natural way to proceed in this direction is to build an age of infection model for populations with multiple groups.

Example 2 Consider a two-group endemic SIR model with preferential mixing and group-targeted vaccination:

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu N_i(1-\phi_i) - (\lambda_i(t)+\mu)S_i,\\
\frac{dI_i}{dt} &= \lambda_i(t)S_i - (\alpha+\mu)I_i,\\
\frac{dR_i}{dt} &= \mu N_i\phi_i + \alpha I_i - \mu R_i, \quad i=1,2,
\end{aligned} \tag{5.19}$$

where $N_i = S_i + I_i + R_i$. Here, λ_i represents the force of infection for susceptibles in group i given by

$$\lambda_i = a_i\sigma\sum_{j=1}^{n} p_{ij}\frac{I_j}{N_j}, \tag{5.20}$$

where a_i denotes the average number of contacts an individual in sub-population i has per unit of time (which represents the activity level of group i), and σ is the probability of infection per contact when the contact is with an infectious individual, ϕ_i denotes the proportion of susceptibles in group i vaccinated (and removed) when entering the population. The fraction I_j/N_j gives the probability that a contact is with an infectious individual in sub-population j. The contact matrix (p_{ij}) has the same form as the preferential mixing considered earlier with

$$p_{ij} = \pi_i\delta_{ij} + (1-\pi_i)p_j, \quad i,\ j=1,2. \tag{5.21}$$

The parameter π_i is the fraction of contacts with individuals in the same sub-population, δ_{ij} is the Kronecker delta (i.e., 1 when $i=j$ and 0 otherwise), and

$$p_j = \frac{(1-\pi_j)a_jN_j}{(1-\pi_1)a_1N_1 + (1-\pi_2)a_2N_2}, \quad j=1,2.$$

Clearly, unless all the sub-groups are isolated (i.e., no interactions between the groups), there must be some i with $\pi_i < 1$.

For each sub-population i, if all contacts are with people within the same group (i.e., $p_{ii}=1$ and $p_{ij}=0$ for $i\neq j$), then the basic and control reproduction numbers for group i are, respectively,

$$\mathscr{R}_{0i} = \frac{\sigma a_i}{\mu+\alpha},\ \ \mathscr{R}_{vi} = \mathscr{R}_{0i}(1-\phi_i), \quad i=1,2. \tag{5.22}$$

When there are contacts between sub-populations, i.e., $p_{ii}<1$ or $\pi_i<1$ for some i, we can derive the basic and control reproduction numbers for the metapopulation.

These reproduction numbers will be functions of $\mathscr{R}_{0i}$ or $\mathscr{R}_{vi}$. The next generation matrix K_v (v for vaccination) is

$$K_v = \begin{pmatrix} \mathscr{R}_{v1}p_{11} & \mathscr{R}_{v1}p_{12} \\ \mathscr{R}_{v2}p_{21} & \mathscr{R}_{v2}p_{22} \end{pmatrix}. \tag{5.23}$$

The control reproduction number $\mathscr{R}_v$ for the metapopulation is given by

$$\mathscr{R}_v = \frac{1}{2}\Big[A + D + \sqrt{(A-D)^2 + 4BC}\,\Big], \tag{5.24}$$

where

$$A = \mathscr{R}_{01}p_{11}(1-\phi_1),\ B = \mathscr{R}_{01}p_{12}(1-\phi_1),$$
$$C = \mathscr{R}_{02}p_{21}(1-\phi_2),\ D = \mathscr{R}_{02}p_{22}(1-\phi_2),$$

and $\mathscr{R}_{0i}$ ($i = 1, 2$) are given in (5.22). If $\phi_1 = \phi_2 = 0$, then $\mathscr{R}_v$ reduces to

$$\mathscr{R}_0 = \frac{1}{2}\Big[\mathscr{R}_{01}p_{11} + \mathscr{R}_{02}p_{22} + \sqrt{(\mathscr{R}_{01}p_{11} - \mathscr{R}_{02}p_{22})^2 + 4\mathscr{R}_{01}p_{12}\mathscr{R}_{02}p_{21}}\Big].$$

To study the effects of vaccination strategies, assume that $\mathscr{R}_0 > 1$ in the absence of vaccination and

$$\mathscr{R}_{01} > 1, \qquad \mathscr{R}_{02} > 1. \tag{5.25}$$

Let

$$\Omega = \{(\phi_1, \phi_2) |\ 0 \le \phi_1 < 1,\ 0 \le \phi_2 < 1\}. \tag{5.26}$$

Then each point $(\phi_1, \phi_2) \in \Omega$ represents a vaccination strategy.

Because we are interested in the case when the two groups are not isolated, either $\pi_1 < 1$ or $\pi_2 < 1$. This will be assumed for the results below. It can be shown that, for each fixed $(\phi_1, \phi_2) \in \Omega$, $\mathscr{R}_v$ increases with both π_1 and π_2, i.e.,

$$\frac{\partial \mathscr{R}_v}{\partial \pi_1} > 0, \quad \frac{\partial \mathscr{R}_v}{\partial \pi_2} > 0 \qquad \text{for all } (\pi_1, \pi_2) \in \Omega. \tag{5.27}$$

For each fixed (π_1, π_2), there are different combinations of ϕ_1 and ϕ_2 that can reduce $\mathscr{R}_v$ to be below 1. For ease of presentation, consider the simpler case in which

$$\pi_1 = \pi_2 = \pi,$$

and consider $\mathscr{R}_v = \mathscr{R}_v(\pi)$ as a function of π. Then, for each fixed $\pi \in [0, 1)$, the curve determined by $\mathscr{R}_v(\pi) = 1$ divides the region Ω into two parts: one is the region

$$\Omega_\pi = \{(\phi_1, \phi_2) |\, 0 \le \mathscr{R}_v(\pi) < 1,\ (\phi_1, \phi_2) \in \Omega,\ \ 0 \le \pi < 1\},$$

which includes all points above the curve (see Fig. 5.1), and another is the region

$$D_\pi = \{(\phi_1, \phi_2) |\, \mathscr{R}_v(\pi) > 1,\ (\phi_1, \phi_2) \in \Omega,\ \ 0 \le \pi < 1\},$$

which includes all points below the curve. It can be shown that the region Ω_π decreases as π increases and reduces to the region Ω^* as $\pi \to 0$, while the region D_π decreases as π decreases and reduces to the region D^* as $\pi \to 1$ (see Fig. 5.1). All these curves intersect at a single point (ϕ_{1c}, ϕ_{2c}) with

$$\phi_{1c} = 1 - \frac{1}{\mathscr{R}_{01}}, \quad \phi_{2c} = 1 - \frac{1}{\mathscr{R}_{02}}. \tag{5.28}$$

We observe from Fig. 5.1 that the region Ω^* (lighter-shaded) is determined by the two inequalities

$$\phi_{1c} < \phi_1 < 1, \quad \phi_{2c} < \phi_2 < 1, \tag{5.29}$$

where ϕ_{1c} and ϕ_{2c} are defined in (5.28). For region D^* (darker-shaded), the upper bound is determined by the line

$$\phi_2 = -\mathscr{A}\phi_1 + \mathscr{B}, \tag{5.30}$$

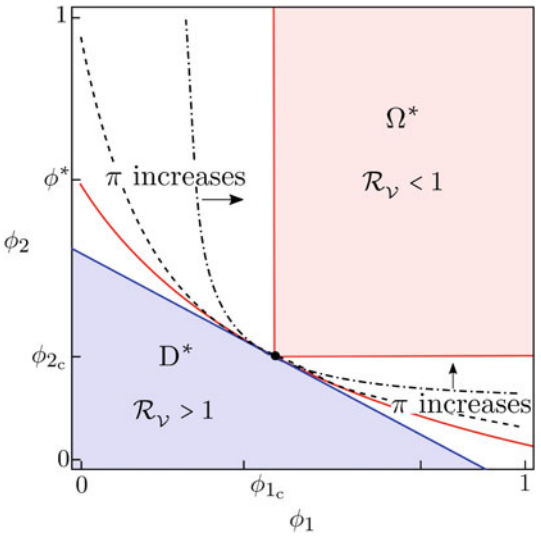

Fig. 5.1 Plot showing the regions Ω^* and D^*. Several curves of $\mathscr{R}_v(\pi) = 1$ for different π values are also shown, with the dashed curves corresponding to $0 < \pi < 1$, the thin solid lines (boundary of Ω^*) corresponding to $\pi = 1$, and the thick line corresponding to $\pi = 0$ (the upper bound of the region D^*). The arrows indicate the direction of change of the curve $\mathscr{R}_v(\pi) = 1$ as π increases from 0 to 1. All of the $\mathscr{R}_v(\pi) = 1$ curves intersect at the single point (ϕ_{1c}, ϕ_{2c}). *Source*: [15]

where

$$\begin{aligned}\mathscr{A} &= \frac{\mathscr{R}_{01}a_1N_1}{\mathscr{R}_{02}a_2N_2},\\ \mathscr{B} &= \frac{(\mathscr{R}_{01}-1)a_1N_1+(R_{02}-1)a_2N_2}{\mathscr{R}_{02}a_2N_2}.\end{aligned} \tag{5.31}$$

The two regions intersect at the point (ϕ_{1c}, ϕ_{2c}).

This result suggests that there is a "lower bound" for vaccination efforts (ϕ_1, ϕ_2), above which the infection can be eradicated regardless of mixing patterns. Similarly, it provides an "upper bound" for vaccination efforts (ϕ_1, ϕ_2), below which the infection cannot be eradicated regardless of mixing patterns (see the definitions for ϕ_1^* and ϕ_2^* defined in (5.32) and see Fig. 5.2 for an illustration of the lower and upper bound). For an "intermediate level" vaccination strategy (ϕ_1, ϕ_2), mixing parameters π_1 and π_2 can play an important role in influencing the effect of vaccination strategies on reducing $\mathscr{R}_v$. Thus, when designing vaccination strategies, one should take into consideration mixing patterns within and between sub-populations.

For given π_i $(i = 1, 2)$, it can be shown that $\frac{\partial \mathscr{R}_v}{\partial \phi_i} < 0$. When the curve $\mathscr{R}_v = 1$ lies between regions D^* and Ω^*, the curve intersects the ϕ_1-axis and ϕ_2-axis at $(\phi_1^*, 0)$ and $(0, \phi_2^*)$, respectively, where

$$\begin{aligned}\phi_1^* &= 1-\frac{1-\mathscr{R}_{02}p_{22}}{\mathscr{R}_{01}p_{11}(1-\mathscr{R}_{02}p_{22})+\mathscr{R}_{01}\mathscr{R}_{02}p_{12}p_{21}},\\ \phi_2^* &= 1-\frac{1-\mathscr{R}_{01}p_{11}}{\mathscr{R}_{02}p_{22}(1-\mathscr{R}_{01}p_{11})+\mathscr{R}_{22}\mathscr{R}_{01}p_{12}p_{21}}.\end{aligned} \tag{5.32}$$

Because $\mathscr{R}_{0i} > 1$ for $i = 1, 2$, it is possible that $\mathscr{R}_{01}p_{11} > 1$ and/or $\mathscr{R}_{02}p_{22} > 1$. Thus, it is possible that $\phi_1^* > 1$ and/or $\phi_2^* > 1$. When $\phi_1^* > 1$, we know from $\partial\mathscr{R}_v/\partial\phi_1 < 1$ that $\mathscr{R}_v > 1$ for any vaccination strategy $(\phi_1, 0)$. Thus, it is impossible to eradicate the infection if only sub-population 1 is vaccinated.

The results described above are based on the control reproduction number. Figure 5.2 shows some simulation results illustrating the effect of vaccination on the prevalence of infection. Different preference levels are used: $\pi_1 = 0.2$ and $\pi_2 = 0.4$, i.e., group 2 has a higher preference contacting people in its own group. Other parameter values used are $\sigma = 0.03$, $\alpha = 0.15$ (an infective period of about 6 days), and $a_1 = 12$, $a_2 = 8$, $\mu = 0.00016$ (a duration of 17 years in school). These values correspond to $\mathscr{R}_{01} = 2.4$ and $\mathscr{R}_{02} = 1.6$. The initial conditions are $x_1(0) = S_1(0)/N_1(0) = 0.4$, $y_1(0) = I_1(0)/N_1(0) = 0.00002$, $x_2(0) = S_2(0)/N_2(0) = 0.6$, $y_2(0) = I_2(0)/N_2(0) = 0.00002$. For this set of parameters, $\phi_1^* = 0.77$ and $\phi_2^* \gg 1$. Figure 5.2a is for a vaccination strategy $(\phi_1, 0)$ with $\phi_1 = 0.2 < \phi_1^*$, for which the infection persists $(\mathscr{R}_v = 1.8 > 1)$, while Fig. 5.2b is for a vaccination strategy $(\phi_1, 0)$ with $\phi_1 = 0.8 > \phi_1^*$, in which case the infection dies out $(\mathscr{R}_v = 0.97 < 1)$.

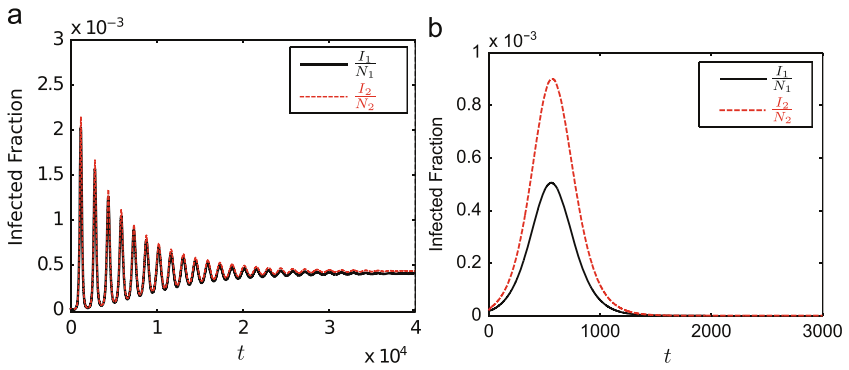

Fig. 5.2 When $\phi_1^* < 1$ and $\phi_2^* > 1$, the disease is eventually eradicated if the vaccination is applied to sub-population 1 alone at a level above ϕ_1^*. (**a**) $(\phi_1, \phi_2) = (0.2, 0)$ and $\phi_1 < \phi_1^* = 0.77$, the disease persists ($\mathscr{R}_v = 1.8$); (**b**) $(\phi_1, \phi_2) = (0.8, 0)$ and $\phi_1 > \phi_1^* = 0.77$, the disease eventually disappears ($\mathscr{R}_v = 0.97$). *Source*: [15]

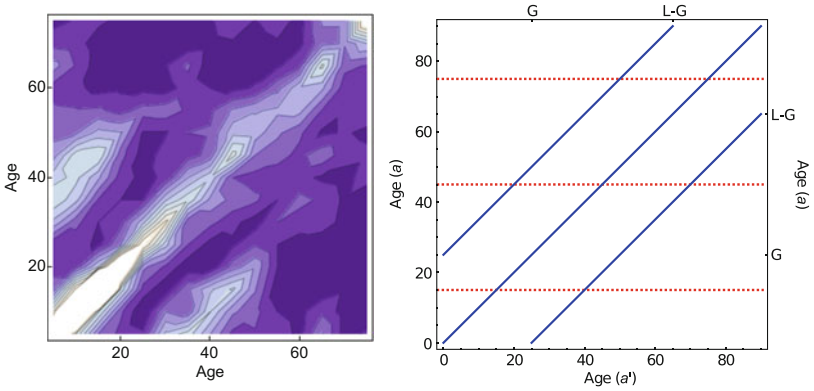

Fig. 5.3 The figure on the left shows observed patterns of contacts between age groups [41]. The lighter areas correspond to higher level of contacts. The figure on the right shows schematic contact matrices illustrating the main and off-diagonals representing contacts among contemporaries, between children and parents, and vice versa

Example 3 This example considers a more general preferential mixing function than the one given in (5.21). This is motivated by the observed mixing pattern shown in Fig. 5.3. Figure 5.3 illustrates the recently collected data reported in [41], which reveals the preferential mixing between parents and children in addition to that among contemporaries.

In the case of n groups, the similar function as (5.21) can be written as

$$p_{ij} = \pi_i \delta_{ij} + (1 - \pi_i) p_j, \quad i, j = 1, 2, \cdots, n, \quad j \neq i \tag{5.33}$$

with

$$p_i = \frac{(1-\pi_i)a_i N_i}{\sum_{k=1}^{n}(1-\pi_k)a_k N_k}, \quad i = 1, 2, \cdots, n.$$

Here, a_i's are the contact rates, δ_{ij}'s are the Kronecker delta (i.e., $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$).

To capture the feature shown in Fig. 5.3 (left), Glasser et al. in [25] extended the mixing function (5.33) to include not only the preference along the diagonal but also along the sub- and super-diagonals as indicated in Fig. 5.3 (right). The function is described by

$$p_{ij} = \phi_{ij} + \Big(1 - \sum_{l=1}^{3}\pi_{li}\Big)p_j, \quad p_j = \frac{\Big(1-\sum_{l=1}^{3}\pi_{li}\Big)a_j N_j}{\sum_{k=1}^{n}\Big(1-\sum_{l=1}^{3}\pi_{lk}\Big)a_k N_k} \tag{5.34}$$

with

$$\phi_{ij} = \begin{cases} \delta_{ij}\pi_{1i} + \delta_{i(j+G)}\pi_{2i}, & i \geq G, \\ \delta_{ij}\pi_{1i} + \delta_{i(j-G)}\pi_{3i}, & i \leq L - G. \end{cases} \tag{5.35}$$

G is the generation time (i.e., average age at which women bear children), L is longevity (i.e., average expectation of life at birth), and $L > G$. The parameters ε_{1i}–ε_{3i} represent the fractions of contacts reserved for contemporaries, children $(j-G)$, and parents $(j+G)$, respectively, and the corresponding delta function is defined as

$$\delta_{i(j\pm G)} = \begin{cases} 1 & \text{if } i = j \pm G, \\ 0 & \text{otherwise.} \end{cases} \tag{5.36}$$

Only people whose ages equal or exceed G can have children, and only those whose ages equal to or less than $L - G$ can have parents, but people aged at least G but not more than $L - G$ can have both children and parents.

Denote the preferential vectors by $\Pi_l = (\pi_{l1}, \pi_{l2}, \cdots, \pi_{ln})$, $l = 1, 2, 3$. When $\Pi_2 = \Pi_3 = 0$, the expression (5.34) reduces to the formula (5.33). For ease of notation, we mix indices and real numbers, but if age classes are 0–4, 5–9, $\cdots$ and $G = 25$ years, for example, by $i > G$ we mean $i >$ class 5. Notice that the non-zero elements of Π_2 and Π_3 are related. If $G = 25$ years, for example, then

$$a_i N_i \pi_{2i} = a_j N_j \pi_{3j}, \quad i = 6, 7, \cdots, \quad j = i - 5.$$

Notice also that $0 \leq \sum_{l=1}^{3}\pi_{li} < 1$.

The mixing function was further extended in [25] to replace the delta functions in (5.36) by Gaussian kernels. Delta formulations are convenient mathematically, but do not allow the age range of one's contemporaries to vary as one ages (e.g., the range narrows perceptively among adolescents), much less differences between the age ranges of one's contemporaries and one's parents or children. By virtue of secular patterns in childbearing, moreover, the age ranges of parents and children may change with age. But the Gaussian formulation allows such variation, reproducing the essential features of these observations. Figure 5.4 demonstrates the comparison between observed mixing patterns (left panel) and the model outcomes (right panel) using the extended mixing functions.

5.3.1 *Optimal Vaccine Allocation in Heterogeneous Populations

One of the significant benefits of models with heterogeneous mixing is that it provides an approach to identifying optimal allocation of vaccines in a metapopulation, particularly when resources are limited. Consider a metapopulation with n sub-populations connected by heterogeneous mixing, which is described by an $n \times n$ matrix $P = (p_{ij})$ for $i, j = 1, 2, \cdots, n$. The following model is studied in [21, 26, 43]:

$$\begin{aligned}
\frac{dS_i}{dt} &= (1-\phi_i)\theta N_i - (\lambda_i + \theta)S_i \\
\frac{dI_i}{dt} &= \lambda_i S_i - (\gamma + \theta)I_i \\
\frac{dR_i}{dt} &= \phi_i \theta N_i + \gamma I_i - \theta R_i \\
N_i &= S_i + I_i + R_i \\
\lambda_i &= \sigma a_i \sum_{j=1}^{n} p_{ij} I_j / N_j, \quad i = 1, 2, \ldots, n,
\end{aligned} \tag{5.37}$$

where ϕ_i are proportions immunized at entry into sub-population i, γ is the *per-capita* recovery rate, θ is the *per-capita* rate for entering and leaving sub-population i so that the population size N_i remains constant. The function λ_i is the force of infection, i.e., *per-capita* hazard rate of infection of susceptible individuals in sub-population i, in which σ is the probability of infection upon contacting an infectious person, a_i is the average contact rate (activity) in sub-population i, p_{ij} is the proportion of ith sub-population's contacts that are with members of jth sub-population, and I_j/N_j is the probability that a randomly encountered member of sub-population j is infectious. A similar model with two-level mixing is studied in [22].

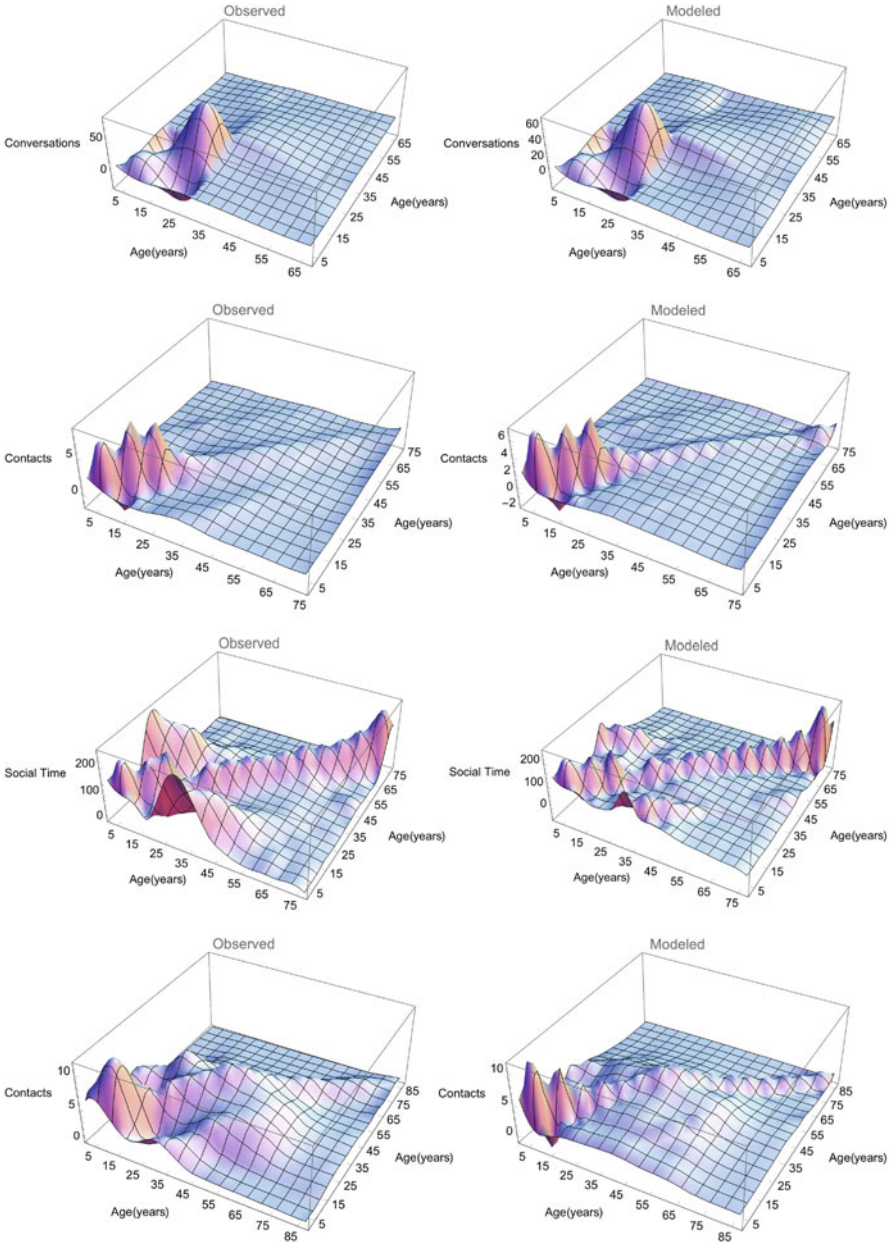

Fig. 5.4 Comparison of the mixing patterns generated by the four empirical studies (top to bottom): [17, 41, 47, 49] (left panel) and the fitted model (5.34) and (5.35) (right panel). Interpolating functions are fitted to geometric means of corresponding row- and column-elements of the mixing matrix *Source*: [25]

The mixing matrix P can incorporate heterogeneities in contact rate (a_i), population size (N_i), preference for mixing within the same sub-population (π_i), etc. Because of these heterogeneities, the optimal combination of vaccination coverages ϕ_i is unlikely to be homogeneous. Typically, the matrix P has to satisfy the following conditions of [14]:

$$\begin{aligned} &p_{ij} \geq 0, \quad i, j = 1, \ldots, n, \\ &\sum_{j=1}^{n} p_{ij} = 1, \quad i = 1, \ldots, n, \\ &a_i N_i p_{ij} = a_j N_j p_{ji}, \quad i, j = 1, \ldots, n. \end{aligned} \tag{5.38}$$

A commonly used non-homogeneous mixing function that satisfies conditions in (5.38) is the preferred mixing function of [30] given by

$$p_{ij} = \pi_i \delta_{ij} + (1 - \pi_i) \frac{(1 - \pi_j) a_j N_j}{\sum_{k=1}^{n} (1 - \pi_k) a_k N_k}, \quad i, j = 1, \ldots, n, \tag{5.39}$$

where $\pi_i \in [0, 1]$ is the fraction of contacts of group i that is reserved for itself (preferential mixing), whereas the complement $(1 - \pi_i)$ is distributed among all sub-populations in proportion to the unreserved contacts, including i (proportionate mixing). We refer to the mixing given by (5.39) as Jacquez-type preferred mixing. Two extreme cases of (5.39) are the *proportionate mixing* when $\pi_i = 0$ and the *isolated mixing* (i.e., no interactions between sub-populations) when $\pi_i = 1$.

For model (5.37), the basic and effective sub-population reproduction numbers, denoted, respectively, by $\mathscr{R}_{0i}$ and $\mathscr{R}_{vi}$, for sub-population i ($i = 1, 2, \ldots, n$) are given by

$$\mathscr{R}_{0i} = \rho a_i, \quad \mathscr{R}_{vi} = \mathscr{R}_{0i}(1 - \phi_i), \quad i = 1, 2, \ldots, n, \tag{5.40}$$

where

$$\rho = \frac{\sigma}{\gamma + \theta}.$$

The next generation matrix (NGM) corresponding to this metapopulation model is

$$K_v = \begin{pmatrix} \mathscr{R}_{v1} p_{11} & \mathscr{R}_{v1} p_{12} & \cdots & \mathscr{R}_{v1} p_{1n} \\ \mathscr{R}_{v2} p_{21} & \mathscr{R}_{v2} p_{22} & \cdots & \mathscr{R}_{v2} p_{2n} \\ \vdots & \ddots & & \vdots \\ \mathscr{R}_{vn} p_{n1} & \mathscr{R}_{vn} p_{n2} & \cdots & \mathscr{R}_{vn} p_{nn} \end{pmatrix}. \tag{5.41}$$

Then the effective reproduction number for the metapopulation is given as

$$\mathscr{R}_v = r(K_v),$$

which is the spectral radius (and the dominant eigenvalue, by the Perron–Frobenius Theorem) of the non-negative matrix K_v. Let $\phi = (\phi_1, \phi_2, \cdots, \phi_n)$. Naturally, $\mathscr{R}_v = \mathscr{R}_v(\phi)$ is a function of ϕ. The total number of vaccine doses, denoted by η, is $\eta = \sum_{i=1}^{n} \phi_i N_i$. For demonstration purposes, we will assume that vaccine efficacy is 100%. We focus on identifying the most efficient allocation of vaccine $\phi = (\phi_1, \phi_2, \cdots, \phi_n) \in [0, 1]^n$ for reducing $\mathscr{R}_v$ with limited vaccine doses η or using fewest doses to achieve $\mathscr{R}_v < 1$ (to prevent outbreaks). More specifically, we consider the following two constrained optimization problems:

(I) Minimize $\mathscr{R}_v = \mathscr{R}_v(\phi)$, subject to $\ell(\phi) := \sum_{i=1}^{n} \phi_i N_i = \eta$, for $\phi \in [0, 1]^n$.

(II) Minimize $\eta = \sum_{i=1}^{n} \phi_i N_i$, subject to $\mathscr{R}_v(\phi) \leq 1$, for $\phi \in [0, 1]^n$.

Consider the optimization problem only for the case of $\mathscr{R}_0 = \mathscr{R}_v(0) \geq 1$, as there will be no outbreak if $\mathscr{R}_0 < 1$. If a solution to Problem (I) exists for a given value of η, let $\Phi^* = \Phi^*(\eta)$ and $\mathscr{R}_{v\{min\}}(\eta)$ denote the optimal vaccination allocation and corresponding minimum reproduction number, , respectively. Let

$$\Omega_{\phi}^{(n)}(\eta) := \Big\{(\phi_1, \phi_2, \cdots, \phi_n) : \ell(\phi) = \eta\Big\}.$$

Then, for the solution Φ^* to be feasible, we need to have

$$\begin{aligned} &\Phi^*(\eta) = (\phi_1^*(\eta), \phi_2^*(\eta), \cdots, \phi_n^*(\eta)) \in [0, 1]^n, \\ &\mathscr{R}_{v\{min\}}(\eta) = \min_{\Omega_{\phi}^{(n)}(\eta) \cap [0,1]^n} \mathscr{R}_v = \mathscr{R}_v\big|_{\Phi^*(\eta)}. \end{aligned} \tag{5.42}$$

An optimal solution $\Phi^*(\eta)$ to Problem (I) that lies in the interior of the unit hypercube must satisfy the following equations:

$$\begin{aligned} \nabla \mathscr{R}_v\big|_{\Phi^*(\eta)} &= \widetilde{\lambda} \nabla \ell = \widetilde{\lambda}(N_1, \cdots, N_n), \quad \Phi^*(\eta) \in (0, 1)^n, \\ \ell\big|_{\Phi^*(\eta)} &= \sum_{i=1}^{n} \phi_i^*(\eta) N_i = \eta, \end{aligned} \tag{5.43}$$

where the constant $\widetilde{\lambda}$ is the Lagrange multiplier.

Similarly, if an optimal solution to Problem (II) exists, it is useful practically to have an explicit expression or estimate of the bounds for the minimum vaccine doses needed, which we denote by η_*. The significance of η_* is that it is the smallest number of vaccination doses that can prevent outbreaks under an optimal vaccination policy.

To find η_*, notice that $\mathscr{R}_v(\phi)$ is a monotonically decreasing function of ϕ_i, and thus, a decreasing function of $\eta = \sum_{i=1}^n \phi_i N_i$. Therefore, the inequality constraint $\mathscr{R}_v(\phi) \le 1$ can be replaced by an equality constraint $\mathscr{R}_v(\phi) = 1$, and thus,

$$\eta_* = \min_{\{\mathscr{R}_v(\phi)=1\}\cap[0,1]^n} \ell(\phi).$$

It follows that η_* is the minimum of $\eta \in [0, N]$ such that $\mathscr{R}_{v\{min\}}(\eta) = 1$ and can be found by solving the equation:

$$\mathscr{R}_{v\{min\}}(\eta_*) = \mathscr{R}_v\big|_{P^*(\eta_*)} = 1. \tag{5.44}$$

Before we present the results of optimal solutions for Problems (I) and (II), we state the following important results, which is proved in [43], regarding the bounds for the effective reproduction number $\mathscr{R}_v$ for general mixing matrices (not just Jacquez-type preferred mixing given in (5.39)).

Theorem 5.2 (Bounds for $\mathscr{R}_v(\phi)$) *Let P be a non-negative, invertible, irreducible matrix such that $-P^{-1}$ is essentially non-negative and the conditions (5.38) are satisfied. Then*

(a) The lower and upper bounds of $\mathscr{R}_v(\phi)$ are

$$\sum_{i=1}^n \omega_i \mathscr{R}_{vi} \le \mathscr{R}_v \le \max\{\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn}\}, \quad \text{where } \omega_i = \frac{a_i N_i}{\sum_{k=1}^n a_k N_k}. \tag{5.45}$$

(b) The lower and upper bounds of $\mathscr{R}_v(\phi)$ correspond to the cases of proportionate mixing and isolated mixing, respectively.

For the optimal solutions of Problems (I) and (II), it is shown in [43] that for the case of $n = 2$ explicit expressions for Φ^* and η_* are possible, and for the case of $n > 2$ their upper and lower bounds can be obtained.

For ease of presentation, introduce the following notation:

$$\begin{aligned} \kappa_1 &:= p_{22}\sqrt{N_1 N_2}\mathscr{R}_{02} - N_2\sqrt{p_{12}p_{21}\mathscr{R}_{01}\mathscr{R}_{02}}, \\ \kappa_2 &:= p_{11}\sqrt{N_1 N_2}\mathscr{R}_{01} - N_1\sqrt{p_{12}p_{21}\mathscr{R}_{01}\mathscr{R}_{02}}, \\ \eta_0 &:= N - \frac{\kappa_1 N_1 + \kappa_2 N_2}{\max\{\kappa_1, \kappa_2\}}. \end{aligned} \tag{5.46}$$

For the mixing given in (5.39), it is easy to verify the following fact:

$$|P| = \begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix} = \pi_1\pi_2 + \frac{\pi_1(1-\pi_2)^2 a_2 N_2 + \pi_2(1-\pi_1)^2 a_1 N_1}{(1-\pi_1)a_1 N_1 + (1-\pi_2)a_2 N_2} > 0, \tag{5.47}$$

provided that $\pi_i \in (0, 1)$, $a_i > 0$, and $N_i > 0$, $i = 1, 2$.

Theorem 5.3 (Optimal Solution to Problem (I) When $n = 2$) *Consider $\mathscr{R}_v = \mathscr{R}_v(\phi_1, \phi_2)$ as a function of ϕ_1 and ϕ_2, and let η_0 and κ_i be given in (5.46). Assume that condition (5.47) holds.*

(a) For a given value of η, the optimal point $\Phi^*(\eta)$ exists and lies in the interior of the unit square if and only if

$$\eta_0 < \eta < N \quad \text{and} \quad \kappa_i > 0 \text{ for } i = 1, 2. \tag{5.48}$$

(b) For each $\eta_0 < \eta < N$, the explicit formulae for $P^*(\eta)$ and $\mathscr{R}_{v\{min\}}(\eta)$ are given by

$$\Phi^*(\eta) = (1, 1) - \frac{N - \eta}{\kappa_1 N_1 + \kappa_2 N_2}(\kappa_1, \kappa_2) \tag{5.49}$$

$$\mathscr{R}_{v\{min\}}(\eta) = |P|\mathscr{R}_{01}\mathscr{R}_{02}\sqrt{N_1 N_2}\frac{N - \eta}{\kappa_1 N_1 + \kappa_2 N_2}. \tag{5.50}$$

(c) If $0 < \eta < \eta_0$, the minimum point $\Phi^*(\eta)$ is one of the boundary points $(\eta/N_1, 0)$ or $(0, \eta/N_2)$ and hence

$$\mathscr{R}_{v\{min\}}(\eta) = \min\{\mathscr{R}_v(\eta/N_1, 0), \mathscr{R}_v(0, \eta/N_2)\}.$$

For the general case of $n > 2$, Theorem 5.2 can be used to derive the lower and upper bounds for the minimum reproduction number $\mathscr{R}_{v\{min\}}(\eta)$. To facilitate biological interpretations, introduce the following notation:

$$\begin{aligned}
&f_i := N_i/N,\ 1 \le i \le n && \text{Population fraction of sub-population } i;\\
&\mathscr{U} := \sum_{i=1}^{n}(1 - \phi_i) f_i && \text{Population fraction unvaccinated;}\\
&\widehat{\mathscr{R}}_0 := \sum_{i=1}^{n}\mathscr{R}_{0i} f_i && \text{Population weighted reproduction number;}\\
&\mathscr{R}_0^{\diamond} := \Big(\sum_{i=1}^{n}\frac{1}{\mathscr{R}_{0i}} f_i\Big)^{-1} && \text{Harmonic mean of } \mathscr{R}_{0i} \text{ weighted by}\\
& && \text{sub-population fractions } f_i;\\
&\widetilde{\mathscr{R}}_0 := \min_i \mathscr{R}_{0i}^2/\widehat{\mathscr{R}}_0 && \text{Analogous to a scaled reproduction number.}
\end{aligned} \tag{5.51}$$

The following results provide the lower and upper bounds for the minimum $\mathscr{R}_{v\{min\}}(\eta)$ in Problem (I):

Theorem 5.4 *Assume that the conditions of Theorem 5.2 hold. Let $\eta < N$, and let $\mathscr{U}$, $\widehat{\mathscr{R}}_0$, $\mathscr{R}_0^{\diamond}$, and $\widetilde{\mathscr{R}}_0$ be defined in (5.51).*

(a) The bounds of $\mathscr{R}_{v\{min\}}(\eta)$ *for* $\phi \in \Omega_p^{(n)}(\eta) \cap [0,1]^n$ *are*

$$\widetilde{\mathscr{R}}_0 \mathscr{U} \le \mathscr{R}_{v\{min\}}(\eta) \le \mathscr{R}_0^{\diamond} \mathscr{U}. \tag{5.52}$$

(b) If $\mathscr{R}_{0i} > 1$ *for all* i*, then*

$$\frac{\eta_*}{N} \le 1 - \frac{1}{\mathscr{R}_0^{\diamond}}. \tag{5.53}$$

Remarks The bounds for the optimal solutions have clear biological meanings based on the biological interpretations of the quantities in (5.51). (i) Note that $\mathscr{R}_0^{\diamond}$ and $\widetilde{\mathscr{R}}_0$ are weighted basic reproduction numbers, and the factor $\mathscr{U}$ is the fraction of the overall population that remains susceptible. In light of this, we see that the lower and upper bounds for $\mathscr{R}_{v\{min\}}(\eta)$ in (5.52) take the familiar form of an effective reproduction number. (ii) For the upper bound of η_*, if $a_i = a$ are all the same, we have $\mathscr{R}_0^{\diamond} = \mathscr{R}_0$, in which case the upper bound in (5.53) becomes $1 - 1/\mathscr{R}_0$. This is similar to the usual formula for the critical vaccination fraction $\phi_c = 1 - 1/\mathscr{R}_0$, for which the number of vaccinated is $\eta_c = \phi_c N = N(1 - 1/\mathscr{R}_0)$.

Although various observations about the effect of mixing on reproduction numbers have been made in previous studies, the result stated in Theorem 5.2 provided a definitive lower and upper bounds corresponding to the proportionate and the isolated mixing (a rigorous proof can be found in [43]) for a large class of mixing matrix P (not just Jacquez-type). Using a model metapopulation composed of a city and several villages, May and Anderson [36, 37] showed that heterogeneity in relevant sub-population characteristics also increased $\mathscr{R}_v$. Hethcote and van Ark [28] argued that person-to-person contact rates in densely populated urban areas should be no more than twice those in sparsely populated rural ones. This change in parameter values diminished the apparent effect of heterogeneity. The facts that population heterogeneities tend to increase $\mathscr{R}_0$ and that models assuming proportionate mixing generate lower values of $\mathscr{R}_0$ have been suggested by other researchers [1, 3, 20].

Example 4 Figure 5.5 illustrates an example from [21], which extends May and Anderson's [36] conclusion that "under a uniformly applied immunization program, the overall fraction that must be immunized is larger than would be estimated by (incorrectly) assuming the population to be homogeneously mixed." Consider $\mathscr{R}_v = \mathscr{R}_v(\phi_1, \phi_2)$ as a function of vaccination coverage (ϕ_1, ϕ_2). The two contour plots of $\mathscr{R}_v$ are for the cases of (a) homogeneous contacts ($a_1 = a_2 = 10$) and (b) heterogeneous contacts ($a_1 = 8, a_2 = 12$), while other parameters are the same for the two sub-populations ($N_1 = N_2$, $\pi_1 = \pi_2 = 0.6$, $\sigma = 0.05$, $\gamma = 1/7$, $\theta = 1/(365 \times 70)$). Note that the total number of vaccine doses is given by $\phi_1 N_1 + \phi_2 N_2$. Because $N_1 = N_2$, a vaccination pair (ϕ_1, ϕ_2) that minimizes the total doses if and only if it minimizes the quantity $\phi_1 + \phi_2$. The thicker curve is the contour for $\mathscr{R}_v = 1$. The thick dashed line with slope -1 is $\phi_1 + \phi_2 = c$ for a

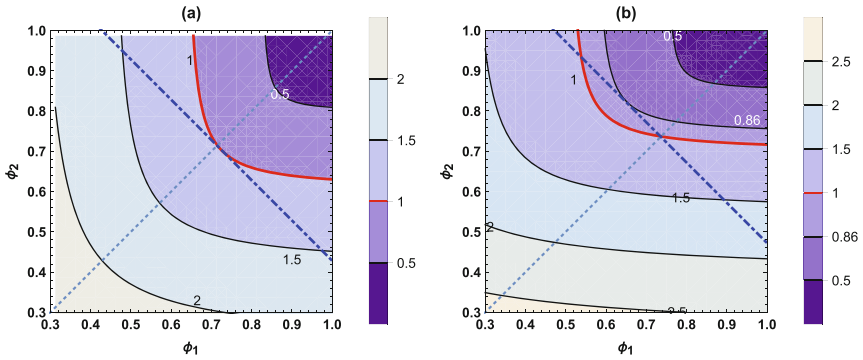

Fig. 5.5 Contour plots of $\mathscr{R}_v$ as a function of ϕ_1 and ϕ_2 for (**a**) homogeneous population ($a_1 = a_2$) and (**b**) heterogeneous population ($a_1 \neq a_2$). For both plots, $\pi_1 = \pi_2$ and $N_1 = N_2$. The thick solid curve represents the contour $\mathscr{R}_v = 1$. The thin and thick dashed lines correspond to $\phi_1 = \phi_2$ and $\phi + \phi_2 = 2 \times 0.74$, respectively. All points (ϕ_1, ϕ_2) on the line $\phi_1 + \phi_2 = c$ for a constant $c > 0$ correspond to the same total vaccination doses. *Source*: [21]

constant $c > 0$. It shows in (a) that $\mathscr{R}_v(0.74, 0.74) = 1$ and that $\mathscr{R}_v(\phi_1, \phi_2) > 1$ for all other pairs (ϕ_1, ϕ_2) with $\phi_1 + \phi_2 = 2 \times 0.74$. This suggests that the optimal allocation is the homogeneous coverage $\phi_1 = \phi_2$. However, the plot in (b) shows a very different result. Particularly, among all pairs (ϕ_1, ϕ_2) on the line $\phi_1 + \phi_2 = 2 \times 0.74$, some can make $\mathscr{R}_v(\phi_1, \phi_2) < 1$. In fact, there is one point (ϕ_{1c}, ϕ_{2c}) at which the minimum $\mathscr{R}_v(\phi_{1c}, \phi_{2c}) = 0.86$ is achieved. This suggests that, in a non-homogeneous population, uniform coverage (equal ϕ_i for all i) may not be the most efficient.

5.4 A Heterogeneous Mixing Age of Infection Model

The basic age of infection model extends the simple SIR epidemic model by allowing an arbitrary number of stages in the model and arbitrary distributions of stay in each stage. However, it does not include the possibility of subgroups with different activity levels and heterogeneous mixing between subgroups. This possibility can be included in a heterogeneous mixing age of infection model. As in homogeneous mixing models, the age of infection approach is more general than simpler models in several respects. Age of infection models allow arbitrary distributions of stay in compartments and arbitrary sequences of compartments. In addition, they allow variable infectivity. This can be included in the kernel $A(s)$ which leads to the infectivity function $\varphi(t)$ describing infectivity rather than simply counting the number of infectives.

As in the previous section, we consider two sub-populations of sizes N_1, N_2, respectively, each divided into susceptibles and infected members with subscripts to identify the sub-population. Suppose that $A_i(s)$ is the mean infectivity of individuals who have been infected s time units previously, and that a_1, a_2 are the contact rates of the two sub-populations. It is necessary to describe also the mixing between the two groups. Suppose that the fraction of contacts made by a member of group i that is with a member of group j is p_{ij}, $i, j = 1, 2$. Then

$$p_{11} + p_{12} = p_{21} + p_{22} = 1.$$

A two-group model may describe a population with groups differing by activity levels and possibly by vulnerability to infection, so that $a_1 \neq a_2$ but $A_1(s) = A_2(s)$. It may also describe a population with one group which has been vaccinated against infection, so that the two groups have the same activity level but different disease model parameters. In this case, $a_1 = a_2$ but $A_1(\tau) \neq A_2(\tau)$. In this model, any differences between groups in susceptibility or infectivity are included in the factors $A_1(s)$, $A_2(s)$.

An age of infection model with two subgroups is

$$S_i' = -a_i S_i \left[\frac{p_{ii}}{N_i} \varphi_i + \frac{p_{ij}}{N_j} \varphi_j \right]$$

$$\varphi_i(t) = \int_0^\infty [-S_i'(t-\tau)] A_i(\tau) d\tau, \quad i, j = 1, 2, \quad i \neq j.$$

Here, $\varphi_i(t)$ is the total infectivity of infected members of group i ($i = 1, 2$).

As for the homogeneous mixing model, we may write this model using only the equations for S_i,

$$S_i'(t) = -a_i S_i(t) \left[\frac{p_{ii}}{N_i} \int_0^\infty A_i(s) S_i'(t-s) ds + \frac{p_{ij}}{N_j} \int_0^\infty A_j(s) S_j'(t-s) ds \right], \quad i, j = 1, 2, \quad i \neq j. \tag{5.54}$$

The next generation matrix is

$$P = \begin{bmatrix} a_1 p_{11} \int_0^\infty A_1(\tau) ds & a_1 p_{12} \frac{N_1}{N_2} \int_0^\infty A_2(s) ds \\ a_2 p_{21} \frac{N_2}{N_1} \int_0^\infty A_1(s) ds & a_2 p_{22} \int_0^\infty A_2(s) ds \end{bmatrix}.$$

The matrix P is similar to the matrix $Q = R^{-1} P R$, with

$$R = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$$

and

$$Q = \begin{bmatrix} a_1 p_{11} \int_0^\infty A_1(s)ds & a_1 p_{12} \int_0^\infty A_2(s)ds \\ a_2 p_{21} \int_0^\infty A_1(s)ds & a_2 p_{22} \int_0^\infty A_2(s)ds \end{bmatrix}.$$

Thus $\mathscr{R}_0$ is the largest root of

$$\det \begin{bmatrix} a_1 p_{11} \int_0^\infty A_1(s)ds - \lambda & a_1 p_{12} \int_0^\infty A_2(s)ds \\ a_2 p_{21} \int_0^\infty A_1(s)ds & a_2 p_{22} \int_0^\infty A_2(s)ds - \lambda \end{bmatrix} = 0. \tag{5.55}$$

In order to obtain an invasion criterion, initially when $S_1(t)$ is close to $S_1(0) = N_1$ and $S_2(t)$ is close to $S_2(0) = N_2$, we replace $S_1(t)$ and $S_2(t)$ by N_1, N_2, respectively, to give a linear system, and the condition that this linear system has a solution $S_i(t) = N_i e^{rt}$ $(i = 1, 2)$ is

$$1 = a_i p_{i1} \int_0^\infty e^{-rs} A_1(s)ds + a_i p_{i2} \int_0^\infty e^{-rs} A_2(s)ds, \quad i = 1, 2. \tag{5.56}$$

The initial exponential growth rate is the solution r of the equation

$$\det \begin{bmatrix} a_1 p_{11} \int_0^\infty e^{-rs} A_1(s)ds - 1 & a_1 p_{12} \int_0^\infty e^{-rs} A_2(s)ds \\ a_2 p_{21} \int_0^\infty e^{-rs} A_1(s)ds & a_2 p_{22} \int_0^\infty e^{-rs} A_2(s)ds - 1 \end{bmatrix} = 0. \tag{5.57}$$

In the special case of proportionate mixing, in which $p_{11} = p_{21}$, $p_{12} = p_{22}$, so that $p_{12} p_{21} = p_{11} p_{22}$, the basic reproduction number is given by

$$\mathscr{R}_0 = a_1 p_{11} \int_0^\infty A_1(s)ds + a_2 p_{22} \int_0^\infty A_2(s)ds,$$

and Eq. (5.57) reduces to

$$a_1 p_{11} \int_0^\infty e^{-rs} A_1(s)ds + a_2 p_{22} \int_0^\infty e^{-rs} A_i(s)ds = 1. \tag{5.58}$$

There is an epidemic if and only if $\mathscr{R}_0 > 1$.

In the special case in which the two groups have the same infectivity distribution but may have different activity levels and possibly vulnerability to infection, so that $A_1(s) = A_2(s) = A(s)$, $\mathscr{R}_0$ is the largest root of

$$\det \begin{bmatrix} a_1 p_{11} \int_0^\infty A(s)ds - \lambda & a_1 p_{12} \int_0^\infty A(s)ds \\ a_2 p_{21} \int_0^\infty A(s)ds & a_2 p_{22} \int_0^\infty A(s)ds - \lambda \end{bmatrix}$$

and the initial exponential growth rate is the solution r of the equation

$$\det\begin{bmatrix} a_1p_{11}\int_0^\infty e^{-rs}A(s)ds - 1 & a_1p_{12}\int_0^\infty e^{-rs}A(s)ds \\ a_2p_{21}\int_0^\infty e^{-rs}A(s)ds & a_2p_{22}\int_0^\infty e^{-rs}A(s)ds - 1 \end{bmatrix} = 0. \tag{5.59}$$

Comparing Eqs. (5.55) and (5.59), we see that each of $\mathscr{R}_0/\int_0^\infty A(\tau)d\tau$ and $1/\int_0^\infty e^{-r\tau}A(\tau)d\tau$ is the largest eigenvalue of the matrix

$$\begin{bmatrix} a_1p_{11} & a_1p_{12} \\ a_2p_{21} & a_2p_{22} \end{bmatrix},$$

the largest root of the equation

$$x^2 - (a_1p_{11} + a_2p_{22})x + a_1a_2(p_{11}p_{22} - p_{12}p_{21}) = 0.$$

Thus

$$\frac{\mathscr{R}_0}{\int_0^\infty A(s)ds} = \frac{1}{\int_0^\infty e^{-rs}A(s)ds},$$

which implies the same relation as for the homogeneous mixing model. Thus, if we assume heterogeneous mixing, we obtain the same estimate of the reproduction number from observation of the initial exponential growth rate, and this conclusion remains valid for an arbitrary number of groups with different contact rates. The estimate of the basic reproduction number from the initial exponential growth rate does not depend on heterogeneity of the model. This result does not generalize to the case $A_1(s) \neq A_2(s)$, but it does remain valid for an arbitrary number of groups with different contact rates.

5.4.1 *The Final Size of a Heterogeneous Mixing Epidemic*

With homogeneous mixing, knowledge of the basic reproduction number translates into knowledge of the final size of the epidemic. However, with heterogeneous mixing, even in the simplest case of proportionate mixing, the size of the epidemic is not determined uniquely by the basic reproduction number.

For the heterogeneous mixing model (5.54) there is a pair of final size relations. We divide the equation for S_1 in (5.54) by $S_i(t)$ and integrate with respect to t from 0 to ∞. Much as in the derivation of the final size relation for the homogeneous

mixing model we obtain a pair of final size relations which may be solved for $S_1(\infty)$ and $S_2(\infty)$:

$$\log \frac{S_i(0)}{S_i(\infty)} = \sum_{j=1}^{2} \left[a_i \frac{p_{ij}}{N_j} \big(N_j - S_j(\infty)\big) \int_0^\infty A_j(s)ds \right], \quad i = 1, 2. \tag{5.60}$$

The system of equations (5.60) has a unique solution $(S_1(\infty), S_2(\infty))$. In order to prove this, we define

$$g_i(x_1, x_2) = \log \frac{S_i(0)}{x_i} - a_i \sum_{j=1}^{2} p_{ij} \left[1 - \frac{x_j}{N_j} \right] \int_0^\infty A_j(s)ds.$$

A solution of (5.60) is a solution (x_1, x_2) of the system

$$g_i(x_1, x_2) = 0, \quad i = 1, 2.$$

For each x_2, $g_1(0^+, x_2) > 0$, $g_1(S_1(0), x_2) < 0$. Also, as a function of x_1, $g_1(x_1, x_2)$ either decreases or decreases initially and then increases to a negative value when $x_1 = S_1(0)$. Thus for each $x_2 < S_2(0)$, there is a unique $x_1(x_2)$ such that $g_1(x_1(x_2), x_2) = 0$. Also, since $g_1(x_1, x_2)$ is an increasing function of x_2, the function $x_1(x_2)$ is increasing. Now, since $g_2(x_1, 0^+) > 0$, $g_2(x_1, S_2(0)) < 0$, there exists x_2 such that $g_2(x_1(x_2), x_2) = 0$. Also, $g_2(x_1(x_2), x_2)$ either decreases monotonically or decreases initially and then increases to a negative value when $x_2 = S_2(0)$. Therefore this solution is also unique. This implies that

$$(x_1(x_2), x_2)$$

is the unique solution of the final size relations.

Numerical simulations indicate that models with heterogeneous mixing may give very different epidemic sizes than models with the same basic reproduction number and homogeneous mixing. The reproduction number of an epidemic model is not sufficient to determine the size of the epidemic if there is heterogeneity in the model. We conjecture that for a given value of the basic reproduction number the maximum epidemic size for any mixing is obtained with homogeneous mixing.

Assume that the parameters $N_1, N_2, \int_0^\infty A(\tau)d\tau$ remain fixed and attempt to minimize $S_1(\infty) + S_2(\infty)$ as a function of a_1, a_2 (with a_1, a_2 constrained to keep $p_1a_1 + p_2a_2 = k$ fixed and p_1, p_2 as specified by proportionate mixing). Homogeneous mixing corresponds to $a_1 = a_2$.

The constraint relating a_1, a_2 implies that

$$\frac{da_2}{da_1} = \frac{2a_1 - k}{k - 2a_2} \cdot \frac{N_1}{N_2};$$

when $a_1 = a_2 = k$ we have

$$\frac{da_2}{da_1} = -\frac{N_1}{N_2}.$$

Also, when $a_1 = a_2$,

$$p_1 = \frac{N_1}{N_1 + N_2}, \quad p_2 = \frac{N_2}{N_1 + N_2}, \quad \frac{S_1}{N_1} = \frac{S_2}{N_2}, \quad \frac{dp_1}{da_1} = \frac{N_1}{kN}.$$

If we differentiate with respect to a_1, we can calculate that

$$\frac{d[S_1(\infty) + S_2(\infty)]}{da_1} = 0$$

when $a_1 = a_2$. We believe that $a_1 = a_2$ is the only critical point of $S_1(\infty) + S_2(\infty)$, although we have not been able to verify this analytically. If $a_1 = a_2$ is the only critical point of $S_1(\infty) + S_2(\infty)$, this critical point must be a minimum. We conjecture that this result is also valid if we allow arbitrary mixing, that is, we conjecture that for a given value of the basic reproduction number the maximum epidemic size for any mixing is obtained with homogeneous mixing.

While we have confined the description of the heterogeneous mixing situation to a two-group model, the extension to an arbitrary number of groups is straightforward. We suggest that in advance planning for a pandemic, the number of groups to be considered for different treatment rates should determine the number of groups to be used in the model. On the other hand, the number of groups to be considered should also depend on the amount and reliability of data, and these two criteria may be contradictory. A model with fewer groups and parameters chosen as weighted averages of the parameters for a model with more groups may give predictions that are quite similar to those of the more detailed models. We suggest also that use of the final size relations for a model with total population size assumed constant is a good time-saving procedure for making predictions if the disease death rate is small.

We have seen that in the case of homogeneous mixing, knowledge of the initial exponential growth rate and the infective period distribution is sufficient to determine the basic reproduction number and thence the final size of an epidemic. In the case of heterogeneous mixing, knowledge of the initial exponential growth rate and the infective period distribution is sufficient to determine the basic reproduction number, but not to determine the final size of the epidemic.

This raises the question of what additional information that may be measured at the start of a disease outbreak would suffice to determine the epidemic final size if the mixing is heterogeneous.

We assume that $A_1(s)$, $A_2(s)$, and the mixing matrix

$$M = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

are known. The next generation matrix is

$$K = \begin{bmatrix} a_1 \frac{p_{11}}{N_1} \int_0^\infty A_1(s)ds & a_1 \frac{p_{12}}{N_2} \int_0^\infty A_2(s)ds \\ a_2 \frac{p_{21}}{N_1} \int_0^\infty A_1(s)ds & a_2 \frac{p_{22}}{N_2} \int_0^\infty A_2(s)ds \end{bmatrix},$$

and $\mathscr{R}_0$ is the largest (positive) eigenvalue of this matrix. There is a corresponding eigenvector with positive components

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Since the components of this eigenvector give the proportions of infective cases in the two groups initially, it is reasonable to hope to be able to determine this eigenvector from early outbreak data.

The general final size relation is

$$\log \frac{S_i(0)}{S_i(\infty)} = \sum_{j=1}^{2} \left[a_i p_{ij} \Big(1 - \frac{S_j(\infty)}{N_j} \Big) \int_0^\infty A_j(s)ds \right], \quad i = 1, 2.$$

These equations may be solved for $S_1(\infty)$, $S_2(\infty)$ if the contact rates a_1, a_2 can be determined from the available information.

The condition that the vector $\mathbf{u}$ with components (u_1, u_2) is an eigenvector of the next generation matrix corresponding to the eigenvalue $\mathscr{R}_0$ is

$$a_i \big(p_{i1} u_1 + p_{i2} u_2 \big) \int_0^\infty A_i(s)ds = \mathscr{R}_0 u_i, \quad i = 1, 2,$$

and since it is assumed that the function $A(\tau)$, the vector $\mathbf{u}$, and the mixing matrix (p_{ij}) are known these equations determine a_1 and a_2.

In vector notation, if we define the column vector

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

and the row vectors

$$M_j = \begin{bmatrix} p_{j1} & p_{j2} \end{bmatrix},$$

we have

$$a_j = \frac{\mathscr{R}_0}{M_j \mathbf{u} \int_0^\infty A_j(s)ds} u_j.$$

When these values are substituted into the final size system, $S_1(\infty)$ and $S_2(\infty)$ may be determined. This argument extends easily to models with an arbitrary number of activity groups.

In real-life applications, there are usually many groups, and the final size of an epidemic is obtained most efficiently by numerical simulations. The results obtained here are more likely to be useful in theoretical applications, such as comparisons of different control strategies.

One may think of the case $a_1 \neq a_2$, $A_1(s) = A_2(s)$ as a model for a disease with heterogeneous mixing but not treatment and the case $a_1 = a_2$, $A_1(s) \neq A_2(s)$ as a model for a disease in which the mixing is homogeneous but treatment that changes the infective period distribution has been applied to a part of the population. Of course, if the treatment also includes quarantine that also changes the contact rate, the case $a_1 \neq a_2$, $A_1(s) \neq A_2(s)$ would be appropriate.

We suggest that in advance planning for a pandemic, the number of groups to be considered for different treatment rates should determine the number of groups to be used in the model. On the other hand, the number of groups to be considered should also depend on the amount and reliability of data, and these two criteria may be contradictory. A model with fewer groups and parameters chosen as weighted averages of the parameters for a model with more groups may give predictions that are quite similar to those of the more detailed models. We suggest also that use of the final size relations for a model with total population size assumed constant is a good time-saving procedure for making predictions if the disease death rate is small.

5.5 Some Warnings

An actual epidemic differs considerably from the idealized models such as (5.1) as well as the extensions considered later. Some notable differences are:

1. When it is realized that an epidemic has begun, individuals are likely to modify their behavior by avoiding crowds to reduce their contacts and by being more careful about hygiene to reduce the risk that a contact will produce infection.
2. If a vaccine is available for the disease which has broken out, public health measures will include vaccination of part of the population. Various vaccination strategies are possible, including vaccination of health care workers and other first line responders to the epidemic, vaccination of members of the population who have been in contact with diagnosed infectives, or vaccination of members of the population who live in close proximity to diagnosed infectives.
3. Diagnosed infectives may be hospitalized, both for treatment and to isolate them from the rest of the population. Isolation may be imperfect; in-hospital transmission of infection was a major problem in the SARS epidemic.
4. Contact tracing of diagnosed infectives may identify people at risk of becoming infective, who may be quarantined (instructed to remain at home and avoid

contacts) and monitored so that they may be isolated immediately if and when they become infective.
5. In some diseases, exposed members who have not yet developed symptoms may already be infective, and this would require inclusion in the model of new infections caused by contacts between susceptibles and asymptomatic infectives from the exposed class.
6. In the SARS epidemic of 2002–2003 in-hospital transmission of disease from patients to health care workers or visitors because of imperfect isolation accounted for many of the cases. This points to an essential heterogeneity in disease transmission which must be included whenever there is any risk of such transmission.

5.6 *Projects: Reproduction Numbers for Discrete Models

This project concerns the computation of the reproduction number for discrete models using the approach of the next generation matrix. A formula for the reproduction number $\mathscr{R}$ (either $\mathscr{R}_0$ or $\mathscr{R}_C$) is derived by adopting the method used in [2] based on the next generation matrix approach. That is, in the discrete-time case

$$\mathscr{R} = \varrho(F(I-T)^{-1}), \tag{5.61}$$

where ϱ represents the spectral radius, F is the matrix associated with new infections, and T is the matrix of transitions with $\varrho(T) < 1$ (see [2, 16, 31, 48]). Here F and T are calculated on the infected variables only evaluated at the disease-free equilibrium, and the Jacobian on these variables is $F+T$, which is assumed to be irreducible.

Consider the simple discrete SEIR model with geometric distributions for the latent and infectious period with parameters α and δ ($\alpha < 1, \delta < 1$), respectively. This is equivalent to assuming constant transition probabilities $1-\alpha$ and $1-\delta$ per unit time from E to I and from I to R, respectively. The model reads

$$\begin{aligned} S_{n+1} &= S_n e^{-\beta \frac{I_n}{N}}, \\ E_{n+1} &= S_n(1-e^{-\beta \frac{I_n}{N}}) + \alpha E_n \\ I_{n+1} &= (1-\alpha)E_n + \delta I_n, \quad n = 1, 2, \cdots. \end{aligned} \tag{5.62}$$

For system (5.62), the matrices associated with new infections and transitions are

$$F = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} \alpha & 0 \\ 1-\alpha & \delta \end{bmatrix},$$

respectively. Then $\varrho(T) = \max\{\alpha, \delta\} < 1$; thus,

$$\mathscr{R} = \varrho(F(I-T)^{-1}) = \frac{\beta}{1-\delta}. \tag{5.63}$$

Question 1 Extend the model (5.62) by incorporating isolation or hospitalization of infectious individuals. For the extended model, compute the reproduction number using the formula (5.61).

Question 2 Extend the model (5.62) by considering different transmission rates β_i for individuals in different infectious stages I_i ($i = 1, 2, \cdots$). Use the formula (5.61) to compute the reproduction number for the extended model.

Project 2 Consider next the case when the infective period follows an arbitrary discrete (bounded) distribution, which is denoted by Y. Let $q_i = \mathbb{P}(Y > i)$ and $\mathbb{P}(Y = i) = q_{i-1} - q_i$. It is easy to see that q_i is a decreasing function, i.e., $q_i \geq q_{i+1}$. In fact, $q_0 = 1$ and $q_m = 0$ for all $m \geq M$, where M is the maximum number of units of time that an individual takes to recover.

Because the geometric is the only memoryless discrete distribution, when other distributions are considered it is necessary to keep track of the past in order to know the values at the present. In fact, it is impossible to use the next generation matrix approach directly because the disease stages (S, E, and I) at time $n+1$ cannot be written in the form

$$[\ E_{n+1},\ I_{n+1},\ S_{n+1}\ ]^T = \mathscr{M}\left([\ E_n,\ I_n,\ S_n\ ]^T\right),$$

where $\mathscr{M} : \mathbb{R}^3 \to \mathbb{R}^3$. To overcome this difficulty we can consider multiple I stages, similar to the approach known as the "linear chain trick" used in continuous models to convert a gamma distribution to a sequence of exponential distributions. Thus, we introduce the subclasses $I^{(1)}, I^{(2)}, \cdots, I^{(M)}$ (see Fig. 5.6). The superscript i corresponds to the time since becoming infectious. Notice that these subclasses $I^{(i)}$ are different from those in the negative binomial model because

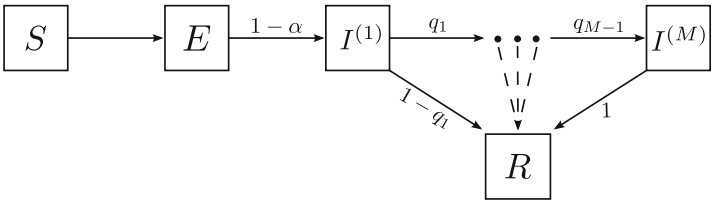

Fig. 5.6 A transition diagram for the case when the stage duration of the infective period has an arbitrary bounded distribution with upper bound M. The superscript i is the stage age in the infectious period and individuals in $I^{(i)}$ (for all i) can enter the recovered class R with a certain probability

here an individual can only stay in $I^{(i)}$ for one unit of time, and must go to either the $I^{(i+1)}$ class with probability q_i or the recovered class R with probability $1-q_i$.

From Fig. 5.6 the model equations can be written as

$$\begin{aligned} S_{n+1} &= S_n e^{-\sum_{i=1}^{M}\beta_i \frac{I_n^{(i)}}{N}}, \\ E_{n+1} &= S_n\left[1-e^{-\sum_{i=1}^{M}\beta_i \frac{I_n^{(i)}}{N}}\right]+\alpha E_n, \\ I_{n+1}^{(1)} &= (1-\alpha)E_n, \quad I_{n+1}^{(2)} = q_1 I_n^{(1)}, \quad I_{n+1}^{(j)} = \frac{q_{j-1}}{q_{j-2}} I_n^{(j-1)}, \quad 3 \le j \le M, \end{aligned} \tag{5.64}$$

where β_i denote the transmission rates at the infective stage i, $1 \le i \le M$. As q_i is the probability that an infective individual remains infective i time units after becoming infective, the transition probability from $I_n^{(2)}$ to $I_{n+1}^{(3)}$ is given by the probability that an infective individual is still infective two time units after becoming infectious given that the person remained infective one time unit ago, i.e., q_2/q_1. This explains the $I_{n+1}^{(3)}$ equation and similarly $I_{n+1}^{(j)}$ equations for $3 \le j \le M$.

Question 1 For the case when transmission rates β_i are stage-dependent, show that

$$\mathscr{R}_0 = \varrho(F(I-T)^{-1}) = \sum_{i=1}^{M} \beta_i q_{i-1}. \tag{5.65}$$

Question 2 Derive $\mathscr{R}_0$ from its biological definition. **Hint:** Using the fact that the distribution Y has an upper bound M and that for a given function f

$$\sum_{m=1}^{M} \mathbb{P}(Y=m) f(m) = \mathbb{E}(f(Y)).$$

The reproduction number from the biological definition (with $f(m) = \sum_{i=1}^{m} \beta_i$) is $\mathscr{R} = \sum_{i=0}^{M-1} \beta_{i+1} q_i$.

The formula (5.63) can also be applied to models with various heterogeneities. Consider a model that includes two sub-populations, female and male populations, with heterogeneous mixing (i.e., no sexual contacts between individuals of the same sex). Assume that the infective periods for female and male populations follow arbitrary discrete (bounded) distributions denoted by Y_f and Y_m, respectively. Here the subscripts f and m stand for female and male, respectively. Let $q_{f,i} = \mathbb{P}(Y_f > i)$ and $q_{m,i} = \mathbb{P}(Y_m > i)$ with $q_{f,0} = q_{m,0} = 1$, and the upper bounds for the two distributions (i.e., the maximum numbers of units of time that an individual takes to recover) be M_w for $w = f, m$.

The model equations are

$$
\begin{aligned}
S_{w,n+1} &= S_{w,n} e^{-\sum_{i=1}^{M_{\tilde{w}}} \beta_{\tilde{w},i} I^{(i)}_{\tilde{w},n}/N},\\
E_{w,n+1} &= S_{w,n}\left[1 - e^{-\sum_{i=1}^{M_{\tilde{w}}} \beta_{\tilde{w},i} I^{(i)}_{\tilde{w},n}/N}\right] + \alpha_w E_{w,n},\\
I^{(1)}_{w,n+1} &= (1-\alpha_w) E_{w,n}, \quad I^{(2)}_{w,n+1} = q_{w,1} I^{(1)}_{w,n},\\
I^{(j)}_{w,n+1} &= \frac{q_{w,j-1}}{q_{w,j-2}} I^{(j-1)}_{w,n}, \qquad 3 \le j \le M_w, \quad \text{for } w = f, m.
\end{aligned}
$$

Here $\tilde{w}$ represents the opposite sex of w, i.e., $\tilde{f} = m$, $\tilde{m} = f$. The constant $\beta_{\tilde{f},i}$ ($\beta_{\tilde{m},i}$) represents the infection rate to a female (male) transmitted by infectious male (female) individuals with stage age i.

Question 3 Show that the reproduction number is given by

$$
\mathscr{R} = \varrho(F(I-T)^{-1}) = \sqrt{\left(\sum_{i=1}^{M_f} \beta_{f,i}\, q_{f,i-1}\right)\left(\sum_{i=1}^{M_m} \beta_{m,i} q_{m,i-1}\right)}. \tag{5.66}
$$

The square root in (5.66) is a consequence of the fact that the secondary infections need to be computed from one female (male) to other females (males) through the male (female) population.

Hint: Consider the order of variables:

$$
(E_{f,n}, I^{(1)}_{f,n}, I^{(2)}_{f,n}, \cdots, I^{(M_f)}_{f,n}, E_{m,n}, I^{(1)}_{m,n}, I^{(2)}_{m,n}, \cdots, I^{(M_m)}_{m,n}).
$$

First show that

$$
F(I-T)^{-1} = \begin{bmatrix} 0 & F_m(I-T_m)^{-1} \\ F_f(I-T_f)^{-1} & 0 \end{bmatrix},
$$

where

$$
F_w(I-T_w)^{-1} = \begin{bmatrix} \sum_{i=1}^{M_w} \beta_{w,i} q_{w,i-1} & \sum_{i=1}^{M_w} \beta_{w,i} q_{w,i-1} & \cdots & \beta_{w,M_w} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad w = f, m.
$$

References: [2, 16, 31, 48].

5.7 *Project: Modeling the Synergy Between HIV and HSV-2

Consider the following model for HSV-2, which includes a male population (specified by subscript m) and a female population with two sub-groups representing low-risk and high-risk groups (specified by subscripts f_1 and f_2, respectively):

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu_i N_i - \lambda_i(t) S_i - \mu_i S_i, \\
\frac{dA_i}{dt} &= \lambda_i(t) S_i + \gamma_i(\theta_i) L_i - (\omega_i + \theta_i + \mu_i) A_i, \\
\frac{dL_i}{dt} &= (\omega_i + \theta_i) A_i - \big(\gamma_i(\theta_i) + \mu_i\big) L_i, \qquad i = m, f_1, f_2,
\end{aligned} \tag{5.67}$$

where $\lambda_i(t)$ $(i = m, f_1, f_2)$ are the force of infection functions given by

$$\begin{aligned}
\lambda_m(t) &= \sum_{i=1}^{2} b_m c_i \beta_{f_i m} \frac{A_{f_i}}{N_{f_i}}, \\
\lambda_{f_j}(t) &= b_{f_j} \beta_{m f_j} \frac{A_m}{N_m}, \quad j = 1, 2,
\end{aligned} \tag{5.68}$$

and $N_i = S_i + A_i + L_i$, $i = m, f_1, f_2$. Each group i $(i = m, f_1, f_2)$ is divided into three subgroups: susceptible (S_i), infected with acute HSV-2 only (A_i), infected with latent HSV-2 only (L_i). The population within each group is assumed to be homogeneous in the sense that individuals have the same infective period, duration of immunity, contact rate, and so on. A transition diagram between these epidemiological classes within group i is depicted in Fig. 5.7.

For each sub-population i $(i = f_1, f_2, m)$ there is a per-capita recruitment rate μ_i into the susceptible group. For all classes there is a constant per-capita rate μ_i of exiting the sexually active population. Thus, the total population N_i in group i remains constant for all time. Susceptible people in group i acquire infection with HSV-2 at the rate $\lambda_i(t)$. Upon being infected with HSV-2, people in group i enter the class A_i. These individuals become latent L_i at the constant rate ω_i (an average duration in A_i is $1/\omega_i$). Following an appropriate stimulus in individuals with latent HSV-2, reactivation may occur at the rate γ_i. Finally, the antiviral treatment rate for the A_i individuals is denoted by θ_i. Because antiviral medications will also suppress

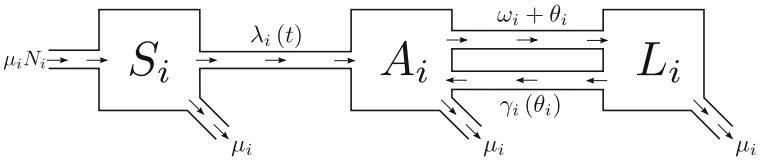

Fig. 5.7 A transition diagram for HSV-2 for subgroup i $(i = m, f_1, f_2)$

reactivation of latent HSV-2, we assume that the reactivation rate of people with latent HSV-2 γ_i is a decreasing function of θ_i, denoted by $\gamma_i(\theta_i)$.

For the forces of infection $\lambda_i(t)$ $(i = m, f_1, f_2)$, b_i is the rate at which individuals in group i acquire new sexual partners (also referred to as contact rates), and c_j denotes the probability that a male chooses a female partner in group j $(j = f_1, f_2)$. Then $c_1 + c_2 = 1$. For ease of notation, let

$$c_1 = c, \quad c_2 = 1 - c.$$

Overall, the number of female partners in groups j $(j = 1, 2)$ that males acquire should be equal to the number of male partners that females in groups j acquire. These observations lead to the following balance conditions:

$$b_m c N_m = b_{f_1} N_{f_1}, \qquad b_m (1 - c) N_m = b_{f_2} N_{f_2}. \tag{5.69}$$

To ensure that constraints in (5.69) are satisfied, we assume in numerical simulations that b_m and c are fixed constants with b_{f_1} and b_{f_2} being varied according to N_m, N_{f_1}, and N_{f_2}. The parameters $\beta_{im}(\beta_{mi})$, $i = f_1, f_2$ are the HSV-2 transmission probabilities per partner between females infected with acute HSV-2 in group i and susceptible males (between males infected with acute HSV-2 and susceptible females in group i).

Question 1 Let R_{mf_jm} denote the average number of secondary male infections generated by one male individual through females in group f_j $(j = 1, 2)$. Show that

$$\mathscr{R}_{mf_jm} = \sqrt{\frac{b_{f_j}\beta_{mf_j}}{\omega_m + \theta_m + \mu_m} \cdot P_m \cdot \frac{b_m c_j \beta_{f_jm}}{\omega_{f_j} + \theta_{f_j} + \mu_{f_j}} \cdot P_{f_j}}, \quad j = 1, 2$$

with P_i $(i = m, f_1, f_2)$ representing the probability that an individual of group i is in the acute stage (A), which is given by

$$P_i = \frac{(\omega_i + \theta_i + \mu_i)(\gamma_i(\theta_i) + \mu_i)}{[\gamma_i^L(\theta_i) + \omega_i + \theta_i + \mu_i]\mu_i}, \quad i = m, f_1, f_2. \tag{5.70}$$

Question 2 Let $\mathscr{R}$ denote the overall reproduction number for the entire population.

(a) Show that

$$\mathscr{R} = \sqrt{\left(\mathscr{R}_{mf_1m}\right)^2 + \left(\mathscr{R}_{mf_2m}\right)^2}, \tag{5.71}$$

where $\mathscr{R}_{mf_jm}$ $(j = 1, 2)$ are given in Question 1.

(b) Provide a biological interpretation of the expression on the right-hand side of Eq. (5.71).

Question 3 Let E_0 denote the disease-free equilibrium of the system (5.67), and let E^* denote an endemic equilibrium.

(a) Show that E_0 is locally asymptotically stable when $\mathscr{R} < 1$ and unstable when $\mathscr{R} > 1$.
(b) Choose the function for $\gamma(\theta)$ to be in the following form: $\gamma_i(\theta_i) = \gamma_i(0)\alpha_i/(\alpha_i + \theta_i)$. Show via numerical simulations that E^* exists and is locally asymptotically stable when $\mathscr{R} > 1$. One case to consider is when $c > 0.5$, e.g., $c = 0.9$ (90% of male contacts are with the low-risk female group). Because of the constraint (5.69), b_i and N_i are not independent. Choose $b_m = 0.1$, $b_{f_1} = 0.0901, b_{f_2} = 9.01$ (so that $b_{f_2}/b_{f_1} = 100$), $N_m = N_{f_1} + N_{f_2} = 10^7$ (e.g., $N_{f_1} = 9.9889 \times 10^6$, $N_{f_2} = 1.1099 \times 10^4$). Consider the case when treatment is absent, i.e., $\theta = 0$. Other parameter values are $\omega = 2.5$, $\gamma_m(0) = 0.436$, $\gamma_{f_1} = \gamma_{f_2} = 0.339$, $\alpha_i = 2$. The time unit is month.
(c) Explore numerically the effect of treatment θ. Consider various scenarios such as treatment in only one subgroup (male, low-risk female or high-risk female group). Summarize the observed outcomes in terms of effect of treatment on the prevalence of HSV-2.

5.8 Project: Effect of Heterogeneities on Reproduction Numbers

Consider the metapopulation model (5.37), which includes vaccination coverage $\phi = (\phi_1, \phi_2, \cdots, \phi_n)$ in the n sub-populations. As pointed out in Sect. 5.3.1 that several types of heterogeneities including the activity (a_i), sub-population size (N_i), and preference for mixing within the sub-population (π_i) may affect the optimal vaccination strategy. In this project, we examine in more details how these heterogeneities may affect $\mathscr{R}_v$ for the case of $n = 2$, and how to choose (ϕ_1, ϕ_2) to reduce $\mathscr{R}_v$ below a certain level. For example, Fig. 5.8 illustrates the different parameter regions in the (ϕ_1, ϕ_2) plane in which $\mathscr{R}_v < 1$ in the cases of (a) proportionate mixing ($\pi_1 = \pi_2 = 0$) and (b) preferential mixing ($\pi_i > 0$).

For Questions 1–3 below, let $\sigma = 0.05$, $\gamma = 1/7$.

Question 1 For the cases in (a) and (b) below, determine the values of the basic reproduction number $\mathscr{R}_0$ for each case. Describe how preferential mixing and heterogeneous activity may influence the effect of heterogeneity in activity on $\mathscr{R}_0$. Let $N_1 = N_2 = 500$. Determine the values of $\mathscr{R}_0$ for each case.

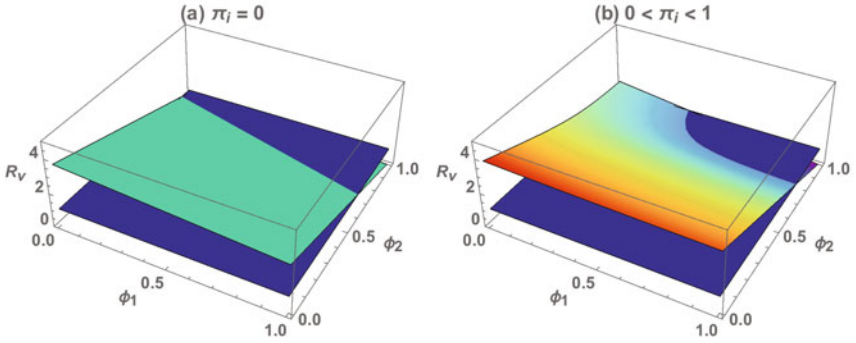

Fig. 5.8 Plots of $\mathscr{R}_v$ as a function of ϕ_1 and ϕ_2 for (**a**) proportionate mixing ($\pi_i = 0$, no preference) and (**b**) preferential mixing ($\pi_i > 0$). Values at or below their intersection with the dark plane, $\mathscr{R}_v = 1$, are combinations of $\phi_i (i = 1, 2)$ at which population-immunity attains or exceeds this threshold

(a) Homogeneous activity: $a_1 = a_2 = 10$ ($a_1 + a_2 = 20$).

(i) No preference: $\pi_1 = \pi_2 = 0$.
(ii) Homogeneous preference: $\pi_1 = \pi_2 = 0.5$.
(iii) Heterogeneous preference: $\pi_1 = 0.25, \pi_2 = 0.75$ and $\pi_1 = 0.75, \pi_2 = 0.25$.

(b) Heterogeneous activity: $a_1 = 8$ and $a_2 = 12$ ($a_1 + a_2 = 20$).
Repeat (i)–(iii) in (a).

Question 2 Same as in *Question 1* but consider heterogeneities in both activity a_i and population size N_i. Given that $N_1 + N_2 = N = 1000$ and $a_1 = 5$, $a_2 = 10$. Repeat (i)–(iii) in *Question 1*(a) for the following two cases:

(a) Homogeneous population size: $N_1 = N_2 = 0.5N$.
(b) Heterogeneous population size: $N_1 = 0.1N$ and $N_2 = 0.9N$.

Question 3 Consider the effective reproduction number $\mathscr{R}_v(\phi_1, \phi_2)$ as a function of vaccination coverage (ϕ_1, ϕ_2).

(a) Find the optimal solution $\Phi^* = (\phi_1^*, \phi_2^*)$ for the following set of parameter values: $\pi_1 = \pi_2 = 0.3$, $a_1 = 15$, $a_2 = 12$, $N_1 = 1100$, $N_2 = 900$. The given number of vaccine doses is $\phi_1 N_1 + \phi_2 N_2 = 990$.
(b) What is the minimum value $\mathscr{R}_{v\{min\}} = \mathscr{R}_v(\Phi^*)$?

References

1. Adler, F.R. (1992) The effects of averaging on the basic reproduction ratio, Math Biosci. **111**(1): 89–98.
2. Allen,L.J. and P. van den Driessche (2008) The basic reproduction number in some discrete-time epidemic models, J. Diff. Equ. App. **14**: 1127–1147.
3. Andersson, H. & Britton, T. (1998) Heterogeneity in epidemic models and its effect on the spread of infection, J. Appl. Prob. **35**: 651–661.
4. Arino, J., F. Brauer, P. van den Driessche, J. Watmough & J. Wu (2006) Simple models for containment of a pandemic, J. Roy. Soc. Interface, **3**: 453–457.
5. Arino, J., F. Brauer, P. van den Driessche, J. Watmough & J. Wu (2007) A final size relation for epidemic models, Math. Biosc. & Eng. **4**: 159–176.
6. Arino, J., F. Brauer, P. van den Driessche, J. Watmough & J. Wu (2008) A model for influenza with vaccination and antiviral treatment, Theor. Pop. Biol. **2**53: 118–130.
7. Berman, A. & R.J. Plemmons (1994) Nonnegative Matrices in the Mathematical Sciences, SIAM, Vol. **9**, 1994.
8. Blythe, S.P., S. Busenberg & C. Castillo-Chavez (1995) Affinity and paired-event probability, Math. Biosc. **1**28: 265–84 .
9. Blythe, S.P. , C. Castillo-Chavez, J. Palmer & M. Cheng (1991) Towards a unified theory of mixing and pair formation, Math. Biosc. **1**07: 379–405.
10. Brauer, F. (2005) The Kermack–McKendrick epidemic model revisited, Math. Biosc. **1**98: 119–131.
11. Brauer, F. (2008) Epidemic models with treatment and heterogeneous mixing, Bull. Math. Biol. **70**: 1869–1885.
12. Brauer, F. & J. Watmough (2009) Age of infection epidemic models with heterogeneous mixing, J. Biol. Dynamics **3**: 324–330.
13. Busenberg, S. & C. Castillo-Chavez (1989) Interaction, pair formation and force of infection terms in sexually transmitted diseases, In Mathematical and Statistical Approaches to AIDS Epidemiology, Lect. Notes Biomath. **8**3, C. Castillo-Chavez (ed.), Springer-Verlag, Berlin-Heidelberg-New York, 289–300.
14. Busenberg, S. & C. Castillo-Chavez (1991) A general solution of the problem of mixing of sub-populations and its application to risk- and age-structured epidemic models for the spread of AIDS, IMA J. Math. Appl. Med. Biol., **8**: 1–29.
15. Chow, L., M. Fan, and Z. Feng (2011) Dynamics of a multi-group epidemiological model with group-targeted vaccination strategies, J. Theor. Biol. **29**: 56–64.
16. De Jong, M., O. Diekmann and J.A.P. Heesterbeek (1994) The computation of R0 for discrete-time epidemic models with dynamic heterogeneity, Math. Biosci. **119**(1): 97–114.
17. Del Valle, S. Y., J.M. Hyman, H.W. Hethcote & S.G. Eubank (2007) Mixing patterns between age groups in social networks, *Social Networks*, **29**(4): 539–554.
18. Diekmann, O. & J.A.P. Heesterbeek (2000) Mathematical Epidemiology of Infectious Diseases. Wiley, Chichester (2000)
19. Diekmann, O., J.A.P. Heesterbeek & J.A.J. Metz (1990) On the definition and the computation of the basic reproductive ratio $\mathscr{R}_0$ in models for infectious diseases in heterogeneous populations, J. Math. Biol. **2**8: 365–382.
20. Diekmann, O., J.A.P. Heesterbeek, & T. Britton (2012) Mathematical tools for understanding infectious disease dynamics, 2012, Princeton University Press.
21. Feng, Z., A.N. Hill, P.J. Smith, J.W. Glasser (2015) An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing, J. Theor. Biol. **386**: 177–187.
22. Feng, Z., A.N. Hill, A.T. Curns, J.W. Glasser (2007) Evaluating targeted interventions via meta-population models with multi-level mixing, Math. Biosci. **287**: 93–104.
23. Ferguson, N.M., D.A.T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, & D.S. Burke (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia, Nature, **4**37: 209–214.

24. Gani, R., H. Hughes, T. Griffin, J. Medlock, & S. Leach (2005) Potential impact of antiviral use on hospitalizations during influenza pandemic, Emerg. Infect. Dis. **1**1: 1355–1362.
25. Glasser, J., Z. Feng, A. Moylan, S. Del Valle & C. Castillo-Chavez (2012) Mixing in age-structured population models of infectious diseases, Math. Biosci., **235**(1): 1–7.
26. Glasser, J.W., Z. Feng, S.B. Omer, P.J. Smith, L.E. Rodewald (2016) The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study, **16**(5): 599–605.
27. Hethcote, H.W. & J.A. Yorke (1984) *Gonorrhea Transmission Dynamics and Control*, Lect. Notes in Biomath. **56**, Springer-Verlag, Berlin-Heidelberg-New York (1984).
28. Hethcote, H.W. & J.W. van Ark (1987) Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation and immunization programs. Math. Biosci., **84**(1): 85–118.
29. Hirsch, M.W. & S. Smale, Differential Equations (1974) Dynamical Systems, and Linear Algebra, Academic Press, Orlando, FL (1974).
30. Jacquez, J.A., C.P. Simon, J. Koopman, L. Sattenspiel, & T. Perry (1988) Modeling and analyzing HIV transmission: the effect of contact patterns, Math. Biosci., **92**: 119–199.
31. Lewis, M.A., J. Renclawowicz, P. van Den Driessche, and M. Wonham (2006) A comparison of continuous and discrete-time West Nile Virus models, Bull. Math. Biol. **68**(3): 491–509.
32. Longini, I.M., M.E. Halloran, A. Nizam, & Y. Yang (2004) Containing pandemic influenza with antiviral agents, Am. J. Epidem. **1**59: 623–633.
33. Longini, I.M., A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D.A.T. Cummings, & M.E. Halloran (2005) Containing pandemic influenza at the source, Science **3**09, 1083–1087.
34. Longini, I. M. & M. E. Halloran (2005) Strategy for distribution of influenza vaccine to high - risk groups and children, Am. J. Epidem. **1**61: 303–306.
35. Ma, J. & D.J.D. Earn (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease, Bull. Math. Biol. **6**8: 679–702.
36. May, R.M. & R.M. Anderson (1984) Spatial heterogeneity and the design of immunization programs. Math Biosci. **72**(1): 83–111.
37. May, R.M. & R.M. Anderson (1984) Spatial, temporal and genetic heterogeneity in host populations and the design of immunization programmes. IMA J Math Appl Math Biol. **1**(3): 233–66.
38. Meyers, L.A. (2007) Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, Bull. Am. Math. Soc. **4**4, 63–86.
39. Meyers, L.A. , M.E.J. Newman & B. Pourbohloul (2006) Predicting epidemics on directed contact networks, J. Theor. Biol. **2**40, 400–418.
40. Meyers, L.A., B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, & R.C. Brunham (2005) Network theory and SARS: predicting outbreak diversity. J. Theor. Biol., **2**32: 71–81.
41. Mossong, J., N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, ... & W. J. Edmunds (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS Medicine, **5**(3): e74.
42. Nold, A. (1980) Heterogeneity in disease transmission modeling, Math. Biosci, **5**2: 227–240.
43. Poghotanyan, G., Z. Feng, J.W. Glasser, A.N. Hill (2018) Constrained minimization problems for the reproduction number in meta-population models, J. Math. Biol. https:/doi.org/10.1007/s00285-018-1216-z.
44. Pourbohloul, B. & J. Miller (2008) Network theory and the spread of communicable diseases, Center for Disease Modeling Preprint 2008-03, www.cdm.yorku.ca/cdmprint03.pdf
45. van den Driessche, P. & J. Watmough (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosc. **1**80: 29–48.
46. van den Driessche, P. & J. Watmough (2002) Further notes on the basic reproduction number, in Mathematical Epidemiology, F. Brauer, P. van den Driessche, & J. Wu (eds.) Springer Lecture Notes, Vol. **1945**.

47. Wallinga, J., P. Teunis, & M. Kretzschmar (2006) Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents, American Journal of Epi, **164**(10): 936–944.
48. Wesley, C. L., L.J. Allen, C.B. Jonsson, Y.-K. Chu, R.D. Owen (2009) A discrete-time rodent-hantavirus model structured by infection and developmental stages, Adv. Stu. Pure Math. **53**: 387–398.
49. Zagheni, E., F.C. Billari, P. Manfredi, A. Melegaro, J. Mossong & W.J. Edmunds (2008) Using time-use data to parameterize models for the spread of close-contact infectious diseases, Am. J. Epi., **168**(9), 1082–1090.