J. Math. Biol. (2013) 67:963–987
DOI 10.1007/s00285-012-0578-x

Mathematical Biology

# The effect of population heterogeneities upon spread of infection

**Damian Clancy · Christopher J. Pearce**

Received: 16 January 2012 / Revised: 25 July 2012 / Published online: 2 September 2012


**Abstract** It has often been observed that population heterogeneities can lead to outbreaks of infection being less frequent and less severe than homogeneous population models would suggest. We address this issue by comparing a model incorporating various forms of heterogeneity with a homogenised model matched according to the value of the basic reproduction number $R_0$. We mainly focus upon heterogeneity in individuals' infectivity and susceptibility, though with some allowance also for heterogeneous patterns of mixing. The measures of infectious spread we consider are (i) the probability of a major outbreak; (ii) the mean outbreak size; (iii) the mean endemic prevalence level; and (iv) the persistence time. For each measure, we establish conditions under which heterogeneity leads to a reduction in infectious spread. We also demonstrate that if such conditions are not satisfied, the reverse may occur. As well as comparison with a homogeneous population, we investigate comparisons between two heterogeneous populations of differing degrees of heterogeneity. All of our results are derived under the assumption that the susceptible population is sufficiently large.

**Keywords** Basic reproduction number · SIR epidemic · SIS epidemic · Outbreak size · Endemic prevalence · Fade-out of infection

**Mathematics Subject Classification (2000)** 92D30 · 60J85 · 60J28

## 1 Introduction

In the simplest models for infectious spread, the population is assumed to consist of identical individuals who mix homogeneously with one another. Clearly these

D. Clancy (✉)· C. J. Pearce
Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZL, UK
e-mail: d.clancy@liv.ac.uk

assumptions are over-simplistic, and more realistic models incorporate a variety of heterogeneities—see, for example, Hethcote (1996), Keeling and Rohani (2007) chapter 3, and references therein. Most simply, individuals in different geographical regions will interact more weakly than individuals living in close proximity; a population consisting of spatially separated groups is sometimes referred to as a 'metapopulation' or 'patchy environment' model (Hagenaars et al. 2004). For many diseases (e.g. childhood infections) the population should be stratified according to age group; for sexually transmitted diseases, stratification according to gender and sexual behaviour is also important, and certain 'core groups' may be chiefly responsible for maintaining infection in the population (Hethcote and Yorke 1984); and for certain infections (e.g. SARS), it has been hypothesised that there exists a sub-group of 'super-spreaders' within the population. Such heterogeneities have important implications for control strategies (see, for instance, Lloyd-Smith et al. 2005). The question then arises as to how, and to what extent, such heterogeneities affect the spread of infection. Comparisons between heterogenous population models and appropriately matched homogeneous populations are thus a topic of long-standing interest in the literature, e.g. Ball (1985), Lefèvre and Malice (1988), Becker and Marschner (1990), Marschner (1992), Adler (1992), Andersson and Britton (1998).

Given a specified heterogeneous population model, it is not always obvious how to construct the corresponding homogeneous population model for comparison, and in general there may be several candidates. For a population in which individuals have different degrees of susceptibility, Ball (1985) took the arithmetic mean of the relevant contact rate parameters, whereas Andersson and Britton (1998) worked instead with the harmonic mean. In analysing a model for assortative/dissortative mixing (individuals have a preference for either within-group or between-group infectious contacts), Marschner (1992) investigated three different averaging methods. Often, the interest is particularly in the effect of averaging upon the basic reproduction number $R_0$, defined to be the average number of new infections caused by a typical infected individual in an otherwise susceptible population. In contrast, recent authors have chosen to compare the heterogeneous population of interest with a homogeneous population having the same $R_0$ value, in terms of measures such as the probability of a major outbreak or the total outbreak size. Thus in modelling Salmonella transmission within a dairy herd, Xiao et al. (2006) presented numerical results indicating that herd heterogeneity reduced the probability of a major outbreak, compared to a homogeneous population with the same $R_0$ value. Yates et al. (2006), working with a more generic heterogeneous population model, found similar results for the major outbreak probability, again based on numerical work. Andreasen (2011) gave sufficient conditions under which the size of a deterministic susceptible-infective-removed (SIR) epidemic model in a heterogeneous population is bounded above by that of the homogeneous population with matched $R_0$ value.

In this paper, we compare populations incorporating various heterogeneities with a homogeneous population matched according to $R_0$ value, in terms of (i) the probability of a major outbreak; (ii) the expected size of such an outbreak, should one occur; (iii) the endemic mean prevalence level, should the infection become endemic in the population; and (iv) the persistence time of infection in the population, should the infection become endemic. Our general model (specified more precisely in Sect. 2)

describes a population stratified into groups, and allows for heterogeneity in individuals' susceptibilities and infectivities, and in group-to-group mixing preferences. For simplicity, we assume no birth or immigration into the population, and no death or emigration. We assume that there is no latent period, and that the infection either confers lifelong immunity or no immunity (that is, no temporary or partial immunity). We use approximation methods valid in the large population limit throughout. In Sect. 7 we discuss just how restrictive our assumptions are in practice, and the extent to which our results remain robust to their relaxation.

Our main results are as follows. Heterogeneities in individuals' susceptibilities and infectivities do not increase the probability of a major outbreak (under appropriate initial conditions). Provided susceptibility and infectivity are not negatively correlated, then heterogeneities in susceptibility and infectivity do not increase the mean size of a major outbreak or the endemic mean prevalence level. Heterogeneities in either susceptibility alone or infectivity alone do not increase the mean time for infection to die out, starting from the endemic level, provided all groups are of equal size. For each of the measures (i)–(iv) we also establish sufficient conditions under which outbreak probability or severity is greater in the 'more heterogeneous' (in a sense made precise in Sect. 2) of two heterogeneous populations having a common $R_0$ value, and exhibit symmetry conditions under which population heterogeneities have no effect upon the measure in question. Comparisons between populations of different degrees of heterogeneity do not generally seem to have been considered by previous authors.

All vectors are row vectors, and we refer to a matrix as non-negative when all its elements are non-negative. Numerical work was carried out using Matlab on a desktop PC.

## 2 Model specification and majorization theory

Consider a closed population of $N$ individuals divided into $k$ groups, with group $i (i = 1, 2, \ldots, k)$ consisting of $N_i$ individuals of whom $a_i$ are initially infected. Denote by $f_i = N_i/N$ the proportion of the population belonging to group $i$, so that $\sum_i f_i = 1$. When a group $i$ individual becomes infected, it remains so for a time distributed as a non-negative random variable $T$, assumed (for simplicity) to have the same distribution for each group. During this infectious period, the group $i$ infective makes contacts with each individual in each group $j = 1, 2, \ldots, k$ at the points of a Poisson process of rate $\beta_{ij}/N$. These Poisson processes and infectious periods are all mutually independent. If a contacted individual is susceptible, then it becomes infected (and infectious); if the contacted individual is already infected then the contact has no effect. We follow Yates et al. (2006) in writing the infection rate parameters in the form $\beta_{ij} = \beta\lambda_i\pi_{ij}\mu_j$, where $\beta$ is some overall measure of infectiousness, $\lambda_i$ represents the infectivity of group $i$ individuals, $\mu_j$ represents the susceptibility of group $j$ individuals, and $\pi_{ij}$ is a mixing parameter representing the relative preference of group $i$ infectives for group $j$ susceptibles. Following Becker and Marschner (1990) we scale the $\lambda_i$, $\mu_j$ values so that $\sum_i \lambda_i f_i = \sum_j \mu_j f_j = 1$, and we scale time so that $E[T] = 1$. We will assume throughout that $\beta > 0$, that $f_i, \lambda_i, \mu_i > 0$ for all $i$ and that the matrix with entries $\pi_{ij}$ is irreducible (see Seneta 1986, p. 18–22).

We also impose constraints $\sum_i \pi_{ij} = 1$ for each $j$ and $\sum_j \pi_{ij} = 1$ for each $i$ (this is equivalent to requiring the matrix with entries $\beta_{ij}$ to be such that there exists a doubly stochastic matrix with the same pattern of non-zero entries, see Borwein et al. (1994)). If, for example, we take $\pi_{ii} = \rho$ for all $i$ and $\pi_{ij} = (1-\rho)/(k-1)$ for $i \neq j$, then the parameter $\rho$ ($0 \leq \rho < 1$) represents the preference for within-group mixing as opposed to between-group mixing. The cases $\rho > 1/k$ and $\rho < 1/k$ are known as assortative and dissortative mixing, respectively (Yates et al. 2006; Hagenaars et al. 2004).

Particular special cases of our general model are as follows.

1. *The separable case*: $\pi_{ij} = 1/k$ for all $i, j$, so that $\beta_{ij} = (\beta/k)\lambda_i \mu_j$.
2. *Heterogeneous susceptibility*: $\pi_{ij} = 1/k$ for all $i, j$ and $\lambda_i = 1$ for all $i$, so that $\beta_{ij} = (\beta/k)\mu_j$.
3. *Heterogeneous infectivity*: $\pi_{ij} = 1/k$ for all $i, j$ and $\mu_j = 1$ for all $j$, so that $\beta_{ij} = (\beta/k)\lambda_i$.
4. *Heterogeneous mixing*: $\lambda_i = \mu_i = 1$ for all $i$, so that $\beta_{ij} = \beta\pi_{ij}$.

For comparing two heterogeneous populations, we shall use superscripts to specify the population, so that population 1 has infection rate parameters $\beta_{ij}^{(1)} = \beta^{(1)}\lambda_i^{(1)}\pi_{ij}^{(1)}\mu_j^{(1)}$ and so on. To make precise what we mean by "more heterogeneous" we require the following definitions. For a vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)$, denote by $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[k]}$ the elements of $\boldsymbol{x}$ in decreasing order. Then we say $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$, denoted $\boldsymbol{x} \prec \boldsymbol{y}$, if $\sum_{i=1}^{j} x_{[i]} \leq \sum_{i=1}^{j} y_{[i]}$ for $j = 1, 2, \ldots, k-1$ and $\sum_{i=1}^{k} x_{[i]} = \sum_{i=1}^{k} y_{[i]}$ (Marshall et al. 2010, definition 1.A.1). For a vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)$ with non-negative components such that $\sum_{i=1}^{k} p_i = 1$, then $\boldsymbol{x}$ is said to be $\boldsymbol{p}$-majorized by $\boldsymbol{y}$, denoted $\boldsymbol{x} \prec_{\boldsymbol{p}} \boldsymbol{y}$, if there exists a permutation $\sigma$ such that $x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(k)}$ and $y_{\sigma(1)} \geq y_{\sigma(2)} \geq \cdots \geq y_{\sigma(k)}$ with $\sum_{i=1}^{j} p_{\sigma(i)} x_{\sigma(i)} \leq \sum_{i=1}^{j} p_{\sigma(i)} y_{\sigma(i)}$ for $j = 1, 2, \ldots, k-1$ and $\sum_{i=1}^{k} p_i x_i = \sum_{i=1}^{k} p_i y_i$ (Marshall et al. 2010, definition 14.A.2). The intuitive interpretation is that if $\boldsymbol{x} \prec \boldsymbol{y}$ or if $\boldsymbol{x} \prec_{\boldsymbol{p}} \boldsymbol{y}$ for some $\boldsymbol{p}$ then $\boldsymbol{y}$ is more heterogeneous than $\boldsymbol{x}$, see Marshall et al. (2010). In particular, note that $(\sum_i y_i)\mathbf{1} \prec \boldsymbol{y}$ for any $\boldsymbol{y}$ and $(\sum_i p_i y_i)\mathbf{1} \prec_{\boldsymbol{p}} \boldsymbol{y}$ for any $\boldsymbol{y}, \boldsymbol{p}$.

We now collect together some results required in the sequel.

**Lemma 1** (i) *$\boldsymbol{x} \prec \boldsymbol{y}$ if and only if there exists a doubly stochastic matrix $A$ such that $\boldsymbol{x} = \boldsymbol{y}A$ (Marshall et al. 2010, Theorem 2.B.2).*

(ii) *If $\boldsymbol{x} \prec_{\boldsymbol{p}} \boldsymbol{y}$ then there exists a non-negative $k \times k$ matrix $A$ with $\sum_{i=1}^{k} a_{ij} = 1$ for all $j$ such that $\boldsymbol{p}A^T = \boldsymbol{p}$ and $\boldsymbol{x} = \boldsymbol{y}A$ (Marshall et al. 2010, Proposition 14.A.3).*

(iii) *Suppose $\boldsymbol{p}^{(1)}$, $\boldsymbol{p}^{(2)}$ are vectors with non-negative components such that $\sum_{i=1}^{k} p_i^{(1)} = \sum_{i=1}^{k} p_i^{(2)} = 1$. Then $\sum_{i=1}^{k} p_i^{(1)} \xi(x_i) \leq \sum_{i=1}^{k} p_i^{(2)} \xi(y_i)$ for all convex functions $\xi$ if and only if there exists a non-negative $k \times k$ matrix $A$ with $\sum_{i=1}^{k} a_{ij} = 1$ for $j = 1, 2, \ldots, k$ such that $\boldsymbol{p}^{(1)}A^T = \boldsymbol{p}^{(2)}$ and $\boldsymbol{x} = \boldsymbol{y}A$ (Marshall et al. 2010, Proposition 14.A.1, originally due to Blackwell (1951, 1953)).*

Our results are all derived under the assumption that a small number of infectives are introduced into a large susceptible population. That is, we rely upon approximations

valid in the limit as $N \to \infty$ (which implies that $N_i \to \infty$ for every $i$, since $f_i > 0$), while each $a_i$ remains finite.

## 3 Probability of a major outbreak

In the early stages of an outbreak, provided that every group size $N_i$ is sufficiently large, the probability that a contact is with an already-infected individual is low, so we may effectively assume that every contact results in a new infection. That is, the infection process may be approximated by a multi-type branching process in which each group $i$ individual lives for a time distributed as $T$ and during this lifetime gives birth to group $j$ offspring ($j = 1, 2, \ldots, k$) according to a Poisson process of rate $\beta\lambda_i\pi_{ij}\mu_j f_j$. For full justification of this approximation see Metz (1978), Ball (1983). Denoting by $G_{ij}$ the number of type $j$ offspring of a typical type $i$ individual, then the total number of progeny of the process follows the same distribution as for a multi-type Galton-Watson process with offspring distributions $G_{ij}$. Denoting by $\psi(\theta) = E[\exp(\theta T)]$ the moment generating function of $T$, then the offspring distributions are determined by the generating functions

$$\phi_i(s_1, s_2, \ldots, s_k) = E\left[\prod_{j=1}^{k} s_j^{G_{ij}}\right] = \psi\left(-\beta\sum_{j=1}^{k}\lambda_i\pi_{ij}\mu_j f_j\left(1 - s_j\right)\right) \quad (1)$$

for $0 \le s_i \le 1, i = 1, 2, \ldots, k$.

The mean offspring matrix $M$ with entries $m_{ij} = \beta\lambda_i\pi_{ij}\mu_j f_j$, often called the next generation matrix, is irreducible (by assumption), and the basic reproduction number $R_0$ is equal to the Perron–Frobenius eigenvalue of $M$. If $R_0 \le 1$ then (with probability 1) the branching process produces only a finite total number of progeny, corresponding to a minor outbreak of infection (see e.g. Mode 1971). If $R_0 > 1$ then the branching process may produce an infinite number of progeny, corresponding to a major outbreak of infection, with probability $1 - \prod_{i=1}^{k} q_i^{a_i}$ where $q_i$ is the extinction probability of the branching process started with a single individual of type $i$, and $\boldsymbol{q} = (q_1, q_2 \ldots, q_k)$ is the unique solution with $0 \le q_i < 1$ of

$$q_i = \phi_i(q_1, q_2, \ldots, q_k) \text{ for } i = 1, 2, \ldots, k. \quad (2)$$

We assume from now on that $R_0 > 1$.

We shall consider only outbreaks initiated by a single infective, the initial infective belonging to group $i$ with probability $\nu_i$ for some probability distribution $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_k)$, so that the probability of a major outbreak is given by $1 - q$ where $q = \sum_{i=1}^{k}\nu_i q_i$. The corresponding homogeneous population model, matched to have the same $R_0$ value, has offspring generating function $\phi_0(s) = \psi(-R_0(1-s))$, and for $R_0 > 1$ the probability of a major outbreak in the homogeneous model is $1 - q_0$ where $q_0$ is the unique solution in $0 \le q_0 < 1$ of $q_0 = \phi_0(q_0)$.

First, we exhibit a symmetry condition under which the major outbreak probability is unaffected by heterogeneities.

**Theorem 1** *If the rows of the mean offspring matrix $M$ all sum to the same value, then the major outbreak probability is the same as for a homogeneous population with the same $R_0$ value. In particular,*

(i) *heterogeneous susceptibility alone does not affect the major outbreak probability;*
(ii) *when all group sizes are equal, heterogeneous mixing alone does not affect the major outbreak probability.*

*Proof* A non-negative irreducible matrix with all row sums equal has Perron–Frobenius eigenvalue equal to the common row sum, so $R_0 = \beta \sum_j \lambda_i \pi_{ij} \mu_j f_j$ (for any $i$). Now for each $i$, Eq. (1) gives $\phi_i(q_0, q_0, \ldots, q_0) = \psi(-R_0(1-q_0)) = q_0$, so that $\boldsymbol{q} = (q_0, q_0, \ldots, q_0)$ provides the solution to Eq. (2), and hence $q = q_0$ for any initial distribution $\boldsymbol{\nu}$. Parts (i) and (ii), corresponding to $m_{ij} = (\beta/k)\mu_j f_j$ and $m_{ij} = (\beta/k)\pi_{ij}$ respectively, follow immediately. □

Theorem 1(i) was previously proved by Becker and Marschner (1990).

Theorem 1 applies irrespective of the initial distribution $\boldsymbol{\nu}$. The natural initial condition is to take the probability that the initial infective belongs to group $i$ to be proportional to the number of individuals in group $i$ and to the susceptibility of group $i$ individuals, so that $\nu_i = \mu_i f_i$ (Becker and Marschner 1990; Yates et al. 2006). In the separable case, this form for $\boldsymbol{\nu}$ provides the eigenvector of $M$ corresponding to the Perron–Frobenius eigenvalue $R_0 = (\beta/k)\sum_{i=1}^{k} \lambda_i \mu_i f_i$. For our general model, we have the following result, the proof of which is adapted from Section 4.3 of Becker and Marschner (1990).

**Theorem 2** *For an outbreak initiated by a single infective introduced into group $i$ with probability $\nu_i$, where the probability distribution $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_k)$ is the left eigenvector of the mean offspring matrix $M$ corresponding to the Perron–Frobenius eigenvalue $R_0 > 1$, then $q \geq q_0$. That is, the probability of a major outbreak in a heterogeneous population is no greater than the corresponding probability in a homogeneous population with the same $R_0$ value.*

*Proof* Since the generating function $\psi$ is convex, an application of Jensen's inequality yields

$$
\begin{aligned}
q &= \sum_{i=1}^{k} \nu_i \, \psi\left(-\sum_{j=1}^{k} m_{ij}\left(1-q_j\right)\right) \\
&\geq \psi\left(-\sum_{i,j} \nu_i m_{ij}\left(1-q_j\right)\right) \\
&= \psi\left(-R_0 \sum_{j=1}^{k} \nu_j \left(1-q_j\right)\right) = \psi\left(-R_0\left(1-q\right)\right) = \phi_0\left(q\right).
\end{aligned}
$$

The function $\phi_0(s)$ is convex with $\phi_0(s) > s$ for $0 \leq s < q_0$ and $\phi_0(s) < s$ for $q_0 < s < 1$, so that $q \geq q_0$ as required. □

Other than in the separable case, the initial condition required by Theorem 2 is not particularly natural. The eigenvector of $M$ with eigenvalue $R_0$ does have a biological interpretation, but in terms of the long-term behaviour of the process. Specifically, if we denote by $Z_i^{(n)}$ the number of group $i$ individuals in generation $n$, then conditional upon the branching process producing an infinite number of offspring,

$$\lim_{n\to\infty} \frac{Z_i^{(n)}}{\sum_{j=1}^{k} Z_j^{(n)}} = \alpha_i \text{ for } i = 1, 2, \ldots, k,$$

where $\boldsymbol{\alpha}$ is the normalised left eigenvector of $M$ with eigenvalue $R_0$ (Jagers 1975, p. 95). In practice, convergence is often quite rapid; that is, conditional upon non-extinction, the proportion of generation $n$ individuals belonging to group $i$ is close to $\alpha_i$ within a few generations. Consequently, Theorem 2 seems reasonably robust to different initial conditions. For instance, Figure 1 of Yates et al. (2006) shows a variety of examples, each with initial condition $\nu_i = \mu_i f_i$, where in each case the major outbreak probability is found to be bounded above by the homogeneous population value, although in most cases $\boldsymbol{\nu}$ is not an eigenvector of $M$. Nishiura et al. (2011) studied a model for Influenza transmission with $k = 2$ groups representing children and adults. Using parameter values based upon data from Influenza A (H1N1-2009) in Mexico, figure 4 of Nishiura et al. (2011) shows that for an outbreak initiated by an adult case ($\boldsymbol{\nu} = (0, 1)$) the major outbreak probability is bounded above by the homogeneous population value, whereas for an outbreak initiated by a child case ($\boldsymbol{\nu} = (1, 0)$) the major outbreak probability can exceed the homogeneous population value, although only slightly; neither $(0, 1)$ nor $(1, 0)$ is an eigenvector of the next generation matrix.

It has been previously shown (Becker and Marschner 1990) that heterogeneity in infectivity alone reduces the probability of a major outbreak, and observed in numerical examples (Yates et al. 2006) that heterogeneity in susceptibility in combination with other forms of heterogeneity can reduce the major outbreak probability. Theorem 2 shows that when heterogeneity in infectivity and susceptibility are combined, the combined heterogeneities do not increase the major outbreak probability.

Having compared heterogeneous with homogeneous populations, we now consider comparisons between two heterogeneous populations. For the remainder of this section we restrict ourselves to the separable case with initial condition $\nu_i = \mu_i f_i$. This allows us to investigate the major outbreak probability using a single-type, rather than multi-type, branching process as follows. Denote by $G_i$ the total number of offspring of a typical type $i$ individual in our multi-type branching process, $G_i = G_{i1} + G_{i2} + \cdots + G_{ik}$. Define a random variable $G$ to be distributed as $G_i$ with probability $\mu_i f_i$ for $i = 1, 2, \ldots, k$, so that $G$ has probability generating function

$$\phi(s) = E\left[s^G\right] = \sum_{i=1}^{k} \mu_i f_i \psi\left(-(\beta/k)\lambda_i(1-s)\right).$$

With initial condition $\nu_i = \mu_i f_i$ it is clear that the total progeny of the multi-type branching process has the same distribution as that of a single-type branching process,

initiated by a single individual, in which each individual has number of offspring distributed as $G$. Hence for $R_0 > 1$ the minor outbreak probability $q$ is the unique solution in $0 \leq q < 1$ of $q = \phi(q)$.

Our general result regarding comparison of two heterogeneous populations is as follows. Note that in Theorem 3, population 2 is in some sense 'more heterogeneous' than population 1, as is made clearer by the succeeding corollaries.

**Theorem 3** *Consider two heterogeneous populations in the separable case, so* $m_{ij}^{(1)} = (\beta^{(1)}/k)\lambda_i^{(1)}\mu_j^{(1)}f_j^{(1)}$ *and* $m_{ij}^{(2)} = (\beta^{(2)}/k)\lambda_i^{(2)}\mu_j^{(2)}f_j^{(2)}$. *For an outbreak initiated by a single infective belonging to group* $i$ *with probability* $\nu_i = \mu_i f_i$, *then* $q^{(1)} \leq q^{(2)}$ *provided there exists a non-negative* $k \times k$ *matrix* $A$ *with* $\sum_{i=1}^{k} a_{ij} = 1$ *for* $j = 1, 2, \ldots, k$ *such that* $\boldsymbol{\mu}^{(1)}F^{(1)}A^T = \boldsymbol{\mu}^{(2)}F^{(2)}$ *and* $\beta^{(1)}\boldsymbol{\lambda}^{(1)} = \beta^{(2)}\boldsymbol{\lambda}^{(2)}A$, *where* $F = diag(f_1, f_2, \ldots, f_k)$.

Notice that in the conditions of Theorem 3, the matrix $A$ appears in the left hand side of the susceptibility condition and the right hand side of the infectivity condition, giving the requisite compensation to keep $R_0$ constant, i.e. these conditions imply $R_0^{(1)} = R_0^{(2)}$.

*Proof* Set $\nu_i^{(1)} = \mu_i^{(1)}f_i^{(1)}$ and $\nu_i^{(2)} = \mu_i^{(2)}f_i^{(2)}$. Then since $\psi$ is convex, applying Lemma 1(iii) yields

$$\begin{aligned} q^{(1)} = \phi^{(1)}\left(q^{(1)}\right) &= \sum_{i=1}^{k} \nu_i^{(1)} \psi\left(-\left(\beta^{(1)}/k\right)\lambda_i^{(1)}\left(1-q^{(1)}\right)\right) \\ &\leq \sum_{i=1}^{k} \nu_i^{(2)} \psi\left(-\left(\beta^{(2)}/k\right)\lambda_i^{(2)}\left(1-q^{(1)}\right)\right) = \phi^{(2)}\left(q^{(1)}\right). \end{aligned}$$

Arguing as in the proof of Theorem 2, it follows that $q^{(1)}$ is less than or equal to the fixed point of the function $\phi^{(2)}$. That is, $q^{(1)} \leq q^{(2)}$. □

The following corollaries are immediate from Theorem 3 and Lemma 1(i), (ii).

**Corollary 1** *For two populations in the separable case with* $\mu_j^{(1)}f_j^{(1)} = \mu_j^{(2)}f_j^{(2)} = \nu_j$, *then for an outbreak initiated by a single infective belonging to group* $i$ *with probability* $\nu_i$ *we have that* $\beta^{(1)}\boldsymbol{\lambda}^{(1)} \prec_{\boldsymbol{\nu}} \beta^{(2)}\boldsymbol{\lambda}^{(2)} \Rightarrow q^{(1)} \leq q^{(2)}$.

**Corollary 2** *With heterogeneous infectivity alone and equal group sizes, then for an outbreak initiated by a single infective whose group is chosen uniformly at random from* $\{1, 2, \ldots, k\}$ *we have that* $\boldsymbol{\lambda}^{(1)} \prec \boldsymbol{\lambda}^{(2)} \Rightarrow q^{(1)} \leq q^{(2)}$.

Notice that we assume a common $\beta$ value in Corollary 2, as otherwise we would have the condition $\beta^{(1)}\boldsymbol{\lambda}^{(1)} \prec \beta^{(2)}\boldsymbol{\lambda}^{(2)}$, implying $\beta^{(1)} = \beta^{(2)}$. Corollary 2 extends a result of Becker and Marschner (1990), that with heterogeneity in infectivity alone the major outbreak probability is maximised in a homogeneous population.

Some numerical examples are shown in Fig. 1 illustrating that the more heterogeneous the population, the smaller the probability of a major outbreak. In previous numerical work, Yates et al. (2006) assumed a constant (non-random) infectious

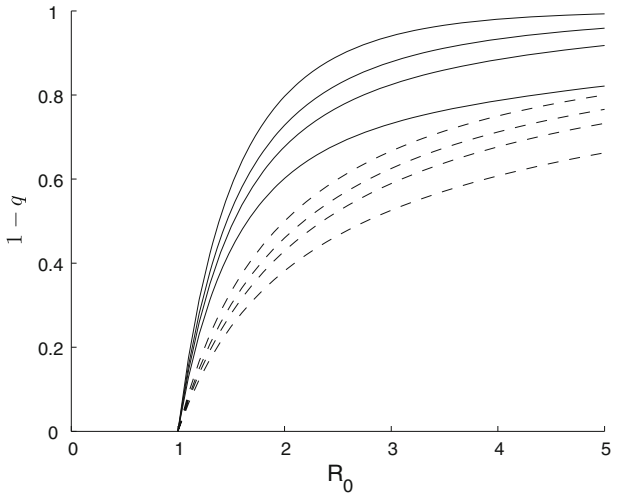


**Fig. 1** The effect of heterogeneity in both infectivity and susceptibility on the major outbreak probability $1-q$. Fixed parameter values $k=2$, $f_1=f_2=1/2$, $\pi_{ij}=1/2$ for all $i, j$. *Solid lines* computed with constant infectious period $T=1$; *dashed lines* with $T$ exponentially distributed with mean 1. For each infectious period distribution, the *curves* from *top* to *bottom* correspond to $\boldsymbol{\mu}=\boldsymbol{\lambda}=(1,1)$; $\boldsymbol{\mu}=(1.5,0.5)$, $\boldsymbol{\lambda}=(1.512,0.488)$; $\boldsymbol{\mu}=(1.4,0.6)$, $\boldsymbol{\lambda}=(1.64,0.36)$; $\boldsymbol{\mu}=(1.32,0.68)$, $\boldsymbol{\lambda}=(1.8,0.2)$. For comparing successive non-homogeneous cases, the matrix $A$ required by Theorem 3 is $A=\begin{pmatrix}0.9 & 0.1\\ 0.1 & 0.9\end{pmatrix}$

period, while Xiao et al. (2006) assumed an exponentially distributed infectious period. Our results apply for any specified distribution of infectious period, and so we illustrate our results for the constant and exponentially distributed cases. Results are shown for $k=2$ equally sized groups. Note that a constant infectious period consistently leads to a higher probability of a major outbreak than an exponentially distributed infectious period, in line with the findings of Vergu (2010) and Britton and Lindenstrand (2009), although the effect of the infectious period distribution is not the focus of the current work.

## 4 Outbreak size

In considering the major outbreak probability, we did not need to specify how an individual behaves after the end of its infectious period—in particular, whether infection is followed by a period of immunity, or by an immediate return to susceptibility. This is because in the large population limit, with high probability no individual is contacted more than once during the early stage of an outbreak, so that whether or not a previously-infected individual is susceptible to re-infection is irrelevant when computing the probability of a major outbreak. Suppose now that infection is followed by lifelong immunity, the model thus obtained being conventionally referred to as a multi-group susceptible-infective-removed (SIR) model. It is then natural to consider the distribution of the final size vector $\boldsymbol{N}^*=(N_1^*, N_2^*, \ldots, N_k^*)$, where $N_i^*$ is the total number of group i individuals to become infected during the course of the outbreak. Denote by $\bar{N}_i^*=N_i^*/N_i$ the proportion of group $i$ individuals ever infected. In a

large population, then conditional upon the occurrence of a large outbreak, it is known (Ball and Clancy 1993) that the distribution of $\bar{\boldsymbol{N}}^* = (\bar{N}_1^*, \bar{N}_2^*, \ldots, \bar{N}_k^*)$ is approximately multivariate normal with mean vector $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_k)$ satisfying

$$\tau_i = 1 - \exp\left(-\sum_{j=1}^{k} \tau_j f_j \beta_{ji}\right) \text{ for } i = 1, 2, \ldots, k. \tag{3}$$

The elements of the corresponding variance matrix can also be written down in terms of the parameters of the process (see Ball and Clancy 1993), but since fluctuations about $\boldsymbol{\tau}$ are of order $1/\sqrt{N}$ in the limit, we shall consider only mean behaviour.

Equation (3) has been analysed by Scalia-Tomba (1986) in the specific case of a constant infectious period $T = 1$, known as the multitype Reed–Frost process, but Eq. (3) is valid for any infectious period distribution of mean 1, see Ball and Clancy (1993). Equation (3) clearly admits the solution $\boldsymbol{\tau} = \mathbf{0}$, corresponding to no major outbreak. Lemma 1(i) of Scalia-Tomba (1986) shows that (provided the matrix with entries $f_j \beta_{ji}$ is irreducible) for $R_0 \leq 1$ there is no non-zero solution in $[0, 1]^k$, while for $R_0 > 1$ there is a unique non-zero solution in $[0, 1]^k$ corresponding to a major outbreak.

Denote by $\tau = \sum_{i=1}^{k} f_i \tau_i$ the mean proportion of the entire population who become infected (conditional upon a major outbreak), and by $\tau_0$ the corresponding quantity in a homogeneous population. Then for $R_0 > 1$, $\tau_0$ is the unique solution in $0 < \tau_0 \leq 1$ of $\tau_0 = 1 - \exp(-R_0 \tau_0)$. Similarly to Theorem 1 we have first the following result.

**Theorem 4** *If the columns of the matrix $FMF^{-1}$ all sum to the same value, then the mean size of a major outbreak is the same as for a homogeneous population with the same $R_0$ value. In particular,*

(i) *heterogeneous infectivity alone does not affect the mean major outbreak size;*
(ii) *when all groups sizes are equal, heterogeneous mixing alone does not affect the mean major outbreak size.*

*Proof* If $\sum_j f_j \lambda_j \pi_{ji} \mu_i$ takes the same value for every $i$, then $R_0 = \beta \sum_j f_j \lambda_j \pi_{ji} \mu_i$ (for any $i$) and Eq. (3) admit the symmetrical solution $\tau_i = \tau_0$ for $i = 1, 2, \ldots, k$. Hence $\tau = \sum_i f_i \tau_i = \tau_0$. Parts (i), (ii) follow immediately. □

Theorem 4(i) was previously observed by Ma and Earn (2006).

For the separable case, Theorem 4 of Andreasen (2011) shows that if $\lambda_1 < \lambda_2 < \cdots < \lambda_k$ and $\mu_1 < \mu_2 < \cdots < \mu_k$ then $\tau \leq \tau_0$. The following result extends the conditions under which this can be shown to hold true.

**Theorem 5** *For the separable case, if $\sum_i \lambda_i \mu_i f_i \geq 1$ then $\tau \leq \tau_0$. That is, such heterogeneity does not increase the mean size of a major outbreak, compared to the homogeneous situation.*

Note: Regarding the group label $i$ as a random variable with probability mass function $\boldsymbol{f}$, the condition $\sum_i \lambda_i \mu_i f_i \geq 1$ may be written as Covariance$(\boldsymbol{\lambda}, \boldsymbol{\mu}) \geq 0$. The condition that elements of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ be similarly ordered clearly implies positive correlation between $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, so that Theorem 4 of Andreasen (2011) is a special case of our result.

*Proof* In the separable case we have $\beta_{ij} = (\beta/k)\lambda_i\mu_j$, so writing $C = \sum_i \tau_i\lambda_i f_i$ then Eq. (3) reduces to

$$\tau_i = 1 - \exp\left(-(\beta/k)C\mu_i\right) \quad \text{for } i = 1, 2, \ldots, k. \tag{4}$$

Substituting back into the definition of $C$ yields

$$C = 1 - \sum_{i=1}^{k} \lambda_i f_i \exp\left(-(\beta/k)C\mu_i\right).$$

Defining the function $g$ to be

$$g(\theta) = 1 - \sum_{i=1}^{k} \lambda_i f_i \exp\left(-(\beta/k)\theta\mu_i\right) \quad \text{for } \theta \in \mathbb{R},$$

then $g$ is a strictly increasing concave function with $g(0) = 0$, $g(\theta) \to 1$ as $\theta \to \infty$, and $g'(0) = R_0$. So for $R_0 > 1$ there is a unique $C \in (0, \infty)$ with $g(C) = C$, and Eq. (4) with this value of $C$ provides the unique non-zero solution in $[0, 1]^k$ of Eq. (3), in agreement with Andreasen (2003).

By Jensen's inequality,

$$\begin{aligned} C &= 1 - \sum_{i=1}^{k} f_i\lambda_i \exp\left(-(\beta/k)C\mu_i\right) \\ &\le 1 - \exp\left(-(\beta/k)C\sum_{i=1}^{k} f_i\lambda_i\mu_i\right) = 1 - \exp\left(-R_0C\right), \end{aligned}$$

so that $C \le \tau_0$. Another application of Jensen's inequality gives

$$\begin{aligned} \tau = \sum_{i=1}^{k} f_i\tau_i &= 1 - \sum_{i=1}^{k} f_i \exp\left(-(\beta/k)C\mu_i\right) \\ &\le 1 - \exp\left(-(\beta/k)C\sum_{i=1}^{k} f_i\mu_i\right) \\ &= 1 - \exp(-(\beta/k)C) \\ &\le 1 - \exp\left(-(\beta/k)\tau_0\right). \end{aligned}$$

The condition $\sum_i \lambda_i\mu_i f_i \ge 1$ is equivalent to $\beta/k \le R_0$, and so $\tau \le 1 - \exp\left(-R_0\tau_0\right) = \tau_0$, as required. □

Theorem 5 shows that with positive correlation between susceptibility and infectivity, heterogeneity leads to reduced infectious spread. It is then natural to conjecture that negative correlation will lead to an increase in mean outbreak size, but in fact

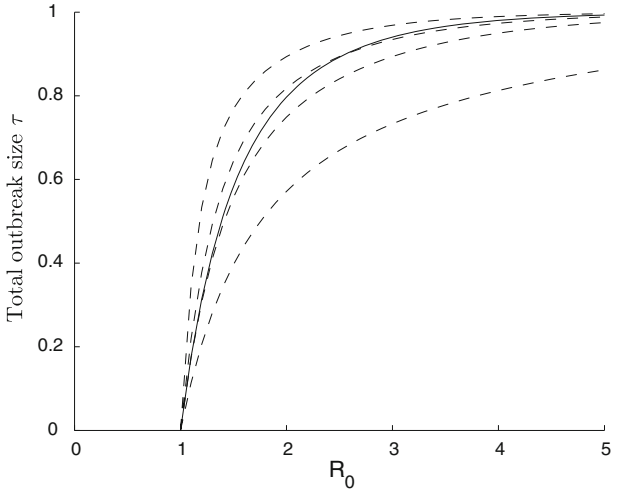


**Fig. 2** The effect of heterogeneity in both infectivity and susceptibility on the mean major outbreak size $\tau$. Fixed parameter values $k = 2$, $f_1 = f_2 = 1/2, \pi_{ij} = 1/2$ for all $i, j$. *Solid line* represents the homogeneous case $\boldsymbol{\mu} = \boldsymbol{\lambda} = (1, 1)$. *Dashed lines* all have $\boldsymbol{\mu} = (1.6, 0.4)$, and from *top* to *bottom* correspond to $\boldsymbol{\lambda} = (0.01, 1.99), (0.2, 1.8), (0.4, 1.6), (1.8, 0.2)$

this is not always the case, as pointed out by Andreasen (2011). Figure 2 illustrates the range of possible effects, with $k = 2$ equally sized groups. For all heterogeneous cases we take susceptibility vector $\boldsymbol{\mu} = (1.6, 0.4)$, so that with $\boldsymbol{\lambda} = (1.8, 0.2)$ there is a positive correlation between susceptibility and infectivity, Theorem 5 applies, and mean outbreak size is reduced in comparison to the homogeneous case. For the other three cases shown there is negative correlation between susceptibility and infectivity, and we see a range of possible behaviours. With $\boldsymbol{\lambda} = (0.01, 1.99)$, mean outbreak size is greater than in the homogeneous case across the range of $R_0$ values shown; with $\boldsymbol{\lambda} = (0.2, 1.8)$, mean outbreak size is greater than the homogeneous case for small $R_0$ but lower for large $R_0$; and for $\boldsymbol{\lambda} = (0.4, 1.6)$ the mean outbreak size is less than in the homogeneous case across the range of $R_0$ values plotted, even though susceptibility and infectivity are negatively correlated. Note that there is no distinction to be made here between different infectious period distributions, since Eq. (3) depends only upon the mean $E[T]$.

Although mean outbreak size is unaffected by heterogeneity in infectivity alone, when heterogeneity is in both susceptibility and infectivity, we see from Fig. 2 that different infectivity vectors $\boldsymbol{\lambda}$ do give rise to differences. In terms of rigorous comparison between different heterogeneous populations, we can show that increasing heterogeneity in susceptibility leads to reduced mean major outbreak size, as follows.

**Theorem 6** *Consider two heterogeneous populations with heterogeneity in susceptibility alone, each having the same group structure, i.e.* $\boldsymbol{f}^{(1)} = \boldsymbol{f}^{(2)} = \boldsymbol{f}$. *Then* $\boldsymbol{\mu}^{(1)} \prec_f \boldsymbol{\mu}^{(2)} \Rightarrow \tau^{(1)} \geq \tau^{(2)}$.

*Proof* With $\boldsymbol{\lambda} = \mathbf{1}$ then $\tau = C$. From Lemma 1(ii),(iii) we have that $\boldsymbol{\mu}^{(1)} \prec_f \boldsymbol{\mu}^{(2)} \Rightarrow g^{(1)}(\theta) \geq g^{(2)}(\theta)$ for all $\theta$, so that $C^{(1)} \geq C^{(2)}$, and the result follows. □

## 5 Endemic prevalence level

In order to study long-term endemic behaviour in the simplest possible context, we now suppose that infection is followed by an immediate return to susceptibility, so we have a multi-type susceptible-infective-susceptible (SIS) model. Provided the population is sufficiently large (with every group large) the infection process may be approximated by a deterministic system as follows. Denote by $x_i(t)$ the proportion of group $i$ individuals who are infected at time $t$, by $x_i^0(t)$ the proportion of group $i$ individuals who were infective at time 0 and have not recovered by time $t$, and by $\bar{F}(u) = \Pr(T > u)$ the survival function of the infectious period $T$. Then the approximating deterministic system is

$$x_i(t) = x_i^0(t) + \int_0^t \beta\,(1 - x_i(u)) \sum_{j=1}^{k} x_j(u) f_j \lambda_j \pi_{ji} \mu_i \bar{F}(t-u)\,du \text{ for } i = 1, 2, \ldots, k. \tag{5}$$

Clearly the system (5) has a disease-free equilibrium point at $\mathbf{x} = (0, 0, \ldots, 0)$. Suppose now that the process is initiated from a non-zero equilibrium point $\mathbf{x}^*$, at time $t = -\infty$ rather than $t = 0$. Then provided $\Pr(T < \infty) = 1$, Eq. (5) simplifies to

$$x_i^* = \int_{-\infty}^{t} \beta\left(1 - x_i^*\right) \sum_{j=1}^{k} x_j^* f_j \lambda_j \pi_{ji} \mu_i \bar{F}(t-u)\,du.$$

Substituting $v = t - u$ we have $\int_{-\infty}^{t} \bar{F}(t-u)\,du = \int_0^{\infty} \bar{F}(v)\,dv = E[T] = 1$, and so

$$x_i^* = \beta\left(1 - x_i^*\right) \sum_{j=1}^{k} x_j^* f_j \lambda_j \pi_{ji} \mu_i \text{ for } i = 1, 2, \ldots, k. \tag{6}$$

When infectious periods are exponentially distributed, Eq. (5) may be written as

$$\frac{dx_i}{dt} = \beta\,(1 - x_i) \sum_{j=1}^{k} x_j f_j \lambda_j \pi_{ji} \mu_i - x_i \text{ for } i = 1, 2, \ldots, k. \tag{7}$$

For the system (7), Lajmanovich and Yorke (1976) showed that if $M$ is irreducible then (i) for $R_0 \leq 1$ the disease-free equilibrium is globally asymptotically stable (and hence the unique feasible equilibrium point); (ii) for $R_0 > 1$ there exists a globally asymptotically stable non-zero equilibrium point $\boldsymbol{x}^*$ in $[0, 1]^k$. Although their analysis relied upon the assumption that infectious periods are exponentially distributed, Eq. (6) depends only upon the mean of the infectious period. It follows that for $R_0 > 1$, for infectious period $T$ having any distribution such that $E[T] = 1$ and $\Pr(T < \infty) = 1$, there exists a unique non-zero equilibrium point $\mathbf{x}^*$, the endemic equilibrium.

Equation (6) is sufficiently similar in structure to Eq. (3) that the analysis can proceed in a closely parallel manner. We denote by $x^* = \sum_{i=1}^{k} f_i x_i^*$ the overall endemic prevalence level, and note that in a homogeneous population the endemic prevalence level is given by $x_0^* = 1 - (1/R_0)$ for $R_0 > 1$.

**Theorem 7** *If the columns of the matrix $FMF^{-1}$ all sum to the same value, then the endemic mean prevalence of infection is the same as for a homogeneous population with the same $R_0$ value. In particular,*

(i) *heterogeneous infectivity alone does not affect the endemic prevalence level;*
(ii) *when all groups sizes are equal, heterogeneity in mixing alone does not affect the endemic prevalence level.*

*Proof* Since $R_0 = \beta \sum_j f_j \lambda_j \pi_{ji} \mu_i$ (for any $i$), Eq. (6) admits the symmetrical solution $x_i^* = x_0^*$ for $i = 1, 2, \ldots, k$. Hence $x^* = \sum_i f_i x_i^* = x_0^*$. Parts (i),(ii) follow immediately. □

The following result is analogous to Theorem 5.

**Theorem 8** *For the separable case, if $\sum_i \lambda_i \mu_i f_i \geq 1$ then $x^* \leq 1 - (1/R_0)$. That is, such heterogeneity does not increase the endemic mean prevalence of infection, compared to the homogeneous situation.*

*Proof* With $\beta_{ij} = (\beta/k)\lambda_i \mu_j$ Eq. (6) becomes

$$x_i = \frac{\beta}{k}(1 - x_i)\mu_i \sum_{j=1}^{k} x_j f_j \lambda_j \text{ for } i = 1, 2, \ldots, k.$$

Following Nold (1980) we set $D = (\beta/k) \sum_j x_j f_j \lambda_j$. The elements of the endemic equilibrium point $\mathbf{x}^*$ are then given by

$$x_i^* = \frac{D\mu_i}{1 + D\mu_i} \text{ for } i = 1, 2, \ldots, k. \tag{8}$$

The value of $D$ may be determined by substituting back from (8) into the definition of $D$, giving either $D = 0$ (corresponding to the disease-free equilibrium) or

$$\frac{\beta}{k} \sum_{j=1}^{k} \frac{\mu_j f_j \lambda_j}{1 + D\mu_j} = 1,$$

given as equation (5.3) in Nold (1980). Defining the function $h$ by

$$h(\theta) = \frac{\beta}{k} \sum_{j=1}^{k} \frac{\mu_j f_j \lambda_j}{1 + \theta\mu_j} \text{ for } \theta \in \mathbb{R}, \tag{9}$$

then $h$ is a strictly decreasing continuous function with $h(0) = R_0$ and $h(\theta) \to 0$ as $\theta \to \infty$. Consequently, for $R_0 > 1$ there is a unique $D \in (0, \infty)$ satisfying $h(D) = 1$.

By Jensen's inequality,

$$h(\theta) = \frac{\beta}{k}\sum_{j=1}^{k} f_j\lambda_j\left(\frac{\mu_j}{1+\theta\mu_j}\right) \le \frac{\beta}{k}\left(\frac{\sum_j f_j\lambda_j\mu_j}{1+\theta\sum_j f_j\lambda_j\mu_j}\right) = \frac{R_0}{1+\theta\sum_j f_j\lambda_j\mu_j}.$$

Now $h(D) = 1$, and so

$$\begin{aligned} \frac{R_0}{1+D\sum_j f_j\lambda_j\mu_j} &\ge 1 \\ \Rightarrow D\sum_j f_j\lambda_j\mu_j &\le R_0 - 1 \\ \Rightarrow D &\le R_0 - 1, \end{aligned} \tag{10}$$

since $\sum_j f_j\lambda_j\mu_j \ge 1$ by assumption. Applying Jensen's inequality once more,

$$\begin{aligned} x^* = \sum_{i=1}^{k} f_i x_i^* &= D\sum_{i=1}^{k} f_i\left(\frac{\mu_i}{1+D\mu_i}\right) \\ &\le D\left(\frac{\sum_i f_i\mu_i}{1+D\sum_i f_i\mu_i}\right) \\ &= \frac{D}{1+D} \\ &\le \frac{R_0-1}{R_0}, \end{aligned}$$

as required. □

Theorem 8 provides a sufficient condition, but not a necessary condition, for heterogeneities to result in a decrease in endemic prevalence. An alternative sufficient condition, and a necessary (but not sufficient) condition, are provided by the following result. The proof is based upon that of theorem 4 of Andreasen (2011).

**Theorem 9** *Consider a heterogeneous population in the separable case and label the groups such that* $\mu_1 \le \mu_2 \le \cdots \le \mu_k$. *If* $\lambda_1\mu_1 \le \lambda_2\mu_2 \le \cdots \le \lambda_k\mu_k$, *then* $x^* \le 1 - (1/R_0)$. *Conversely, if* $\lambda_1\mu_1 \ge \lambda_2\mu_2 \ge \cdots \ge \lambda_k\mu_k$, *then* $x^* \ge 1 - (1/R_0)$.

*Proof* From formula (8) and definition (9), the equation $h(D) = 1$ may be written as $(\beta/k)\sum_{j=1}^{k}\mu_j f_j\lambda_j(1-x_j^*) = 1$, or equivalently

$$\frac{\beta}{k}\sum_{j=1}^{k}\mu_j f_j\lambda_j x_j^* = R_0 - 1. \tag{11}$$

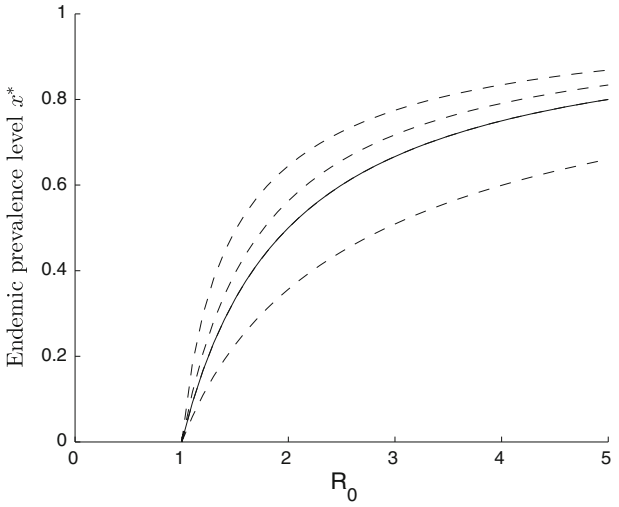


**Fig. 3** The effect of heterogeneity in both infectivity and susceptibility on the endemic prevalence level $x^*$. Fixed parameter values $k = 2$, $f_1 = f_2 = 1/2$, $\pi_{ij} = 1/2$ for all $i, j$. *Solid line* represents the homogeneous case $\boldsymbol{\mu} = \boldsymbol{\lambda} = (1, 1)$. *Dashed lines* all have $\boldsymbol{\mu} = (1.6, 0.4)$, and from *top* to *bottom* correspond to $\boldsymbol{\lambda} = (0.01, 1.99)$, $(0.2, 1.8)$, $(1.8, 0.2)$

Regarding the group label $j$ as a random variable with probability mass function $\boldsymbol{f}$, and denoting $\Lambda = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_k)$, then

$$\begin{aligned}\text{Covariance}\left(\boldsymbol{\mu}\Lambda, \boldsymbol{x}^*\right) &= \sum_{j=1}^{k} f_j \lambda_j \mu_j x_j^* - \left(\sum_{j=1}^{k} f_j \lambda_j \mu_j\right)\left(\sum_{j=1}^{k} f_j x_j^*\right) \\ &= \frac{k R_0}{\beta}\left(1 - \frac{1}{R_0} - x^*\right).\end{aligned}$$

From formula (8) we see that $x_j^*$ is an increasing function of $\mu_j$, so that provided the elements of $\boldsymbol{\mu}\Lambda$ and $\boldsymbol{\mu}$ are similarly ordered, it follows that the variables $\boldsymbol{\mu}\Lambda$ and $\boldsymbol{x}^*$ have positive correlation with respect to any probability mass function $\boldsymbol{f}$, and the result follows. The converse follows similarly, with opposite ordering of $\boldsymbol{\mu}\Lambda$ and $\boldsymbol{\mu}$ implying a negative correlation. □

For $k = 2$ groups, Theorem 9 covers all possibilities, although not for $k \geq 3$. Figure 3 shows some numerical examples for $k = 2$ equally sized groups.

As for mean outbreak size, we see from Fig. 3 that when heterogeneity is in both susceptibility and infectivity then different infectivity vectors $\boldsymbol{\lambda}$ can give rise to differences in endemic prevalence level. The following result, analogous to Theorem 6, shows that increasing heterogeneity in susceptibility leads to a reduced endemic prevalence level.

**Theorem 10** *Consider two heterogeneous populations with heterogeneity in susceptibility alone, each having the same group structure* $\boldsymbol{f}$. *Then* $\boldsymbol{\mu}^{(1)} \prec_{\boldsymbol{f}} \boldsymbol{\mu}^{(2)} \Rightarrow x^{*(1)} \geq x^{*(2)}$.

*Proof* With $\boldsymbol{\lambda} = \mathbf{1}$ then $D = (\beta/k)x^*$. From Lemma 1(ii),(iii), since $\xi(\mu) = -\mu/(1+\theta\mu)$ is a convex function for any $\theta > 0$, we have that $\boldsymbol{\mu}^{(1)} \prec_f \boldsymbol{\mu}^{(2)}$ implies

$$h^{(1)}(\theta) = \frac{\beta}{k}\sum_{j=1}^{k} f_j\left(\frac{\mu_j^{(1)}}{1+\theta\mu_j^{(1)}}\right) \geq \frac{\beta}{k}\sum_{j=1}^{k} f_j\left(\frac{\mu_j^{(2)}}{1+\theta\mu_j^{(2)}}\right) = h^{(2)}(\theta),$$

so that $D^{(1)} \geq D^{(2)}$ and the result follows. □

## 6 Time to fade-out of infection

When an infection becomes endemic in a population, beyond the endemic prevalence level the next characteristic of interest is the persistence time of the infection, often referred to as the time to fade-out. For the multi-type SIS model of the preceding section, provided infectious periods are exponentially distributed, the infectives process $\boldsymbol{X}(t)$ is a Markov chain on a finite state-space with a single absorbing state at $\boldsymbol{X} = \mathbf{0}$, all other states forming a single communicating class. It follows that with probability 1 the process will be absorbed at $\boldsymbol{X} = \mathbf{0}$ within finite time; that is, infection is certain to die out eventually. If $R_0 > 1$ then the time to absorption can be very long, and in the meantime the process will settle to a (unique) quasi-stationary distribution (Darroch and Seneta 1967). Denoting by $q_{\boldsymbol{x}}$ the quasi-stationary probability of being in state $\boldsymbol{x}$, and by $\boldsymbol{e}_i$ the vector with 1 as the $i$th component and zeros elsewhere, then starting from quasi-stationarity the (constant) hazard rate for absorption is $\sum_{i=1}^{k} q_{\boldsymbol{e}_i}$, so that the time to extinction is exponentially distributed with mean $(\sum_{i=1}^{k} q_{\boldsymbol{e}_i})^{-1}$ (see, for instance, Nåsell 1999). Since we are interested in the time to extinction given that the infection has become established in the population, it is natural to consider the process to be initiated from quasi-stationarity. The quasi-stationary distribution may be evaluated as an eigenvector of the transition rate matrix of the process, but since there is no explicit form for the elements $q_{\boldsymbol{x}}$ this is not very useful for investigating the effects of heterogeneities, so instead we proceed as follows.

For $R_0 > 1$, the process $(\boldsymbol{X}(t) - NF\boldsymbol{x}^*)(NF)^{-1/2}$, where $\boldsymbol{x}^*$ is the deterministic endemic equilibrium point of Sect. 4 and $F = diag\ (f_1, f_2, \ldots, f_k)$, may be approximated by a $k$-dimensional Ornstein–Uhlenbeck process, at least so long as the process remains in the neighbourhood of the point $\boldsymbol{X} = NF\boldsymbol{x}^*$, and so the quasi-stationary distribution can be approximated using the stationary distribution of this Ornstein–Uhlenbeck process. The method is based on results such as those of Barbour (1972, 1976) and Section 11.2 of Ethier and Kurtz (1986), and has previously been used to investigate quasi-stationary behaviour of a variety of epidemic models, for instance see Nåsell (1999, 2002, 2005), Andersson and Britton (1998), Lindholm (2008), Clancy and Mendy (2011). The approximation becomes precise in the limit as $N \to \infty$.

The approximating Ornstein–Uhlenbeck process has local drift matrix given by the Jacobian $J$ of the differential equation system (7) at $\boldsymbol{x}^*$, with elements

$$\begin{aligned} J_{ij} &= \beta\left(1-x_i^*\right) f_j \lambda_j \pi_{ji} \mu_i - \delta_{ij}\left(\beta \sum_{r=1}^{k} x_r^* f_r \lambda_r \pi_{ri} \mu_i + 1\right) \\ &= \beta\left(1-x_i^*\right) f_j \lambda_j \pi_{ji} \mu_i - \delta_{ij}\left(\frac{x_i^*}{1-x_i^*} + 1\right) \text{ (due to Eq. (6))} \\ &= \beta\left(1-x_i^*\right) f_j \lambda_j \pi_{ji} \mu_i - \frac{\delta_{ij}}{1-x_i^*}. \end{aligned}$$

where $\delta_{ij}$ is the Kronecker delta symbol. The local variance matrix of the Ornstein–Uhlenbeck process is a diagonal matrix $G$ with diagonal entries

$$G_{ii} = \beta\left(1-x_i^*\right) \sum_{r=1}^{k} x_r f_r \lambda_r \pi_{ri} \mu_i + x_i^* = 2x_i^*.$$

The stationary distribution of the Ornstein–Uhlenbeck process is (Gardiner 2009, Section 4.5.6) a $k$-dimensional normal distribution with mean zero and variance matrix $S$ satisfying the Lyapunov equation

$$JS + SJ^T + G = 0. \tag{12}$$

From Laub (2005), Theorem 13.21, we know that since $\boldsymbol{x}^*$ is locally stable equation (12) has a unique solution $S$, and since $G$ is symmetric it follows that $S$ is also symmetric.

Infection dies out when the process makes an excursion away from $NF\boldsymbol{x}^*$ to $\mathbf{0}$, the chance of which depends upon the variance of the quasi-stationary distribution. Denoting by $\sigma_{Total}$ the standard deviation (in the normal approximation) of the total number of infectives in the population, then an indication of typical time to fade-out of infection is given by the reciprocal of the coefficient of variation $CV = \sigma_{Total}/Nx^*$, where $x^* = \sum_i f_i x_i^*$. That is to say, a small value of $CV$ suggests a large expected persistence time (Hagenaars et al. 2004; Clancy and Mendy 2011). This treatment of persistence time is rather approximate, but gives some indication of the effects of model parameters.

For simplicity, for the remainder of this section we consider only the case of equally-sized groups. In this case, $f_i = 1/k$ for all $i$ and

$$CV^2 = \frac{\sigma_{Total}^2}{(Nx^*)^2} = \frac{k \sum_{i,j} S_{ij}}{N\left(\sum_i x_i^*\right)^2}. \tag{13}$$

A general algebraic solution to equation (12) is rather complicated, so we restrict ourselves to considering the three types of heterogeneity (in mixing, in infectivity, and in susceptibility) separately.

**Theorem 11** *For a population of k equally-sized groups, consider the coefficient of variation of the total number of infectives present in quasi-stationarity, as approximated via Eq.* (13).

(i) *With heterogeneity in mixing alone, $CV$ takes the same value as in the homogeneous case.*
(ii) *With heterogeneity in infectivity alone, then for two population with infectivity vectors $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}$ we have $\boldsymbol{\lambda}^{(1)} \prec \boldsymbol{\lambda}^{(2)} \Rightarrow CV^{(1)} \leq CV^{(2)}$. In particular, $CV$ is minimised in the homogeneous case.*
(iii) *With heterogeneity in susceptibility alone, $CV$ is at least as great as in the homogeneous case.*

*Proof* (i) With $f_i = 1/k$ and $\mu_i = \lambda_i = 1$ for all $i$, then $x_i^* = 1 - (1/R_0)$ for all $i$ where $R_0 = \beta/k$, and Eq. (12) reduces to

$$2R_0 S_{ij} - \sum_{r=1}^{k} \left(\pi_{ri} S_{rj} + \pi_{rj} S_{ir}\right) = 2\left(1 - \frac{1}{R_0}\right)\delta_{ij} \text{ for } i, j = 1, 2, \ldots, k. \quad (14)$$

Summing equations (14) over $i$ and $j$ and making use of the constraints $\sum_j \pi_{ij} = 1$, we find that $\sum_{i,j} S_{ij} = k/R_0$, and so formula (13) reduces to

$$CV^2 = \frac{R_0}{N(R_0 - 1)^2}.$$

In particular, the value of $CV$ does not depend upon the parameters $\pi_{ij}$.
(ii) With $f_i = 1/k$, $\mu_i = 1$ and $\pi_{ij} = 1/k$ for all $i, j$, then once again $x_i^* = 1 - (1/R_0)$ for all $i$ where $R_0 = \beta/k$, and Eq. (12) reduces to

$$2R_0 S_{ij} - \frac{1}{k}\sum_{r=1}^{k} \lambda_r \left(S_{rj} + S_{ir}\right) = 2\left(1 - \frac{1}{R_0}\right)\delta_{ij} \text{ for } i, j = 1, 2, \ldots, k. \quad (15)$$

With the aid of the Maple symbolic algebra package, the solution is found to have elements

$$S_{ij} = \frac{k\,(R_0 - 1)\left(\lambda_i + \lambda_j\right) + \sum_r \lambda_r^2}{k^2 R_0^2\,(2R_0 - 1)} + \left(\frac{R_0 - 1}{R_0^2}\right)\delta_{ij} \quad (16)$$

for $i, j = 1, 2, \ldots, k$. Note that Maple was not able to directly solve Eq. (15), but rather was used to solve for the cases $k = 2, 3$. It was then possible to guess the form of the solution for general $k$, and straightforward to check that the form (16) does indeed satisfy Eq. (15). Formulae (16) imply that

$$\sum_{i,j} S_{ij} = \frac{k}{R_0} + \frac{\left(\sum_r \lambda_r^2\right) - 1}{R_0^2\,(2R_0 - 1)}$$

and so

$$CV^2 = \frac{1}{N\,(R_0 - 1)^2}\left(R_0 + \frac{(1/k)\left(\sum_r \lambda_r^2\right) - 1}{(2R_0 - 1)}\right).$$

Now since $\xi(\lambda_r) = \lambda_r^2$ is a convex function, it follows from Lemma 1(i),(iii) that $\boldsymbol{\lambda}^{(1)} \prec \boldsymbol{\lambda}^{(2)} \Rightarrow CV^{(1)} \leq CV^{(2)}$.

(iii) With $f_i = 1/k$, $\lambda_i = 1$ and $\pi_{ij} = 1/k$ for all $i, j$, then $R_0 = \beta/k$ and $x_i^* = D\mu_i/(1 + D\mu_i)$ for $i = 1, 2, \ldots, k$ with $D$ defined as in Sect. 5. Equation (12) in this case can be written as

$$\left(\frac{1}{1-x_i^*} + \frac{1}{1-x_j^*}\right) S_{ij} - \frac{R_0}{kD}\left(x_i^* \sum_{r=1}^{k} S_{rj} + x_j^* \sum_{r=1}^{k} S_{ir}\right) = 2x_i^*\,\delta_{ij}$$

for $i, j = 1, 2, \ldots, k$. With the aid of Maple, the solution is found to have elements

$$S_{ij} = \frac{x_i^*\left(1-x_i^*\right) x_j^*\left(1-x_j^*\right) R_0}{kD - R_0 \sum_{r=1}^{k} x_r^*\left(1-x_r^*\right)} + x_i^*\left(1-x_i^*\right)\delta_{ij} \text{ for } i, j = 1, 2, \ldots, k,$$

so that, after a little simplifying algebra, we have

$$\sum_{i,j} S_{ij} = \frac{kD \sum_r x_r^*\left(1-x_r^*\right)}{kD - R_0 \sum_{r=1}^{k} x_r^*\left(1-x_r^*\right)}.$$

The definition of $D$ from Sect. 5 reduces here to $D = (R_0/k)\sum_r x_r^*$, so

$$\begin{aligned} CV^2 &= \frac{k\left(\sum_r x_r^*\right)\left(\sum_r x_r^*\left(1-x_r^*\right)\right)}{\sum_r x_r^* - \sum_{r=1}^{k} x_r^*\left(1-x_r^*\right)} \Bigg/ N\left(\sum_r x_r^*\right)^2 \\ &= \frac{k}{N}\left(\frac{1}{\sum_r \left(x_r^*\right)^2} - \frac{1}{\sum_r x_r^*}\right). \end{aligned} \tag{17}$$

Since $x_i^*$ is an increasing function of $\mu_i$ and the ratio $x_i^*/\mu_i$ is a decreasing function of $\mu_i$, then with respect to the probability mass function $\boldsymbol{\mu}/k$,

$$\text{Covariance}\left(\boldsymbol{x}^*, \boldsymbol{x}^*\, diag\,(1/\mu_1, 1/\mu_2, \ldots, 1/\mu_k)\right) \leq 0.$$

That is,

$$\begin{aligned} \sum_r (\mu_r/k) x_r^* (x_r^*/\mu_r) &\leq \left(\sum_r (\mu_r/k) x_r^*\right)\left(\sum_r (\mu_r/k)(x_r^*/\mu_r)\right) \\ \sum_r \left(x_r^*\right)^2 &\leq \frac{1}{k}\left(\sum_r \mu_r x_r^*\right)\left(\sum_r x_r^*\right). \end{aligned}$$

Equation (11) with $f_i = 1/k$ and $\lambda_i = 1$ gives $\sum_r \mu_r x_r^* = k(R_0 - 1)/R_0$, and we also have $\sum_r x_r^* = kD/R_0$, so that

$$\sum_r \left(x_r^*\right)^2 \leq kD\left(R_0-1\right)/R_0^2.$$

Equation (17) now yields

$$CV^2 \geq \frac{k}{N}\left(\frac{R_0^2}{kD\left(R_0-1\right)}-\frac{R_0}{kD}\right)=\frac{R_0}{ND\left(R_0-1\right)},$$

but from inequality (10) we know that $D \leq R_0 - 1$, and so finally

$$CV^2 \geq \frac{R_0}{N\left(R_0-1\right)^2}.$$

The right hand side of this last inequality gives the value of $CV^2$ in the homogeneous case, and the result follows.

□

Given the link between coefficient of variation and mean persistence time, Theorem 11 indicates that for a population of $k$ equally-sized groups, (i) heterogeneity in mixing alone does not affect (to first order) the mean time to fade-out of infection; (ii) with heterogeneity in infectivity alone, the greater the heterogeneity the more rapidly fade-out will occur, on average; (iii) heterogeneity in susceptibility alone reduces the mean time to fade-out of infection, compared to the homogeneous case. It is interesting to note that heterogeneity in susceptibility alone and heterogeneity in infectivity alone can both affect persistence times, in contrast to the results of Sects. 3, 4, 5.

## 7 Discussion

In general, our results suggest that in many circumstances population heterogeneities tend to result in rarer, less severe and less prolonged outbreaks of infection, compared to the homogeneous situation, although when different types of heterogeneities are combined this is not always the case. The most obvious interpretation is that if we know the true value of $R_0$, and mistakenly assume the population to be homogeneous, we may overestimate the likely severity of an outbreak. However, the question then arises as to how we can know the true value of $R_0$, since using an estimation procedure based upon an incorrect assumption of homogeneity will lead to an incorrect $R_0$ estimate. This suggests the following alternative interpretation of our results: if one attempts to estimate $R_0$ based upon (for instance) the observed size of a major outbreak or observed endemic prevalence level, then under the assumptions of Theorems 5, 9, an incorrect assumption of homogeneity is likely to lead to underestimation of $R_0$. This is in line with observations reported in Hethcote (1996).

It has previously been observed that heterogeneity in susceptibility alone does not affect the probability of a major outbreak (Becker and Marschner 1990), while heterogeneity in infectivity alone does not affect the mean size of a major outbreak (Ma and Earn 2006). We have shown that more generally, the major outbreak probability is unaffected by heterogeneities for which the value of $\sum_j \lambda_i \pi_{ij} \mu_j f_j$ does

not depend upon $i$; that is, the average total number of infectious contacts emanating from an infected individual does not depend upon the group to which the individual belongs. Mean outbreak size, and likewise the endemic mean prevalence level, are unaffected by heterogeneities for which $\sum_i f_i \lambda_i \pi_{ij} \mu_j$ does not depend upon $j$; that is, the average number of infectious contacts received by a susceptible individual from an infective individual chosen uniformly at random from the whole population does not depend upon the group to which the susceptible individual belongs. In contrast, Corollary 2 shows that increasing heterogeneity in infectivity leads to a reduction in the major outbreak probability, while Theorems 6, 10 show that increasing heterogeneity in susceptibility leads to a reduction in mean outbreak size and in endemic mean prevalence level. The mean time to fade-out of infection is affected by heterogeneities in either susceptibility alone or infectivity alone, with either form of heterogeneity leading to more rapid fade-out (Theorem 11).

In studying major outbreak probability and persistence time, a stochastic model is essential. In studying major outbreak size and endemic prevalence level, we were justified in adopting a deterministic approach. Nevertheless, it is important to be aware of the underlying stochastic model, to avoid over-interpretation of the deterministic approximation. In particular, our results of Sects. 4, 5 rely upon the numbers of individuals in each group being sufficiently large that the relevant approximations hold. If instead we are interested in a population consisting of a large number of small groups, such as households, then entirely different deterministic approximations are required, and even the characterisation of the basic reproduction number $R_0$ as the maximal eigenvalue of the next generation matrix is no longer appropriate—see Ball (1999), Ball et al. (2004), Neal (2006) for details.

Results similar to our Theorems 2, 5 have been obtained in the context of an infection spreading across a network by Miller (2008); it is assumed that individuals are exchangeable, and so the issue of choice of the initial infected individual in Theorem 2 does not arise; and an individual's infectivity and susceptibility are assumed independent of one another, so that the non-negative correlation condition of Theorem 5 is automatically satisfied. A network model in which an individual's infectivity and susceptibility may be correlated is studied in Meester and Trapman (2011).

The modelling assumptions we have adopted in order to prove our results become more stringent as we move from earlier to later stages of the infectious process. We now discuss the extent to which our results would be expected to remain robust to relaxation of the various assumptions.

We have assumed throughout that the infectious period distribution $T$ is the same for every group. It is straightforward to extend the results of Sects. 4, 5, 6 to allow for differing infectious period distributions in different groups (although in Sect. 6 we still require infectious periods to be exponentially distributed); for the results of Sect. 3, this remains true provided that the distributions differ only by linear scaling.

In investigating the initial outbreak of infection, whether in terms of the probability of a major outbreak or mean outbreak size, although we did not explicitly incorporate a latent period in our model the results of Sects. 3 and 4 remain valid for a latent period having any specified (almost surely finite) distribution, since the inclusion of such a latent period does not alter which individuals become infected but only the times at which infections occur (see Ball 1986). The fact that our model neglects demographic

processes (birth, death, immigration, emigration) is not a serious restriction in the initial outbreak phase, provided only that any outbreak takes place over a relatively short time period, during which the effects of such demographic processes are negligible. In computing the probability of a major outbreak, we did not need to make any assumptions regarding immune response—whether infection is followed by complete lifelong immunity, by an immediate return to susceptibility, or by some form of temporary or partial immunity, the branching process approximation valid in this early stage remains unaltered. For mean outbreak size, although we assumed that infection is followed by lifelong immunity, in practice our results remain valid provided the immune response is sufficient that the number of individuals becoming infected twice during a single outbreak remains negligible.

In studying endemicity, our assumptions become a little more restrictive. So far as demographic processes are concerned, one would not expect their effect to be negligible over the time scale relevant to an infection in endemic equilibrium. However, provided the population size in each group is approximately constant, with births and deaths balancing out, since the mean effect is zero there is some justification for neglecting demographic processes in the formulation of model (7). This is not an entirely convincing argument, because deaths of infective individuals would need to be balanced by births of new infected individuals, and deaths of susceptibles by births of new susceptibles. In fact, the mechanism allowing an infection to maintain itself in endemic equilibrium will often be recruitment of new susceptible individuals through birth or immigration, rather than the return of previously-infected individuals to a susceptible state. A model explicitly incorporating such demographic processes will be less amenable to rigorous analysis than our multigroup SIS model of Sects. 4, 5; we have deliberately chosen to focus upon the simplest possible model to allow a clear exposition of the effects of heterogeneities. Nevertheless, our multigroup SIS model is not unrealistic for certain infections, and in particular such a model has been used in practice to study the spread of gonorrhea (Hethcote and Yorke 1984; Lajmanovich and Yorke 1976; Nold 1980).

Our modelling assumptions become less tenable when we move on to consider persistence time of the infection, partly because we must now consider a longer time scale, and partly because quantification of variability is crucial to the results of Sect. 6. Thus Theorem 11 is concerned explicitly with variances computed based upon a form of central limit theorem. The assumption of exponentially distributed infectious periods is thus a significant restriction, since one would expect not only the mean, but also the variance of the infectious period distribution to be relevant. Demographic processes are important at this stage, not only because we have a longer time scale to deal with, but also because even if births and deaths balance each other on average, the amount of variability is now relevant. It is worth noting that Theorem 11(i) is somewhat at odds with the conclusions of Hagenaars et al. (2004), who considered an SIR model incorporating demographic processes and found that increasing heterogeneity typically led to more rapid fade-out of infection. The model of Hagenaars et al. (2004) allows for heterogeneity in mixing only, specifically taking $\pi_{ii} = \rho$ for all $i$ and $\pi_{ij} = (1-\rho)/(k-1)$ for $i \neq j$ with $1/k \leq \rho \leq 1$ (assortative mixing). Clearly there is much scope for further work in this area.

**Acknowledgments** Chris Pearce was supported by a PhD studentship from the Engineering and Physical Sciences Research Council.

## References

Adler FR (1992) The effects of averaging on the basic reproduction ratio. Math Biosci 111:89–98
Andersson H, Britton T (1998) Heterogeneity in epidemic models and its effect on the spread of infection. J Appl Probab 35:651–661
Andersson H, Britton T (2000) Stochastic epidemics in dynamic populations: quasi-stationarity and extinction. J Math Biol 41:559–580
Andreasen V (2003) Dynamics of annual influenza A epidemics with immuno-selection. J Math Biol 46:504–536
Andreasen V (2011) The final size of an epidemic and its relation to the basic reproduction number. Bull Math Biol 73:2305–2321
Ball FG (1983) The threshold behaviour of epidemic models. J Appl Probab 20:227–241
Ball FG (1985) Deterministic and stochastic epidemics with several kinds of susceptibles. Adv Appl Probab 17:1–22
Ball FG (1986) A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. Adv Appl Probab 18:289–310
Ball FG (1999) Stochastic and deterministic models for SIS epidemics among a population partitioned into households. Math Biosci 156:41–68
Ball FG, Britton T, Lyne OD (2004) Stochastic multitype epidemics in a community of households: estimation of threshold parameter and secure vaccination coverage. Biometrika 91:345–362
Ball FG, Clancy D (1993) The final size and severity of a generalised stochastic multitype epidemic model. Adv Appl Probab 25:721–736
Barbour AD (1972) The principle of the diffusion of arbitrary constants. J Appl Probab 9:519–541
Barbour AD (1976) Quasi-stationary distributions in Markov population processes. Adv Appl Probab 8: 296–314
Becker N, Marschner I (1990) The effect of heterogeneity on the spread of disease. Lect Notes Biomath 86:90–103
Blackwell D (1951) Comparisons of experiments. In: Proceedings of the second Berkeley symposium on mathematical statistics and probability. University of California Press, Berkeley, pp 93–102
Blackwell D (1953) Equivalent comparisons of experiments. Ann Math Stat 24:265–272
Borwein JM, Lewis AS, Nussbaum RD (1994) Entropy minimization, DAD problems, and doubly stochastic kernels. J Funct Anal 123:264–307
Britton T, Lindenstrand D (2009) Epidemic modelling: aspects where stochasticity matters. Math Biosci 222:109–116
Clancy D, Mendy ST (2011) The effect of waning immunity on long-term behaviour of stochastic models for the spread of infection. J Math Biol 61:527–544
Darroch J, Seneta E (1967) On quasi-stationary distributions in absorbing continuous-time finite Markov chains. J Appl Probab 4:192–196
Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. Wiley, New Jersey
Gardiner CG (2009) Stochastic methods: a handbook for the natural and social sciences. Springer, Berlin
Hagenaars TJ, Donnelly CA, Ferguson NM (2004) Spatial heterogeneity and the persistence of infectious diseases. J Theor Biol 229:349–359
Hethcote HW (1996) Modeling heterogeneous mixing in infectious disease dynamics. In: Isham V, Medley GFH (eds) Models for infectious human diseases. Cambridge University Press, Cambridge, pp 215–238
Hethcote HW, Yorke JA (1984) Gonorrhea transmission dynamics and control. Springer, Berlin
Jagers P (1975) Branching processes with biological applications. Wiley, London
Keeling MJ, Rohani P (2007) Modeling infectious diseases in humans and animals. Princeton University Press, Princeton
Laub AJ (2005) Matrix analysis for scientists and engineers. SIAM publications, Philadelphia
Lajmanovich A, Yorke JA (1976) A deterministic model for gonorrhea in a nonhomogeneous population. Math Biosci 28:221–236
Lefèvre C, Malice M-P (1988) Comparisons for carrier-borne epidemics in heterogeneous and homogeneous populations. J Appl Probab 25:663–674

Lindholm M (2008) On the time to extinction for a two-type version of Bartlett's epidemic model. Math Biosci 212:99–108
Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. Nature 438:355–359
Ma J, Earn DJD (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease. Bull Math Biol 68:679–702
Marschner IC (1992) The effect of preferential mixing on the growth of an epidemic. Math Biosci 109:39–67
Marshall AW, Olkin I, Arnold BC (2010) Inequalities: theory of majorization and its applications. Springer, Berlin
Meester R, Trapman P (2011) Bounding basic characteristics of spatial epidemics with a new percolation model. Adv Appl Probab 43:335–347
Metz JAJ (1978) The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. Acta Biother 27:75–123
Miller JC (2008) Bounding the size and probability of epidemics on networks. J Appl Probab 45:498–512
Mode CJ (1971) Multitype branching processes. Elsevier, New York
Nåsell I (1999) On the time to extinction in recurrent epidemics. J Roy Stat Soc B 61:309–330
Nåsell I (2002) Stochastic models of some endemic infections. Math Biosci 179:1–19
Nåsell I (2005) A new look at the critical community size for childhood infections. Theor Popul Biol 67:203–216
Neal PJ (2006) Stochastic and deterministic analysis of SIS household epidemics. Adv Appl Probab 38: 943–968
Nishiura H, Cook AR, Cowling BJ (2011) Assortativity and the probability of epidemic extinction: a case study of pandemic Influenza A (H1N1-2009). Interdiscip Perspect Infect Dis 2011, Article ID 194507
Nold A (1980) Heterogeneity in disease-transmission modelling. Math Biosci 52:227–240
Scalia-Tomba (1986) Asymptotic final size distribution of the multitype Reed–Frost process. J Appl Probab 23:563–584
Seneta E (1986) Non-negative matrices and Markov chains. Springer, New York
Vergu E, Busson H, Ezanno P (2010) Impact of the infection period distribution on the epidemic spread in a metapopulation model. PLoS ONE 5:e9371
Xiao Y, Clancy D, French NP, Bowers RG (2006) A semi-stochastic model for Salmonella infection in a multi-group herd. Math Biosci 200:214–233
Yates A, Antia R, Regoes RR (2006) How do pathogen evolution and host heterogeneity interact in disease emergence? Proc Roy Soc B 273:3075–3083