How population heterogeneity in susceptibility and infectivity influences epidemic dynamics

R.I. Hickson a,*, M.G. Roberts b

a School of Mathematical & Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia
b Institute of Natural & Mathematical Sciences, New Zealand Institute for Advanced Study and Infectious Disease Research Centre, Massey University, Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand

HIGHLIGHTS

• We present a new method for analysing epidemic models.
• We consider heterogeneous distributions of susceptibility and infectivity.
• We solve the generalisation of the final size equation.
• We consider the effects of pre-epidemic immunity on the mortality distribution.
• We find the smallest final size and mortality if “children” are vaccinated.

ARTICLE INFO

Article history:
Received 8 July 2013
Received in revised form
22 November 2013
Accepted 8 January 2014
Available online 17 January 2014

Keywords:
Infectious diseases
Heterogeneous population
Final size
Epidemic
Vaccination

ABSTRACT

An important concern in public health is what population group should be prioritised for vaccination. To this end, we present an epidemic model with arbitrary initial distributions for population susceptibility, and corresponding infectivity distributions. We consider four scenarios: first, a population with heterogeneous susceptibility with a uniform distribution, but homogeneous infectivity; second, a heterogeneously susceptible population with linear heterogeneous infectivity functions, where the most susceptible are either the most or least infectious; third, a bimodal distribution for susceptibility, with all combinations of infectivity functions; finally, we consider the effects of additional pre-epidemic immunity, ostensibly through vaccination, on the epidemic dynamics. For a seasonal influenza-like infectious disease, we find the smallest final size and overall number of deaths due to the epidemic to occur if the most susceptible are vaccinated, corresponding to targeting children.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Annual epidemics of influenza occur in all temperate regions of the world (Finkelman et al., 2007). Questions of interest to public health practitioners are twofold. For whom should we prioritise vaccination: the most susceptible, least susceptible, the most infectious or least infectious? Then, how do these combinations interplay? We present a model to address these questions that is applicable to any pathogen for which the classic SIR structure is appropriate, such as influenza or other influenza like illnesses (ILI). In particular, we assume the individuals in a population to have different susceptibilities to infection prior to the beginning of the epidemic. We also investigate the additional effects of vaccination (or increased immunity achieved by other means such as previous exposure to a similar pathogen) prior to the epidemic.

Perhaps the best-known model for the spread of an epidemic is the so-called SIR model. The host population is of constant size, of which proportion S are susceptible to infection, I are infectious, and R are removed either through immunity or death. The proportion infectious, I, is also referred to as the prevalence (of infection). The dynamics may be specified by the scaled (in time) differential equations:

\[
\frac{dS}{dt} = -R_0SI,
\]

\[
\frac{dI}{dt} = R_0SI - I,
\]

\[
\frac{dR}{dt} = I,
\]

where \(R_0\) is the basic reproduction number (Diekmann et al., 2013; Roberts, 2007).

* Corresponding author. Tel.: +61 2 4921 6081; fax: +61 2 4921 6898.
E-mail addresses: r.hickson@unswalumni.com, rihickson@gmail.com (R.I. Hickson), m.g.roberts@massey.ac.nz (M.G. Roberts).
The effects of heterogeneity in populations have been well explored in the literature. Several authors have modelled the effect of heterogeneity in susceptibility by either dividing the population into multiple compartments (Hyman and Li, 2005; Bonzi et al., 2011; Ball, 1985; Hsu Schmitz, 2002), or by considering continuous distributions (Katriel, 2012; Boylan, 1991; Dwyer et al., 1997, 2000, 2002; Veliov, 2005). In general, they have found that when susceptibility is the only variable property, the final size of the epidemic is always smaller for a heterogeneous population than for a homogeneous population with the same reproduction number (Katriel, 2012; Boylan, 1991). The final size is the total proportion of the population infected throughout the epidemic (see, for example, Diekmann et al., 2013). Dwyer et al. (1997) found that heterogeneity in susceptibility is important at large and small spatial scales, but that allowing susceptibility to vary randomly over time has no effect on the pathogen dynamics: their model simplifies to the classic SIR model. That is, if the population randomly changes its susceptibility over the course of an epidemic, the classic SIR model is appropriate.

Other authors have also considered heterogeneity in both susceptibility and infectivity, again by either dividing the population into discrete compartments (Hyman and Li, 2006; Andreasen, 2011; Clancy and Pearce, 2012; Dushoff and Levin, 1995) or by using continuous distributions (Diekmann et al., 1990; Novozhilov, 2008, 2012). The main findings are that for the same reproduction number, heterogeneity in infectivity alone does not change the mean final size of an epidemic (Clancy and Pearce, 2012), but when susceptibility and infectivity are negatively correlated, the final size is larger than that for the homogeneous case (Andreasen, 2011; Clancy and Pearce, 2012). However, if the reproduction number changes, Novozhilov (2012) found that a variation in infectivity may result in larger epidemics. Katriel (2012) developed a Kermack–McKendrick model with heterogeneous susceptibility measured by a single parameter with a continuous distribution, resulting in an equivalent model and similar results to Novozhilov (2008, 2012). Katriel derived a final size equation based on the mean and variance of the susceptibility distribution, and found the upper and lower bounds. A major result was that the largest attack rate (final size) for a given mean is found for a homogeneous population: in other words when the variability is smallest. However, infectivity was assumed to be homogeneous. Katriel used the final size equation to determine the effects of vaccination on a population prior to an outbreak, exploring the outcome of complete vaccination and “leaky vaccination” (where the susceptibility of those vaccinated is reduced by a factor $0 < r < 1$) and a proportion of the population vaccinated. Katriel determined the threshold conditions for these scenarios, showing that the same reduction in the reproduction number can lead to very different attack rates. Katriel also explored the effect of a recurring epidemic, where a proportion of the population had been infected in the previous year, and susceptibility modified accordingly, and found the attack rate to be lower in the heterogeneous population.

Novozhilov (2008, 2012) presented a model similar to the one considered in the present paper, and used moment generating functions to find analytical expressions for a final size. Novozhilov (2012) explored the effect of different variances in an initial gamma distribution of susceptibility on the transient epidemic dynamics, finding that for the same reproduction number, increasing the variance of the distribution results in decreasing the final size, whereas heterogeneity in infectivity exhibits the opposite effect. That is, increasing the variance of the distribution of infectivity (with homogeneous susceptibility) may result in a larger final size. Novozhilov numerically explored heterogeneity in both susceptibility and infectivity, and concluded that since the interaction is nonlinear, the results cannot be predicted by considering heterogeneity in susceptibility and infectivity separately. Novozhilov did not explore the effect of different initial distributions on the transient dynamics, show the effect of the epidemic on the distribution of susceptibility and infectivity over time, or explore the effect of a heterogeneous mortality due to the pathogen. These aspects, together with the vaccination scenario, are the main thrust of the present paper.

Although the final size result has been found by others for similar models (see, for example, Novozhilov, 2008, 2012; Katriel, 2012), those models incorporated heterogeneity in susceptibility only. We present an analytic expression for the final size with heterogeneity in both susceptibility and infectivity, which also gives the resulting distribution of the population.

There are extensive literature on the effect of heterogeneity in contact rates (Novozhilov, 2012; Dushoff, 1999; Andreasen, 2011; Clancy and Pearce, 2012; Dushoff and Levin, 1995; Hethcote and Van Ark, 1987; May and Anderson, 1988; Nold, 1980; Veliov, 2005; Glass et al., 2011) inter alios, and it has been found that the final size of an epidemic is larger when mixing is heterogeneous (Dushoff and Levin, 1995). In the present paper we do not explicitly consider variation in contact rates, although these differences are important determinants of the dynamics. However, the transmission rate for the classic SIR model incorporates contact rates, and we are altering this with our heterogeneity parameter. Hence there are equivalences with the differences in susceptibility and infectivity that we do explore.

We extend the SIR model (1) by allowing susceptibility to vary in the population as a function of a variable $\theta$ according to some (initial) distribution. Other authors (Novozhilov, 2012; Hyman and Li, 2006) considered models where there were different parameters for the heterogeneity of susceptibility and infectivity. We assume a single attribute of the individual that dictates both their susceptibility and infectivity. For an I.I.I. parameter $\theta$ could serve as a proxy for age, with younger people being more susceptible and infectivity then reducing with age (Glass et al., 2012; Lopez and Huang, 2013; Kimura et al., 2011), but this will not be appropriate for all pathogens. We investigate a single outbreak, so waning immunity and population demography are not factors of interest. We compare results from the heterogeneous model to results from the SIR model, where $R_0$ from Eq. (1) is equal to the basic reproduction number, $R_0$, of the heterogeneous model, and where $R_0$ is equal to the parameter $\beta$, which is the median value of $R$ when the entire population is initially susceptible. It is expected that $\beta$ would equal the value of $R_0$ estimated using the homogeneous model, whereas $R$ would be estimated using the heterogeneous model. Therefore, for a consistent comparison, we investigate both cases.

Roberts (2013) used the methodology outlined here: separation of variables followed by decomposition using basis functions to analyse an SIR model with uncertainty about the value of $R_0$. Here we consider the effect of variation in susceptibility and infectivity in the population, and although we use the language of probability throughout, the model is deterministic. This method is capable of utilising arbitrary distributions, and is illustrated by some numerical examples. A selection of distributions are then used to explore the interaction between susceptibility and infectivity, and the effect of this heterogeneity on immunity gained prior to the epidemic.

Generally, our results agree with others in the literature (see, for example, Katriel, 2012; Boylan, 1991): when only heterogeneity in susceptibility is considered the final size is smaller than that for the homogeneous case. However, we found an exception to this: when the least susceptible are vaccinated prior to the epidemic, the final size for the heterogeneous population (19%) is larger than that for the homogeneous population (15.5%) for both $R_0 = R$ and $R_0 = \beta$. When both susceptibility and infectivity are heterogeneous...
the final size may be larger than that for the homogeneous case with \( R_0 = \mathcal{R} \). We find that for all but the cases where either the most susceptible or least susceptible are vaccinated, when there is a positive correlation between infectivity and susceptibility the final size for the heterogeneous model is larger than that for the SIR model with \( R_0 = \beta \), and a negative correlation leads to a heterogeneous final size larger than that for the SIR model. For all distributions explored, except the most susceptible being vaccinated, the final size was largest when infectivity was positively correlated with susceptibility. This agrees with intuition, as the most susceptible are infected first, and if they are also the most infectious the epidemic size will be maximised. However, of more interest is the resulting difference in the timing of the epidemic (that is, is the peak earlier or later, larger or smaller than the homogeneous case?) with the various combinations of susceptibility and infectivity, and the resulting population distribution.

Indeed, by considering a mortality rate due to infection that is dependent on \( \theta \), we can determine vaccination schemes that reduce either or both the final size or the number of deaths during the epidemic.

The model is described in Section 2.1, analytical expressions for the final size are found in Section 2.2, the reproduction number is considered in Section 2.3, and the numerical analysis of the model is described in Section 2.4. The results are presented in Section 3, and discussed in Section 4. It must be emphasized that a uniform distribution of susceptibility does not imply homogeneity; it implies a heterogeneous population with equal numbers across the spectrum of susceptibility values. On the other hand, a constant infectivity function implies homogeneity of infectivity.

2. Method

2.1. The model

We now analyse an extension of the SIR model to include heterogeneous susceptibility and infectivity. We assume that each individual in the population has a fixed susceptibility to infection, indicated by the value of the variable \( \theta \in [-1, 1] \). Since \( \theta \) must be finite, the range \([-1, 1]\) is used to ensure that \( \theta = 0 \) corresponds to the median value of \( \mathcal{R} \), and \( \rho \) is used to scale the extent of the heterogeneity. If an individual is susceptible and contacts an infectious individual, then their probability of becoming infected is \( 1 - \rho \theta / \beta \) times the probability that an individual with \( \theta = 0 \) would become infected, where \( \beta \) is the median value of \( \mathcal{R} \). Since \( x(t, \theta) \) is the density of susceptibles at time \( t \), then the proportion of the population susceptible is

\[
S(t) = \langle x \rangle = w \int_{-1}^{1} x(t, \theta) \, d\theta,
\]

where \( w \) is the appropriate weight to maintain \( S + I + R = 1 \). Since \( \theta \in [-1, 1] \), \( w = 1/2 \). The proportions of the population infectious, \( I(t) = \langle y \rangle \), and removed, \( R(t) = \langle z \rangle \), are similarly determined. We allow for the fact that there may be some dependence of infectivity (once infectious) on an individual’s value of \( \theta \) by introducing a function \( c(\theta) \), such that the force of infection on an individual with \( \theta = 0 \) is \( \beta \langle cy \rangle \). The methodology applies to any general function, \( c(\theta) \), though we later use a specific example. The density dynamics are governed by

\[
\begin{align*}
\frac{dx}{dt} &= -\beta_x + \rho \langle xy \rangle, \\
\frac{dy}{dt} &= \beta_x - \rho \langle xy \rangle - y, \\
\frac{dz}{dt} &= y.
\end{align*}
\]

This is a system of nonlinear partial integro-differential equations that cannot be solved analytically. Therefore, we investigate the system using the method of separation of variables.

Since there are two dependent variables, \( t \) and \( \theta \), we expand the state variables as

\[
\begin{align*}
x(t, \theta) &= \sum_{i=1}^{\infty} x_i(t) \phi_i(\theta), \\
y(t, \theta) &= \sum_{i=1}^{\infty} y_i(t) \phi_i(\theta), \\
z(t, \theta) &= \sum_{i=1}^{\infty} z_i(t) \phi_i(\theta),
\end{align*}
\]

(3)

where \( \{\phi_i\} \) is a set of basis functions defined on \([-1, 1]\), such that \( \langle \phi_i \phi_j \rangle = 0 \) if \( i \neq j \). Substituting the expansions (3) into Eq. (2), multiplying by \( \phi_i(\theta) \), and integrating over \( \theta \) we obtain

\[
\begin{align*}
x_j(t) &= -\langle cy \rangle \left( \beta_{x_j} + \rho \sum_{i=1}^{\infty} x_i(t) \langle \phi_i \phi_j \rangle \right), \\
y_j(t) &= \langle cy \rangle \left( \beta_{x_j} + \rho \sum_{i=1}^{\infty} x_i(t) \langle \phi_i \phi_j \rangle \right) - y_j, \\
z_j(t) &= y_j.
\end{align*}
\]

After truncation at \( M \) terms, these equations simplify to

\[
\begin{align*}
x &= -c \cdot y/(1 + \rho \mathbf{A} \mathbf{x}), \\
y &= c \cdot y/(1 + \rho \mathbf{A}) \mathbf{x} - y, \\
z &= y,
\end{align*}
\]

(4)

with \( \mathbf{x} = (x_1, x_2, \ldots, x_M)' \), and similar expressions for \( \mathbf{y} \) and \( \mathbf{z} \) (prime is transpose). The \( M \times M \) matrix \( \mathbf{A} \) and \( M \)-dimensional vector \( \mathbf{c} \) have components

\[
A_{ij} &= \langle \phi_i \phi_j \rangle / \langle \phi_j^2 \rangle, \quad c_i = \langle c(\theta) \phi_i \rangle.
\]

Eqs. (4) comprise a set of \( 3M \) ordinary differential equations for \( (\mathbf{x}, \mathbf{y}, \mathbf{z}) \).

The appropriate basis \( \{\phi_i\} \) is the set of Legendre polynomials which are orthogonal on \([-1, 1]\) (see Abramowitz and Stegun, 1970; Xiu, 2010). We take \( \phi_{1/2}(\theta) = P_0(\theta) \), defining \( \phi_1(\theta) = P_0(\theta) = 1 \), \( \phi_2(\theta) = P_1(\theta) = \theta \) and

\[
\phi_{i+1}(\theta) = \frac{2i+1}{i+1} \phi_{i+1}(\theta) - \frac{i}{i+1} \phi_i(\theta), \quad i = 1, 2, \ldots
\]

The orthogonality relationship is \( \langle \phi_i \phi_j \rangle = \delta_{ij} / (2i - 1) \), where \( \delta_{ij} = 0 \) for \( i \neq j \) and \( \delta_{ij} = 1 \). Note also that \( \langle \phi_1 \rangle = 0 \) for \( i \neq 1 \). The initial conditions are \( x(0, \theta) = \tilde{x}(\theta) \), \( y(0, \theta) = \tilde{y}(\theta) \), and \( z(0, \theta) = \tilde{z}(\theta) \). The initial proportions of the population infected are \( \tilde{S} = \langle \tilde{x} \rangle \), \( \tilde{I} = \langle \tilde{y} \rangle \), and \( \tilde{R} = \langle \tilde{z} \rangle \). An example of an appropriate initial condition would be \( \tilde{x}(\theta) = 2 \tilde{S}(\theta) \), \( \tilde{y}(\theta) = 2 \tilde{I}(\theta) \), and \( \tilde{z}(\theta) = 2 \tilde{I}(\theta) - \tilde{x}(\theta) - \tilde{y}(\theta) \), where \( \tilde{F}(\theta) = (\tilde{x} + \tilde{y} + \tilde{z})/2 \) is a probability density function.

We have not explicitly specified the function \( c(\theta) \) in order to illustrate the generality of the method. However, for numerical calculations, we assume \( c(\theta) = 1 + \alpha \theta \) where \( \alpha = m/2 \) and \( m = 0, \pm 1 \). This is an example function that accounts for either a positive or a negative correlation between susceptibility and infectivity, where those more susceptible are either more or less infectious respectively. Other functions could be used, and should reflect the situation and research questions being explored. For this function, using expansions (3) and the properties of Legendre polynomials, we can write \( \langle cy \rangle = c \cdot y_1(t) + m y_2(t) / 3 \).
2.2. Final size

From Eq. (2) the incidence of infection can be written as
\[ i(t, \theta) = -\frac{\partial x}{\partial t} = (\beta + \rho \theta)x(t, \theta). \]

The second equation in (2) may be solved to obtain
\[ y(t, \theta) = \bar{y}(\theta)e^{-\int 1 + e^{-t}} = \bar{y}(\theta)e^{-t} + \int_0^t i(t, \theta)e^{t - t} \, dt, \]
and combined with the first, assuming \( \bar{y}(\theta) \leq 1 \), to obtain the Kermack–McKendrick form:
\[ \frac{\partial x}{\partial t} = i(t, \theta) = (\beta + \rho \theta)x(t, \theta) \ast e^{-t}. \]

It is then straightforward to obtain the coupled equations
\[ \frac{\partial}{\partial t}(\theta X) = -\langle \theta \rangle, \]
\[ \frac{\partial}{\partial t}(\theta Y) = (\beta + \rho \theta)(\theta X(t, \theta)) \ast e^{-t}. \]

Taking Laplace transforms, and writing \( L(x) = \int_0^\infty x(t, \theta)e^{\theta t} \, dt \),
\( sL'(x) = \langle \theta \rangle (L(x)) = -L(x(t)), \)
\( sL\log x - \log \hat{x}(\theta) = -\frac{(\beta + \rho \theta)^2}{1 - \rho \theta (L(X)) - \langle \theta \rangle (L(X))}. \)

Hence,
\[ sL\log x - \log \hat{x}(\theta) = \frac{(\beta + \rho \theta)^2}{1 + \rho \theta (L(X)) - \langle \theta \rangle (L(X))}. \]

Taking the limit \( s \to 0 \), with \( \lim_{s \to 0} x(t, \theta) = x_{\infty}(\theta) \), we obtain
\[ \log \left( \frac{\hat{x}(\theta)}{x_{\infty}(\theta)} \right) = (\beta + \rho \theta)(\langle \theta \rangle (\hat{x}(\theta)) - x_{\infty}(\theta))). \]

Hence,
\[ x_{\infty}(\theta) = \hat{x}(\theta)e^{-k\theta + \rho \theta}, \] (5)
where \( k \) solves the transcendental equation:
\[ k = \langle \theta \rangle (\hat{x}(\theta)(1 - e^{-k\theta + \rho \theta}))). \]

For the special case of the initial susceptibility having a uniform distribution in \( \theta \), \( \hat{x}(\theta) = 1 \), and \( c(\theta) = 1 + m\theta \) we have
\[ k = 1 - \frac{\rho \theta + m}{\rho \theta + m} e^{-k\theta + \rho \theta} \sinh \rho \theta + \frac{m}{\rho \theta} e^{-k\theta + \rho \theta} \cosh \rho \theta. \]

In general, suppose the graph of \( \hat{x}(\theta) \) consists of three straight lines joining the points \( (\theta_i, \theta_i) \) for \( i = 0, \ldots, 4 \), with \( \theta_0 = -1 \) and \( \theta_4 = 0 \). Then the lines may be represented by \( x = a_i(\theta + b_i) \), where \( a_i = (x_i - x_{i-1})/(\theta_i - \theta_{i-1}) \) and \( b_i = x_i - a_i \theta_i \). We may then evaluate
\[ \langle \theta \rangle (\hat{x}(\theta)(1 - e^{-k\theta + \rho \theta}))), \]
\[ \frac{1}{2} \sum_{i=1}^4 F_i(\theta_i)\hat{x}(\theta_i), \]
where
\[ F_i(\theta_i) = b_i(\theta_i + \frac{a_i + b_i m + \frac{a_i m}{\rho \theta}}{2} + \frac{a_i m}{\rho \theta} \beta_i + \frac{b_i + a_i + b_i m + (1 + k\rho \theta) + 2a_i m}{2} e^{-k\theta_i + \rho \theta}). \]

This example covers all cases considered in Section 3, except the bimodal distribution discussed in Section 3.3.

2.3. The reproduction number

The final size, calculated using (Eq. 6), always has the solution \( x_{\infty}(\theta) = \hat{x}(\theta) \), \( k = 0 \), for which the epidemic fails to take off. The equations also have a positive solution when the basic reproduction number for Eq. (2),
\[ R = \langle (\beta + \rho \theta) \rangle \rangle \hat{x}(\theta), \] (8)
where \( \hat{x}(\theta) = 1 \), is greater than one. For a uniform initial distribution of susceptibility and an infectivity function \( c(\theta) = 1 + m\theta \), by Eq. (8) \( R = (\beta + m\rho)/3 \).

If the population is not completely susceptible due to a prior vaccination event at \( t = t_0 < 0 \), the effective reproduction number is
\[ R_{\text{eff}} = \langle (\beta + \rho \theta) c(\theta) \rangle \hat{x}(\theta), \]
where \( \hat{x}(\theta) = x(t_0, \theta) - v(\theta) \) is the initial distribution of the susceptible population minus those vaccinated prior to the epidemic. The effective reproduction number is a well known quantity in the literature of infectious disease modelling (see, for example, Diekmann et al., 2013).

2.4. Simulations

Eqs. (4) were solved numerically using MATLAB (The MathWorks Inc., 2011) to explore the effect of heterogeneous population structure on epidemic dynamics. A probability distribution function \( F(\theta) \) was first determined, and then decomposed using Legendre polynomials to obtain initial conditions for \( (x, y, z, w) \), with \( x = 25F(\theta), \ y = 2F(\theta), \) and \( z(\theta) = 2F(\theta) - x(\theta) - y(\theta) \). The final sizes were found using both the analytical expression obtained in Section 2.2 and numerically. Excellent agreement was found in all cases. The scenarios explored are described below, and the results are presented in Figs. 1–4 and Table 1.

The simulation parameters were chosen to illustrate a realistic ILI outbreak. For all simulations the values \( \beta = 1.25 \) and \( \rho = 0.2 \) were used, based on the Influenza Surveillance report in New Zealand (Lopez and Huang, 2012), except for scenario 5 which uses \( \rho = 1 \) to further demonstrate the effect heterogeneity is having on the dynamics. The initial conditions of \( I = 10^{-5}, R = 0, \) and \( S = 1 - I - R \) were used for scenarios 1–3, for both the heterogeneous and SIR models. The number of terms used was determined by the stability of the solution, where further terms made a negligible contribution. \( M = 25 \) was found to be sufficient, leading to a set of 75 nonlinear coupled ODEs. For different initial distributions, a different number of terms may be required. However, as long as the appropriate basis is used, the result will converge. The case fatality ratio, \( d \), was taken to be a linear function of \( \theta \):
\[ d = \frac{0.05(1 - \theta)}{2}, \] (9)
hence ranging from 0 to 5% depending on \( \theta \) (Nishiura, 2010). The proportion of the population who died due to the epidemic, \( D \), was determined by
\[ D = \langle d(\theta) x(t_{\text{final}}, \theta) \rangle = 0.025 \left( \frac{\bar{x}(t_{\text{final}})}{3} - Z_2(t_{\text{final}}) \right). \] (10)

As discussed in the Introduction, two different values of \( R_0 \) were explored for the SIR model: \( R_0 = \beta = 1.25 \) and \( R_0 = R \), the basic reproduction number of the heterogeneous model, calculated using Eq. (8). Note when \( c(\theta) = 1 \) and \( \bar{x} = 1 \), \( R = \beta \).

Five different scenarios were examined as described below:

1 Heterogeneous susceptibility and homogeneous infectivity: For this scenario we used a uniform initial distribution for susceptibility, and a constant infectivity function, \( c(\theta) = 1 \). The constant infectivity function is equivalent to the standard assumption of a homogeneous infectivity distribution for the population. This scenario was used to demonstrate the effect of heterogeneity in susceptibility on the distribution of \( \theta \) in the population through time, and on the distribution of deaths due to the epidemic.
Heterogeneous susceptibility and infectivity: In this scenario, we explore the situation where both susceptibility and infectivity are distributed within the population. A uniform initial distribution was used for susceptibility. Two options were considered for the infectivity function: a straight line with a positive gradient, which corresponds to a positive correlation between susceptibility and infectivity, and a straight line with a negative gradient, which corresponds to a negative correlation. The function for the positive correlation, \( c(\theta) = 1 + \theta \), implies that the most susceptible are also the most infectious. The function with the negative correlation, \( c(\theta) = 1 - \theta \), implies that the least susceptible are the most infectious. Both options were explored to give an insight into the effect of the various combinations on the pathogen dynamics: specifically, the size and timing of the epidemic peak, final size, and deaths due to the pathogen.

A bimodal susceptibility distribution: Instead of a uniform initial distribution of susceptibility, as in the previous two scenarios, here a bimodal distribution was used. The distribution was chosen to illustrate the ability of the method to accept arbitrary distributions, to clearly depict the effect of the initial distribution over the course of the epidemic, and to show how heterogeneity in fatality ratios affects the distribution of deaths due to the epidemic. The distribution was

\[
F(\theta) = \frac{\exp\left(-\frac{(\theta - \mu_1)^2}{2\sigma^2}\right) + \exp\left(-\frac{(\theta - \mu_2)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}},
\]

where \( \mu_1 = 1/2 \), \( \mu_2 = -1/2 \), and \( \sigma = 1/5 \). This effectively divides the population into two groups: the least and the most susceptible. The constant, positively correlated and negatively correlated infectivity functions described in the previous scenarios were used in conjunction with this susceptibility distribution.

The effect of vaccination: To consider the effect of pre-epidemic vaccination, the model was updated to an SIRV model. However, the vaccinated compartment \( V \) was not time dependent as vaccination was applied prior to the epidemic. Hence the initial condition was chosen so that \( S + I + R + V = 1 \), where \( V \) was the initial proportion of the population vaccinated. The ESR report (Lopez and Huang, 2013, p. 43) gives us the vaccination coverage by age group, where taking into account the 50+ age groups, approximately 12% of the population was vaccinated. Therefore \( V = 0.12 \) was used for the simulations. The three linear infectivity functions previously discussed were explored for this scenario. The initial susceptibility distribution was modified in one of the three ways:

(a) vaccinating a proportion of the most susceptible (right-hand corner lost),
(b) vaccinating a proportion of the least susceptible (left-hand corner lost),
(c) a combination of the above (some of both corners lost).

Care was taken to ensure that the same proportion of the population was vaccinated for all three options. Schematics of these options are shown in the first column of Table 1. The final sizes and distributions for these three cases can be determined using Eq. (7).

Fig. 1. Results from scenario 1 with heterogeneous susceptibility and homogeneous infectivity. In (a) and (b) ‘Heterogeneous’ refers to our model, and ‘Classic, \( R_0 = R \)’ and ‘Classic, \( R_0 = \beta \)’ to the SIR model, using the stated values for \( R_0 \). Note the lines overlap. In (d) ‘Fatality’ refers to the fatality ratio as a function of \( \theta \), Eq. (9), and ‘Deaths’ refers to the proportion of the population that dies due to the epidemic, Eq. (10).
5 More heterogeneity: For this scenario we used ρ = 1 to demonstrate what effect a larger range of heterogeneity in the population would have on the dynamics. For a uniform initial susceptibility distribution, three different infectivity options were explored: constant, positively and negatively correlated. For a bimodal initial susceptibility distribution a positively correlated infectivity function was used.

3. Results

The results of the simulations outlined in Section 2.4 are presented and discussed in the following subsections. The results are summarised in Table 1, where the final size and percentage of population deaths due to an epidemic, calculated using (Eqs. (6) and (10)) respectively, are given for different combinations of susceptibility distributions and infectivity functions. In Figs. 1–4, ‘Heterogeneous’ refers to our model, Eq. (2), ‘Classic, R₀ = R’ refers to the SIR model with R₀ equal to the basic reproduction number of the heterogeneous model, given by Eq. (8), and ‘Classic, R₀ = β’ refers to the SIR model with R₀ equal to β from the heterogeneous model. For scenarios 3 and 4, we concentrate on the results with the positively correlated infectivity function as this is the most appropriate for an ILL.

3.1. Heterogeneous susceptibility and homogeneous infectivity

The results for scenario 1, with a uniform initial distribution of susceptibility and homogeneous infectivity, are depicted in Fig. 1. The differences between our heterogeneous model and the SIR model are negligible. This result agrees with those found by others (Katriel, 2012; Boylan, 1991): for the SIR model with both R₀ = β and R₀ = R the final size of the epidemic is 37.1% of the population, adding the small amount of heterogeneity to the model, ρ = 0.2, has had a negligible effect and reduced the final size to 36.9%. Of interest is the distribution of those infected during the course of the epidemic as shown in Fig. 1(c). The expected result of the most susceptible being infected more than the least susceptible is observed, but this can now be quantified.

Fig. 2. Results from scenario 2 with heterogeneous susceptibility and infectivity. The susceptibility is initially uniformly distributed, and the infectivity is positively correlated in (a) and (b) (the most susceptible are the most infectious) and negatively correlated in (c) and (d) (the least susceptible are the most infectious). ‘Heterogeneous’ refers to our model, and ‘Classic, R₀ = R’ and ‘Classic, R₀ = β’ to the SIR model, using the stated values for R₀.
most susceptible are also the most infectious. When comparing the heterogeneous result with the SIR model for $R_0 = R$, the peak of the epidemic has the same timing but is marginally larger for the SIR model, resulting in a final size of 43.9% compared to the 42.0% of the heterogeneous model. More interestingly, the peak of the heterogeneous model is both earlier (at 30 vs. 37 time units) and larger compared to the SIR model with $R_0 = \beta$. When comparing the heterogeneous result with the SIR model for $R_0 = \beta$, the peak of the epidemic has the same timing but is marginally larger for the heterogeneous model. This allows for an easy comparison of the results from different infectivity functions, with the largest final size at 37.1%.

Fig. 2(c) and (d) show the results for the negatively correlated infectivity function, where the least susceptible are the most infectious. In this case, for the SIR model with $R_0 = R$, the peak is marginally smaller than the heterogeneous model, resulting in comparable final sizes at 29.3% for the SIR model and 30.5% for the heterogeneous model. However, for the SIR model with $R_0 = \beta$, the peak of the epidemic is larger and earlier (at 37 vs. 46 time units) than the heterogeneous model, and the final size is notably larger at 37.1%.

The final sizes and mortalities due to the pathogen, for an initial uniform susceptibility distribution and the three infectivity functions explored, as well as the final size calculated from the SIR model with $R_0 = \beta$, are summarised in the first row of Table 1. This allows for an easy comparison of the results from different infectivity functions, with the largest final size and proportion of deaths occurring for a positively correlated function, and the smallest for a negatively correlated function.

3.3. A bimodal susceptibility distribution

An initial uniform susceptibility distribution is not always appropriate, hence for scenario 3 a bimodal distribution was specified, and the results with a positively correlated infectivity function are shown in Fig. 3. The epidemic curve for infection looks identical to that with the initial uniform susceptibility distribution (compare Figs. 2(b) and 3(b)). Fig. 3(c) is a striking visual representation of the most susceptible group being infected first, and more of the group being infected over the course of the epidemic. Furthermore, Fig. 3(d) shows how the distribution of deaths has changed, and although fewer of the least susceptible group were infected, a higher proportion of the population that died due to the epidemic is from this group. This has clear repercussions for intervention planning for ILIs as discussed in Section 4. It also illustrates the strength of the method, which can utilize arbitrary initial distributions of susceptibility.

3.4. The effect of vaccination

In scenario 4, we explored the effect of vaccination, or immunity otherwise gained prior to the epidemic, modifying the already heterogeneous distribution of susceptibility in the population. To correspond to the proportion of the population over 50 years of age vaccinated in New Zealand in 2012 (Lopez and Huang, 2013), 12% of the population was considered to be vaccinated. The results are shown in Fig. 4, where a positively correlated infectivity is considered since it is the most likely scenario for an ILI outbreak (Glass et al., 2012). In Fig. 4(a) and (b) the ‘right-hand corner’ were immune, corresponding to the most susceptible, which are children for ILIs. At 12% of the population vaccinated, the heterogeneous model shows that the epidemic has been prevented, although the SIR models with either $R_0 = R$ or $R_0 = \beta$ still predict outbreaks, resulting in final sizes of 23.0% or 15.5% respectively. In Fig. 4(c) and (d) the ‘left-hand corner’
3.5. More heterogeneity

Scenario 5 demonstrates the effect of a larger range of heterogeneity, with \( \rho = 1 \). The results are depicted in Fig. 5, for an initial uniform susceptibility distribution and the three different infectivity functions in Fig. 5(a)–(c), and for a bimodal distribution with a positively correlated infectivity function in Fig. 5(d). For Fig. 5(a), \( R = 1.25 \) and the final sizes are 31.2% for the heterogeneous model, and 37.1% for the SIR model with both \( R_0 = R \) and \( R_0 = \beta \). For Fig. 5(b), \( R = 1.59 \) and the final sizes are 49.5% for the heterogeneous model, 63.3% for the SIR model with \( R_0 = R \), and 37.1% for the SIR model with \( R_0 = \beta \). For Fig. 5(c), \( R = 0.92 \) and the final sizes are 0.02% for the heterogeneous model, 0.01% for the SIR model with \( R_0 = R \), and 37.1% for the SIR model with \( R_0 = \beta \). For Fig. 5(d), \( R = 1.53 \) and the final sizes are 48.4% for the heterogeneous model, 60.5% for the SIR model with \( R_0 = R \), and 37.1% for the SIR model with \( R_0 = \beta \).

4. Discussion and conclusion

There are three different issues surrounding epidemics that are considered in this paper. First, the effect of heterogeneity in the susceptibility and infectivity of a population on an epidemic, characterised by the size and timing of the peak, the value and distribution of both the final size of the epidemic, and the mortality due to the pathogen. Second, how these characteristics are affected by vaccination, or immunity otherwise gained, prior to the epidemic in a heterogeneous population, compared to what would be predicted by the SIR model. Third, how this methodology and our results could be used to determine the optimal vaccination strategy in terms of both the quantity of vaccine used and the section of the population to target.

The effects of population heterogeneity on epidemic dynamics are demonstrated by Figs. 1–3 and Table 1. In particular, for both infectivity correlations shown in Fig. 2, the peak of the epidemic is shifted in time and size from the SIR model with \( R_0 = \beta \), which has important implications for epidemic planning. For example, when determining hospital load during an outbreak, the positively correlated infectivity function results in more patients than expected, arriving earlier than the expected peak load, which could be problematic, for example if insufficient resources are available. However for the negatively correlated infectivity function, since the outcome is not as severe more resources could be acquired than necessary, and it may appear as though authorities overreacted. Although the SIR model with \( R_0 = R \) has the same
timed for the peak, and negligible differences in the final size, $R_{\text{eff}}$ is calculated using the heterogeneous model, and hence if only the homogeneous model is used, the $R_{\text{eff}} = \beta$ result would be predicted.

Although the difference in final size was previously known, the timing of the peak was less well documented, and the methodology presented here also result in the distribution of the final size and mortality due to the pathogen. To measure how much those most susceptible are infected relative to the least susceptible, we define relative susceptibility to be the susceptibility of the heterogeneous population divided by that of a homogeneous population, which simplifies to

\[ S_{\text{rel}} = \frac{\frac{\beta + \rho(\theta)\psi}{\beta(X)}}{\frac{\beta x_{1}(t)}{\beta x_{1}(t)}} = 1 + \frac{\rho x_{1}(t)}{\frac{\beta x_{1}(t)}{\beta x_{1}(t)}} \]

This changes during the course of an epidemic, and the size of the change is a measure of how much those most susceptible are infected relative to those least susceptible. The change is relatively small with $R_{\text{eff}} \approx 1.25$, but is more apparent for results with a higher $R_{\text{eff}}$ than those explored here (not shown). The changes would also be much higher for more extreme heterogeneities such as those illustrated in Fig. 5.

The effects of vaccination in a heterogeneous population on the epidemic are depicted in Fig. 4 with more detail in Table 1. For a typical seasonal influenza outbreak, a positively correlated infectivity distribution would be the most appropriate as children (corresponding to $\theta \rightarrow 1$) tend to be both the most susceptible to infection and the most able to infect others (Glass et al., 2012). When comparing results from the initial uniform susceptibility distribution, Fig. 3, with those from different vaccination schemes for a positively correlated infectivity distribution, shown in Fig. 4 and Table 1, it is important to remember the assumption that the same proportion of the population is vaccinated for all three vaccination options. In both cases, the $SIR$ model results in significantly different timings of the epidemic peak, with vaccinating the most susceptible resulting in no outbreak for the homogeneous model, but outbreaks for the $SIR$ model. When vaccinating the least susceptible an earlier, larger peak results for the heterogeneous model. These results are summarised by the notable differences in the final sizes as shown in Table 1.

We expected the results in Table 1 to show that the final size and mortalities were reduced due to immunisation, whichever subpopulation group is targeted, when comparing the bottom three rows with the top row and accounting for the respective infectivity functions. However, in the case of a pathogen where there is a negative correlation between susceptibility and infectivity, if the most susceptible group is targeted for vaccination the mortality increases compared to the initial uniform susceptibility case (0.78% vs. 0.73%). Although this is only by a small margin, for
larger reproduction numbers the difference would be greater. More importantly, there are large differences that result from targeting the correct group, which results in no outbreak occurring. That is, the most susceptible group for a positively correlated infectivity, and the least susceptible for the negatively correlated infectivity. The ability for vaccination of just 12% of the population to stop the outbreak entirely is surprising, given that the infectivity, and the least susceptible for the negatively correlated ring. That is, the most susceptible group for a positively correlated targeting the correct group, which results in no outbreak occurring. More importantly, there are large differences that result from larger reproduction numbers the difference would be greater.

When infectivity is positively correlated with susceptibility, the results clearly show that the ‘right-hand corner lost’ vaccination strategy is the most effective, having prevented an outbreak from occurring. In fact, the model shows that vaccinating the least susceptible, ‘left-hand corner lost’, results in a final size that is approximately twice that predicted by the SIR model, and notably earlier! Indeed, when comparing this vaccination scheme to the corresponding result without vaccination, the final size has only been decreased from 42.0% to 39.9%. A vaccination scheme that targets the same proportion of the population, 12%, but equal numbers of the most and least susceptible (both corners lost) also results in a higher final size (20.8%) than that predicted by the SIR model with $R_0 = \beta$ (15.5%), though it is similar to the final size of the SIR model with $R_0 = \beta$ (23.0%). Therefore, for any ILL outbreak similar to that explored here, targeting children for vaccination rather than the older generation would result in a smaller final size of the epidemic and less mortality. To properly optimise the vaccination strategy, the target would be to make $R_{\text{eff}} < 1$ while minimising $\bar{v}(\theta)$, the proportion of the population vaccinated.

The differences between all scenarios in the population mortality due to the epidemic are small, but this is due to the small number of deaths expected in an ILL outbreak, with a maximum fatality ratio of 5%. Given the small numbers, of more importance, than the absolute number is the distribution of mortality. This distribution of the deaths in the population is important when determining the cost-effectiveness of interventions, using health-economics measures such as quality adjusted life years. We note that the quality of life is negligibly affected by ILLs, however this measure takes into account the age-dependent cost of the deaths, and our methodology provides the information required in the form of the distribution of deaths due to the epidemic.

In a recent study, the application of a vaccine that resulted in a non-uniform reduction in susceptibility was considered (Gomes et al., in press). While the overall effect on infection dynamics was described, the methodology did not explicitly determine how an epidemic would change the distribution of susceptibility in the population. The methodology described here could be used to address that question.

The model developed in this paper and the results derived from it provide insight into the effect of heterogeneous population structure on the dynamics of the epidemic, specifically the size and timing of the epidemic peak, the final size and its distribution, and the distribution of deaths due to the pathogen. Furthermore, our method accepts arbitrary population distributions for susceptibility and infectivity, and hence empirically derived distributions could be utilised in our model. For an ILL it is often the case that a proportion of the population has immunity due to prior exposure,
perhaps to a genetically similar virus. Hence $R_0$ would be higher than the value assumed here for illustration, but $R$ would be non-zero. The combination of empirical distributions and actual mortality data would provide a more accurate understanding and prediction of the epidemic dynamics in heterogeneous populations than the classic $SIR$ model. Future work will use this methodology to consider vaccination during the epidemic, and the effects of a time-dependent $\beta$.

Acknowledgements

The authors would like to acknowledge insightful conversations with Prof. Geoffrey Mercer. Prof. Gabriela Gomes kindly made a preprint of Gomes et al. (in press) available.

References


