Bull Math Biol (2011) 73:2305–2321
DOI 10.1007/s11538-010-9623-3

ORIGINAL ARTICLE

# The Final Size of an Epidemic and Its Relation to the Basic Reproduction Number

**Viggo Andreasen**

Received: 31 August 2010 / Accepted: 20 December 2010 / Published online: 6 January 2011


**Abstract** We study the final size equation for an epidemic in a subdivided population with general mixing patterns among subgroups. The equation is determined by a matrix with the same spectrum as the next generation matrix and it exhibits a threshold controlled by the common dominant eigenvalue, the basic reproduction number $\mathcal{R}_0$: There is a unique positive solution giving the size of the epidemic if and only if $\mathcal{R}_0$ exceeds unity. When mixing heterogeneities arise only from variation in contact rates and proportionate mixing, the final size of the epidemic in a heterogeneously mixing population is always smaller than that in a homogeneously mixing population with the same basic reproduction number $\mathcal{R}_0$. For other mixing patterns, the relation may be reversed.

**Keywords** Heterogeneous mixing · Final size relation

## 1 Introduction

Recent focus on pandemic planning and control of emerging diseases has spawned renewed interest in the final size of epidemics, and in particular in the impact of behavioral heterogeneities (Wallinga et al. 2006; Mossong et al. 2008; Kretzschmar and Mikolajczyk 2009). Classical epidemic models with homogeneous mixing predicts an unrealistically large final size of a single epidemic and many modelers have included detailed accounts for heterogeneities in contact structure to obtain a more realistic outcome (Elveback et al. 1976; Eubank et al. 2004;

V. Andreasen (✉)
Department of Science, Roskilde University, 4000 Roskilde, Denmark
e-mail: viggo@ruc.dk

V. Andreasen
Fogarty International Center, NIH, Bethesda, MA, USA

Ferguson et al. 2005; Longini et al. 2005; Arinaminpathy and McLean 2008). Interest in the final size problem stems from other contexts as well. For epidemics on networks, the size of the giant cluster corresponds to the final size and behavioral heterogeneities, reflected in the degree-distribution, are in fact the core of many network studies (Keeling 1999; Pastor-Satorras and Vespignani 2001; Eames and Keeling 2002; Newman 2002). The final size problem is also pivotal in the "burn-out approximation" that separates the fast (epidemic) time scale from slower processes such as host genetics (Gillespie 1975), host demography (May 1985; Andreasen and Frommelt 2005), or influenza-drift (Andreasen 2003; Boni et al. 2004; Andreasen and Sasaki 2006).

For a deterministic epidemic in a closed, homogeneous population, the final size equation gives the number (or frequency) of susceptible hosts at the end of the epidemic and involves only a single parameter, the basic reproduction number $\mathcal{R}_0$. The equation may be derived by studying how the size of the infectious class varies with the size of the susceptible class through the epidemic—an approach that dates back to Kermack and McKendrick (1927). The original Kermack–McKendrik-equation, and hence the final epidemic size is not affected by heterogeneities in the temporal or host-specific intensity of infectivity, in the sense that these quantities only change the equation through their effect on $\mathcal{R}_0$ (Ma and Earn 2006). Heterogeneity in susceptibility, however, introduces nonlinearities that qualitatively change the structure of the equation and epidemic size. A generalization of the final size relation to heterogeneous susceptibility was first obtained by Gart (1968) for the case where only susceptibility varies. Ball (1985) generalized Gart's result to stochastic epidemics and proved that for fixed $\mathcal{R}_0$ the final size of a Gart epidemic is always smaller than the epidemic in a homogeneous population. Modern reformulations of Gart's results extending them to proportionate mixing may be found in Diekmann and Heesterbeek (2000), Sect. 6.4 and in Dwyer et al. (2000), Andreasen (2003), Arino et al. (2007), Brauer (2008), Volz (2008a, 2008b). Diekmann and Heesterbeek (2000), Sect. 6.3 derive the final size relation for general mixing patterns but do not analyze its structure in detail. Corless et al. (1996) relate the final size equation to the Lambert W-function.

For heterogeneous populations the basic reproduction number $\mathcal{R}_0$ is defined as the dominant eigenvalue of the "next generation matrix" describing disease spread during the initial phase of an epidemic in a fully susceptible population (Diekmann et al. 1990). Though defined in terms of the transmission conditions at the onset of an epidemic, the basic reproduction number must characterize the final epidemic size in homogeneously as well as heterogeneously mixing populations at least in the trivial sense that the magnitude of $\mathcal{R}_0$ controls whether an epidemic occurs or not. We explore two aspects of this observation in detail. Firstly, we discuss how $\mathcal{R}_0$ structures the final size equation. Secondly, we show how heterogeneities can alter the epidemic size for fixed $\mathcal{R}_0$.

The advantage of using $\mathcal{R}_0$ in the characterization of epidemic size is that much effort has been devoted to quantifying the transmissibility and $\mathcal{R}_0$ for many important pathogens at the onset of pandemics or other disease out-breaks (Anderson et al. 1996; Ferguson et al. 2001; Riley et al. 2003; Mills et al. 2004). The disadvantage is that the empirical relation between $\mathcal{R}_0$ and the final size is poorly understood (Tildesley and Keeling 2009).

In the next section, we review what is known about the general final size equation and discuss its connection to the next generation matrix and $\mathcal{R}_0$. We then prove—by three quite different methods—that the final size equation determines a unique value for the epidemic size. From a purely mathematical stand point, one proof would suffice. However, the proofs allow us to discuss different aspects of the final size relation. The subsequent section presents our second main finding, namely a generalization of Ball's comparison between the final size in homogeneously and heterogeneously mixing populations. In particular, we show that Ball's result holds for a wider class of proportionate mixing patterns including situations where variation in susceptibility and infectivity is due only to heterogeneity in contact rates. We conclude with an example showing that for other types of proportionate mixing, heterogeneity may increase the epidemic size.

## 2 The Final Size of an Epidemic

The essential features of the heterogeneous mixing problem are captured in an epidemic model with $n$ groups of size $N_k$, $k = 1, \ldots, n$. For simplicity, we study only the case where hosts become infectious immediately upon infection and recover at a fixed rate which we may set to unity with out loss of generality and we express all population sizes relative to the total population such that $\sum N_k = 1$. We assume that host mortality and recruitment during the epidemic period may be neglected, so that the state of the population is characterized by the fraction of the population in each group that is susceptible and infectious, $S_k$, $I_k$, $k = 1, \ldots, n$. Letting $b_{kj} I_j$ denote the rate at which infectious individuals in group $j$ infect hosts in group $k$, the epidemic model becomes

$$\dot{S}_k = -S_k \sum_j b_{kj} I_j, \tag{1}$$

$$\dot{I}_k = S_k \sum_j b_{kj} I_j - I_k. \tag{2}$$

It is straightforward to see that the matrix

$$B = \begin{pmatrix} b_{11}N_1 & \ldots & b_{1n}N_1 \\ \vdots & & \vdots \\ b_{n1}N_n & \ldots & b_{nn}N_n \end{pmatrix} \tag{3}$$

is the co-called next-generation matrix for the model, and hence that the basic reproduction number $\mathcal{R}_0$ by definition is the dominant eigenvalue of $B$. For a general treatment of the problem, see Diekmann and Heesterbeek (2000) or Arino et al. (2007) and for an extension to diseases where infectivity varies during the infection period see Diekmann et al. (2010).

In this paper, we will focus on directly transmitted diseases such as influenza and measles so it is natural to assume that $B$ is positive—or at least nonnegative and primitive—allowing us to use Perron–Frobenius theory.

To reflect the impact of a single epidemic in a totally susceptible population, we set $S_k(0) = N_k$ and assume that $I_k(0)$ is positive with $I_k(0) \approx 0$, but a generalization to a situation with preexisting immunity is straightforward (Andreasen 2003). Following Brauer (2008), one can show that $I_k(t) \to 0$ for $t \to \infty$ while there exists some number $S_k(\infty)$ such that $S_k \to S_k(\infty)$ for $t \to \infty$.

The value of $S_k(\infty)$ may be determined by integration over the entire epidemic period, which yields

$$\log S_k(\infty) - \log S_k(0) = \int_0^\infty \dot{S}/S \, dt = -\sum_j b_{kj} \int_0^\infty I_j \, dt, \tag{4}$$

$$S_k(\infty) - S_k(0) = \int_0^\infty \dot{S}_k + \dot{I}_k \, dt = -\int_0^\infty I_k \, dt. \tag{5}$$

We can now characterize the outcome of the epidemic in terms of the fraction of susceptible hosts that did not get infected during the epidemic $\sigma_k = S_k(\infty)/S_k(0) = S_k(\infty)/N_k$. The quantities $\sigma = (\sigma_1, \ldots, \sigma_n)$ give the common expressions for the size of the epidemic in that the attack rate in group $k$ is $x_k = 1 - \sigma_k$ and the final size of the epidemic in group $k$ is $(1 - \sigma_k)N_k$. Letting $x$ and $N$ denote the vectors $(x_1, \ldots, x_n)$ and $(N_1, \ldots, N_n)$, the final size of the epidemic is $x^T N$ in the whole population. Here and throughout the paper, all vectors are considered to be column vectors and $T$ denotes the transpose.

In terms of $\sigma$, the size of the epidemic is a solution to the coupled implicit equations

$$0 = \sum_j b_{kj} N_j (1 - \sigma_j) + \log \sigma_k = F_k(\sigma), \quad k = 1, \ldots, n.$$

In matrix notation with

$$A = \begin{pmatrix} b_{11}N_1 & \ldots & b_{1n}N_n \\ \vdots & & \vdots \\ b_{n1}N_1 & \ldots & b_{nn}N_n \end{pmatrix}, \tag{6}$$

log meaning the coordinatewise log-function, and $0 = (0, \ldots, 0)$, the final size equation becomes

$$0 = A(1 - \sigma) + \log \sigma = F(\sigma). \tag{7}$$

We will refer to the $kj$'th element in $A$ as $a_{kj}$. Clearly, the matrix $A$ may be obtained by "transposing" the population sizes in the $B$-matrix. In mathematical terms, the matrix $A$ is *similar* to $B$ because with the diagonal matrix $\Delta = \mathrm{diag}(N_1, \ldots, N_n)$ we have

$$A = \Delta^{-1} B \Delta.$$

Thus, the two matrices have the same spectrum and in particular $\mathcal{R}_0$ is the dominant eigenvalue of $A$. Since $B$ is assumed to be positive (nonnegative and primitive) so is $A$. It is remarkable that the linear part of the final size equation is so closely related

to the next generation matrix $B$ although $B$ only describes the conditions at the onset of the epidemic.

Taking the coordinatewise exp of (7) yields the alternative version of the final size equation in $x = 1 - \sigma$

$$x = 1 - \exp(-Ax). \tag{8}$$

Equation (8) may be interpreted as a probabilistic identity as $x_k$ is the probability that an individual in group $k$ becomes infected during the epidemic while $\exp(-\sum a_{kj}x_j)$ gives the probability of remaining susceptible during the entire epidemic (Diekmann and Heesterbeek 2000; Wallinga et al. 2006). With the probabilistic interpretation it is clear how the coordinate transformation $\Delta$ arises because the initial rate of infection in group $k$ scales the size of the population in group $k$ that can be *infected* while the intensities $a_{kj}x_j$ scale with the size of the *infecting* group.

## 3 The Solutions to the Final Size Equation

A direct computation shows that the equation has a trivial solution at $(1, \ldots, 1)$ corresponding to the situation where no epidemic has occurred and the population remains in the disease-free state. From a mathematical view point, it is not obvious how many other feasible solutions there may exist to (7) but biological intuition suggest the following.

**Conjecture 1** *Equation* (7) *has a single solution in the open unit cube* $(0, 1)^n$ *if* $\mathcal{R}_0 > 1$ *and none if* $\mathcal{R}_0 < 1$.

We will prove three theorems specifying additional conditions to the conjecture. The first theorem is proved by a simple and illustrative geometric argument generalizing the usual analysis of the homogeneous final size equation to the case of two mixing groups. The second theorem is general. The proof builds on bifurcation theory applied to the dynamical system $x \mapsto 1 - \exp(-Ax)$ and shows the similarity with the bifurcation structure for models of endemic diseases. The last version of the theorem which goes back to Gart (1968) and Ball (1985) applies only to the case of proportionate mixing, but the proof is constructive and includes a characterization of the dynamics during the epidemic.

**Theorem 1** *The final size equation for two mixing groups* (7) *has a unique solution* $(\sigma_1, \sigma_2)$ *in the open unit square* $(0, 1)^2$ *if and only if* $\mathcal{R}_0 > 1$.

*Proof* For $n = 2$, we can rearrange the final size equations as

$$s_2(\sigma_1) = \sigma_2 = \frac{1}{a_{12}} \log \sigma_1 + \frac{a_{11}}{a_{12}}(1 - \sigma_1) + 1,$$

$$s_1(\sigma_2) = \sigma_1 = \frac{1}{a_{21}} \log \sigma_2 + \frac{a_{22}}{a_{21}}(1 - \sigma_2) + 1,$$

so geometrically we are looking for intersections of the two curves $\sigma_1 = s_1(\sigma_2)$ and $\sigma_2 = s_2(\sigma_1)$ in the open unit square.

The following properties hold for $s_k$:

1. $s_k(1) = 1$. (The existence of a disease-free state).
2. $s_l'(1) = (1 - a_{kk})/a_{kl}$, $k = 1, 2$, $l = 2, 1$.
3. $s_k''(\sigma_l) < 0$ for $0 < \sigma_l < 1$.
4. $s_k(\sigma_l) \to -\infty$ for $\sigma_l \to 0^+$.

To see the existence of a solution for $\mathcal{R}_0 > 1$, we distinguish between the case (i) where at least one diagonal term $a_{kk}$ exceeds unity and the case (ii) where both diagonal terms are smaller than one.

Case (i) $a_{kk} > 1$ for $k = 1$ or 2, is the situation where transmission in at least one of the subpopulations alone can sustain the epidemic. As the dominant eigenvalue must exceed the magnitude of its diagonal elements in a positive matrix, case (i) implies $\mathcal{R}_0 > 1$. Assume $a_{11} > 1$ so that $s_2'(1) < 0$. The curve $\sigma_2 = s_2$ has a negative slope at $(1, 1)$ and by Property (4) combined with the continuity of $s_2$ we conclude that the curve will cross the entire unit square from bottom to top. This ensures the existence of an internal intersection of the two curves; see Fig. 1.

In case (ii) where $a_{kk} < 1$ for $k = 1, 2$, the two curves will intersect if the tangent of $s_2$ lies above the tangent of $s_1$ at the point of intersection $(1, 1)$; see Fig. 1. The tangent condition translates to

$$1 > s_1'(1)s_2'(1) = \frac{1 - a_{11}}{a_{12}} \frac{1 - a_{22}}{a_{21}}.$$

This condition may be rewritten as $p_A(1) < 0$, where $p_A(u)$ is the characteristic polynomial of the matrix $A$. The characteristic polynomial $p_A(u)$ is negative for $\operatorname{tr} A/2 < u < \mathcal{R}_0$ and since we have that $2 > a_{11} + a_{22} = \operatorname{tr} A$, we conclude that the condition is satisfied iff $\mathcal{R}_0 > 1$.

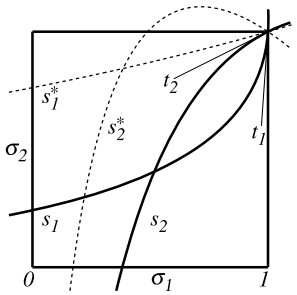


**Fig. 1** Geometric solution of the final size equation for an epidemic model with general mixing between two subgroups. The condition for the final size is expressed in terms of $\sigma_k$, the fraction of the susceptible population in group $k$ that remains susceptible at the end of the epidemic. The curve $\sigma_2 = s_1(\sigma_1)$ determines the final size condition for group 2 for known $\sigma_1$ and does not depend directly on the transmission into group 2. The intersection of $s_1$ and $s_2$ gives the final size in the two groups. An intersection exists iff the tangent $t_2$ lies above the tangent $t_1$. The tangent condition is met if and only if the reproduction number $\mathcal{R}_0$ exceeds unity. The *curves* $s_1$ and $s_2$ represent a case where none of the subpopulations alone can support an epidemic. The *curve* $s_2^*$ shows a case where group 1 alone can support an epidemic, while $s_1^*$ combined with $s_2$ shows a situation where no epidemic will occur

To see that the two curves have at most one intersection in addition to the trivial intersection at $(1, 1)$, observe that by the (down) convexity of $s_2(\sigma_1)$ all points on the curve must lie above the secant between the intersection and the point $(1, 1)$, while all points on $s_1(\sigma_2)$ lie below the secant. This excludes the existence of multiple intersections.

If $\mathcal{R}_0 < 1$, the two curves cannot intersect in the interior of the unit square. By its convexity, the curve $s_2(\sigma_1)$ will remain below its tangent at $(1, 1)$ while $s_1(\sigma_2)$ remains above its tangent. This completes the proof of Theorem 1. □

A generalization of the proof to higher dimension $n > 2$ is not straightforward and a more promising avenue is to apply bifurcation theory to (7) treating $\mathcal{R}_0$ as a bifurcation parameter. In other words for a final size matrix $A_1$ with dominant eigenvalue 1, we consider as $\mathcal{R}_0$ changes the structure of the solutions to (7) where $A = \mathcal{R}_0 A_1$. We first note that the Jacobian of $F$ at the no-epidemic point $DF(1, \ldots, 1)$ is singular for $\mathcal{R}_0 = 1$ suggesting that the interior root bifurcates off the trivial root when $\mathcal{R}_0$ passes through 1 in a manner similar to the transcritical bifurcation that is known from classical epidemic models with endemic equilibria.

The analysis proceeds through a series of lemmas and we postpone the statement the theorem until we can motivate its exact formulation. The first step along the analysis is the following lemma.

**Lemma 1** *If $\mathcal{R}_0 < 1$, (7) has no solutions in the interior of $(0, 1)^n$.*

*Proof* Assume—to obtain a contradiction—that $\sigma \in (0, 1)^n$ solves (7). By the Perron–Frobenius theorem, the left-hand eigenvector $w$ corresponding to the dominant eigenvalue of $A$ may be chosen to be positive in all coordinates and satisfy $w^T 1 = 1$. Taking the inner product of $w$ and (7) gives

$$\begin{aligned} 0 = w^T 0 &= w^T \log\sigma + w^T A(1-\sigma) \\ &= w^T \log\sigma + \mathcal{R}_0 w^T (1-\sigma) \\ &\le \log w^T\sigma + \mathcal{R}_0\big(1 - w^T\sigma\big) \quad \text{by Jensen's inequality} \\ &\le (\mathcal{R}_0 - 1)\big(1 - w^T\sigma\big) < 0 \quad \text{by the inequality } \log y \le y - 1, \end{aligned}$$

which is a contradiction, and hence proves Lemma 1. □

We next show the existence of a root in $(0, 1)^n$ for $\mathcal{R}_0 > 1$ and small. The existence of a nontrivial root crossing $(1, \ldots, 1)$ at $\mathcal{R}_0 = 1$ follows naturally from the singularity of $DF$ so the main issue is to demonstrate that the root be positive. We need to impose a condition that ensures that we are in "the generic case."

**Lemma 2** *Let $v_1$ be a positive (right-hand) eigenvector of $A$ corresponding to the dominant eigenvalue $\mathcal{R}_0$ and let $u_1$ be the vector $v_1$ squared coordinatewise. If $u_1$ is linearly independent of the set of subdominant eigenvectors $v_2, \ldots, v_n$, then for $\mathcal{R}_0 > 1$ and $\mathcal{R}_0$ sufficiently small there exists a solution to (7) in the open unit cube $(0, 1)^n$.*

*Remark* The condition on $u_1$ is satisfied for almost all final size matrices $A$ and ensures that we are in the generic case. If $A$ does not have a full spectrum, generalized eigenvectors should be added to the subdominant eigenvectors to form a full $n-1$ dimensional basis.

*Proof* To see the existence of a feasible solution for $\mathcal{R}_0 > 1$, we perform an asymptotic expansion of (7) in $(\mathcal{R}_0 - 1)$ and show that there exists a solution of the form

$$1 - \sigma = x = (\mathcal{R}_0 - 1)\alpha_0 v_1 + (\mathcal{R}_0 - 1)^2 \sum_k \alpha_k v_k + O\big((\mathcal{R}_0 - 1)^3\big),$$

where the coefficients $\alpha_0 \neq 0$ and $\alpha_2, \ldots, \alpha_n$ are to be determined. Letting $x^2$ and $\log x$ denote the coordinatewise operations on the vector and $\lambda_k$ the subdominant eigenvectors of $A$, (7) simplifies to

$$\begin{aligned} 0 &= Ax + \log(1-x) \\ &= Ax - x - \frac{1}{2}x^2 + \text{h.o.t.} \\ &= (\mathcal{R}_0 - 1)^2 \alpha_0 v_1 - \frac{1}{2}(\mathcal{R}_0 - 1)^2 \alpha_0^2 u_1 + (\mathcal{R}_0 - 1)^2 \sum_k \alpha_k (\lambda_k - 1) v_k \\ &\quad + O\big((\mathcal{R}_0 - 1)^3\big). \end{aligned}$$

Retaining only terms of order $(\mathcal{R}_0 - 1)^2$ and setting $\alpha'_k = \alpha_k/\alpha_0$, we observe that $\alpha_0$ and $\alpha'_k$ must solve the linear equation

$$v_1 = \frac{1}{2}\alpha_0 u_1 + \sum \alpha'_k (1 - \lambda_k) v_k.$$

Our assumption that the vectors on the right-hand side be linearly independent ensures the existence and uniqueness of a solution. Since $v_1$ is independent of $v_2, \ldots, v_n$, we know that $\alpha_0 \neq 0$ and from Lemma 1 we conclude that $\alpha_0 > 0$. This shows the existence of a small, positive root in $x$, and hence of a root $\sigma = 1 - x \in (0,1)^n$ for $0 < \mathcal{R}_0 - 1 \ll 1$.

We next note

**Lemma 3** *Additional roots to* (7) *can not arise through bifurcations in the interior of the unit cube* $(0,1)^n$.

To see this observe that except for the trivial root, (7) has exactly the same roots as the vector-equation

$$0 = a_{kk} + \sum_{j \neq k} a_{kj} \frac{1 - \sigma_j}{1 - \sigma_k} + \frac{\log \sigma_k}{1 - \sigma_k}, \quad k = 1, \ldots, n. \tag{9}$$

Consider the left-hand sides of (9) as a vector valued function $G(\sigma_1, \ldots, \sigma_n)$, set $g(y) = \log y/(1-y)$, $g'_k = g'(\sigma_k)$, and $\bar{\sigma}_k = 1 - \sigma_k$. The Jacobian of $G$ has determi-

nant

$$\det DG$$
$$= \det \begin{pmatrix} \sum_{k\neq 1} \frac{a_{1k}\bar{\sigma}_k}{\bar{\sigma}_1^2} + g_1' & -\frac{a_{12}}{\bar{\sigma}_1} & \cdots & -\frac{a_{1n}}{\bar{\sigma}_1} \\ -\frac{a_{21}}{\bar{\sigma}_2} & \sum_{k\neq 2} \frac{a_{2k}\bar{\sigma}_k}{\bar{\sigma}_2^2} + g_2' & \cdots & -\frac{a_{2n}}{\bar{\sigma}_2} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{\bar{\sigma}_n} & -\frac{a_{n2}}{\bar{\sigma}_n} & \cdots & \sum_{k\neq n} \frac{a_{nk}\bar{\sigma}_k}{\bar{\sigma}_n^2} + g_n' \end{pmatrix}$$

$$= \Pi \bar{\sigma}_k^{-1} \det \begin{pmatrix} \sum_{k\neq 1} \frac{a_{1k}\bar{\sigma}_k}{\bar{\sigma}_1} + g_1'\bar{\sigma}_1 & -\frac{a_{12}\bar{\sigma}_2}{\bar{\sigma}_1} & \cdots & -\frac{a_{1n}\bar{\sigma}_n}{\bar{\sigma}_1} \\ -\frac{a_{21}\bar{\sigma}_1}{\bar{\sigma}_2} & \sum_{k\neq 2} \frac{a_{2k}\bar{\sigma}_k}{\bar{\sigma}_2} + g_2'\bar{\sigma}_2 & \cdots & -\frac{a_{2n}\bar{\sigma}_n}{\bar{\sigma}_2} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}\bar{\sigma}_1}{\bar{\sigma}_n} & -\frac{a_{n2}\bar{\sigma}_2}{\bar{\sigma}_n} & \cdots & \sum_{k\neq n} \frac{a_{nk}\bar{\sigma}_k}{\bar{\sigma}_n} + g_n'\bar{\sigma}_n \end{pmatrix}.$$

Since $g'(y) > 0$ by the convexity of $\log y$, the matrix in the last line is diagonally dominant and hence regular. It follows that $\det DG \neq 0$.

Since $DG$ is everywhere regular the implicit function theorem shows that the number of solutions to $G = 0$ cannot change in the open unit-cube. This completes the proof of Lemma 3. □

We next need to exclude the possibility of roots entering through the boundaries.

**Lemma 4** *Equation* (7) *has no nontrivial roots on the boundary of the unit cube* $(0, 1)^n$.

*Proof* Assume that $\sigma$ is a point on the boundary of the unit cube and solves (7). Clearly, $\sigma_k \neq 0$ for all $k$. So, we must have $\sigma_k = 1$ for some coordinate $k$. Inspection of the $k$th row in the equation shows that since $\log \sigma_k = 0$, we have $\sum a_{kl}(1-\sigma_l) = 0$, implying that $\sigma_l = 1$ for all $l$.

Finally, we need to exclude the possibility of additional roots crossing through the trivial root. Additional bifurcations at $(1, \ldots, 1)$ may occur only as subdominant eigenvalues pass through unity, however, from the Perron–Frobenius theorem we know that the associated eigenvectors can not be positive on all coordinates. Applying the same method as in the proof of Lemma 2 to the subdominant eigenvector, we find that such roots cannot enter the unit cube $(0, 1)^n$.

This concludes our analysis of the general case and we have proved the following theorem.

**Theorem 2** *Let* $v_1, \ldots, v_m$ *denote the set of eigenvectors and generalized eigenvectors of the final size matrix* $A$ *and* $u_1, \ldots, u_m$ *the set of these vectors squared coordinatewise. If each* $u_k$ *is linearly independent of the set of all eigenvectors and generalized eigenvectors excluding* $v_k$, *then* (7) *has a single solution in the open unit cube* $(0, 1)^n$ *if* $\mathcal{R}_0 > 1$ *and none if* $\mathcal{R}_0 < 1$.

The condition on the eigenvectors is satisfied for almost all final size matrices $A$ and may be considered the generic case. It is unclear if the theorem will also hold in the non-generic case. Nongeneric cases must be of dimension $n \geq 3$ because in the two-dimensional case the positive vector $u_1$ cannot be proportional to the subdominant eigenvector $v_2$ as the coordinates of $v_2$ have opposite sign.

For the special case of proportionate mixing $b_{kj} = p_k q_j$, the situation is less complex and the interior root can be found in a simple way as the final size equation (7) separates to

$$\begin{aligned} 0 &= p_k \log \sigma_1 - p_1 \log \sigma_k, \quad k = 2, \ldots, n, \\ 0 &= \sum_k q_k N_k \left(1 - \sigma_1^{p_k/p_1}\right) + \log \sigma_1. \end{aligned} \tag{10}$$

Thus, the final size of the Gart epidemic is determined by the roots of (10) (Gart 1968; Ball 1985). Formally this separation arises because the linear terms in (7) form a matrix of rank 1. From a dynamical perspective, the separation is due to a simple power relationship between the sizes of the susceptible subpopulations in that $S_k(t)/S_k(0) = (S_1(t)/S_1(0))^{p_k/p_1}$. This observation, which was the key to Gart's original analysis, allows an explicit description of the dynamics during the epidemic in terms of $\sigma_1(t) = S_1(t)/S_1(0)$. Clearly, $\sigma_1$ is a monotonically decreasing variable throughout the epidemic; for details see Andreasen (2003).

Gart's characterization of the final size offers a simpler proof of our conjecture as it suffices to show the following theorem.

**Theorem 3** *The Gart-equation* (10) *has a unique nontrivial root in the open interval* $(0, 1)$, *if the reproduction number exceeds unity, i.e.*,

$$\mathcal{R}_0 = \sum_k q_k p_k N_k > 1. \tag{11}$$

*If* $\mathcal{R}_0 < 1$, *then the Gart-equation has no nontrivial roots in* $(0, 1)$.

For a proof, see Andreasen (2003).

## 4 The Effect of Heterogeneity on the Size of the Epidemic

For our comparison of epidemic size, it turns out to be convenient to express the epidemic size in terms of attack rates $x_k$ rather than susceptibles remaining uninfected, $x_k = 1 - \sigma_k$.

To see the effect of heterogeneous disease transmission on the final size of the epidemic, we compare the final size of the epidemic in the heterogeneous population to that of an epidemic with the same $\mathcal{R}_0$ in a homogeneously mixed population.

The final size in the homogeneously mixing population $\xi$ is determined by

$$\mathcal{R}_0 \xi + \log(1 - \xi) = 0 \tag{12}$$

while the final size in the heterogeneously mixing population is $x^T N = \sum x_k N_k$. The attack rates $x$ solve

$$Ax = h(x), \tag{13}$$

where $h(x) = -\log(1-x)$ coordinatewise.

We focus exclusively on the case of proportionate mixing $b_{kj} = p_k q_j$ in a fully susceptible population. Here, $\mathcal{R}_0 = \sum p_k q_k N_k$ and using Gart's characterization we can express the attack rate in group $k$ in terms of $\sigma_1$, the fraction of hosts in group 1 that are still susceptible after the epidemic, as $x_k = 1 - \sigma_1^{p_k/p_1}$.

For two extreme cases, the result is known from previous studies. For the case where only susceptibility varies $q_1 = \cdots = q_n$, Ball (1985) showed that heterogeneities in disease transmission decreases the final size of the epidemic $x^T N \leq \xi$. Conversely, Ma and Earn (2006) observed that if susceptibility is constant $p_1 = \cdots p_n$, then the final size of the epidemic depends only on the magnitude of $\mathcal{R}_0$, i.e., $\xi = x^T N$.

These two results cannot be combined directly. In fact, we shall give an example where heterogeneity in both susceptibility and infectivity gives a *larger* epidemic than one would observe in a homogeneously mixing population with the same transmissibility $\mathcal{R}_0$, so an additional constraint on $p$ and $q$ is needed. To motivate this constraint, we first discuss the situation where host heterogeneity is attributable to variation in contact rates (Hethcote and Yorke 1984). Assume that each host in group $k$ makes $c_k$ contacts per time unit and that these contacts are distributed among all groups in proportion to that group's contribution to the total number of contacts. Each host in group $k$ then makes

$$c_k \frac{c_j N_j}{\sum_i c_i N_i}$$

contacts with hosts in group $j$ so we have that

$$b_{kj} I_j = c_k c_j \frac{\beta I_j}{\sum_i c_i N_i},$$

where $\beta$ is the probability that a susceptible host become infected after a contact with an infected host. Host heterogeneity that arises in this way clearly is of the proportionate mixing type $q_k p_j$ but it in addition has the property that $p$ and $q$ are proportional vectors. We need a slightly weaker version of this property to prove the following theorem. □

**Theorem 4** *Let $\mathcal{R}_0$ denote the dominant eigenvalue of the next generation matrix*

$$B = \begin{pmatrix} p_1 q_1 N_1 & \dots & p_1 q_n N_1 \\ \vdots & & \vdots \\ p_n q_1 N_n & \dots & p_n q_n N_n \end{pmatrix}.$$

*and $A$ the corresponding final size matrix* (6)*, and let $x$ be the solution to* (13) *and $\xi$ the solution to* (12)*. If*

$$p_1 < p_2 \cdots < p_n \quad \textit{and} \quad q_1 < q_2 < \cdots < q_n,$$

*then for all positive vectors $N$ with $\sum N_k = 1$, the inequality $N^T x \le \xi$ will hold.*

*Proof* We first note that since $\mathcal{R}_0 \xi = h(\xi)$, by the (up) convexity of $h$ it suffices to show that $\mathcal{R}_0 N^T x \ge h(N^T x)$, since this shows that the number $N^T x$ must lie to the left of $\xi$.

Taking the inner product with $N$ on both sides of (13) and applying Jensen's inequality to $h$ gives

$$N^T A x = N^T h(x) \ge h\big(N^T x\big),$$

and it remains to see that $\mathcal{R}_0 N^T x \ge N^T A x$.

$$\begin{aligned} N^T A x - \mathcal{R}_0 N^T x &= \sum_k p_k N_k \sum_j q_j x_j N_j - \sum_j p_j q_j N_j \sum_k x_k N_k \\ &= \sum\sum_{kj} N_k N_j p_k p_j (q_j x_j / p_j - q_j x_k / p_k) \\ &= \operatorname{cov}_{pN}(q, x/p) \Big( \sum p_k N_k \Big)^2, \end{aligned}$$

where $\operatorname{cov}_{pN}$ is the covariance with respect to the normalized distribution $p_k N_k$.

The sign of the last expression depends on how the ratio of the attack rate to the susceptibility $x_k/p_k$ depends on the magnitude of the infectivity $q_j$. Since we are assuming proportionate mixing, the analysis of the Gart equation shows that $x_k = 1 - s^{p_k}$ for some $s$. Specifically, we have that $s = \sqrt[p_1]{\sigma_1}$ where $\sigma_1$ solves (10) but it suffices to observe that $0 < s < 1$. The function $\phi(y) = (1 - s^y)/y$ is decreasing due to the concavity (down) of the log-function. By assumption, we now have that $q_1 < \cdots < q_n$ and $\phi(p_1) = x_1/p_1 > \cdots > \phi(p_n) = x_n/p_n$. If follows that $\operatorname{cov}_{pN}(q, x/p) < 0$. We conclude that $\mathcal{R}_0 N^T x \ge N^T A x$ which completes our proof. □

## 5 A Numerical Example

As a specific example, we study how the susceptibility $p$ and the infectivity $q$ affect the ratio $\rho$ of the total attack rate in a homogeneous population $1 - \xi$ to the attack rate in a heterogeneous population $x^T N = \sum (1 - \sigma^{p_k}) N_k$ with the same reproduction number, i.e.,

$$\rho = \frac{\sum (1 - \sigma^{p_k}) N_k}{(1 - \xi)}. \tag{14}$$

For simplicity, we focus on a population consisting of two subpopulations of equal size $N_1 = N_2 = \frac{1}{2}$ and we will assume that subpopulation 1 has the smallest susceptibility and that it is fixed at unity so that $p = (1, p_2)$. For fixed reproduction ratio $\mathcal{R}_0$,

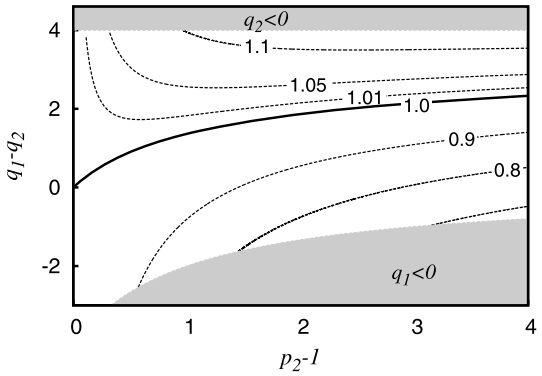


**Fig. 2** The final size of an epidemic in a heterogeneously mixing population relative to the size in a homogeneous population with the same basic reproduction number $\mathcal{R}_0 = 2$. The population consists of two subpopulation of the same size $N_1 = N_2 = \frac{1}{2}$. Mixing is of the proportionate mixing type and in subpopulation 1 infectivity and susceptibility are $p_1 = 1$ and $q_1$, respectively. In subpopulation 2, the corresponding quantities are $p_2$ and $q_2$. The *heavy curve* gives those parameter values where heterogeneity does not change the final size. In the *upper part* of the figure, heterogeneity leads to an increased epidemic size. In the *lower part* of diagram, the heterogeneous epidemic is the smaller than a epidemic in a homogeneous population. Only parameter values with $p_2 > 1$ and $q_1 - q_2 < 0$ can arise from heterogeneities that are solely caused by variation in contact rates

the mixing pattern is now determined by two parameters which we take to be $p_2 \geq 1$ and $q_1 - q_2$ with the additional requirement that since $q_2 \geq 0$ and $q_1 \geq 0$, we must have $\mathcal{R}_0/N_1 \geq q_1 - q_2 \geq \mathcal{R}_0/(p_2 N_2)$.

Level curves of $\rho$ in the $(p_2, q_1 - q_2)$-plane are characterized by the Gart-equation (10) combined with the size of $\mathcal{R}_0$, (11), and the level condition (14). The three conditions allow us to parameterize the level curves explicitly in terms of $\sigma$

$$p_2(\sigma) = \frac{\log(1 - [\rho(1-\xi) + (1-\sigma)N_1]/N_2)}{\log \sigma},$$

$$q_2(\sigma) = \frac{\log \sigma + \mathcal{R}_0(1-\sigma)}{p_2(\sigma)(1-\sigma)N_2 - (1-\sigma^{p_2})N_2},$$

$$q_1(\sigma) = \frac{\mathcal{R}_0 - q_2(\sigma)p_2(\sigma)N_2}{N_1},$$

which produces Fig. 2 for the case of $\mathcal{R}_0 = 2$. The heavy line $\rho = 1$ shows those parameter values where the final size of the epidemic is unaffected by mixing heterogeneities. From Ma and Earn (2006) we know that the vertical axis $p_2 = 1$ corresponding to constant susceptibility lie on $\rho = 1$ as well, and from Ball (1985) that $\rho \leq 1$ at $q_1 - q_2 = 0$ while we have just extended Ball's result to $\rho \leq 1$ for $q_1 - q_2 \leq 0$.

The maximal value of $\rho$ will depend on the structure of the population. In our example, it occurs in the limit where subpopulation 2 is highly susceptible $p_2 \to \infty$ while incapable of spreading the disease $q_2 = 0$. As we have fixed $\mathcal{R}_0 = 2$, we have $q_1 = \mathcal{R}_0/N_1$ and since subpopulation 2 does not contribute to disease-spread, the final attack-rate $1 - s$ in population 1 is determined the homogeneous final size equa-

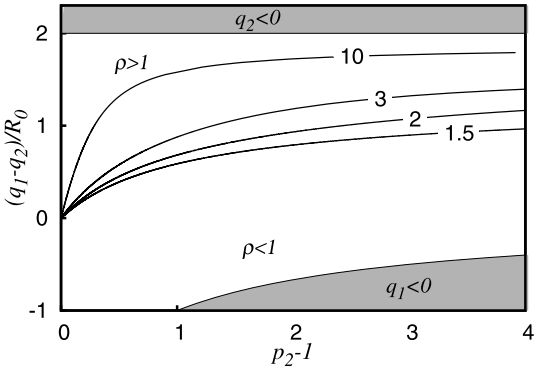


**Fig. 3** The effect of the basic reproduction number on the final size in a heterogeneously mixing population. The *curves* show those parameter-values where heterogeneity does not change the size of the epidemic ($\rho = 1$) for $\mathcal{R}_0 = 1.5, \ldots, 10$. All parameter values except $\mathcal{R}_0$ are the same as in Fig. 2

tion $\log s + q_1 N_1(1-s) = 0$, which gives $s = \xi$. The total attack-rate in the heterogeneous populations is now $(1-\xi)N_1 + N_2$ giving $\rho_{\max} = N_1 + N_2/(1-\xi)$. For $N_1 = N_2 = \frac{1}{2}$ and $\mathcal{R}_0 = 2$ we find $\rho_{\max} = 1.127$.

The minimal $\rho$ arises in the limit where all disease transmission is concentrated in the most susceptible population; since we have fixed $p_1 = 1$ this corresponds to the situation where $p_2 \to \infty$ while $p_2 q_2 N_2 \to \mathcal{R}_0$ and $q_1 = 0$, leading to an attack rate of $1-\xi$ in population 2 and vanishing attack rate in population 1. Thus, $\rho_{\min} = N_2 = 0.5$.

For increasing $\mathcal{R}_0$, the reduction in the final size gets more pronounced—this may be seen in Fig. 3 and by observing that $\rho \to 1$ as $1-\xi \to 1$.

## 6 Discussion

This paper presents new results on the structure of the final size equation for general epidemics and on the effect of heterogeneity on the size of epidemics with proportionate mixing.

The structure of the final size equation for an epidemic in a heterogeneous population consists of a fixed nonlinear term plus a linear part that depends on disease transmission. The linear part is similar to the next generation matrix describing the conditions at the onset of the epidemic, but while the rows of the next generation matrix scale with the sizes of the subpopulations, for the final size matrix it is the columns that scale. Since the final size matrix can be obtained from the next generation matrix by a coordinate transformation, the two matrices have the same spectrum. In particular, the common dominant eigenvalue $\mathcal{R}_0$ controls the epidemic threshold for the onset of the epidemic as well as the number of roots of the final size equation. This generalizes the dual role played by $\mathcal{R}_0$ in the classical homogeneous epidemic model.

Intuitively, one would expect that the final size equation uniquely specifies the size of the epidemic and the intuition is confirmed except for the well-known caveat that the final size equation will also yield a trivial solution corresponding to no epidemic. Since the final size equation does not contain information about the dynamical aspects of the system there is no natural way to distinguish between the two solutions. The

uniqueness of the internal solution shows that the magnitude of the epidemic and the distribution of infected hosts are independent of how the initial cases are distributed.

The existence of the nontrivial solution is governed by a threshold condition on the basic reproduction number $\mathcal{R}_0$ and the solution arises through the nondynamical component of a transcritical bifurcation off the no-epidemic root, in the sense that a nontrivial solution crosses through the trivial solution and into the biologically feasible region exactly when $\mathcal{R}_0$ passes unity.

The analysis of the final epidemic size and the impact of heterogeneity carries over to situations where the transmission classes are described by a discrete or continuous distribution of transmission types—or degree distribution in the terminology of epidemic networks. In particular, the final size of an epidemic on a scale-free network as analyzed by Pastor-Satorras and Vespignani (2001) may be seen as a special case of our analysis (Lloyd and May 2001; Kiss et al. 2006). The model of Pastor-Satorras and Vespignani as well as the present approach ignores the local depletion of susceptible hosts occurring in true network-models providing a way to distinguish between the effect of heterogeneous contact rates as such and local depletion in network-models (Keeling 1999).

To assess the impact of host heterogeneity on epidemic size, we focus on proportionate mixing and further require that the magnitude of susceptibility and infectivity in the sub-populations are ranked in the same order. This case includes heterogeneity that is solely due to variation in contact rates combined with proportionate mixing, thus it includes most models of venereal diseases and many models of air-borne infections. We show that for this case the final size in the heterogeneously mixing population is always smaller than what it would have been in a homogeneously mixing population with the same basic reproduction number in analogy with the results for endemic diseases in heterogeneously mixing populations (Hethcote and Yorke 1984). However, if the ordering of susceptibilities is the reverse of the ordering of infectivities, heterogeneity may lead to an increase in epidemic size. Mathematically mixing patterns with such negative correlation between susceptibility and infectivity are known to cause unusual results. They can, for example, give rise to sustained oscillations in an age-structured SIR-model (Andreasen 1995), but their biological importance remains to be seen.

It is unclear how our size comparison may be extended to transmission patterns other than proportionate mixing for example preferred mixing. While the final size equation generalizes naturally, there is no obvious generalization of the condition that the magnitude of susceptibility and infectivity are ordered in the same sequence. The technical condition in the proof relates to the covariance of infectivity and total attack rate over susceptibility, a condition that is somewhat awkward to work with in situations where the total attack rate is poorly characterized. Alternatively, one may explore the differential inequalities used by Ball (1985).

Our analysis have assumed that the infection does not affect host mortality and that the epidemic does not change population size. This assumption may exclude many outbreaks particularly in animal populations. An extension of our results to a deadly disease is far from obvious. The present formulation is formally the same for density and frequency dependent transmission, but a varying population size requires that the density dependent aspect of disease transmission be addressed explicitly (Getz

and Pickering 1983). Furthermore, many of our results including the derivation of the final size condition as well as the similarity between the next generation and final size matrices rest on the assumption of constant population size. Arino et al. (2007) and Brauer (2008) provide a possible starting point but the characterization of the final epidemic size in a varying population remains an open question.

**Acknowledgements** This research was supported in part by a grant 271-07-0555 from the Danish Medical Research Council and by the RAPIDD programme of the Science & Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. V.A. thanks Bryan Grenfell for his hospitality during this project.

## References

Anderson, R. M., Donnelly, C. A., Ferguson, N. M., Woolhouse, M. E. J., Watt, C. J. et al. (1996). Transmission dynamics and epidemiology of BSE in British cattle. *Nature (London)*, *382*, 779–788.

Andreasen, V. (1995). Instability in an SIR-model with age-dependent susceptibility. In O. Arino, D. Axelrod, M. Kimmel, & M. Langlais (Eds.), *Mathematical population dynamics* (Vol. 1, pp. 3–14). Winnipeg: Wuerz Publ.

Andreasen, V. (2003). Dynamics of annual influenza A epidemics with immuno-selection. *J. Math. Biol.*, *46*, 504–536.

Andreasen, V., & Frommelt, T. (2005). A school-oriented, age-structured epidemic model. *SIAM J. Appl. Math.*, *65*, 1870–1887.

Andreasen, V., & Sasaki, A. (2006). Shaping the phylogenetic tree of influenza by cross-immunity. *Theor. Popul. Biol.*, *70*, 164–173.

Arinaminpathy, N., & McLean, A. R. (2008). Antiviral treatment for the control of pandemic influenza: some logistical constraints. *J. Roy. Soc. Interfaces*, *5*, 545–553.

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

Ball, F. (1985). Deterministic and stochastic epidemics with several kinds of susceptibles. *Adv. Appl. Probab.*, *17*, 1–22.

Boni, M. F., Gog, J. R., Andreasen, V., & Christiansen, F. B. (2004). Influenza drift and epidemic size: the race between generating and escaping immunity. *Theor. Popul. Biol.*, *65*, 179–191.

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

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., & Knuth, D. E. (1996). On the Lambert W function. *Adv. Comput. Math.*, *5*, 329–359.

Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases*. Chichester: Wiley.

Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R0 in models for infectious-diseases in heterogeneous populations. *J. Math. Biol.*, *28*, 365–382.

Diekmann, O., Heesterbeek, J. A. P., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface*, *7*(47), 873–885.

Dwyer, G., Dushoff, J., Elkinton, J. S., & Levin, S. A. (2000). Pathogen driven outbreaks in forest defoliators revisited: Building models from experimental data. *Am. Nat.*, *156*, 105–120.

Eames, K. T. D., & Keeling, M. J. (2002). Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc. Natl. Acad. Sci. USA*, *99*, 13330–13335.

Elveback, L. R., Fox, J. P., Ackerman, E., Langworthy, A., Boyd, M., & Gatewood, L. (1976). Influenza simulation-model for immunization studies. *Am. J. Epidemiol.*, *103*, 152–165.

Eubank, S., Guclu, H., Kumar, V. S. A., Marathe, M. V., Srinivasan, A. et al. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature (London)*, *429*, 180–184.

Ferguson, N. M., Donnelly, C. A., & Anderson, R. M. (2001). Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature (London)*, *413*, 542–548.

Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., Fraser, C., Riley, S. et al. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature (London)*, *437*, 209–214.

Gart, J. J. (1968). The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics*, *24*, 557–565.

Getz, W. M., & Pickering, J. (1983). Epidemic models—thresholds and population regulation. *Am. Nat.*, *121*, 892–898.

Gillespie, J. H. (1975). Natural selection for resistance to epidemics. *Ecology*, *56*, 493–495.

Hethcote, H. W., & Yorke, J. A. (1984). *Gonorrhea transmission dynamics and control*. Berlin: Springer.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. Lond. B*, *266*, 859–867.

Kermack, W. O., & McKendrick, A. G. (1927). Contributions to the mathematical theory of epidemics, 1. *Proc. R.. Soc. A*, *115*, 700–721. (Reprinted in *Bull. Math. Biol., 53*, 33–55, 1991)

Kiss, I. Z., Green, D. M., & Kao, R. R. (2006). The effect of contact heterogeneity and multiple routes of transmission on final epidemic size. *Math. Biosci.*, *203*, 124–136.

Kretzschmar, M., & Mikolajczyk, R. T. (2009). Contact profiles in eight European countries and implications for modelling the spread of airborne infectious diseases. *PLOS One*, *4*.

Lloyd, A. L., & May, R. M. (2001). How viruses spread among computers and people. *Science*, *292*, 1316–1317.

Longini, I. M., Nizam, A., Xu, S. F., Ungchusak, K., Hanshaoworakul, W. et al. (2005). Containing pandemic influenza at the source. *Science*, *309*, 1083–1087.

Ma, J. L., & Earn, D. J. D. (2006). Generality of the final size formula for an epidemic of a newly invading infectious disease. *Bull. Math. Biol.*, *68*, 679–702.

May, R. M. (1985). Regulation of populations with nonoverlapping generations by microparasites—a purely chaotic system. *Am. Nat.*, *125*, 573–584.

Mills, C. E., Robins, J. M., & Lipsitch, M. (2004). Transmissibility of 1918 pandemic influenza. *Nature (London)*, *432*, 904–906.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K. et al. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.*, *5*, 381–391.

Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E*, *66*, 016128.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, *86*(14), 3200–3203.

Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J. et al. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science*, *300*, 1961–1966.

Tildesley, M. J., & Keeling, M. J. (2009). Is R-0 a good predictor of final epidemic size: Foot-and-mouth disease in the UK. *J. Theor. Biol.*, *258*, 623–629.

Volz, E. (2008a). Susceptible-infected-recovered epidemics in populations with heterogeneous contact rates. *Eur. Phys. J. B*, *63*, 381–386.

Volz, E. (2008b). SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.*, *56*, 293–310.

Wallinga, J., Teunis, P., & Kretzschmar, M. (2006). Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.*, *164*, 936–944.