Available online at www.sciencedirect.com
ScienceDirect
Mathematical Biosciences 203 (2006) 124–136
Mathematical Biosciences
www.elsevier.com/locate/mbs

The effect of contact heterogeneity and multiple routes of transmission on final epidemic size

Istvan Z. Kiss *, Darren M. Green, Rowland R. Kao

Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK

Received 12 August 2005; received in revised form 28 February 2006; accepted 2 March 2006
Available online 19 April 2006

Abstract

Heterogeneity in the number of potentially infectious contacts amongst members of a population increases the basic reproduction ratio (R_0) and markedly alters disease dynamics compared to traditional mean-field models. Most models describing transmission on contact networks only account for one specific route of transmission. However, for many infectious diseases multiple routes of transmission exist. The model presented here captures transmission through a well defined network of contacts, complemented by mean-field type transmission amongst the nodes of the network that accounts for alternative routes of transmission. The impact of these combined transmission mechanisms on the final epidemic size is investigated analytically. The analytic predictions for the purely mean-field case and the transmission through the network-only case are confirmed by individual-based network simulations. There is a critical transmission potential above which an increased contribution of the mean-field type transmission increases the final epidemic size while an increased contribution of the transmission through the network decreases it. Below the critical transmission potential the opposite effect is observed.

Keywords: Epidemiology; Contact heterogeneity; Scale-free networks; Transmission; Final epidemic size

* Corresponding author. Tel.: +44 186 528 1092; fax: +44 186 531 0447.
E-mail address: istvan.kiss@zoo.ox.ac.uk (I.Z. Kiss).

0025-5564/$ - see front matter
doi:10.1016/j.mbs.2006.03.002

1. Introduction

Homogeneous random mixing is a common assumption of many epidemiological models [1]. If we assume a single closed population of N identical individuals, homogeneous random mixing implies that all individuals are equally likely to contact each other, and therefore if infected, are equally likely to infect susceptible members of the population. In a compartmental *SIR* model for disease transmission, the individuals in the population are divided into different compartments denoted by S (susceptible), I (infectious) and R (recovered and immune, or removed). Let us define a transmission rate β to be the number of infectious contacts per individual per unit time. Infectious nodes recover and become immune at rate g. In this case, the threshold for disease persistence is given by $\beta/g = 1$. This quantity is identified as the basic reproduction number or R_0 and is commonly defined as the number of secondary infections generated by a single infectious individual introduced into a totally susceptible population [1].

If there is a well-defined contact structure [2–4], then contact between any two individuals is not equally likely. Heterogeneity in the contacts amongst individuals has an important effect on R_0 and disease dynamics [1,5–7]. Let us assume that the population is divided into n distinct groups of sizes N_k ($k = 1,2,\ldots,n$) such that each individual in group k has exactly k contacts. If the population size is N ($N = N_1 + N_2 + \cdots + N_n$), then the probability that a uniformly chosen individual has k contacts is $p(k) = N_k/N$. Empirical studies have shown that many real networks have scale-free (SF) degree distributions $p(k) \approx k^{-\gamma}$ with $2 \leqslant \gamma \leqslant 3$ (see Ref. [2] for a review). While there are problems with interpreting empirical data describing real networks [8,9], and scale-free properties are only approximately relevant for finite populations, nevertheless the relevance of scale-free contact structures has been highlighted in different contexts, such as in the case of human sexual contacts [8,10], and for livestock trading networks within Great Britain [11,12]. Recently, interest in individual-based network models that incorporate the effect of contact heterogeneity has increased, at least partially due to increasing amounts of data describing real networks [10,13,14] and available computational power [7,14–18].

When comparing theoretical contact networks to real disease transmission, the definition of a potentially infectious contact or link between two individuals is crucial. If the contact network is imprecisely identified, or there are multiple modes of disease propagation (e.g. blood transfusion, shared needle use and sexual contact in the case of HIV [8,19]), then these factors must be considered by changing the contact structure and therefore the disease dynamics. To account for these factors, we study the effect of transmission that occurs through a combination of a well-defined contact structure (in this case, a scale-free network) and mean-field-type transmission amongst the network nodes that is independent of the number of network contacts. This may represent either two complementary transmission mechanisms, or a single transmission mechanism that is only approximated by the assumed scale-free contact structure. A control parameter λ is used to vary the contribution of the two different transmission mechanisms to the overall transmission. The impact of this alternative transmission mechanism on the final epidemic size in a simple *SIR* model is explored using analytic and numeric calculations. Theoretical predictions for the final epidemic size are validated using simulations in the cases of homogeneous and scale-free networks when transmission occurs through the links of the networks only.

2. Disease transmission model

If the members of a population are considered to be nodes in a network each with their own degree (i.e. the number of potential contacts with other members of the population), an undirected network of size N with node degree distribution $p(k)$ is obtained. The value $\langle k\rangle = \sum_l lp(l)$ is the average number of contacts per node. In general $\langle f(k)\rangle = \sum_l f(l)p(l)$. The network structure may not capture all transmission between individuals. This may be because the network is imprecisely defined, or because there are other transmission routes not captured by the network structure. In order to explore this, we consider additional transmission that can be approximated by mean-field terms.

In the context of the *SIR* model, each node in the network at any time can be either susceptible (S), infectious (I) or removed (R). If S_k and I_k represent the number of susceptible and infectious individuals within group k, respectively (where $S_k + I_k + R_k = N_k$) the following system of differential equations captures disease spread for arbitrarily large networks ($N \to \infty$), for both transmission through the network and the mean-field type transmission

$$\begin{cases} \dot{S}_k(t) = -(1-\lambda)\tau k S_k(t) \sum_l \frac{l-1}{l} p(l|k) \frac{I_l(t)}{N_l} - \lambda\beta \frac{S_k(t)}{N} \sum_l I_l(t), & \text{(1a)} \\ \dot{I}_k(t) = (1-\lambda)\tau k S_k(t) \sum_l \frac{l-1}{l} p(l|k) \frac{I_l(t)}{N_l} + \lambda\beta \frac{S_k(t)}{N} \sum_l I_l(t) - gI_k(t), & \text{(1b)} \end{cases} \quad \text{for } k = 1,\ldots,n. \tag{1}$$

Here β and g are as defined above and τ is the transmission rate across a network contact between an infected and a susceptible node. The parameter λ ($0 \leqslant \lambda \leqslant 1$) varies the contribution of the two different transmission mechanisms to the overall transmission. The first term describing the creation of new infections through the network is proportional to the transmission rate τ, the degree k of the susceptible nodes being considered, the number of susceptible nodes with k connections, and the probability that any given neighbour of a susceptible node with k connections is infected. The second term describes a simple mean-field transmission, independent of node degree. The term $p(l|k)$ is the probability that a node of degree k is connected to a node of degree l, and $\sum_l p(l|k) = 1$.

The formulation for heterogeneous contact follows that of Anderson and May for transmission of sexually transmitted diseases [1], however it replaces $p(l|k)$ in the transmission term with $\frac{l-1}{l}p(l|k)$, since here we assume the network is static, and thus infection over the network results in loss of one susceptible partner. For intermediate λ, this correction slightly underestimates the number of infections, since nodes infected via the mean-field route are unlikely to have lost a potentially infectious connection when they themselves are infected, as is also the case when the network connections change over time.

If the network of contacts is neither assorative nor disassortative [3,4] (i.e. there are no correlations between the degree of connected nodes), then $p(l|k)$ simply depends on the degree of node l and $p(l)$, and is given by $p(l|k) = lp(l)/\langle k\rangle$. Therefore Eqs. (1a) and (1b) can be rewritten to give

$$\begin{cases} \dot{s}_k(t) = -s_k(t) \sum_l \left((1-\lambda)\tau \frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right) i_l(t), & \text{(2a)} \\ \dot{i}_k(t) = s_k(t) \sum_l \left((1-\lambda)\tau \frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right) i_l(t) - g i_k(t), & \text{(2b)} \end{cases} \quad \text{for } k = 1,\ldots,n, \tag{2}$$

where $s_k = S_k/N$ and $i_k = I_k/N$. Using Eqs. (2a) and (2b) the overall R_0 can be calculated as (see Appendix A for details)

$$R_0 = \frac{1}{2}\left[\rho_r + \rho_0\left(\frac{\langle k^2\rangle}{\langle k\rangle^2} - \frac{1}{\langle k\rangle}\right)\right] + \frac{1}{2}\left\{\left[\rho_r - \rho_0\left(\frac{\langle k^2\rangle}{\langle k\rangle^2} - \frac{1}{\langle k\rangle}\right)\right]^2 + 4\rho_r\rho_0\left(1 - \frac{1}{\langle k\rangle}\right)\right\}^{1/2}, \tag{3}$$

where $\rho_r = \lambda\beta/g$ and $\rho_0 = (1-\lambda)\tau\langle k\rangle/g$ are the transmission potentials of the two different transmission mechanisms. These both represent the number of secondary infections generated throughout the infectious period of an infectious node chosen uniformly at random within an entirely susceptible population. The value of R_0 has a similar definition to that of ρ_r and ρ_0, but differs in that the node is not chosen entirely at random, but weighted according to the probability of it itself being infected. Both are important, however the value of $R_0 = 1$ remains the threshold for disease invasion [1,20].

If $\lambda = 1$ then $\rho_0 = 0$, and the purely mean-field case is recovered with $R_0 = \rho_r = \beta/g$. In the $\lambda = 0$ case $\rho_r = 0$, and the threshold for epidemic outbreaks is given by $R_0 = \rho_0(\langle k^2\rangle/\langle k\rangle^2 - 1/\langle k\rangle) = 1$ [20]. If the network of contacts has a scale-free degree distribution, $p(k) \propto k^{-\gamma}$, then for values of $\gamma > 3$ the second moment of the degree distribution ($\langle k^2\rangle$) is finite. If $\gamma \leqslant 3$, then the second moment of the nodes' degree distribution diverges ($\langle k^2\rangle \to \infty$) and according to Eq. (3) R_0 diverges too ($R_0 \to \infty$). For sufficiently high heterogeneity, even infinitesimally small transmission rates can result in an epidemic [6,7].

3. The final epidemic size

The final epidemic size ($r(\infty) = 1 - \sum_l s_l(\infty)$), defined as the total number of individuals affected by the disease by the end of the epidemic, is an important indicator of the severity of an epidemic outbreak or of control strategy efficacy. Following Anderson and May [1], we identify a set of parametric equations defining $r(\infty)$ in terms of the transmission potentials of the two different transmission mechanisms (see Appendix B for details). The equation for $r(\infty)$ is

$$r(\infty) = \langle 1 - \exp(-k\alpha - \alpha_r)\rangle, \tag{4}$$

where ρ_r and ρ_0 are determined by

$$\rho_r = \alpha_r/\langle 1 - \exp(-k\alpha - \alpha_r)\rangle, \tag{5}$$

$$\rho_0 = \alpha\langle k\rangle^2/\langle (k-1)(1 - \exp(-k\alpha - \alpha_r))\rangle. \tag{6}$$

The averages in the parametric Eqs. (4)–(6) are evaluated for various values of α and α_r by approximating the sums with integrals. The following calculations assume a scale-free network of contacts corresponding to the preferential attachment model of Barabási and Albert [2,21] with probability density $p(k) = 2m^2/k^3$ ($k \geqslant m$), for continuous k, and an average of $\langle k\rangle = 2m$ contacts per node. If the transformations $x = k/m$ and $\Phi = m\alpha$ are applied to the resulting integrals the following family of parametric equations are obtained

$$r(\infty) = 2\int_1^\infty \frac{1 - \exp(-\Phi x - \alpha_r)}{x^3}\,\mathrm{d}x, \tag{7}$$

$$\rho_r = \left(\frac{\alpha_r}{2}\right) \Big/ \int_1^\infty \frac{1-\exp(-\Phi x-\alpha_r)}{x^3}\,\mathrm{d}x, \tag{8}$$

$$\rho_0 = 2\Phi \Big/ \int_1^\infty \frac{mx-1}{mx^3}[1-\exp(-\Phi x-\alpha_r)]\,\mathrm{d}x. \tag{9}$$

For each fixed value of $m(=\langle k\rangle/2)$ a contour plot of $r(\infty)$ (not shown) as a function of ρ_r and ρ_0 can be obtained if the integrals are numerically evaluated for various values of Φ and α_r. The qualitative behaviour of $r(\infty)$ does not changes with m. Moreover as m is increased, $r(\infty)$ converges towards the contour plot given by Eqs. (7)–(9) where Eq. (9) is modified to

$$\rho_0 = 2\Phi \Big/ \int_1^\infty \frac{1-\exp(-\Phi x-\alpha_r)}{x^2}\,\mathrm{d}x. \tag{10}$$

Eqs. (7), (8) and (10) define a unique contour plot since m only appears through the $\Phi(=m\alpha)$ term. The contour plot of the final epidemic size $r(\infty)$ (solid lines), as given by Eqs. (7), (8) and (10), is presented in Fig. 1 as a function of ρ_r and ρ_0. Eqs. (7), (8) and (10) correspond to the case where the $(l-1)p(l|k)/l$ term is replaced with the $p(l|k)$ term in Eqs. (1a–2b). The former formulation is especially important if most of the transmission occurs through the links of a network with a small $\langle k\rangle$ value, however even these two extreme cases have very similar qualitative behaviour.

The intersection of fixed overall transmission potential lines ($\rho_r + \rho_0 = \rho_{\text{total}}$, where ρ_{total} is a constant; dashed lines in Fig. 1) with the $r(\infty)$ contour lines illustrates the effect of changing the contributions of the different transmission mechanisms on the final epidemic size. A critical value, $\rho_{\text{crit}} \approx 1.4$, of the overall transmission potential separates two distinct regimes. For transmission potential higher than ρ_{crit}, an increased contribution of mean-field type transmission will increase the final epidemic size, while an increased proportion of transmission through the network decrease it. For transmission potentials lower than ρ_{crit}, the opposite response is

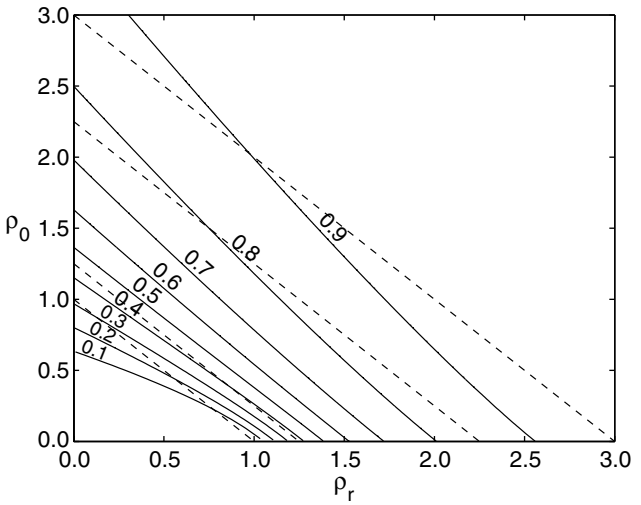

Fig. 1. Contour plot of the final epidemic size ($r(\infty)$) (solid lines) and lines of fixed overall transmission potential (dashed lines, $\rho_0 + \rho_r = \rho_{\text{total}}$) as a function of the transmission potentials of the mean-field type transmission and of the transmission through the links of the network. The transition between the two regimes occurs at $\rho_r + \rho_0 = \rho_{\text{crit}} \approx 1.4$.

observed. Fig. 1 also shows that for fixed ρ_{total} no intermediate value of λ results in a higher final epidemic size than the higher $r(\infty)$ value of the two extreme cases $\lambda = 1$ or 0.

The extreme cases of $\lambda = 1$ and 0 are considered in detail. If $\lambda = 1$ then Eqs. (4)–(6) reduce to

$$r(\infty) = 1 - \exp(-R_0 r(\infty)) = 1 - \exp(-\rho_r r(\infty)), \tag{11}$$

which is the well known result for the mean-field *SIR* model. Barbour and Mollison [22] use random graphs to construct different realizations of the Reed–Frost chain-binomial epidemic process [23]. They present an alternative derivation of Eq. (11) for the final epidemic size by calculating the size of the giant component (i.e. the largest connected subset of nodes) in large random graphs. Their study also emphasises the important role of explicitly considering the contact structure when modelling disease spread.

In the case of $\lambda = 0$, transmission occurs only through the defined network structure and Eqs. (4)–(6) reduce to

$$r(\infty) = \langle 1 - \exp(-k\alpha)\rangle, \tag{12}$$

$$\rho_0 = \alpha\langle k\rangle^2 / \langle (k-1)(1-\exp(k\alpha))\rangle, \tag{13}$$

which approaches the result of May and Lloyd [6], if the preferential attachment model, with $m = \langle k\rangle/2 \gg 1$ is used. These predictions are presented in Fig. 2 as a function of the transmission potential $\rho(=\rho_0 = \rho_r)$. For smaller values of ρ, the final epidemic size on scale-free networks is considerably higher than that for the mean-field model. However, as ρ increases, the critical value of $\rho_{\text{crit}} \approx 1.4$ is reached, where the final epidemic sizes are the same for both types of transmission mechanisms. For large ρ, the final epidemic size corresponding to the mean-field model approaches its asymptote (total population size) more rapidly than for scale-free networks.

Newman [24] studied the effect of clustering on the size of the giant component and final epidemic size in networks with tunable degree distribution and tunable clustering coefficient. Cluster-

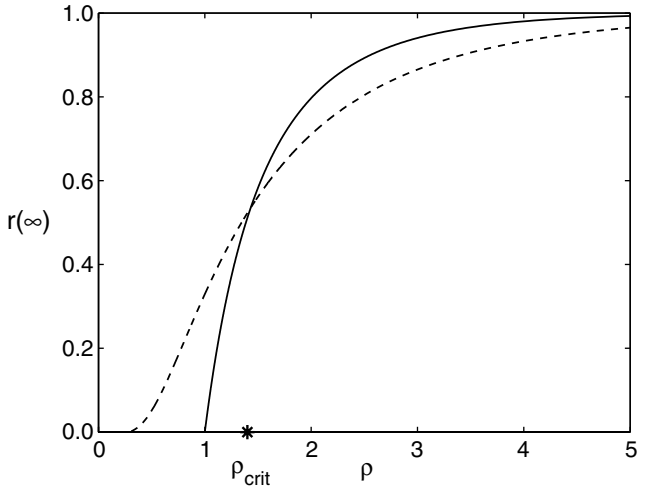

Fig. 2. The theoretical predictions of the final epidemic size as a function of the transmission potential $\rho = \rho_r = \rho_0$. The final epidemic size, $r(\infty)$, for the mean-field *SIR* model (solid line) is plotted using Eq. (11). For the scale-free network case (dashed line) $r(\infty)$ is plotted using the preferential attachment model in Eqs. (12) and (13) in the $m = \langle k\rangle/2 \to \infty$ limit. The final epidemic sizes are equal at $\rho_{\text{crit}} \approx 1.4$.

ing and network heterogeneity have similar effects on the final epidemic size. High levels of clustering lower the epidemic threshold, and there is a critical transmissibility separating two distinct regimes for different levels of clustering. For transmissibility lower than the critical value, final epidemic size on clustered networks is higher than that on unclustered networks. For transmissibility higher than the critical value, the opposite effect is observed.

Studies of disease spread using individual-based network models have shown that disease on scale-free networks initially spreads preferentially to nodes with high degree [15]. As ρ increases on scale-free networks, the disease depletes the population of highly connected nodes, and must propagate to the poorly connected, less accessible nodes. In homogeneous networks this depletion never occurs, and the final epidemic size is larger for large ρ than on equivalent SF networks.

4. Comparison of theoretical predictions with simulation results

The theoretical predictions for $r(\infty)$ given by Eqs. (4)–(6) are compared to an individual-based stochastic model that simulates the spread of disease on computer-generated homogeneous and scale-free networks. It is assumed that transmission occurs through the links of the networks only ($\lambda = 0$). Homogeneous networks are generated such that each network node has the same number of links. This distribution of contacts is expected to approximate well the mean-field case where the number of contacts per node is large; in extremis, each node can potentially contact any other node. Scale-free networks are obtained using the preferential attachment model of Barabási and Albert [2,21]. Their model accounts for the continuous addition of new nodes seen in real networks and the preferential attachment of these to nodes already present in the network. The network construction algorithm starts with a small number (m_0) of connected nodes. At every step, a new node with $m(\leqslant m_0)$ links, is added to the network, connecting to already existing nodes. The probability Π that a new node connects to an existing node u depends on the degree u_k of that node with $\Pi(u_k) = u_k/\sum_l u_l$. Numerical simulations of the Barabási and Albert model produce networks with scale-free degree distribution, $p(k) \propto k^{-\gamma}$, with an exponent $\gamma = 2.9 \pm 0.1$.

All the networks used in the simulations were built using $N = 10\,000$ nodes. The epidemics are seeded with ten randomly chosen nodes to avoid stochastic extinction. The probability of a susceptible node with k infectious neighbours becoming infectious in a small interval of time Δt is directly related to $\tau k \Delta t$. An infectious node recovers at rate g, and without loss of generality, $g = 1$ is used for all the simulations. The results from all the simulations are averaged over 50 different network realizations and 50 epidemic realizations on each individual network.

The theoretical prediction for $r(\infty)$ in the case of homogeneous networks is obtained from Eqs. (4) and (6) upon using the probability density function $p(k) = \delta(k - \langle k \rangle)$, where δ is the Dirac delta function. In this case $r(\infty)$ is given by

$$r(\infty) = 1 - \exp(-\rho_0(1 - 1/\langle k \rangle) r(\infty)). \tag{14}$$

In Fig. 3, the final epidemic size is plotted according to the theoretical prediction (Eq. (14)) and simulation results using different values of $\langle k \rangle$. In the $\langle k \rangle \to \infty$ limit Eq. (14) is equivalent to Eq. (11), and represents the purely mean-field case.

For scale-free networks $p(k) = 2m^2/k^3$ ($k \geqslant m$) is used in Eqs. (4) and (6) leading to the following parametric equations:

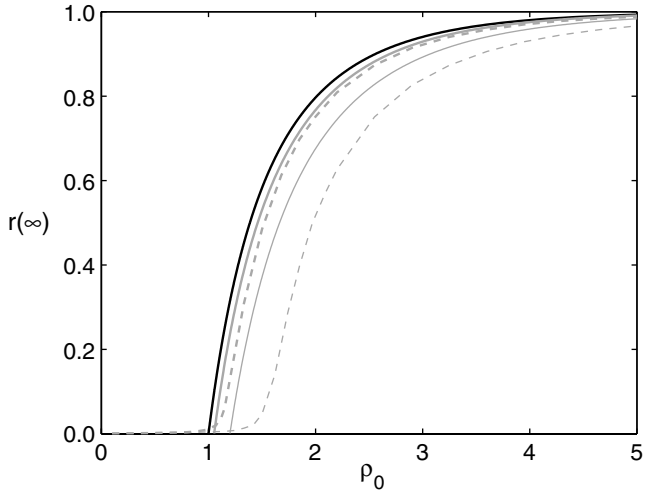

Fig. 3. Comparison between the theoretical predictions using Eq. (14) for homogeneous networks (solid grey lines, thick for $\langle k\rangle = 20$ and thin for $\langle k\rangle = 6$) and computer simulations of the final epidemic size $r(\infty)$ (dashed grey lines, thick for $\langle k\rangle = 20$ and thin for $\langle k\rangle = 6$). The solid black line represents the $\langle k\rangle \to \infty$ limit equivalent to the mean-field case given by Eq. (11).

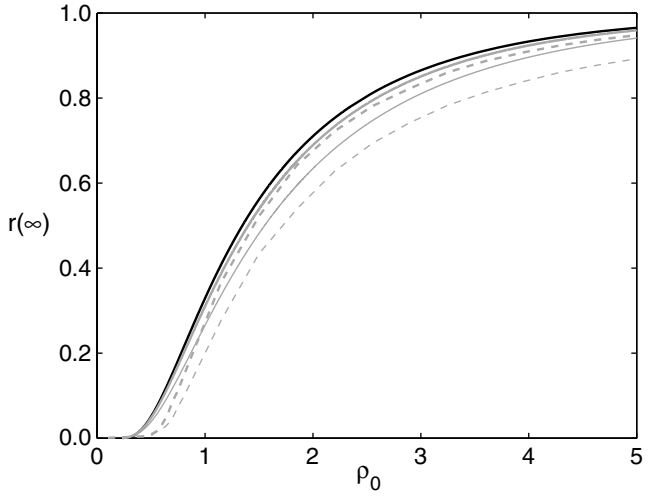

Fig. 4. Comparison between the theoretical predictions using Eqs. (15) and (16) for scale-free networks (solid grey lines, thick for $\langle k\rangle = 20$ and thin for $\langle k\rangle = 6$) and computer simulations of the final epidemic size $r(\infty)$ (dashed grey lines, thick for $\langle k\rangle = 20$ and thin for $\langle k\rangle = 6$). The solid black line represents the $m = \langle k\rangle/2 \to \infty$ limit and is identical to the result of May and Lloyd [6].

$$r(\infty) = 2\int_1^\infty \frac{1-\exp(-\Phi x)}{x^3}\mathrm{d}x, \tag{15}$$

$$\rho_0 = 2\Phi \Big/ \int_1^\infty \frac{mx-1}{mx^3}[1-\exp(-\Phi x)]\mathrm{d}x. \tag{16}$$

In Fig. 4, the final epidemic size is plotted for scale-free networks using the theoretical prediction (Eqs. (15) and (16), with $m = \langle k \rangle/2$) along with simulation results for different values of $\langle k \rangle$. In the limit of $m \to \infty$, Eqs. (15) and (16) are equivalent to those obtained by May and Lloyd [6].

In both cases, for large values of $\langle k \rangle$ the agreement is good, but less so for smaller values of $\langle k \rangle$. Comparing Figs. 3 and 4 show that the difference between the theoretical and numerical results, for smaller values of $\langle k \rangle$, is less significant for highly heterogeneous networks, where the highly connected nodes dominate the dynamics. The differences for smaller values of $\langle k \rangle$ can be accounted for by the local depletion of susceptible neighbours [25].

5. Discussion

The importance of contact heterogeneity and multiple transmission mechanisms has long been recognised. For example, Diekmann et al. [26] explicitly incorporated the reduction of infectious output due to repeated contacts between individuals into the expression for R_0, showing how it reduces the effective infectious output. Here, we have investigated the role of contact heterogeneity and multiple routes of transmission in determining final epidemic size. We have used as a basis for comparison scale-free networks, as these are relevant for a number of disease contact structures, particularly in the case of sexual contact networks. However, many of the concepts derived in the scale-free context are important in the simple context of population heterogeneity, provided that heterogeneity incorporates correlation between susceptibility and transmissibility. Transmission through the network leads to a 'hierarchical' spread [15] that ensures rapid infection of highly connected nodes ('superspreaders'), and this has consequences for reactive disease control through contact tracing [27].

Here, we have shown that for small values of ρ, heterogeneity ensures a larger final epidemic size than on homogeneous networks. For higher ρ values, the depletion of highly connected nodes means that the local 'susceptible neighbourhood' of infectious individuals will be rapidly depleted, as the average number of connections in remaining susceptible nodes goes down more quickly than in more homogeneous networks. On homogeneous networks, the initial spread is slower, but may result in a larger epidemic even though R_0 is lower. Interestingly, mixed strategies (where the overall transmission potential is the same, but heterogeneity is only partially exploited) are never favoured – the final epidemic size is always lower than that at one of the two extreme cases of $\lambda = 0$ or 1 (Fig. 1).

An important aim of most epidemic control programmes is to minimise the number of individuals affected by disease (i.e. final epidemic size). This can be achieved by reducing transmissibility, shortening the duration of the infectious period and reducing the contact rate between susceptible and infected individuals. Critical to the use of mathematical models in determining the most effective policy is good parameter estimation, and 'real-time' use of mathematical models for disease control requires that these be developed rapidly [28,29]. Knowledge about both the transmissibility of the disease and the network of contacts are needed if the estimates of the transmission potential are used to identify effective control strategies promptly in the event of an epidemic [27,30]. Here, we have shown that an estimate of the transmission potential of a disease is only a good indicator of the final epidemic size (and therefore the most effective control strategy to take) if there is a priori knowledge of the underlying contact structure that drives the epidemic.

Acknowledgements

R.R.K. and I.Z.K. are funded by the Wellcome Trust. D.M.G. is funded by DEFRA. The authors wish to thank Professor R.M. May for useful discussions on this topic and for pointing out the calculation of the threshold criterion.

Appendix A. Calculation of the threshold criterion

To derive an overall epidemic threshold (when $0 < \lambda < 1$), we assume that in the initial stages $i_k \ll s_k$ and $s_k \approx p(k)$. Following these assumptions Eq. (2b) now reads

$$\dot{i}_k(t) = p(k)\sum_l \left((1-\lambda)\tau\frac{k(l-1)}{\langle k\rangle} + \lambda\beta - \frac{g\delta_{lk}}{p(k)}\right) i_l(t). \tag{A.1}$$

Now define

$$I(t) = \sum_k i_k(t) \tag{A.2a}$$

and

$$J(t) = \sum_k k i_k(t). \tag{A.2b}$$

Combining Eqs. (A.1) and (A.2) the following pair of linear equations is obtained:

$$\dot{I}(t) = [\lambda\beta - g - (1-\lambda)\tau]I(t) + (1-\lambda)\tau J(t), \tag{A.3}$$

$$\dot{J}(t) = \left[\lambda\beta\langle k\rangle - (1-\lambda)\tau\frac{\langle k^2\rangle}{\langle k\rangle}\right]I(t) + \left[(1-\lambda)\tau\frac{\langle k^2\rangle}{\langle k\rangle} - g\right]J(t). \tag{A.4}$$

A linear stability analysis of the disease free state $(I,J) = (0,0)$ is performed. By substituting $I(t), J(t) \to I, J\exp(\Lambda t)$ in the already linearised equations (A.3) and (A.4) a quadratic equation for the eigenvalues Λ is obtained, with the transition from a negative to a positive eigenvalue given by

$$\frac{1}{2}\left[\lambda\beta + (1-\lambda)\tau\left(\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right)\right] + \frac{1}{2}\left\{\left[\lambda\beta - (1-\lambda)\tau\left(\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right)\right]^2 + 4\lambda(1-\lambda)\beta\tau(\langle k\rangle - 1)\right\}^{1/2} = g. \tag{A.5}$$

Rewriting Eq. (A.5) in function of $\rho_r = \lambda\beta/g$ and $\rho_0 = (1-\lambda)\tau\langle k\rangle/g$ (ρ_r and ρ_0 defined in the paper) we obtain

$$R_0 = \frac{1}{2}\left[\rho_r + \rho_0\left(\frac{\langle k^2\rangle}{\langle k\rangle^2} - \frac{1}{\langle k\rangle}\right)\right] + \frac{1}{2}\left\{\left[\rho_r - \rho_0\left(\frac{\langle k^2\rangle}{\langle k\rangle^2} - \frac{1}{\langle k\rangle}\right)\right]^2 + 4\rho_r\rho_0\left(1 - \frac{1}{\langle k\rangle}\right)\right\}^{1/2}. \tag{A.6}$$

Appendix B. The calculation of the final epidemic size ($r(\infty)$)

Let us start form Eqs. (2a) and (2b)

$$\dot{s}_k(t) = -s_k(t)\sum_l \left((1-\lambda)\tau\frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right) i_l(t), \tag{B.1}$$

$$\dot{i}_k(t) = s_k(t)\sum_l \left((1-\lambda)\tau\frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right) i_l(t) - g i_k(t), \tag{B.2}$$

and assume that

$$\lambda_k = \sum_l \left((1-\lambda)\tau\frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right) i_l(t). \tag{B.3}$$

Integrating Eq. (B.1) we obtain

$s_k(t) = s_k(0)\exp(-\Phi_k(t))$ with $s_k(0) = N_k/N$

where

$$\Phi_k(t) = \int_0^t \lambda_k(s)\mathrm{d}s. \tag{B.4}$$

If Eqs. (B.1) and (B.2) are added together, and integrated from 0 to t, and the boundary conditions applied at 0 and $t \to \infty$, the following expression is obtained for the final size of the epidemic of nodes with k connections ($r_k = r_k(\infty)$):

$$\frac{r_k}{g} = \int_0^\infty i_k(s)\mathrm{d}s. \tag{B.5}$$

If in Eq. (B.4) for λ_k Eq. (B.3) is used and $t \to \infty$ is assumed then

$$\Phi_k(\infty) = \sum_l \left((1-\lambda)\tau\frac{k(l-1)}{\langle k\rangle} + \lambda\beta\right)\frac{r_l}{g}, \tag{B.6}$$

however $s_k(\infty) = N_k/N - r_k = (N_k/N)\exp(-\Phi_k(\infty))$ which gives

$$r_k(\infty) = N_k/N(1-\exp(-\Phi_k(\infty))). \tag{B.7}$$

If in Eq. (B.6) for r_k Eq. (B.7) is used then

$$\Phi_k(\infty) = \sum_l \left(\frac{(1-\lambda)\tau}{g}\frac{k(l-1)}{\langle k\rangle} + \frac{\lambda\beta}{g}\right)(1-\exp(-\Phi_l(\infty)))\frac{N_l}{N}. \tag{B.8}$$

Expanding the sum in Eq. (B.8) we obtain

$\Phi_k(\infty) = k\alpha + \alpha_r$

where

$$\alpha = \sum_l \frac{(1-\lambda)\tau}{g\langle k\rangle}(1-\exp(-l\alpha-\alpha_r))(l-1)\frac{N_l}{N}, \tag{B.9}$$

$$\alpha_r = \sum_l \frac{\lambda\beta}{g}(1-\exp(-l\alpha-\alpha_r))\frac{N_l}{N}. \tag{B.10}$$

The final epidemic size for the whole population is given by

$$r(\infty) = \sum_k r_k(\infty) = \sum_k N_k/N(1 - \exp(-k\alpha - \alpha_r)) = \sum_k p(k)(1 - \exp(-k\alpha - \alpha_r)), \quad \text{(B.11)}$$

and if in general $\langle f(k)\rangle = \sum_k f(k)p(k)$ the following parametric equations are obtained:

$$r(\infty) = \langle 1 - \exp(-k\alpha - \alpha_r)\rangle, \quad \text{(B.12)}$$

where α and α_r are determined by

$$\rho_r = \alpha_r/\langle 1 - \exp(-k\alpha - \alpha_r)\rangle, \quad \text{(B.13)}$$

$$\rho_0 = \alpha\langle k\rangle^2/\langle (k-1)(1 - \exp(-k\alpha - \alpha_r))\rangle, \quad \text{(B.14)}$$

with $\rho_r = \lambda\beta/g$ and $\rho_0 = (1-\lambda)\tau\langle k\rangle/g$ the transmission potentials of the two different transmission mechanisms, as previously defined. It is worth noting that Eqs. (B.9), (B.10) and Eqs. (B.13), (B.14) are equivalent, and for the degree distribution ($p(k) = 2m^2/k^3$ ($k \geqslant m$)) used in this paper, the sums in Eqs. (B.13) and (B.14) are convergent. Therefore, various values of α and α_r generate well defined finite values for ρ_r, ρ_0 and $r(\infty)$.

References

[1] R.M. Anderson, R.M. May, Infectious Diseases of Humans, Oxford Science, Oxford, 1991.
[2] R. Albert, A.L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47.
[3] M.E.J. Newman, Assortative mixing in networks, Phys. Rev. Lett. 89 (2002) 208701.
[4] M.E.J. Newman, Mixing patterns in networks, Phys. Rev. E 67 (2003) 026126.
[5] R.M. May, R.M. Anderson, Transmission dynamics of HIV infection, Nature 326 (1987) 137.
[6] R.M. May, A.L. Lloyd, Infection dynamics on scale-free networks, Phys. Rev. E 64 (2001) 066112.
[7] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (2001) 3200.
[8] H.J. Jones, M.S. Handcock, An assessment of preferential attachment as a mechanism for human sexual network formation, Proc. R. Soc. Lond. B 270 (2003) 1123.
[9] M.P.H. Stumpf, C. Wiuf, R.M. May, Subnets of scale-free networks are not scale-free: sampling properties of networks, Proc. Natl. Acad. Sci. USA 102 (2005) 4221.
[10] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, Y. Aberg, The web of human sexual contacts, Nature 411 (2001) 907.
[11] R.M. Christley, S.E. Robinson, R. Lysons, N.P. French, Network analysis of cattle movement in Great Britain, in: Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine, Nairn, Scotland, 2005, p. 234.
[12] R.R. Kao, L. Danon, D.M. Green, I.Z. Kiss, Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain, Proc. R. Soc. Lond. B, in press, doi:10.1098/rspb.2006.3505.
[13] R. Albert, H. Jeong, A.-L. Barabási, Diameter of the World-Wide Web, Nature 401 (1999) 130.
[14] D.J. Watts, S.H. Strogtaz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440.
[15] M. Barthélemy, A. Barrat, R. Pastoras-Satorras, A. Vespignani, Velocity and hierarchical spread of epidemic outbreaks in scale-free networks, Phys. Rev. Lett. 92 (2004) 178701.
[16] N. Madar, T. Kalisky, R. Cohen, D. ben-Avraham, S. Havlin, Immunization and epidemic dynamics in complex networks, Eur. Phys. J. B 38 (2004) 269.
[17] M.J. Keeling, The effect of local spatial structure on epidemiological invasions, Proc. R. Soc. Lond. B 266 (1999) 859.
[18] K.T.D. Eames, M.J. Keeling, Contact tracing and disease control, Proc. R. Soc. Lond. B 270 (2003) 2565.
[19] J.L. Sorenson, L.A. Wermuth, D.R. Gibson, K.-H. Choi, J.R. Guydish, S.L. Batki, Preventing AIDS in Drugs Users and Their Sexual Partners, Guilford, New York, NY, 1991.

[20] O. Diekmann, J.A.P. Heesterbeek, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, John Wiley & Sons Ltd, Chichester, UK, 2000.
[21] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509.
[22] A. Barbour, D. Mollison, Epidemics and random graphs, in: J.P. Gabriel, C. Lefevre, P. Picard (Eds.), Stochastic Processes in Epidemic Theory, Springer, New York, 1990, p. 86.
[23] N.T.J. Bailey, The Mathematical Theory of Infectious Diseases and Its Applications, Griffin, London, 1975.
[24] M.E.J. Newman, Properties of highly clustered networks, Phys. Rev. E 68 (2003) 026121.
[25] M.J. Keeling, B.T. Grenfell, Individual-based perspectives on R_0, J. Theor. Biol. 203 (2000) 51.
[26] O. Diekmann, M.C.M. De Jong, J.A.J. Metz, A deterministic epidemic model taking account of repeated contacts between the same individuals, J. Appl. Probab. 35 (2) (1998) 448.
[27] I.Z. Kiss, D.M. Green, R.R. Kao, Infectious disease control using contact tracing in random and scale-free networks, J. R. Soc. Interface 3 (2006) 55.
[28] R.R. Kao, The role of mathematical modeling in the control of the 2001 FMD epidemic in the UK, Trends Microbiol. 10 (2002) 279.
[29] R.R. Kao, The impact of local heterogeneity on alternative control strategies for foot-and-mouth disease, Proc. R. Soc. Lond. B 270 (2003) 2557.
[30] I.Z. Kiss, D.M. Green, R.R. Kao, Disease contact tracing in random and clustered networks, Proc. R. Soc. Lond. B 272 (2005) 1407.