Influence of non-homogeneous mixing on final epidemic size in a meta-population model

Jingan Cui$^a$, Yanan Zhang$^a$ and Zhilan Feng$^{a,b}$

$^a$School of Science, Beijing University of Civil Engineering and Architecture, Beijing, People's Republic of China; $^b$Department of Mathematics, Purdue University, West Lafayette, IN, USA

ABSTRACT

In meta-population models for infectious diseases, the basic reproduction number $R_0$ can be as much as 70% larger in the case of preferential mixing than that in homogeneous mixing [J.W. Glasser, Z. Feng, S.B. Omer, P.J. Smith, and L.E. Rodewald, The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: A modelling study, Lancet ID 16 (2016), pp. 599–605. doi:10.1016/S1473-3099(16)00004-9]. This suggests that realistic mixing can be an important factor to consider in order for the models to provide a reliable assessment of intervention strategies. The influence of mixing is more significant when the population is highly heterogeneous. In this paper, another quantity, the final epidemic size ($F$) of an outbreak, is considered to examine the influence of mixing and population heterogeneity. Final size relation is derived for a meta-population model accounting for a general mixing. The results show that $F$ can be influenced by the pattern of mixing in a significant way. Another interesting finding is that, heterogeneity in various sub-population characteristics may have the opposite effect on $R_0$ and $F$.

1. Introduction

Epidemiological models can be used to identify key factors affecting disease outbreaks and to evaluate intervention strategies. However, because the model predictions and evaluations can be influenced by assumptions made to simplify the analysis, it is critical to understand how certain simplifying assumptions may affect the model outcomes, particularly if the models are used to inform public health policy-making.

One of the most commonly used assumptions in epidemiological models is the homogeneous, or random, mixing in the whole population, which allows the model to ignore heterogeneity in various sub-population factors including activity (pertinent to disease transmission), susceptibility, preference in contact, group size, etc. Although these simplifying assumptions are reasonable in many cases depending on the questions to be studied, there are cases where these heterogeneous factors cannot be ignored (see e.g. [11–13]).
Figure 1. Plot of the final size relation given in (18) for the case of a homogeneous population with random mixing. It shows that $F$ is an increasing function of $R_0$.

When epidemic models are used to evaluate strategies for disease control and prevention, the quantities used often include the basic or effective reproduction number ($R_0$ or $R_e$), the final epidemic size ($F$), and the peak size of an outbreak ($P$). Andreasen [2] studied the role of heterogeneity in susceptibility and infectiousness on final epidemic size for the case of proportionate mixing and discussed the relationship between $F$ and $R_0$. Tildesley and Keeling [19] considered a model with a spatial structure to discuss the issue of using $R_0$ as a predictor of final epidemic size for the case of foot-and-mouth disease in the UK. Results concerning final epidemic sizes can also be found in other studies including [1,3,4,10,15,16]. None of these studies considered the influence of heterogeneity on the final size or the relationship between $F$ and $R_0$ in a meta-population with preferential mixing.

In this paper, we consider an epidemic meta-population model that includes a general mixing function (which includes proportionate, preferential, and homogeneous mixing as special cases). We examine how heterogeneity in various sub-population factors may influence $R_0$, $F$, and $P$, and how preferential mixing may affect the relationship between $F$ and $R_0$. Our results show that models that do not account for heterogeneous mixing may produce biased or even incorrect information.

From previous epidemic models with homogeneous mixing, a general understanding is that $F$ and $R_0$ are positively related (see Figure 1). We show in this paper that, when preferential mixing is considered in the model with heterogeneity in activity and/or preference, it is possible that $F$ and $R_0$ are negatively related (see Figures 4 and 5). Moreover, when vaccination is considered, the homogeneous coverage may not be the most effective strategy (see Figures 7 and 8).

The paper is organized as follows. In Section 2, we present the model with $n$ sub-populations and non-homogeneous mixing. Derivations for the reproduction number and a final epidemic size relation are also included in this section. Section 3 focuses on the investigation of the effects of heterogeneity in various sub-population factors and mixing on the final epidemic size and epidemic peak during an outbreak. Two examples of using
the meta-population model to evaluate intervention strategies are also presented in this
section. Section 4 includes discussions of the results.

2. The model and final epidemic size

Consider a meta-population with \(n\) sub-populations or groups with sizes \(N_i (1 \leq i \leq n)\). The population in each group \(i\) is divided into three epidemiological classes denoted by \(S_i\) (susceptible), \(I_i\) (infectious), and \(R_i\) (recovered and immune due to infection). Thus, \(N_i = S_i + I_i + R_i (1 \leq i \leq n)\). Because we are concerned with the final epidemic size of a single disease outbreak, births and deaths are ignored. The epidemic model with \(n\) sub-groups consists of the following ordinary differential equations:

\[
\begin{align*}
S'_i &= -a_i \beta_i S_i \sum_{j=1}^{n} c_{ij} \frac{I_j}{N_j}, \\
I'_i &= a_i \beta_i S_i \sum_{j=1}^{n} c_{ij} \frac{I_j}{N_j} - \gamma I_i, \\
R'_i &= \gamma I_i, \quad 1 \leq i \leq n,
\end{align*}
\]

with initial conditions:

\[
S_i(0) = N_i - I_{i0}, \quad I_i(0) = I_{i0}, \quad R_i(0) = 0, \quad 1 \leq i \leq n. 
\]

In model (1), \(a_i\) is the per capita contact rate (referred to as activity), \(\beta_i\) is the probability that a susceptible person in group \(i\) becomes infected upon contacting an infectious person, \(\gamma\) denotes the per-capita recovery rate. The contacts between sub-groups are described by the mixing matrix \(C = (c_{ij})\), where \(c_{ij}\) is the proportion of the \(i\)th sub-group’s contacts that is with members of the \(j\)th group, and \(I_j/N_j\) is the probability that a randomly encountered member of group \(j\) is infectious. As pointed out in Busenberg and Castillo-Chavez [6], a mixing function needs to satisfy the following basic conditions:

\[
\begin{align*}
(1) & \quad c_{ij} \geq 0, \\
(2) & \quad \sum_{j=1}^{n} c_{ij} = 1, \\
(3) & \quad a_i N_i c_{ij} = a_j N_j c_{ji}, \quad 1 \leq i, j \leq n.
\end{align*}
\]

One of the commonly used mixing functions, as originally considered by Nold [17] and later extended by Jacquez et al. [14], is the preferential mixing with the form:

\[
c_{ij} = \epsilon_i \delta_{ij} + (1 - \epsilon_i) f_j, \quad \text{where} \quad f_j = \frac{(1 - \epsilon_j) a_j N_j}{\sum_k (1 - \epsilon_k) a_k N_k}.
\]

In the preferential mixing (4), \(\epsilon_i\) denotes the fraction of contacts reserved for one’s own group (referred to as preferences), and \(\delta_{ij}\) is the Kronecker delta (1 when \(i = j\) and 0 otherwise). The function \(f_j\) describes mixing that is random, i.e. proportional to unreserved contacts, \((1 - \epsilon_j) a_j N_j\). All parameters are non-negative.
The isolated basic reproduction number for group $i$ is
\[ R_{0i} = \frac{a_i \beta_i}{\gamma}, \quad 1 \leq i \leq n. \] (5)

The next generation matrix is
\[
K = \begin{pmatrix}
R_{01}c_{11} & R_{01}c_{12} & \cdots & R_{01}c_{1n} \\
R_{02}c_{21} & R_{02}c_{22} & \cdots & R_{02}c_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
R_{0n}c_{n1} & R_{0n}c_{n2} & \cdots & R_{0n}c_{nn}
\end{pmatrix} = \text{diag}(R_{01}, R_{02}, \ldots, R_{0n})C, \tag{6}
\]

where $C = (c_{ij})$ is the mixing matrix. The basic reproduction number for the meta-population is
\[ R_0 = \rho(K), \tag{7} \]
when $\rho(K)$ denotes the dominant eigenvalue of $K$. In general, we can only compute $\rho(K)$ numerically when $n > 2$. Two special cases are when the mixing is proportionate (i.e. when $\epsilon_i = 0$) or isolated (i.e. when $\epsilon_i = 1$) for all $i$. In these cases, $\rho(K)$ is given by either the trace of $K$ or the maximum of $R_{0i}$. That is, the analytic expressions are:

\[
R_0 = \sum_{i=1}^{n} R_{0i}c_{ii} \quad \text{when} \quad \epsilon_i = 0, \quad 1 \leq i \leq n, \\
R_0 = \max_{1 \leq i \leq n} \{R_{0i}\} \quad \text{when} \quad \epsilon_i = 1, \quad 1 \leq i \leq n. \tag{8}
\]

If vaccines are available before the epidemic starts, a certain level of population immunity can be achieved via vaccination. Let $p_i$ denote the immunity of sub-population $i$ at time $t=0 (0 \leq p_i \leq 1)$. Then the disease dynamics can be modelled by the equations in (1) with modified initial conditions:
\[
S_i(0) = (1 - p_i)(N_i - I_{i0}), \quad I_i(0) = I_{i0}, \quad R_i(0) = p_i(N_i - I_{i0}), \quad 1 \leq i \leq n. \tag{9}
\]

In this case, the isolated effective reproduction number for group $i$ is $R_{ei} = (1 - p_i)R_{0i}$, and the next generation matrix is $K_e = \text{diag}(R_{c1}, R_{c2}, \ldots, R_{cn})C$. The effective reproduction number for the meta-population is $R_e = \rho(K_e)$.

If the effect of another type of control or intervention programme is to reduce activities $a_i$ and/or probabilities of infection $\beta_i$, let $R_{ei}$ denote the isolated effective reproduction number for group $i$ (same formula as for $R_{0i}$ in (5) but with the new values of $a_i$ and $\beta_i$), then the effective reproduction number for the meta-population $R_e$ is given by the dominant eigenvalue of the next generation matrix $K_e$, which is the same as $K$ in (6) but with $R_{0i}$ being replaced by $R_{ei}$.

We now use (1) with initial condition (2) to derive a relation for the final epidemic size. Note from the $S_i'$ and $I_i'$ equations in (1) that
\[ (S_i + I_i)' = -\gamma I_i, \quad 1 \leq i \leq n. \tag{10} \]

Because $S_i(t) \geq 0$ and $I_i(t) \geq 0$ for all $t > 0$, $S_i(t) + I_i(t)$ is a non-negative decreasing function and hence has a limit as $t \to \infty$ and $\int_{t=0}^{\infty} I_i(t)dt = 0$. It follows that $I_i(t) \to 0$ as
$t \to \infty$ (see [5], Section 2.4). Thus, $S_i(\infty) = \lim_{t \to \infty} S(t)$ exists. Using $I_i(\infty) = 0$ and $S_i(0) + I_i(0) = N_i$, and from (10) we get

$$N_i - S_i(\infty) = \gamma \int_0^\infty I_i(t) \, dt, \quad 1 \leq i \leq n. \quad (11)$$

Dividing the $S_i'$ equation in (1) by $S_i(t)$ on both sides, we get

$$\frac{dS_i}{S_i} = -a_i \beta_i \sum_{j=1}^n c_{ij} \frac{I_j(t)}{N_j} \, dt, \quad 1 \leq i \leq n. \quad (12)$$

Integration of both sides of (12) from 0 to $\infty$ yields

$$\log \frac{S_i(0)}{S_i(\infty)} = \sum_{j=1}^n a_i \beta_i c_{ij} \int_0^\infty \frac{I_j(t)}{N_j} \, dt. \quad (13)$$

From (11), (13), and $R_{0i} = a_i \beta_i / \gamma$, we obtain

$$\log \frac{S_i(0)}{S_i(\infty)} = \sum_{j=1}^n R_{0i} c_{ij} \left( 1 - \frac{S_j(\infty)}{N_j} \right), \quad 1 \leq i \leq n. \quad (14)$$

Let $Z_i = S_i(0) - S_i(\infty)$, then

$$N_i - S_i(\infty) = N_i - (S_i(0) - Z_i) = I_i(0) + Z_i, \quad 1 \leq i \leq n. \quad (15)$$

Thus, from (14) and (15), we know that $Z_i$ can be determined by a solution of the following set of equations:

$$Z_i = S_i(0) \left( 1 - \exp \left[ -\sum_{j=1}^n R_{0i} c_{ij} \frac{Z_j + I_j(0)}{N_j} \right] \right), \quad 1 \leq i \leq n. \quad (16)$$

Let $N = \sum_{i=1}^n N_i$ denote the total population size, and let $F_i = Z_i/N$ denote the fraction of infected in group $i$ (over the total $N$). We define the final epidemic size $F$ by

$$F = \sum_{i=1}^n F_i = \sum_{i=1}^n \frac{Z_i}{N}. \quad (17)$$

In the special case when the population is considered homogeneous with random mixing, i.e. the parameter values for all sub-groups are identical (e.g. $a_i = a$, $\beta_i = \beta$, $N_i = N/n$, $S_i(0) = S(0)/N$, $I_i(0) = I(0)/N$ are all independent of $i$ and $c_i = 0$), the final size relation simplifies to

$$Z = S(0) \left( 1 - \exp \left[ -R_0 \frac{Z + I(0)}{N} \right] \right), \quad (18)$$

where $Z = \sum_{i=1}^n Z_i$ and $R_0 = R_{0i} = a \beta / \gamma$. If we vary $R_0$ by varying $\beta$ while keeping $a$ and $\gamma$ fixed, the final size relation (18) is illustrated in Figure 1, which shows that $F$ is an increasing function of $R_0$. 


As shown in Feng et al. [11] and Glasser et al. [13], in meta-population models with preferential mixing, even when $\beta$ and the mean activity $\sum_{i=1}^{n} a_i/n$ are constant (with all other parameter values fixed), $R_0$ may vary with the level of heterogeneity in $a_i$ and preference $\epsilon_i$. Therefore, the dependence of $F$ on $R_0$ may not be increasing as exhibited in Figure 1. In the following sections, we will investigate how population heterogeneity in models with non-random mixing may influence the final size $F$, the relationship between $F$ and $R_0$, and model evaluations of intervention programmes.

3. Effect of heterogeneity on final size and the relationship between $F$ and $R_0$

A general understanding based on epidemic models for a homogeneous population with random mixing is that $F$ is an increasing function of $R_0$, as shown in Figure 1. Recall that the increase in $R_0$ is caused by increasing $\beta$. When heterogeneity such as activity ($a_i$), preference of contacts ($\epsilon_i$), and sub-population size ($N_i$) are considered in a meta-population model with preferential mixing, $R_0$ may vary by a significantly magnitude even when $\beta_i$ are fixed (see e.g. [13]). In addition, it has been shown in Poghotanyan et al. [18] that, among a large class of mixing matrices ($c_{ij}$) including the preferential mixing in (4), $R_0 = \rho(K)$ is bounded below and above by the basic reproduction numbers corresponding to proportionate mixing ($\epsilon_i = 0$ for all $i$) and isolated mixing ($\epsilon_i = 1$ for all $i$), respectively, i.e.

$$\sum_{i=1}^{n} R_{0i}w_i \leq R_0 \leq \max_{1 \leq i \leq n} R_{0i}, \quad (19)$$

where

$$R_{0i} = \frac{a_i\beta_i}{\gamma}, \quad w_i = \frac{a_iN_i}{\sum_{j=1}^{n} a_jN_j}. \quad \text{(20)}$$

It is clear from (19) that we can keep $\beta_i$ fixed and change $R_0$ by varying $a_i$ and $N_i$ while keeping the same mean activity and mean population size. Particularly, different $\epsilon_i$ values may also affect how $R_0$ and $F$ may be influenced by heterogeneity in $a_i$ or $N_i$. Thus, the relationship between $F$ and $R_0$ may be more complex in meta-population models than the one illustrated in Figure 1. This will be explored next.

Results in this section are for the case of $n = 2$ groups, but they can be extended to the case of $n > 2$. Although the equations for $Z_i$ in (16) cannot be solved analytically, we can explore the dependence of $Z_i$ on other parameters numerically. Rewrite (16) as

$$F(Z_1, Z_2) = 1, \quad G(Z_1, Z_2) = 1, \quad \text{(20)}$$

where

$$F(Z_1, Z_2) = \frac{S_1(0)}{Z_1} \left( 1 - \exp \left[ -\sum_{j=1}^{2} R_{01}c_{1j} \frac{Z_j + I_j(0)}{N_j} \right] \right),$$

$$G(Z_1, Z_2) = \frac{S_2(0)}{Z_2} \left( 1 - \exp \left[ -\sum_{j=1}^{2} R_{02}c_{2j} \frac{Z_j + I_j(0)}{N_j} \right] \right). \quad \text{(21)}$$
Figure 2. (a) The solution $(Z_1, Z_2)$ of the equations (21) is shown as the intersection of the curves $F(Z_1, Z_2) = 1$ and $G(Z_1, Z_2) = 1$. (b) Cumulative infections during an outbreak in the two groups and the total. The plots are for the case of proportionate mixing (i.e. $\epsilon_1 = \epsilon_2 = 0$) with $a_1 = 14, a_2 = 6$, and $N_1 = N_2 = 5000$.

Figure 2(a) illustrates the final size as an intersection of the curves determined by $F = 1$ and $G = 1$.

For demonstration purposes, we used the following parameter values in Figure 2 (the time unit is days): $\beta_1 = \beta_2 = 0.035$, $1/\gamma = 5$ (infectious period is 5 days), $a_1 + a_2 = 20$ with $a_1 = 14$ and $a_2 = 6$ (heterogeneous activity with the mean activity being 10), $\epsilon_1 = \epsilon_2 = 0$ (proportionate mixing), $N_1 + N_2 = 10^4$, $\sigma = N_1/N = 0.5$ (homogeneous sub-population size with $N_1 = N_2 = 5000$), and $I_1(0) = I_2(0) = 1$. For these parameter values, the isolated reproduction numbers for sub-populations 1 and 2 are $R_{01} = 2.45$ and $R_{02} = 1.05$, respectively. The overall effective reproduction number for the metapopulation is $R_0 = 2.03$. From the equations in (16), we have $Z_1 = 4210$ and $Z_2 = 2730$. These values are also consistent with the cumulative infections in groups 1 and 2 presented in Figure 2(b) from numerical simulations of system (1) with the initial condition (2).

3.1. Influence of heterogeneity and non-homogeneous mixing

To explore the effect of mixing on model predictions, we consider preferential mixing (4) and examine heterogeneity in various characteristics affecting the mixing function including activity $a_i$, group size $N_i$, and preference $\epsilon_i$. For demonstration purposes, we fix the same parameter values $\beta_1 = \beta_2 = 0.035$, $N_1 = N_2 = 5000$, and the mean activity $(a_1 + a_2)/2 = 10$. Values for other parameters may vary depending on applications.

3.1.1. Heterogeneity in activity $a_i$

Consider first the case of proportionate mixing ($\epsilon_1 = \epsilon_2 = 0$). The effect of heterogeneity and variability in activity $a_i$ on $F$ is illustrated in Figure 3, which shows the final sizes for three sets of activities $(a_1, a_2)$. The $a_i$ values used in the three rows from top to bottom are
(with increasing variability)

Top : \((a_1, a_2) = (10, 10)\), Middle : \((a_1, a_2) = (14, 6)\), Bottom : \((a_1, a_2) = (18, 2)\).

The left column shows the values of \(Z_i\) determined by equations in (16), as the intersection of \(F = 1\) and \(G = 1\) (marked by the dot). The right column shows the epidemic curves based on the simulations of the system (1) with initial condition (2). Notice that the number of individuals in the \(R\) class represents the cumulative infections at anytime.

In this case, the basic reproduction numbers are \(R_0 = 1.75, 2.03, \text{ and } 2.87\), respectively, while the final sizes are \(F = 0.713, 0.695, \text{ and } 0.599\), respectively. We observe the following changes as the variability in activity increases

- \(F\) decreases;
- \(R_0\) increases;
- \(P\) increases;
- the time to epidemic peak decreases.

To make it more transparent how variability in \(a_i\) affects \(R_0, F\), and their relationship, Figure 4 plots \(F\) and \(R_0\) for several sets of activity \((a_1, a_2)\) that have the same mean. It shows that the homogeneous activity \((a_1 = a_2 = 10)\) corresponds to the smallest \(R_0 = 1.75\) and largest \(F = 0.713\), while the heterogeneous activity with the highest variability \((a_1 = 18, a_2 = 2)\) corresponds to the largest \(R_0 = 2.87\) and smallest \(F = 0.599\). This provides an example that the final size does not vary with \(R_0\) positively. That is, the usual understanding based on models with homogeneous assumption that \(F\) increases with \(R_0\) (see Figure 1) may not be true when more realistic mixing functions are incorporated in meta-population models. Such results may have important implications for public health policy-making. For example, when a control programme is aimed at reducing \(R_0\), it should also check whether or not it may lead to an increased final epidemic size.

### 3.1.2. Preference level in mixing

The examples in Section 3.1.1 are for proportionate mixing (i.e. \(\epsilon_i = 0\)). A more realistic mixing, in general, is when \(\epsilon_i > 0\), representing a preference of contacts with others in the same sub-group. Heterogeneity in factors such as \(a_i\) and \(\epsilon_i\) may have an important influence in \(R_0\) and \(F\), as suggested in (19). We now explore the effect of variability in \(a_i\) and the joint effect of \(a_i\) and \(\epsilon_i\) on \(F\) and \(R_0\). Consider first the simpler case when \(\epsilon_1 = \epsilon_2\). Figure 5 provides a similar information as in Figure 4 but includes cases of \(\epsilon_i > 0\).

We observe in Figure 5 that (i) for each fixed \(\epsilon_i\) value, \(R_0\) increases but \(F\) decreases with the variability in \(a_i\), and the changes are more dramatic for higher \(\epsilon_i\); and (ii) for each fixed heterogeneous pair \((a_1, a_2)\), \(R_0\) increases and \(F\) decreases with \(\epsilon_i\). Because the bar charts in Figure 5(a,b) are for the same sets of \(a_i\) and \(\epsilon_i\), it shows again that \(F\) and \(R_0\) are related negatively when heterogeneity in \(a_i\) and/or \(\epsilon_i\) is varied. This illustrates again an opposite relationship between \(F\) and \(R_0\) to that shown in Figure 1.

### 3.1.3. Sub-population size \(N_i\)

The influence of heterogeneity in \(N_i\) can be examined by fixing all other parameter values. For example, Figure 6 shows the case when \(a_1 = 15, a_2 = 5, \epsilon_i = 0, \text{ and all other parameter values except } N_i \text{ are the same as before. The total } N = N_1 + N_2 = 10^4 \text{ is fixed but}
Figure 3. Comparison of $\mathcal{F}$, $\mathcal{R}_0$, and $\mathcal{P}$ when activity is homogeneous (top) and heterogeneous with increasing variance (middle and bottom). Mixing is proportional ($\epsilon_1 = \epsilon_2 = 0$). The left column shows the final size (marked by the point of intersection of $F = 1$ and $G = 1$ as the solution of (20)). The right column shows the the total sizes of $S$, $I$, and $R$ as the solution of model (1) with initial condition (2).
Figure 4. Plots of $R_0$ and $F$ for contact rates $a_1$ and $a_2$ with the same mean but different variability. When $(a_1, a_2)$ changes from (10,10) to (18,2), $R_0$ increases from 1.75 to 2.87 whereas $F$ decreases from 0.713 to 0.599.

Figure 5. Plots of $F$ and $R_0$ for various preference level $\epsilon_i$ and activity variability in $a_i$ (with equal mean $(a_1 + a_2)/2 = 10$), showing that as $\epsilon_i$ or variability in $a_i$ increases, $F$ decreases while $R_0$ increases.

$\sigma = N_1/N$ varies. It shows that both $F$ and $R_0$ increase with $\sigma$, and thus, in this case, $F$ and $R_0$ are positively related. The values of $F$ and $R_0$ for the homogeneous case ($\sigma = 0.5$ and $a_1 = a_2 = 10$) are also shown by the dot-dashed and dashed lines, respectively. Thus, heterogeneity in characteristics including $a_i$ and $N_i$ can either increase or decrease $F$ and $R_0$.

3.2. Examples

Because population heterogeneity in factors such as $a_i$ or $\epsilon_i$ may affect the final size and the basic reproduction number in significant ways, we investigate in this section how the model evaluations of intervention strategies may be affected. Two examples are discussed below.

3.2.1. Heterogeneity in vaccine coverage $p_i$

Consider the model equations in (1) with the initial condition in (9). Assume that the vaccine efficacy is 100% so that the vaccination coverage in group $i$ is the same as the
Figure 6. Changes in $F$ and $R_0$ with the sub-population sizes $N_i (N_1 = \sigma N$ and $N_2 = (1 - \sigma)N$ for fixed $N)$. The dot-dashed line and dashed line are values of $F$ and $R_0$, respectively, for homogeneous case in which $\sigma = 0.5$ and $a_1 = a_2 = 10$.

Figure 7. Similar to Figure 6, but this shows the dependence of $F$ and $R_e$ on heterogeneity in sub-population immunity $p_i$ with $p_1 + p_2 = 1.32$. The dot-dashed line and dashed line are values of $F^{(hom)}$ and $R_e^{(hom)}$, respectively, for the homogeneous case with $a_1 = a_2 = 10$ and $p_1 = p_2 = 0.66$.

In Figure 7, $F$ and $R_e$ for several sets of heterogeneous coverage with $p_1 + p_2 = 1.32$ are plotted. It is interesting to notice that for $p_1 = 0.55$, both $F$ and $R_e$ are smaller than immunity $p_i$, and the total number of vaccine doses is $p_1 N_1 + p_2 N_2$. When $N_1 = N_2$ with fixed total size $N = N_1 + N_2$, fixing the total number of vaccine doses is equivalent to fixing the value of $p = p_1 + p_2$. Thus, we can examine the effect of heterogeneity in vaccine allocation $(p_1, p_2)$ on $F$ and $R_e$ for fixed $p$. Choose parameter values to be the same as before except that $\epsilon_1 = \epsilon_2 = 0.6$, $\beta = 0.05$, $a_1 = 8$, $a_2 = 12$, and $1/\gamma = 7$. For these parameter values, the basic reproduction numbers for the two groups are $R_{01} = 2.8$ and $R_{02} = 4.2$ with $R_0 = 3.8$. For demonstration purpose, let $p_1 + p_2 = 1.32$. For the homogeneous case ($a_1 = a_2 = 10$ and $p_1 = p_2 = 0.66$), the final size is $F^{(hom)} = 0.1$ and the effective reproduction number is $R_e^{(hom)} = 1.2$, which are labelled in Figure 7 by the dot-dashed and dashed lines, respectively.
Figure 8. Numerical simulations of system (1) with initial conditions (9) showing (a) the cumulative infections and (b) epidemic curves for various vaccine allocation \((p_1, p_2)\) with \(p_1 + p_2 = 1.32\) (the values of \(p_1\) for different curves are labelled in the legend).

their corresponding values in the homogeneous case, \(F^{(\text{hom})}\) and \(R_e^{(\text{hom})}\) (shown by the dot-dashed and dashed lines). But for other \(p_1\) values, \(F > F^{(\text{hom})}\) and \(R_e > R_e^{(\text{hom})}\). This shows that, for a fixed amount of vaccine doses, although the homogeneous coverage \(p_1 = p_2\) can provide a relatively more effective allocation than many heterogeneous vaccine allocations in reducing \(R_e\), it may not be the best choice. This result is consistent with those found in Feng et al. [11], in which it is shown that the optimal vaccine allocation can be determined using a gradient approach based on a Lagrange optimization problem. The results in Feng et al. [11] is formulated in terms of minimizing the effective reproduction number.

In addition to the final size, the peak size \((P)\) and time to peak \((T)\) are also sensitive to the heterogeneity in \(p_i\). This is illustrated in Figure 8 (the \(p_1\) values for different curves are labelled in the legend). All parameter values are the same as in Figure 7.

We observe in Figure 8(a) that, as \(p_1\) increases from 0.35 to 0.55, the number of cumulative infections decreases (see the solid curves). However, as \(p_1\) continues to increase to 0.65 until 0.85, the number of cumulative infections becomes increasing (the dot-dashed to dashed curves). Similarly in Figure 8(b), \(P\) and \(T\) decrease from \(p_1 = 0.35\) to 0.55 (solid curves) but increase from \(p_1 = 0.65\) to 0.85 (dashed curves).

3.2.2. The 2011 acute haemorrhagic conjunctivitis outbreak in China

In this section, we examine how population heterogeneity in various characteristics may affect model evaluations of control and intervention programmes. We demonstrate this by using an example based on the 2011 outbreak of AHC in Changsha, China. Studies on the efficacy of quarantine during this outbreak is presented in [7,8]. By fitting a homogeneous SIR model to the data from a case report for a school population, they obtained an estimate for the transmission coefficient \(\beta\). Based on this, we adopt \(\beta = 0.05\) as a baseline value for the probability of infection on contact with a baseline value for activity to be \(a = 15\). The
infectious period for this AHC outbreak in China is about 7–10 days, so we choose $1/\gamma$ to be 8 days. Under these parameter values, the baseline value for the reproduction number is $R_0 = \beta a / \gamma = 6$. This is much higher than the value of the reproduction number $R^*_0$ presented in [9] for the 2003 AHC outbreak in Mexico, where $R^*_0 = 2.64$. The main reason for this difference is in the value of infectious period, which are about 8 and 3 days for the outbreaks in China and Mexico, respectively.

Consider the case when the population is divided into two groups, and individuals in group 1 have a lower risk of being infected due to a reduced contact rate, better health habit (e.g. washing hands to increase protection against infection) or other behavioural changes as suggested by public health officials. Thus, let $\beta_2 = 0.05$ (baseline value) and $N = N_1 + N_2 = 10^3$ (e.g. a school population) be fixed. Assume that individuals in the low risk group have a reduced activity with $a_1 = 5 (< a_2 = 15)$, and let

$$\beta_1 = \delta \beta_2 \quad \text{with} \quad \delta < 1, \quad \sigma = \frac{N_1}{N} \quad \text{with} \quad 0 \leq \sigma \leq 1.$$ 

The mixing is assumed to be proportionate. We will examine the changes in $F$ due to interventions aimed at decreasing $\delta$ and increasing $\sigma$. The simulation results are presented in Figure 9, which shows the model evaluations for the effect of reducing $\delta$ and increasing $\sigma$ on the final epidemic size $F$.

We observe in Figure 9 that increasing $\sigma$ can be effective in reducing the final size only if $\delta$ is sufficiently small. For example, if 60% of individuals are in group 1 ($\sigma = 0.6$) and the reduction factor for the risk of infection is $\delta = 0.6$, the final size is 0.79. The reason for the high percentage of cumulative infections is because of the high value of the basic reproduction number $R_{20} = 6$ for group 2. In fact, in the absence of control, almost the entire population will be infected during the outbreak. For the same value of $\delta = 0.6$, even when $\sigma = N_1/N = 1$, the final size will still be as high as 32%. The only possibility to have the final size below 10% is when $\delta \leq 0.2$ and $\sigma \geq 0.9$. However, as long as $\sigma < 1$ (or $N_2/N > 0$), there will be a significant number of infections.
Figure 10. The surface is a plot of the final size $F$ versus $\delta$ and $\sigma$, and the plane corresponds to $F = 0.2$. The intersection curve identifies the region for $(\delta, \sigma)$ (where the surface is below the plane) in which the final size is below 20%.

Figure 10 plots the surface of the final size as a function of $\delta$ and $\sigma$, which is obtained by connecting the values of the bar chart. We can identify the region in the $(\delta, \sigma)$ plane in which the final size is below a certain level. For example, the intersection curve of the $F$ surface with the plane with $F = 0.2$ determines the proportion $\sigma$ of individuals in group 1 (for a given $\delta$ value), above which the final size is below 20%.

4. Discussion

The results presented in this paper provide another example that incorporation of population heterogeneity and non-homogeneous mixing in epidemic models can generate outcomes that are qualitatively different from that based on models for populations with homogeneous mixing. Particularly, for two of the most important quantities, reproduction number ($R_0$ or $R_e$) and final epidemic size ($F$), we demonstrated how they may be influenced by population heterogeneity in various characteristics and mixing by comparing the results from a meta-population model with and without heterogeneity and non-homogeneous mixing. The final epidemic size relation and the reproduction number derived from the meta-population model (1) with a general mixing matrix $C = (c_{ij})$ allows us to conduct various comparisons (see Sections 3.1.1–3.1.3). We also presented two examples to illustrate the effect of heterogeneity and mixing in the model evaluation of intervention strategies (Section 3.2). An interesting result is that, although homogeneous vaccine coverage may produce higher reduction in $F$ than many heterogeneous combinations of coverage, it may not be the most effective option (see Figures 7 and 8) if the objective is to maximally reduce $F$ for a given number of vaccine doses.

Another interesting finding of this study is the correlation between $F$ and $R_0$. In models with homogeneous mixing, $F$ and $R_0$ are positively related, as shown in Figure 1. However,
this relation may be reversed when heterogeneity is considered in a meta-population model with non-homogeneous mixing, as illustrated in Figures 4 and 5. Particularly, an increase of variability in activity \((a_1 - a_2)\) or preference level \((\epsilon_i)\) can increase \(R_0\) but decrease \(F\). Thus, in this case, \(F\) and \(R_0\) are negatively related. The finding that \(R_0\) can be increased by population heterogeneity in various factors is also supported by the fact pointed out by Tilman and Kareiva [20] that \(R_0\) is higher in situations with significant spatial or heterogeneity features.

The final size relation derived in this paper from a meta-population model, which considers \(n\) sub-populations, multiple types of heterogeneity, and a general mixing function including preferential mixing, provides a useful evaluation tool for intervention strategies. It also raised the warning issue that, although in many models control measures can be evaluated using the effective reproduction number \(R_e\), it should not be the only quantity to consider, especially if a programme reduces \(R_e\) but increases \(F\).

**Acknowledgements**

We thank John Glasser for reading the manuscript and providing many constructive suggestions that improved the presentation of this paper. We thank also the Reviewer for helpful comments.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported partially by the BUCEA Post Graduate Innovation Projects (PG2018094 and PG2018097).

**References**

[1] R.M. Anderson and R.M. May, *Infectious Diseases of Humans: Dynamics and Control*, Cambridge University Press, New York, 1991.

[2] V. Andreasen, *The final size of an epidemic and its relation to the basic reproduction number*, Bull. Math. Biol. 73 (2011), pp. 2305–2321.

[3] J. Arino, F. Brauer, P. van den Driessche, J. Watmough, and J. Wu, *A final size relation for epidemic models*, Math. Biosci. Eng. 4 (2007), pp. 159–175.

[4] F. Brauer, *Epidemic models with heterogeneous mixing and treatment*, Bull. Math. Biol. 70 (2008), pp. 1869–1885.

[5] F. Brauer, C. Castillo-Chavez, and Z. Feng, *Mathematical Models in Epidemiology*, Springer, 2018, to appear.

[6] S. Busenberg and C. Castillo-Chavez, *A general solution of the problem of mixing of subpopulations and its application to risk-and age-structured epidemic models for the spread of AIDS*, IMA J. Math. Appl. Med. Biol. 8 (1991), pp. 1–29.

[7] T. Chen, R. Liu, and Q. Wang, *Application of susceptible-infected-recovered model in dealing with an outbreak of acute hemorrhagic conjunctivitis on one school*, Chin. J. Epidemiol. 32 (2011), pp. 723–726.

[8] T. Chen and R. Liu, *Study of the efficacy of quarantine during outbreaks of acute hemorrhagic conjunctivitis at schools through the susceptible-infective-quarantine-removal-model*, Chin. J. Epidemiol. 34 (2013), pp. 75–79.

[9] G. Chowell, E. Shim, F. Brauer, P. Diaz-Dueñas, J.M. Hyman, and C. Castillo-Chavez, *Modelling the transmission dynamics of acute haemorrhagic conjunctivitis: Application to the 2003 outbreak in Mexico*, Stat. Med. 25 (2006), pp. 1840–1857.
[10] Z. Feng, *Final and peak epidemic sizes for SEIR models with quarantine and isolation*, Math. Biosci. Eng. 4 (2007), pp. 675–686.
[11] Z. Feng, A.N. Hill, P.J. Smith, and J.W. Glasser, *An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing*, J. Theor. Biol. 386 (2015), pp. 177–187.
[12] Z. Feng, A.N. Hill, A.T. Curns, and J.W. Glasser, *Evaluating targeted interventions via metapopulation models with multi-level mixing*, Math. Biosci. 287 (2017), pp. 93–104.
[13] J.W. Glasser, Z. Feng, S.B. Omer, P.J. Smith, and L.E. Rodewald, *The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: A modelling study*, Lancet ID 16 (2016), pp. 599–605. doi:10.1016/S1473-3099(16)00004-9.
[14] J.A. Jacquez, C.P. Simon, J. Koopman, L. Sattenspiel, and T. Perry, *Modeling and analyzing HIV transmission: The effect of contact patterns*, Math. Biosci. 92 (1988), pp. 119–199.
[15] J. Ma and D.J. Earn, *Generality of the final size formula for an epidemic of a newly invading infectious disease*, Bull. Math. Biol. 68 (2006), pp. 679–702.
[16] P. Magal, O. Seydi, and G. Webb, *Final size of an epidemic for a two-group SIR model*, SIAM J. Appl. Math. 76 (2016), pp. 2042–2059.
[17] A. Nold, *Heterogeneity in disease-transmission modeling*, Math. Biosci. 52 (1980), pp. 227–240.
[18] G. Poghotanyan, Z. Feng, J.W. Glasser, and A.N. Hill, *Constrained minimization problems for the reproduction number in meta-population models*, J. Math. Biol. (2018), pp. 1–37. doi:10.1007/s00285-018-1216-z.
[19] M.J. Tildesley and M.J. Keeling, *Is $R_0$ a good predictor of final epidemic size: Foot-and-mouth disease in the UK*, J. Theor. Biol. 258 (2009), pp. 623–629.
[20] D. Tilman and P. Kareiva, *Spatial Ecology*, Princeton University Press, Princeton, NJ, 1998.