AN EPIDEMIC MODEL WITH INFECTOR-DEPENDENT SEVERITY

FRANK BALL,* *The University of Nottingham*
TOM BRITTON,** *Stockholm University*

Abstract

A stochastic epidemic model is defined in which infected individuals have different severities of disease (e.g. mildly and severely infected) and the severity of an infected individual depends on the severity of the individual he or she was infected by; typically, severe or mild infectives have an increased tendency to infect others severely or, respectively, mildly. Large-population properties of the model are derived, using branching process approximations for the initial stages of an outbreak and density-dependent population processes when a major outbreak occurs. The effects of vaccination are considered, using two distinct models for vaccine action. The consequences of launching a vaccination program are studied in terms of the effect it has on reducing the final size in the event of a major outbreak as a function of the vaccination coverage, and also by determining the critical vaccination coverage above which only small outbreaks can occur.

Keywords: Stochastic epidemic; final size; basic reproduction number; varying severity; branching process; density dependent population process; threshold behaviour

2000 Mathematics Subject Classification: Primary 92D30
Secondary 60K99

1. Introduction

Many infectious diseases have the property that the degree of severity of infected individuals varies, and the degree of severity of an infective is also believed to affect potential future spreading, e.g. measles (Morley and Aaby (1997)), varicella (Parang and Archana (2004)), and dengue fever (Mangada and Igarashi (1998)). In the present paper we study a model which attempts to capture this feature within the class of SIR (susceptible $\rightarrow$ infective $\rightarrow$ removed) epidemic models, describing the spread of an infectious disease in a closed finite community; see, for example, Lefèvre (1990) and Andersson and Britton (2000, Chapter 2).

For the same purpose, Ball and Britton (2005) defined an epidemic model for a homogeneously mixing community, where the degree of severity of an infective depended on the amount of 'infection force' an individual had been exposed to, and where severe infectives typically exposed more force of infection than mildly infected. The new model of the present paper captures the same phenomenon but from a different perspective, also assuming a homogeneously mixing community. Here the severity of an infective depends on who this individual was infected by, in that he or she is typically more likely to become severely infected if infected by a severe infective and mildly infective if infected by a mild infective. Contrary to the model in Ball and

Britton (2005), an infective always remains the same type and cannot become a severe infective by additional exposure to infection. We denote our model the infector-dependent severity (IDS) model.

For a large community starting with few initial infectives, we show a threshold limit theorem for the IDS model, which states that the epidemic has a (strictly) positive chance of becoming established and growing beyond all limits if and only if a suitably defined threshold parameter (reproduction number) is larger than unity. We also show that if the epidemic is initiated by a positive fraction of infectives then the numbers of ultimately mildly and severely infected individuals (i.e. the total size of the epidemic) satisfies a law of large numbers. Furthermore, we make plausible that if either (i) the epidemic is initiated by a positive fraction of infectives or (ii) the epidemic is initiated by few infectives but it becomes established, then the total size of the epidemic satisfies a central limit theorem.

We study two different models for the effect of vaccination. In one model the vaccine reduces the severity of disease, in the sense that an individual who would have become severely infected becomes only mildly infected if he or she has been vaccinated, and someone who would have become mildly infected avoids becoming infected altogether. The other vaccination model is more general and flexible in that it reduces susceptibility and infectivity (if infected) by possibly dependent random factors. For both vaccination models, the effect of vaccination on the reproduction number and the final outbreak size are studied, as is the critical vaccination coverage, i.e. the fraction necessary to vaccinate in order to surely prevent a major outbreak.

In Section 2 we define the IDS model, state the main results, and give some heuristic arguments motivating the results. In Section 3 we define the two vaccine-response models and study their effect on outbreak sizes as a function of the vaccination coverage. Sections 2 and 3 also contain some numerical examples that illustrate the theory. In Section 4 we give the formal theorems and their proofs, together with heuristic arguments for results that are strongly supported by numerical examples but currently lack fully rigorous proofs. The paper ends with a discussion in Section 5.

2. The IDS epidemic model

2.1. Definition of the model

Consider a closed homogeneously mixing population consisting of n initially susceptible individuals, m_M mildly infectious individuals and m_S severely infectious individuals. An individual who becomes mildly infected remains so for a period I_M with distribution F_M. During this period the individual has 'close contacts' with other individuals, chosen uniformly among the initial susceptibles, at rate λ_M (i.e. at the points of a homogeneous Poisson process with rate λ_M). If the contacted person is still susceptible he or she becomes a mild case with probability $p_M^{(M)}$ and a severe case with the remaining probability $p_S^{(M)} = 1 - p_M^{(M)}$. Similarly, an individual who becomes severely infected remains so for a period I_S with distribution F_S, during which he or she has 'close contacts' with others, chosen uniformly among the initially susceptible individuals, at rate λ_S. In this case a contacted person becomes a mild case with probability $p_M^{(S)}$ and a severe case with the remaining probability $p_S^{(S)} = 1 - p_M^{(S)}$. At the end of his or her infectious period an infective individual recovers, becomes immune, and plays no further part in the epidemic—a state called removed. All infectious periods, contact processes, and uniform 'selections' are mutually independent. The infectious period distributions F_M and F_S are arbitrary but specified and have finite means. The epidemic process continues until there is no mild or severe infective in the population, at which point the epidemic stops. The numbers

of individuals who were mildly and severely infected during the outbreak specifies the *final size* of the epidemic.

Some special choices of parameter values of the IDS model give rise to models previously studied in the literature. The case where a contacted susceptible necessarily becomes the same type as his or her infector (i.e. $p_M^{(M)} = 1$ and $p_S^{(S)} = 1$) is identical to the so-called competing epidemic model investigated (with exponentially distributed infectious periods) by Kendall and Saunders (1983) and Scalia-Tomba and Svensson (2001); see also Scalia-Tomba (2005). The case where the infectious type of an individual is independent of that of his or her infector (i.e. $p_M^{(M)} = p_M^{(S)} = p_M$ and $p_S^{(M)} = p_S^{(S)} = p_S$) has been studied by Ball and Clancy (1995) (see also Picard and Lefèvre (1990)), who allowed for more than two different infectious states.

Epidemics with $p_M^{(M)} = 1$ or $p_S^{(S)} = 1$, such as the competing epidemic, have a quite different and more complicated asymptotic behaviour than epidemics with other parameter values, since in the early stages of an epidemic in a large population at least one of the mild and severe outbreaks evolves (almost) independently of the other. We exclude such cases from the analysis in this paper and assume throughout that $p_M^{(M)} < 1$ and $p_S^{(S)} < 1$, or equivalently that $p_S^{(M)} > 0$ and $p_M^{(S)} > 0$.

2.2. Main results and ideas

2.2.1. *Early stages and minor outbreaks.* Suppose that the number of susceptibles n is large and that the initial numbers of mild and severe infectives, m_M and m_S, are both small. Then, during the early stages of the epidemic, the probability that a contact is made with an infective or removed individual is very small, so the process of infectives can be approximated by a branching process, in which every infectious contact gives rise to a new infective. The branching process is two type, with the type being either *mild* or *severe*. A typical mild or severe individual lives for a time having distribution F_M or F_S, respectively, during which he or she has offspring at the points of a homogeneous Poisson process with rate λ_M or λ_S, respectively. Offspring of a mild or severe individual are mild independently with probability $p_M^{(M)}$ or $p_M^{(S)}$, respectively, and severe otherwise.

The approximation of the epidemic process by a branching process is made fully rigorous in Section 4.1 by letting n tend to ∞ and using a coupling argument; see Theorem 4.1. The branching process either goes extinct or it grows unboundedly. A threshold theorem for the epidemic model can be obtained by associating the epidemic becoming established with nonextinction of the branching process.

The mean offspring matrix for the branching process is given by

$$\boldsymbol{M} = \begin{bmatrix} \lambda_M \iota_M p_M^{(M)} & \lambda_M \iota_M p_S^{(M)} \\ \lambda_S \iota_S p_M^{(S)} & \lambda_S \iota_S p_S^{(S)} \end{bmatrix}, \tag{2.1}$$

where $\iota_M = \mathrm{E}[I_M]$ and $\iota_S = \mathrm{E}[I_S]$. The largest eigenvalue of this matrix is denoted by R_0 and called the *basic reproduction number*; see, e.g. Heesterbeek and Dietz (1996). Letting $T = \lambda_M \iota_M p_M^{(M)} + \lambda_S \iota_S p_S^{(S)}$, it is given by

$$R_0 = \tfrac{1}{2}\left(T + \sqrt{T^2 + 4\lambda_M \iota_M \lambda_S \iota_S (1 - p_M^{(M)} - p_S^{(S)})}\right). \tag{2.2}$$

By standard branching process theory (e.g. Haccou *et al.* (2005, Chapter 5)), the epidemic with few initial infectives has a nonzero probability of becoming established if and only if $R_0 > 1$. The probability that an epidemic becomes established can be determined as follows.

For $i, j \in \{M, S\}$, let $R_j^{(i)}$ denote the number of type-j offspring of a typical type-i individual, and define probability generating functions

$$f_M(s_M, s_S) = \mathrm{E}\left[s_M^{R_M^{(M)}} s_S^{R_S^{(M)}} \right] \quad \text{and} \quad f_S(s_M, s_S) = \mathrm{E}\left[s_M^{R_M^{(S)}} s_S^{R_S^{(S)}} \right].$$

Then, provided $R_0 > 1$, the probability that the epidemic becomes established is given by $1 - \pi_M^{m_M} \pi_S^{m_S}$, where (π_M, π_S) is the unique solution in $(0, 1)^2$ of the equations

$$\pi_M = f_M(\pi_M, \pi_S), \qquad \pi_S = f_S(\pi_M, \pi_S). \tag{2.3}$$

(If $R_0 \leq 1$ then $(\pi_M, \pi_S) = (1, 1)$ is the only solution of (2.3) and there is zero probability of an epidemic becoming established.) Note that if $\phi_M(\theta) = \mathrm{E}[\exp(-\theta I_M)]$ and $\phi_S(\theta) = \mathrm{E}[\exp(-\theta I_S)]$ denote the moment generating functions of typical mild and severe infectious periods, respectively, then

$$f_M(s_M, s_S) = \phi_M(\lambda_M(p_M^{(M)}(1 - s_M) + p_S^{(M)}(1 - s_S)))$$

and

$$f_S(s_M, s_S) = \phi_S(\lambda_S(p_M^{(S)}(1 - s_M) + p_S^{(S)}(1 - s_S))).$$

If the epidemic fails to become established, its final outcome can be approximated by that of the corresponding branching process; explicit results for moments are easily obtained using standard branching process theory.

2.2.2. *Large outbreaks.* We now consider the final outcome of epidemics that become established. Suppose that n is large, and let $\mu_M = m_M/n$ and $\mu_S = m_S/n$. Let r_M and r_S denote the proportions of susceptibles that ultimately become mild and severe infectives, respectively. Then the total force of infection exerted on a given susceptible during the entire epidemic is $\lambda_M \iota_M (\mu_M + r_M) + \lambda_S \iota_S (\mu_S + r_S)$. This follows since $n(\mu_M + r_M)$ is the total number of mild infectives, each of which has contact with a given susceptible at the average accumulated rate $\lambda_M \iota_M/n$; similarly, $n(\mu_S + r_S)$ is the total number of severe infectives, each of which has contact with a given susceptible at the average accumulated rate $\lambda_S \iota_S/n$. It follows that the probability that a given susceptible remains uninfected throughout the epidemic is $\exp(-(\lambda_M \iota_M (\mu_M + r_M) + \lambda_S \iota_S (\mu_S + r_S)))$, since infectious individuals make contacts at the points of independent Poisson processes. Thus, (r_M, r_S) satisfies

$$1 - r_M - r_S = \exp(-(\lambda_M \iota_M (\mu_M + r_M) + \lambda_S \iota_S (\mu_S + r_S))). \tag{2.4}$$

In general, it is not possible to derive a second balance equation satisfied by (r_M, r_S) and, hence, to uniquely determine the final outcome of the epidemic. If $\lambda_M \iota_M = \lambda_S \iota_S$ then (2.4) yields an equation for $r_M + r_S$ and, hence, for the final proportion of susceptibles, but the individual r_M and r_S still need to be determined. In the special case studied by Ball and Clancy (1995), where $p_M^{(M)} = p_M^{(S)} = p_M$ and $p_S^{(M)} = p_S^{(S)} = p_S$, contacted susceptibles become mildly infected independently with probability p_M, irrespective of the type of their infectors, so $r_M/r_S = p_M/p_S$ and the final outcome can be determined. We now assume that the infectious periods follow exponential distributions, with rate γ_M for mild infectives and γ_S for severe infectives, so $\iota_M = \gamma_M^{-1}$ and $\iota_S = \gamma_S^{-1}$. If the numbers of susceptibles and initial

infectives are both large, the stochastic model can be approximated by the deterministic model

$$\frac{dx}{dt} = -(\lambda_M y_M + \lambda_S y_S)x, \tag{2.5a}$$

$$\frac{dy_M}{dt} = (\lambda_M p_M^{(M)} y_M + \lambda_S p_M^{(S)} y_S)x - \gamma_M y_M, \tag{2.5b}$$

$$\frac{dy_S}{dt} = (\lambda_M p_S^{(M)} y_M + \lambda_S p_S^{(S)} y_S)x - \gamma_S y_S, \tag{2.5c}$$

$$\frac{dz_M}{dt} = \gamma_M y_M, \tag{2.5d}$$

$$\frac{dz_S}{dt} = \gamma_S y_S, \tag{2.5e}$$

with initial condition

$$(x(0), y_M(0), y_S(0), z_M(0), z_S(0)) = (1, \mu_M, \mu_S, 0, 0). \tag{2.6}$$

Here, $x(t)$, $y_M(t)$, $y_S(t)$, $z_M(t)$, and $z_S(t)$ respectively denote the 'proportions' of susceptible, mild infective, severe infective, mild removed, and severe removed individuals in the population at time t. The final outcome of the deterministic epidemic can be obtained by solving (2.5a)–(2.5e) numerically. Note that $(r_M, r_S) = (z_M(\infty) - \mu_M, z_S(\infty) - \mu_S)$. The balance equation (2.4) and the above observations concerning (r_M, r_S) are easily verified by analysing the differential equations (2.5a)–(2.5e).

The approximation of the stochastic epidemic relies on n, the number of initial susceptibles, becoming large, whilst keeping the infection and removal parameters fixed. We study two distinct cases for the initial number of infectives:

(i) $n^{-1}m_M \to \mu_M$ and $n^{-1}m_S \to \mu_S$, where $\mu_M + \mu_S > 0$, so there are many initial infectives when n is large; and

(ii) (m_M, m_S) is fixed, so the numbers of initial infectives of the two types is held fixed as n tends to ∞ and, hence, there are relatively few initial infectives when n is large.

For case (i), the theory of density dependent population processes described in Ethier and Kurtz (1986, Chapter 11) can be used to show that the stochastic epidemic, suitably normalized, converges as n tends to ∞ to the deterministic model (2.5a)–(2.5e), with fluctuations about the deterministic limit following a zero-mean Gaussian process. The theory also enables a central limit theorem to be derived for the final outcome of the epidemic. Let $R_M^{(n)}$ and $R_S^{(n)}$ denote the number of susceptibles that ultimately become mild and severe infectives, respectively. Then, writing (r_M, r_S) as $(r_M(\mu_M, \mu_S), r_S(\mu_M, \mu_S))$,

$$\sqrt{n} \begin{pmatrix} n^{-1}R_M^{(n)} - r_M(\mu_M, \mu_S) \\ n^{-1}R_S^{(n)} - r_S(\mu_M, \mu_S) \end{pmatrix} \xrightarrow{D} N(\mathbf{0}, \mathbf{\Sigma}(\mu_M, \mu_S)) \quad \text{as } n \to \infty, \tag{2.7}$$

where '$\xrightarrow{D}$' denotes convergence in distribution and $N(\mathbf{0}, \mathbf{\Sigma}(\mu_M, \mu_S))$ denotes a bivariate normal distribution with mean vector zero and variance-covariance matrix $\mathbf{\Sigma}(\mu_M, \mu_S)$. It follows from (2.7) that $(R_M^{(n)}, R_S^{(n)})$ is approximately bivariate normally distributed with mean vector $n(r_M(\mu_M, \mu_S), r_S(\mu_M, \mu_S))^\top$, where '$\top$' denotes transpose, and variance-covariance matrix $n\mathbf{\Sigma}(\mu_M, \mu_S)$. Of course, the mean vector and variance-covariance matrix also depend on

the infection and removal parameters, but that dependence is not shown explicitly. A heuristic proof of (2.7) is given in Section 4.2, where the calculation of $\Sigma(\mu_M, \mu_S)$ is described. A formal proof that $n^{-1}(R_M^{(n)}, R_S^{(n)})$ converges in probability to (r_M, r_S) as n tends to ∞ is given in Section 4.2; see Theorem 4.3. Note that (2.4) and Theorem 4.3 also hold for the competing epidemic model.

Finally, consider case (ii), in which there are few initial infectives. For $t \geq 0$, let $Y_M(t)$ and $Y_S(t)$ respectively denote the numbers of mild and severe individuals alive at time t in the approximating branching process, and let $T(t)$ denote the total number of births (irrespective of type) during $(0, t]$. The theory of exponential growth and asymptotic composition of a branching process (e.g. Haccou et al. (2005, Chapter 6)) implies that there exists a random variable $W \geq 0$ such that, as $t \to \infty$,

$$Y_M(t) \sim We^{\alpha t}v_M, \qquad Y_S(t) \sim We^{\alpha t}v_S, \quad \text{and} \quad T(t) \sim We^{\alpha t}u_T, \tag{2.8}$$

where α is the Malthusian parameter of the branching process, (v_M, v_S) is the left eigenvector (normalized so that $v_M + v_S = 1$) corresponding to the eigenvalue $e^{\alpha t}$ of an associated mean matrix, and $u_T > 0$ is a constant that can be determined. Moreover, $W = 0$ if and only if the branching process goes extinct.

Suppose that $R_0 > 1$ and that the branching process does not go extinct. In Section 4.1 it is shown that, for large n, the process of infectives in the epidemic process and the approximating branching process coincide until order $\sqrt{n}$ susceptibles have been infected in the epidemic, i.e until $T(t)$ is of order $\sqrt{n}$. Let $t_n = \inf\{t > 0: T(t) \geq \log n\}$ be the time elapsing until at least $\log n$ individuals have been born in the branching process. Then, using (2.8), for large n, the numbers of susceptibles, mild infectives, and severe infectives at time t_n in the epidemic process are approximately $n - \log n$, $v_M u_T^{-1} \log n$, and $v_S u_T^{-1} \log n$, respectively. In Section 4.3 it is shown that if the epidemic infects at least $\log n$ susceptibles then there exists $c^* > 0$ such that the probability that the epidemic infects at least $c^* n$ susceptibles tends to 1 as n tends to ∞; see Lemma 4.1, in which $c^* = 1 - \varepsilon_0$. The above suggests that as $c \downarrow 0$, the ratio of mild to severe infectives, when cn susceptibles have been infected, is given by v_M/v_S and, hence, from (2.7), that, conditional upon the epidemic becoming established,

$$\sqrt{n}\begin{pmatrix} R_M^{(n)} - r_M^{(0)} \\ R_S^{(n)} - r_S^{(0)} \end{pmatrix} \xrightarrow{D} N(\mathbf{0}, \Sigma^{(0)}) \quad \text{as } n \to \infty, \tag{2.9}$$

where $r_i^{(0)} = \lim_{c \downarrow 0} r_i(cv_M, cv_S)$, $i \in \{M, S\}$, and $\Sigma^{(0)} = \lim_{c \downarrow 0} \Sigma(cv_M, cv_S)$. Further details and a heuristic proof of (2.9) are given in Section 4.3. Result (2.9) yields a normal approximation for the final outcome of epidemics, with few initial cases, that become established.

2.3. Illustrations

In order to see how well our approximations works in finite populations we have performed simulations. Figure 1 is based on 10 000 simulations of the IDS model for a community consisting of $n = 100\,000$ individuals. The contact parameters were chosen to be $\lambda_M = 2.5$ and $\lambda_S = 1$, which reflect that severely infected people might not be as socially active as mildly infected people. The duration of the infectious period was modelled by a unit-mean exponential distribution for both mild and severe cases, so $\iota_M = \iota_S = 1$. It was further assumed that both mild and severe cases infected new individuals to the same type as themselves with probability 0.8, implying that $p_M^{(M)} = 0.8$, $p_S^{(M)} = 0.2$, $p_M^{(S)} = 0.2$, and $p_S^{(S)} = 0.8$. Using (2.2), $R_0 = 2.0782$ for this set of parameter values, so the epidemic is clearly above threshold. Each
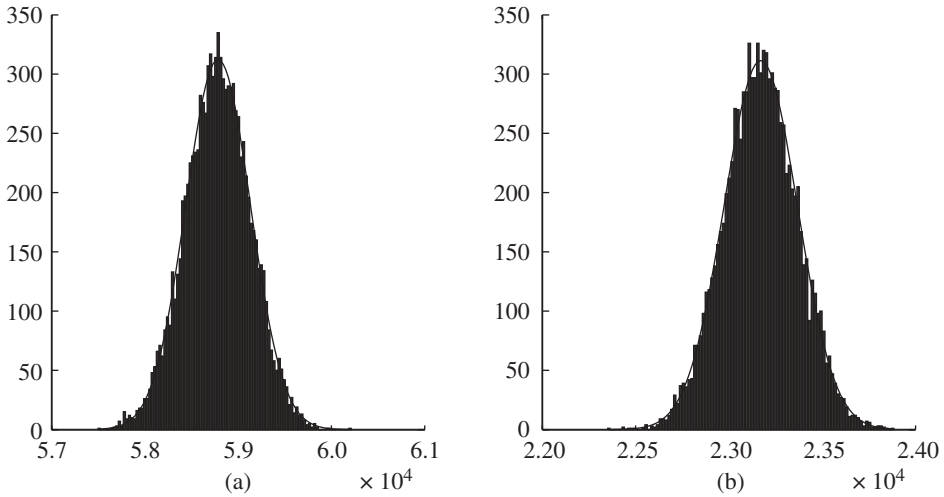
FIGURE 1: Histograms of the final number of (a) mild cases and (b) severe cases from the simulations of the IDS model in a community of 100 000 individuals, with theoretical normal approximation superimposed.

simulation was initiated by 100 mild and 100 severely infectious individuals, thus avoiding minor outbreaks. Figure 1 contains histograms of the final numbers of mild (Figure 1(a)) and severe (Figure 1(b)) cases from the simulations. The average scaled (i.e. divided by the initial number of susceptibles, n) numbers of mildly and severely infected in the simulations, including initial infectives, were 0.5878 and 0.2317, respectively. The corresponding theoretical values $r_M + \mu_M$ and $r_S + \mu_S$ were computed as described in Section 2.2.2 and found to be $r_M + \mu_M = 0.5879$ and $r_S + \mu_S = 0.2317$, i.e. (almost) the same as the empirical means. Similarly, the limiting covariance matrix Σ (see (2.7)) was computed numerically. The limiting scaled variances and covariance (now multiplied by n) for the proportions infected (elements $(1, 1)$, $(2, 2)$, and $(1, 2)$ of Σ in (2.7)) were $\sigma_{11} = 1.2052$, $\sigma_{22} = 0.3871$, and $\sigma_{12} = -0.3780$. The corresponding empirical variances from the simulations were 1.2117, 0.3871, and -0.3712, respectively, also very close to their asymptotic counterparts. In the histograms we have superimposed the limiting normal distributions having the asymptotic means and variances just mentioned. It is seen that the empirical distributions agree remarkably well with the limiting normal distributions.

In order to study how quickly the asymptotics kick in, both for the branching process and the final size approximations, we simulated outbreaks starting with few initially infected for different community sizes. We chose the same values for the contact parameters and removal parameters as above (i.e. $\lambda_M = 2.5$ and $\lambda_S = \gamma_M = \gamma_S = 1$), and always started with one mild and one severe infective, for community sizes $n = 100$, 1000, 10 000, 100 000, and 1 000 000. It turned out that the speed with which the asymptotics kick in depends on how well the two types of infected mix, i.e. on the distance of $p_M^{(M)}$ and $p_S^{(S)}$ from 1. (Recall that when they are equal to 1 we have the competing epidemic model which has a rather different and more complicated asymptotic behaviour.) For this reason, in Table 1 we list the theoretical and empirical probabilities of major outbreaks for different community sizes and choices of $(p_M^{(M)}, p_S^{(S)})$. We chose to let $p_M^{(M)} = p_S^{(S)} = p$ and varied p from $p = 0.5$, so individuals are equally likely to infect someone mildly or severely (independent of their own

TABLE 1: Empirical proportions of major outbreaks for different community sizes, n, and values of $p = p_M^{(M)} = p_S^{(S)}$.

	P(major outbreak)				
n	$p = 0.5$	$p = 0.8$	$p = 0.9$	$p = 0.99$	$p = 0.999$
100	0.6068	0.6350	0.6332	0.6360	0.6348
1000	0.6354	0.6416	0.6474	0.6239	0.6035
10 000	0.6384	0.6423	0.6496	0.6221	0.6060
100 000	0.6336	0.6446	0.6424	0.6307	0.6139
1 000 000	0.6372	0.6426	0.6427	0.6328	0.6103
∞	0.6307	0.6479	0.6475	0.6247	0.6091

TABLE 2: Empirical mean proportions for mild and severe cases among major outbreaks, for different community sizes and values of $p = p_M^{(M)} = p_S^{(S)}$.

	Mean proportions for mild cases				
n	$p = 0.5$	$p = 0.8$	$p = 0.9$	$p = 0.99$	$p = 0.999$
100	0.3580	0.5655	0.6770	0.7929	0.8074
1000	0.3566	0.5853	0.7199	0.8662	0.8795
10 000	0.3562	0.5871	0.7240	0.8731	0.8899
100 000	0.3563	0.5872	0.7243	0.8744	0.8906
1 000 000	0.3563	0.5872	0.7244	0.8746	0.8908
∞	0.3566	0.5871	0.7246	0.8748	0.8910
	Mean proportions for severe cases				
100	0.3552	0.2384	0.1611	0.0667	0.0547
1000	0.3566	0.2321	0.1369	0.0222	0.0086
10 000	0.3564	0.2299	0.1328	0.0160	0.0024
100 000	0.3563	0.2297	0.1325	0.0149	0.0017
1 000 000	0.3563	0.2297	0.1324	0.0147	0.0015
∞	0.3566	0.2296	0.1324	0.0147	0.0015

state), up to $p = 0.999$, where individuals nearly always create infectives of the same type as themselves. For each p and community size n, the results are based on 10 000 simulations. For the asymptotic values ($n = \infty$), the probability of a major outbreak is computed from the branching process approximation.

As seen in Table 1, the probability of a major outbreak, computed as the proportion out of the 10 000 simulations resulting in more than 20% becoming infected, agrees quite well with the branching process approximation. The approximation seems adequate when the community consists of at least a few thousand individuals, more or less irrespective of the value of p.

Using the same simulations as in Table 1, we also calculated the proportions infected of the two types in the major epidemics. In Table 2 we present the means of these proportions for the different community sizes, n, and values of p. The asymptotic ($n = \infty$) proportions infected of the two types are the numerical solutions $r_M^{(0)}$ and $r_S^{(0)}$, obtained as described in Section 2.2.2, for a small value of c. (In practice, c was divided successively by 10 until the change in the total proportion infected was less than 10^{-5}.)

It is seen that the mean proportions infected of the two types agree better and better with their asymptotic counterparts as n grows. Unless p is close to 1, the asymptotics provide a good approximation when n is as small as 1000, whereas when p approaches 1 (making the model close to the competing epidemic), the community size has to be considerably larger for the asymptotics to kick in.

3. Vaccination

3.1. Reduction in susceptibility and infectivity model

In this model it is assumed that vaccinated individuals independently have a random reduction in susceptibility and infectivity described by the pair of (possibly dependent) random variables (A, B), where A denotes the relative susceptibility and B denotes the relative infectivity if infected; see, e.g. Becker and Starczak (1998) and Becker *et al.* (2006). For simplicity we assume that the distribution of (A, B) is discrete: $P(A = a_i, B = b_i) = p_i$, $i = 1, \ldots$; see Becker and Starczak (1998) and Ball and Becker (2006).

By relative susceptibility and relative infectivity we mean the following. A vaccinated, not yet infected, individual having vaccine response (a_i, b_i) that is contacted by an infective, be it a severe or mild contact, becomes infected with probability a_i. A vaccinated individual with vaccine response (a_i, b_i) who still gets infected has a reduced contact rate. If the individual is mildly infected, the contact rate is $b_i \lambda_M$, and if the individual is severely infected, the contact rate is $b_i \lambda_S$. Thus, the smaller A is, the more effective the vaccine is in terms of reducing susceptibility, and the smaller B is, the more effective the vaccine is in terms of reducing infectivity. As reduced infectivity really matters only for individuals who do become infected, the product AB is also a relevant measure of vaccine efficacy. Vaccine efficacy is usually defined to equal 1 for perfect vaccines and 0 for completely useless ones. Thus, we have the following definitions of vaccine efficacies in terms of susceptibility, infectivity, and susceptibility-infectivity (see Becker *et al.* (2006)):

$$V E_S = 1 - E[A], \qquad V E_I = 1 - \left(\frac{E[AB]}{E[A]} \right), \qquad V E_{SI} = 1 - E[AB].$$

Suppose that a proportion v of susceptibles are vaccinated prior to the arrival of the infectious disease, and that the vaccine response is as above. Then the initial stages of the epidemic also admits an approximation by a suitable two-type branching process, in which (mild or severe) individuals correspond to (mild or severe) contacts in the epidemic process, and births in the branching process correspond to contacts emanating from the corresponding contact in the epidemic process. Consider a mild contact. This will be with a vaccinated individual having vaccine response (a_i, b_i) with probability $v p_i$ and with an unvaccinated individual with probability $1 - v$. If the contacted individual has response (a_i, b_i), an infection takes place with probability a_i and, if it does, the expected number of new contacts he or she makes is $b_i \lambda_M \iota_M$. Thus, a mild contact, on average, gives rise to $(1 - v + v E[AB]) \lambda_M \iota_M$ new contacts, of which a fraction $p_M^{(M)}$ are mild and the remaining fraction $p_S^{(M)}$ are severe. The same reasoning can be applied to severe contacts, so the mean offspring matrix is given by

$$M^{(AB)}(v) = (1 - v + v E[AB]) M,$$

where M is the mean offspring matrix without vaccination defined in (2.1). Hence, the largest eigenvalue $R_v^{(AB)}$ of $M^{(AB)}(v)$ satisfies $R_v^{(AB)} = (1 - v + v E[AB]) R_0$, where R_0 is the largest

eigenvalue without vaccination. Thus, if $1 < R_0 \leq (E[AB])^{-1}$ then the critical vaccination coverage, above which the entire community is protected from major outbreaks, is given by

$$v_c^{(AB)} = (1 - E[AB])^{-1}(1 - R_0^{-1}) = (V\,E_{SI})^{-1}(1 - R_0^{-1}); \tag{3.1}$$

cf. Becker and Starczak (1998). If $R_0 > (E[AB])^{-1}$ then vaccination alone cannot protect the community from major outbreaks.

Suppose that the vaccine response is nonrandom, in that $A \equiv a$ and $B \equiv b$ (i.e. $P(A = a, B = b) = 1$) for some $(a, b) \in (0, 1)^2$. Then the behaviour of large-population epidemics that take off can be approximated by deriving a deterministic model for

$$(x_U(t), x_V(t), y_{MU}(t), y_{MV}(t), y_{SU}(t), y_{SV}(t), z_M(t), z_S(t)),$$

where the additional suffices U and V refer to unvaccinated and vaccinated individuals, respectively. The resulting set of ordinary differential equations can be solved numerically to determine the final outcome of the epidemic for any given set of parameter values and initial conditions. This can be extended to the case in which there are finitely many vaccine responses by suitably enlarging the set of differential equations. A vaccine response that has received considerable attention in the literature is the so-called *all-or-nothing* model (Halloran *et al.* (1992)), in which a vaccinated individual is rendered completely immune with probability ε, otherwise the vaccine has no effect. Thus, $P(A = 0, B = 0) = \varepsilon = 1 - P(A = 1, B = 1)$, so $V\,E_S = V\,E_{SI} = \varepsilon$ and $V\,E_I = 0$. The final outcome of a large-population epidemic that takes off when, prior to the arrival of the disease, a proportion v of susceptibles were vaccinated with an all-or-nothing vaccine having $V\,E_S = \varepsilon$, can be approximated by solving the ordinary differential equations (2.5a)–(2.5e) with initial condition $(x(0), y_M(0), y_S(0), z_M(0), z_S(0)) = (1 - v\varepsilon, \mu_M, \mu_S, 0, 0)$.

3.2. Reduction in severity model

In this model, used in Ball and Becker (2006), vaccination reduces the severity of the disease, in that an individual who would have become severely infected had he or she not been vaccinated becomes only mildly infected if vaccinated, and a person who would have become mildly infected had he or she not been vaccinated avoids becoming infected when vaccinated. Consequently, a mild or severe case that has contact with a vaccinated but uninfected individual, infects that person mildly with probability $p_S^{(M)}$ or $p_S^{(S)}$, respectively, and not at all with probability $p_M^{(M)}$ or $p_M^{(S)}$, respectively.

Suppose that a fraction v of susceptibles are vaccinated prior to the arrival of the infectious disease, so a fraction v of all contacts are with vaccinated individuals and the remaining fraction $1 - v$ are with unvaccinated individuals. Just as for the case without vaccination, the initial stages of the epidemic can be approximated by a two-type branching process, but now the mean offspring matrix is

$$\boldsymbol{M}^{(Sev)}(v) = \begin{bmatrix} \lambda_M \iota_M((1-v)p_M^{(M)} + vp_S^{(M)}) & \lambda_M \iota_M(1-v)p_S^{(M)} \\ \lambda_S \iota_S((1-v)p_M^{(S)} + vp_S^{(S)}) & \lambda_S \iota_S(1-v)p_S^{(S)} \end{bmatrix}.$$

Starting with few initial infectives, a major outbreak can occur if and only if the largest eigenvalue $R_v^{(Sev)}$ of this matrix exceeds unity, though now there is no simple relationship between $R_v^{(Sev)}$ and R_0. We omit the details but it is easily shown that $R_v^{(Sev)}$ and the corresponding critical vaccination coverage $v_c^{(Sev)}$ are both given by the roots of quadratic equations. Note

that vaccination alone cannot protect the community from major outbreaks if $\lambda_M \iota_M p_S^{(M)} > 1$, since then $R_1^{(Sev)} > 1$.

As with the first vaccine-action model, the behaviour of large-population epidemics that become established can be approximated by a deterministic model, but now only susceptible individuals need to be classified by their vaccination status.

3.3. Illustrations

We illustrate the two vaccine-action models by studying their application to the example given at the beginning of Section 2.3. Thus, in the absence of vaccination, the infectious periods each follow a unit-mean exponential distribution, $\lambda_M = 2.5$, $\lambda_S = 1$, $p_M^{(M)} = 0.8$, and $p_S^{(S)} = 0.8$; so, as noted previously, $R_0 = 2.078$.

Figure 2 shows the proportions of the population that ultimately become mild and severe cases for the reduction in susceptibility and infectivity model. Graphs are plotted in Figure 2 for both an all-or-nothing vaccine response with $P(A = 1, B = 1) = 0.2 = 1 - P(A = 0, B = 0)$ and a leaky vaccine (Halloran *et al.* (1992)) with $A \equiv 0.2$ and $B \equiv 1$. In both cases $V E_S = V E_{SI} = 0.8$ and $V E_I = 0$, thus, making the two vaccine models comparable. In Figure 2 the limiting proportions of mild and severe cases are plotted as a function of the vaccination coverage. It is seen that both vaccine responses have the same critical vaccination coverage $v_c^{(AB)} = 0.649$, which follows from (3.1). However, for subcritical vaccination coverages, the all-or-nothing vaccine outperforms the leaky vaccine in that fewer people become infected with the all-or-nothing vaccine. This observation holds in general, also for epidemic models without varying severity, and has the following explanation (Ball and Becker (2006)). Among vaccinated individuals the chance that a first infectious contact results in infection is the same in the two models. However, for those who avoid infection at the first contact, the all-or-nothing vaccinated are completely immune whereas the leaky vaccinated only have reduced susceptibility and, hence, may become infected at a subsequent contact.

For the reduction in severity model, the corresponding results are presented in Figure 3(a). The critical vaccination coverage is $v_c^{(Sev)} = 0.738$. In this example, the proportions of the population that ultimately become mild and severe cases both decrease with vaccination
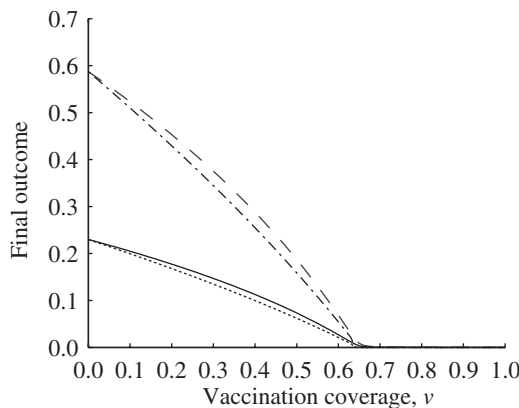

FIGURE 2: Limiting proportions infected as a function of the vaccination coverage under the reduction in susceptibility and infectivity model for both a leaky vaccine (mild (*dashed line*) and severe (*solid line*) cases) and an all-or-nothing vaccine (mild (*dash–dot line*) and severe (*dotted line*) cases).
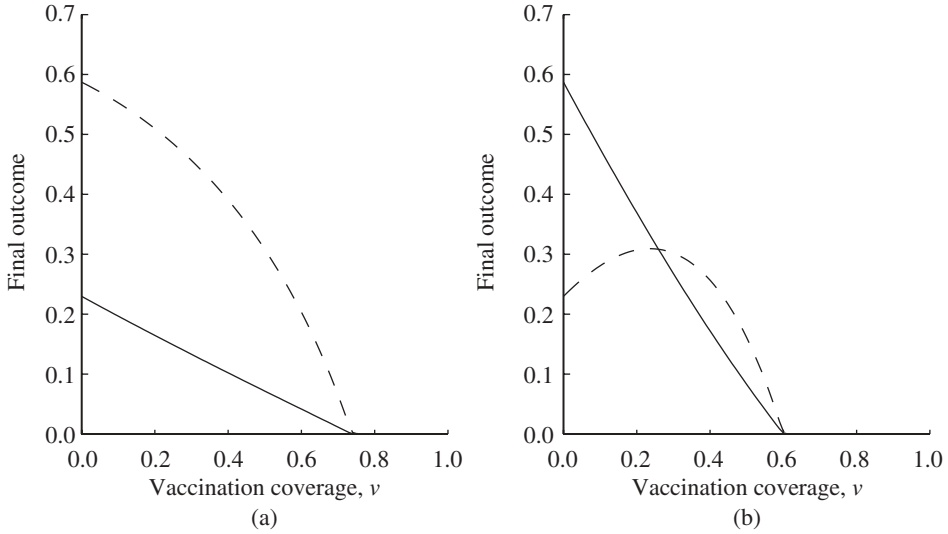
FIGURE 3: Limiting proportions infected as a function of the vaccination coverage for the reduction in severity model, with (a) $(\lambda_M, \lambda_S) = (2.5, 1)$ and (b) $(\lambda_M, \lambda_S) = (1, 2.5)$; mild (*dashed line*) and severe (*solid line*) cases.

coverage v. However, this is not always the case. For some parameter values the proportion mildly infected *increases* with v in certain v-regions. This is illustrated in Figure 3(b), where we have interchanged the contact rates so that $\lambda_M = 1$ and $\lambda_S = 2.5$, keeping all other parameters unchanged. This phenomenon can be explained as follows. In the absence of vaccination there are more severe than mild cases because severe infectives have a higher contact rate and infectives tend to produce subsequent cases having the same type as themselves. When a small proportion of the population is vaccinated, some of the vaccinated individuals, who previously would have become severe cases, will now become mild cases and consequently create further mild cases by contacting unvaccinated individuals. However, the total number of cases will decrease because mild cases make fewer contacts and contacts made by mild cases with vaccinated individuals are not likely to result in disease transmission. As vaccination coverage increases, the overall reduction in cases dominates the greater proportion of cases being mild, and the total number of mild cases eventually decreases.

4. Proofs

4.1. Few initial infectives: branching process approximation

We analyse the initial behaviour of an epidemic with few initial infectives by extending the coupling argument of Ball and Donnelly (1995) to the IDS model. Consider a sequence of IDS epidemics, $\{E^{(n)} : n \geq 1\}$ say, indexed by the initial number of susceptibles, n, with each epidemic having m_M initial mild infectives and m_S initial severe infectives. Let $Y = \{Y(t) : t \geq 0\}$ denote the branching process described in Section 2.2, where, for $t \geq 0$, $Y(t) = (Y_M(t), Y_S(t))$, with $Y_M(t)$ and $Y_S(t)$ respectively denoting the numbers of mild and severe individuals at time t, and $Y(0) = (m_M, m_S)$. Let $(\Omega, \mathcal{F}, P)$ be a probability space on which is defined the branching process Y and, independently of Y, the random

variables $\chi_i^{(n)}$, $i, n = 1, 2, \ldots$, where, for each n, $\chi_1^{(n)}$, $\chi_2^{(n)}$, $\ldots$ are independent and uniformly distributed on $\{1, 2, \ldots, n\}$.

For $n = 1, 2, \ldots$, the epidemic $E^{(n)}$ is constructed on $(\Omega, \mathcal{F}, \mathrm{P})$ as follows. Individuals and births in the branching process correspond respectively to infectives and infectious contacts in the epidemic process. Label the initial susceptibles in $E^{(n)}$, $1, 2, \ldots, n$. For the kth birth in the branching process, the individual contacted in the epidemic is $\chi_k^{(n)}$. If the contacted individual is susceptible then it becomes infected and adopts the same type as the corresponding individual, i^* say, in the branching process. If the contacted individual is not susceptible then that contact, and any descendants of i^* in the branching process, are ignored in $E^{(n)}$. The epidemic stops when there is no infective left in the population.

For $n = 1, 2, \ldots$, let $\boldsymbol{Y}^{(n)} = \{\boldsymbol{Y}^{(n)}(t) : t \geq 0\}$, where, for $t \geq 0$, $\boldsymbol{Y}^{(n)}(t) = (Y_{\mathrm{M}}^{(n)}(t), Y_{\mathrm{S}}^{(n)}(t))$, with $Y_{\mathrm{M}}^{(n)}(t)$ and $Y_{\mathrm{S}}^{(n)}(t)$ respectively denoting the numbers of mild and severe infectives at time t in $E^{(n)}$. Observe that $\boldsymbol{Y}$ and $\boldsymbol{Y}^{(n)}$ coincide up until the first time a contact is made with a previously contacted individual in $E^{(n)}$. For $n = 1, 2, \ldots$, let $\eta^{(n)} = \min\{k \geq 2 : \chi_k^{(n)} = \chi_l^{(n)}$ for some $l < k\}$ denote the number of contacts made in $E^{(n)}$ until a contact is made with a previously contacted individual. Noting the connection with the birthday problem, it is well known (see, e.g. Aldous (1985, p. 96)) that $n^{-1/2}\eta^{(n)} \xrightarrow{\mathrm{D}} \eta$ as $n \to \infty$, where η is a random variable with probability density function

$$f(x) = x \exp(-\tfrac{1}{2}x^2), \qquad x > 0.$$

As in the proof of Theorem 2.1 of Ball and Donnelly (1995), the Skorokhod representation theorem implies that we may assume that the $\chi_i^{(n)}$, $i, n = 1, 2, \ldots$, are constructed so that $n^{-1/2}\eta^{(n)} \to \eta$ almost surely as $n \to \infty$, where η is also defined on $(\Omega, \mathcal{F}, \mathrm{P})$.

For $n = 1, 2, \ldots$, let $Z_{\mathrm{M}}^{(n)}$ and $Z_{\mathrm{S}}^{(n)}$ respectively denote the total numbers of mild and severe infectives removed in the epidemic $E^{(n)}$, and let Z_{M} and Z_{S} respectively denote the total mild and severe progeny, including the initial ancestors, in the branching process $\boldsymbol{Y}$. Let $A_E \in \mathcal{F}$ denote the set on which $\boldsymbol{Y}$ becomes extinct.

Theorem 4.1. (a) *For* P-*almost all* $\omega \in A_E$, $(Z_{\mathrm{M}}^{(n)}, Z_{\mathrm{S}}^{(n)}) \to (Z_{\mathrm{M}}, Z_{\mathrm{S}})$ *as* $n \to \infty$.

(b) *For* P-*almost all* $\omega \in A_E^c = \Omega \setminus A_E$, $(Z_{\mathrm{M}}^{(n)}, Z_{\mathrm{S}}^{(n)}) \to (\infty, \infty)$ *as* $n \to \infty$.

Proof. For P-almost all $\omega \in A_E$, $Z_{\mathrm{M}}(\omega)$ and $Z_{\mathrm{S}}(\omega)$ are both finite, and $\eta^{(n)}(\omega) > \tfrac{1}{2}\sqrt{n}\eta(\omega)$ for all sufficiently large n. Hence, for such ω and n, every birth in the branching process yields an infection in $E^{(n)}$, and part (a) follows. Part (b) is proved by noting that, for P-almost all $\omega \in A_E^c$, the above argument shows that, for any $l \in \mathbb{Z}^+$, if $Z_{\mathrm{M}} + Z_{\mathrm{S}} \geq l$ then $Z_{\mathrm{M}}^{(n)} + Z_{\mathrm{S}}^{(n)} \geq l$ for all sufficiently large n.

The assertions in Section 2.2.1, concerning the probability that an epidemic with few initial infectives becomes established and the final outcome of epidemics that do not become established, follow using Theorem 4.1 and standard branching process theory.

4.2. Many initial infectives: a law of large numbers and central limit theorem for the final outcome

Suppose that the infectious periods follow exponential distributions, with rate γ_{M} for mild infectives and rate γ_{S} for severe infectives, and that initially, i.e. at time $t = 0$, there are n susceptibles, $m_{\mathrm{M}}^{(n)}$ mild infectives, and $m_{\mathrm{S}}^{(n)}$ severe infectives. For $t \geq 0$, let $X^{(n)}(t)$, $Y_{\mathrm{M}}^{(n)}(t)$, $Y_{\mathrm{S}}^{(n)}(t)$, $Z_{\mathrm{M}}^{(n)}(t)$, and $Z_{\mathrm{S}}^{(n)}(t)$ respectively denote the numbers of susceptible, mild infective, severe infective, mild removed, and severe removed individuals at time t. The

epidemic is completely specified by the process $X^{(n)} = \{X^{(n)}(t): t \geq 0\}$, where $X^{(n)}(t) = (X^{(n)}(t), Y_M^{(n)}(t), Y_S^{(n)}(t), Z_M^{(n)}(t))^\top$, since

$$X^{(n)}(t) + Y_M^{(n)}(t) + Y_S^{(n)}(t) + Z_M^{(n)}(t) + Z_S^{(n)}(t) = n + m_M^{(n)} + m_S^{(n)} \quad \text{for all } t \geq 0.$$

The process $X^{(n)}$ is a continuous-time Markov chain with transition intensities

$$q_{(i, j_M, j_S, k_M),(i-1, j_M+1, j_S, k_M)}^{(n)} = n\left(\frac{i}{n}\right)\left(\lambda_M p_M^{(M)}\left(\frac{j_M}{n}\right) + \lambda_S p_M^{(S)}\left(\frac{j_S}{n}\right)\right),$$

$$q_{(i, j_M, j_S, k_M),(i-1, j_M, j_S+1, k_M)}^{(n)} = n\left(\frac{i}{n}\right)\left(\lambda_M p_S^{(M)}\left(\frac{j_M}{n}\right) + \lambda_S p_S^{(S)}\left(\frac{j_S}{n}\right)\right),$$

$$q_{(i, j_M, j_S, k_M),(i, j_M-1, j_S, k_M+1)}^{(n)} = n\gamma_M\left(\frac{j_M}{n}\right),$$

$$q_{(i, j_M, j_S, k_M),(i, j_M, j_S-1, k_M)}^{(n)} = n\gamma_S\left(\frac{j_S}{n}\right),$$

corresponding to a mild infection, a severe infection, a mild removal, and a severe removal, respectively, where (i, j_M, j_S, k_M) is the state of $X^{(n)}$ at a given time.

The transition intensities are expressed in the above form to indicate that $X^{(n)}$ is a density dependent population process, as defined by Ethier and Kurtz (1986, Chapter 11). Consider a sequence of epidemics, indexed by n, and suppose that $n^{-1}m_M^{(n)} \to \mu_M$ and $n^{-1}m_S^{(n)} \to \mu_S$ as $n \to \infty$, where $\mu_M + \mu_S > 0$. It follows, from Theorems 11.2.1 and 11.2.3 of Ethier and Kurtz (1986), that, as $n \to \infty$, $n^{-1}X^{(n)}$ converges almost surely over any finite time interval to $x = \{x(t): t \geq 0\}$, where $x(t) = (x(t), y_M(t), y_S(t), z_M(t))^\top$ is given by the solution of (2.5a)–(2.5e) with initial condition (2.6), and a central limit theorem holds for fluctuations of $X^{(n)}$ about x.

Let $\tau^{(n)} = \inf\{t > 0: Y_M^{(n)}(t) + Y_S^{(n)}(t) = 0\}$ denote the duration of the nth epidemic and let $\tau = \inf\{t > 0: y_M(t) + y_S(t) \leq 0\}$. Then $X^{(n)}(\tau^{(n)})$ is the final outcome of the nth epidemic and Theorem 11.4.1 of Ethier and Kurtz (1986) yields a central limit theorem for $X^{(n)}(\tau^{(n)})$, provided $\tau < \infty$. However, $\tau = \infty$, since τ is clearly greater than the time until all initial infectives have been removed and the latter equals ∞ in the deterministic model as the number of initial infectives decays exponentially. This difficulty can be overcome by using the following random time-scale transformation of $X^{(n)}$; cf. Ethier and Kurtz (1986, p. 467).

For $t \geq 0$, let

$$A^{(n)}(t) = \int_0^t n^{-1}(\lambda_M Y_M^{(n)}(u) + \lambda_S Y_S^{(n)}(u))\, du$$

denote the total force of infection exerted on a given susceptible during $[0, t]$, and let $A^{(n)} = A^{(n)}(\infty) (= A^{(n)}(\tau^{(n)}))$. For $0 \leq t \leq A^{(n)}$, let

$$U^{(n)}(t) = \inf\{u \geq 0: A^{(n)}(u) = t\} \quad \text{and} \quad \tilde{X}^{(n)}(t) = X^{(n)}(U^{(n)}(t)),$$

and write $\tilde{X}^{(n)}(t) = (\tilde{X}^{(n)}(t), \tilde{Y}_M^{(n)}(t), \tilde{Y}_S^{(n)}(t), \tilde{Z}_M^{(n)}(t))^\top$. The process

$$\tilde{X}^{(n)} = \{\tilde{X}^{(n)}(t): 0 \leq t \leq A^{(n)}\}$$

is a random time-scale transformation of $X^{(n)}$, obtained by running the clock at rate

$$n(\lambda_M Y_M^{(n)}(t) + \lambda_S Y_S^{(n)}(t))^{-1}.$$

Hence, $\tilde{X}^{(n)}$ is a continuous-time Markov chain with transition intensities

$$\tilde{q}^{(n)}_{(i,j_M,j_S,k_M),(i-1,j_M+1,j_S,k_M)} = \frac{n(i/n)(\lambda_M p^{(M)}_M(j_M/n) + \lambda_S p^{(S)}_M(j_S/n))}{\lambda_M(j_M/n) + \lambda_S(j_S/n)},$$

$$\tilde{q}^{(n)}_{(i,j_M,j_S,k_M),(i-1,j_M,j_S+1,k_M)} = \frac{n(i/n)(\lambda_M p^{(M)}_S(j_M/n) + \lambda_S p^{(S)}_S(j_S/n))}{\lambda_M(j_M/n) + \lambda_S(j_S/n)},$$

$$\tilde{q}^{(n)}_{(i,j_M,j_S,k_M),(i,j_M-1,j_S,k_M+1)} = \frac{n\gamma_M(j_M/n)}{\lambda_M(j_M/n) + \lambda_S(j_S/n)},$$

$$\tilde{q}^{(n)}_{(i,j_M,j_S,k_M),(i,j_M,j_S-1,k_M)} = \frac{n\gamma_S(j_S/n)}{\lambda_M(j_M/n) + \lambda_S(j_S/n)},$$

and initial condition $\tilde{X}^{(n)}(0) = (n, m^{(n)}_M, m^{(n)}_S, 0)^\top$.

The possible jumps of $\tilde{X}^{(n)}$ from a typical state $\boldsymbol{i} = (i, j_M, j_S, k_M)^\top$ are

$$\boldsymbol{\Delta} = \{(-1, 1, 0, 0)^\top, (-1, 0, 1, 0)^\top, (0, -1, 0, 1)^\top, (0, 0, -1, 0)^\top\}.$$

The intensities of these jumps admit the form $n\tilde{\beta}_l(n^{-1}\boldsymbol{i})$, $\boldsymbol{l} \in \boldsymbol{\Delta}$, with the functions $\tilde{\beta}_l$ ($\boldsymbol{l} \in \boldsymbol{\Delta}$) given by

$$\tilde{\beta}_{(-1,1,0,0)}(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \frac{\tilde{x}(\lambda_M p^{(M)}_M \tilde{y}_M + \lambda_S p^{(S)}_M \tilde{y}_S)}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}, \tag{4.1a}$$

$$\tilde{\beta}_{(-1,0,1,0)}(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \frac{\tilde{x}(\lambda_M p^{(M)}_S \tilde{y}_M + \lambda_S p^{(S)}_S \tilde{y}_S)}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}, \tag{4.1b}$$

$$\tilde{\beta}_{(0,-1,0,1)}(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \frac{\gamma_M \tilde{y}_M}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}, \tag{4.1c}$$

$$\tilde{\beta}_{(0,0,-1,0)}(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \frac{\gamma_S \tilde{y}_S}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}. \tag{4.1d}$$

The processes $\tilde{X}^{(n)}$ ($n \geq 1$) can be defined on a probability space, $(\Omega, \mathcal{F}, P)$ say, using a special case of a construction given in Ethier and Kurtz (1986, Chapter 11). Let

$$N_l = \{N_l(t) : t \geq 0\}, \qquad \boldsymbol{l} \in \boldsymbol{\Delta},$$

be independent unit-rate Poisson processes defined on $(\Omega, \mathcal{F}, P)$. Then, for $n \geq 1$, a realisation of the nth epidemic process is given by

$$\tilde{X}^{(n)}(t) = \tilde{X}^{(n)}(0) + \sum_{l \in \boldsymbol{\Delta}} l N_l\left(n \int_0^t \tilde{\beta}_l(n^{-1}\tilde{X}^{(n)}(u)) \, du\right).$$

Let

$$\tilde{F}(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \sum_{l \in \boldsymbol{\Delta}} l \tilde{\beta}_l(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M),$$

and, for $t \geq 0$, let $\tilde{\boldsymbol{x}}(t) = (\tilde{x}(t), \tilde{y}_M(t), \tilde{y}_S(t), \tilde{z}_M(t))^\top$ be defined by

$$\tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}(0) + \int_0^t \tilde{F}(\tilde{\boldsymbol{x}}(u)) \, du, \tag{4.2}$$

where $\tilde{\boldsymbol{x}}(0) = (1, \mu_M, \mu_S, 0)^\top$. Thus, $\tilde{\boldsymbol{x}}(t)$ satisfies the differential equation

$$\frac{d\tilde{x}}{dt} = -\tilde{x}, \tag{4.3a}$$

$$\frac{d\tilde{y}_M}{dt} = \frac{\tilde{x}(\lambda_M p_M^{(M)} \tilde{y}_M + \lambda_S p_M^{(S)} \tilde{y}_S) - \gamma_M \tilde{y}_M}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}, \tag{4.3b}$$

$$\frac{d\tilde{y}_S}{dt} = \frac{\tilde{x}(\lambda_M p_S^{(M)} \tilde{y}_M + \lambda_S p_S^{(S)} \tilde{y}_S) - \gamma_S \tilde{y}_S}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}, \tag{4.3c}$$

$$\frac{d\tilde{z}_M}{dt} = \frac{\gamma_M \tilde{y}_M}{\lambda_M \tilde{y}_M + \lambda_S \tilde{y}_S}. \tag{4.3d}$$

Note that $\tilde{\boldsymbol{x}}(t)$ is defined for $0 \le t \le \tilde{\tau}$, where $\tilde{\tau} = \inf\{t \ge 0 : \tilde{y}_M + \tilde{y}_S \le 0\}$. Now $\tilde{\tau} < \infty$. For an informal proof, note that $\tilde{\tau}$ is the limit as n tends to ∞ of the total force of infection acting on a given individual in the epidemic $X^{(n)}$, which must be less than the corresponding quantity if everyone were to become infected, which in turn must be less than $(1 + \mu_M + \mu_S) \max(\gamma_M^{-1}, \gamma_S^{-1})$. For a formal proof, (4.3a) implies that $\tilde{x}(t) = \exp(-t)$, and (4.3b) and (4.3c) imply that $(\tilde{y}_M + \tilde{y}_S)'(t) \le \tilde{x} - (\min(\gamma_M, \gamma_S)/\max(\lambda_M, \lambda_S))$, where $'$ denotes the derivative with respect to t. Thus, there exist $t_0 > 0$ and $a < 0$ such that $(\tilde{y}_M + \tilde{y}_S)'(t) \le a$ for $t \ge t_0$; whence $\tilde{\tau} < \infty$.

The process $\tilde{X}^{(n)}$ terminates at time $t = \tilde{\tau}^{(n)}$, where $\tilde{\tau}^{(n)} = \inf\{t > 0 : \tilde{Y}_M^{(n)}(t) + \tilde{Y}_S^{(n)}(t) = 0\}$. In the following we assume that $\tilde{X}^{(n)}(t) = \tilde{X}^{(n)}(\tilde{\tau}^{(n)})$ for $t > \tilde{\tau}^{(n)}$.

Theorem 4.2. *For* $0 \le t \le \tilde{\tau}$,

$$\lim_{n \to \infty} \sup_{0 \le u \le t} |n^{-1}\tilde{X}^{(n)}(u) - \tilde{\boldsymbol{x}}(u)| = 0 \quad \textit{almost surely.} \tag{4.4}$$

Proof. For $0 \le t < \tilde{\tau}$, (4.4) follows directly from Theorem 11.2.1 of Ethier and Kurtz (1986), since $\tilde{F}(\tilde{\boldsymbol{x}})$ is Lipschitz-continuous in some small neighbourhood of $\{\tilde{\boldsymbol{x}}(u) : 0 \le u \le t\}$. There exists $E \in \mathcal{F}$, with $P(E) = 1$, such that, for $\boldsymbol{l} \in \boldsymbol{\Delta}$, $t \ge 0$, and $\omega \in E$, $\lim_{n \to \infty} \sup_{u \le t} |n^{-1}(N_l(nu, \omega) - nu)| = 0$. (This is a key observation in the proof of Theorem 11.2.1 of Ethier and Kurtz (1986) and it implies that, if $\omega \in E$, the limit in (4.4) holds for all $t < \tilde{\tau}$.) Fix an $\omega \in E$ and an $\varepsilon > 0$. The intensity functions $\tilde{\beta}_l$ ($\boldsymbol{l} \in \boldsymbol{\Delta}$) are bounded on $S = [0, 1) \times (0, \infty)^3$, so

$$M_\beta = \sum_{\boldsymbol{l} \in \boldsymbol{\Delta}} |\boldsymbol{l}| \sup_{\tilde{\boldsymbol{x}} \in S} \tilde{\beta}(\tilde{\boldsymbol{x}}) < \infty.$$

There exists $t_1 \in (\tilde{\tau} - \varepsilon/3M_\beta, \tilde{\tau})$ such that $|\tilde{\boldsymbol{x}}(\tilde{\tau}) - \tilde{\boldsymbol{x}}(t_1)| < \varepsilon/3$. Since $\omega \in E$,

$$|n^{-1}\tilde{X}^{(n)}(t_1, \omega) - \tilde{\boldsymbol{x}}(t_1)| < \frac{\varepsilon}{3} \quad \text{and} \quad |n^{-1}\tilde{X}^{(n)}(\tilde{\tau}, \omega) - n^{-1}\tilde{X}^{(n)}(t_1, \omega)| < \frac{\varepsilon}{3}$$

for all sufficiently large n. Thus, $|n^{-1}\tilde{X}^{(n)}(\tilde{\tau}) - \tilde{\boldsymbol{x}}(\tilde{\tau})| < \varepsilon$ for all sufficiently large n, and (4.4) follows with $t = \tilde{\tau}$ since $\varepsilon > 0$ is arbitrary and $P(E) = 1$.

Note that $\tilde{X}^{(n)}(\tilde{\tau}^{(n)})$ yields the final outcome of the nth epidemic. For $0 \le t \le \tilde{\tau}^{(n)}$, let $\tilde{\boldsymbol{Z}}^{(n)}(t) = (\tilde{Z}_M^{(n)}(t), \tilde{Z}_S^{(n)}(t))$, where $\tilde{Z}_S^{(n)}(t) = n + m_M^{(n)} + m_S^{(n)} - \tilde{X}^{(n)}(t) - \tilde{Y}_M^{(n)}(t) - \tilde{Y}_S^{(n)}(t) - \tilde{Z}_M^{(n)}(t)$ denotes the number of severe removals in $\tilde{X}^{(n)}$ during $(0, t]$. For $0 \le t \le \tilde{\tau}$, let $\tilde{\boldsymbol{z}}(t) = (\tilde{z}_M(t), \tilde{z}_S(t))$, where $\tilde{z}_S(t) = 1 + \mu_M + \mu_S - \tilde{x}(t) - \tilde{y}_M(t) - \tilde{y}_S(t) - \tilde{z}_M(t)$. In the notation of Section 2.2.2, $(R_M^{(n)}, R_S^{(n)}) = (\tilde{Z}_M^{(n)}(\tilde{\tau}^{(n)}) - m_M^{(n)}, \tilde{Z}_M^{(n)}(\tilde{\tau}^{(n)}) - m_S^{(n)})$ and $(r_M, r_S) = (\tilde{z}_M(\tilde{\tau}) - \mu_M, \tilde{z}_S(\tilde{\tau}) - \mu_S)$.

Theorem 4.3. *For any $\varepsilon > 0$,*

$$\lim_{n \to \infty} P(|n^{-1}\tilde{\boldsymbol{Z}}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{\boldsymbol{z}}(\tau)| < \varepsilon) = 1, \tag{4.5}$$

i.e. $n^{-1}(R_{\mathrm{M}}^{(n)}, R_{\mathrm{S}}^{(n)}) \xrightarrow{\mathrm{P}} (r_{\mathrm{M}}, r_{\mathrm{S}})$ as $n \to \infty$, where '$\xrightarrow{\mathrm{P}}$' denotes convergence in probability.

Proof. For $0 \leq \xi \leq 1$, let $\boldsymbol{Y}_\xi = \{(Y_{\mathrm{M}\xi}(t), Y_{\mathrm{S}\xi}(t)) : t \geq 0\}$ be the two-type linear birth-and-death process having birth-rate matrix

$$\boldsymbol{\Lambda}(\xi) = \xi \begin{bmatrix} \lambda_{\mathrm{M}} p_{\mathrm{M}}^{(\mathrm{M})} & \lambda_{\mathrm{M}} p_{\mathrm{S}}^{(\mathrm{M})} \\ \lambda_{\mathrm{S}} p_{\mathrm{M}}^{(\mathrm{S})} & \lambda_{\mathrm{S}} p_{\mathrm{S}}^{(\mathrm{S})} \end{bmatrix},$$

and death rates γ_{M} and γ_{S}. The process $\boldsymbol{Y}_\xi$ has basic reproduction number $R_0(\xi) = \xi R_0$, where R_0 is given by (2.2). Note that the elements of $\boldsymbol{\Lambda}(\xi)$ give the rates with which a given mild or severe infective creates further mild and severe infectives when a fraction ξ of the initial susceptibles are still susceptible. Note that there exists $t_0 \in (0, \tilde{\tau})$ such that $R_0(\tilde{x}(t_0)) < 1$. (If this were not the case then the process of infectives in the untransformed deterministic model (2.5a)–(2.5e) would be bounded below by a deterministic linear two-type birth-and-death process having reproduction number greater than or equal to 1 and, hence, $(y_{\mathrm{M}}(t), y_{\mathrm{S}}(t))$ would not tend to $(0, 0)$ as t tended to ∞.) For $0 \leq \xi \leq 1$, let $T^{(\mathrm{M})}(\xi)$ denote the total progeny in $\boldsymbol{Y}_\xi$ if $\boldsymbol{Y}_\xi(0) = (1, 0)$, counting both types and including the initial ancestor, and define $T^{(\mathrm{S})}(\xi)$ similarly when $\boldsymbol{Y}_\xi(0) = (0, 1)$.

For $0 \leq t \leq \tilde{\tau}^{(n)}$, let

$$\tilde{T}_R^{(n)}(t) = \tilde{Z}_{\mathrm{M}}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{Z}_{\mathrm{M}}^{(n)}(t) + \tilde{Z}_{\mathrm{S}}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{Z}_{\mathrm{S}}^{(n)}(t)$$

denote the total number of removals that occur in $\tilde{X}^{(n)}$ after time t. For $0 \leq \xi \leq 1$ and $i \in \{\mathrm{M}, \mathrm{S}\}$, let $T_1^{(i)}(\xi), T_2^{(i)}(\xi), \ldots$ be independent and identically distributed copies of $T^{(i)}(\xi)$. Observe that $T^{(\mathrm{M})}(\xi)$ and $T^{(\mathrm{S})}(\xi)$ are each stochastically increasing in ξ and, since the birth-and-death process $\boldsymbol{Y}_\xi$ provides an upper bound for the epidemic process after a fraction $1 - \xi$ of susceptibles have been infected,

$$\tilde{T}_R^{(n)}(t) \overset{\mathrm{ST}}{\leq} \sum_{i=1}^{\tilde{Y}_{\mathrm{M}}^{(n)}(t)} T_i^{(\mathrm{M})}(n^{-1}\tilde{X}^{(n)}(t)) + \sum_{i=1}^{\tilde{Y}_{\mathrm{S}}^{(n)}(t)} T_i^{(\mathrm{S})}(n^{-1}\tilde{X}^{(n)}(t)) \tag{4.6}$$

for any $t \in [0, \tilde{\tau}^{(n)})$, where '$\overset{\mathrm{ST}}{\leq}$' denotes stochastic ordering. These observations can be proved formally using the coupling described at the beginning of Section 4.3.

Theorem 4.2 implies that, for any $t \in (t_0, \tilde{\tau})$,

$$\lim_{n \to \infty} P(\tilde{Y}_{\mathrm{M}}^{(n)}(t) \leq \tfrac{3}{2}\tilde{y}_{\mathrm{M}}(t)) = 1, \qquad \lim_{n \to \infty} P(\tilde{Y}_{\mathrm{S}}^{(n)}(t) \leq \tfrac{3}{2}\tilde{y}_{\mathrm{S}}(t)) = 1, \tag{4.7}$$

and

$$\lim_{n \to \infty} P(n^{-1}\tilde{X}^{(n)}(t) \leq \tilde{x}(t_0)) = 1. \tag{4.8}$$

For $0 \leq \xi \leq 1$, let

$$\mu_T^{(\mathrm{M})}(\xi) = \mathrm{E}[T^{(\mathrm{M})}(\xi)] \quad \text{and} \quad \mu_T^{(\mathrm{S})}(\xi) = \mathrm{E}[T^{(\mathrm{S})}(\xi)],$$

and note that $\mu_T^{(M)}(\tilde{x}(t_0)) < \infty$ and $\mu_T^{(S)}(\tilde{x}(t_0)) < \infty$ since $R_0(\tilde{x}(t_0)) < 1$. Using, (4.7), (4.8), the stochastic monotonicity of $T^{(M)}(\xi)$ and $T^{(S)}(\xi)$, and the weak law of large numbers, it follows, from (4.6), that, for any $t \in (t_0, \tilde{\tau})$,

$$\lim_{n \to \infty} \mathrm{P}(n^{-1}\tilde{T}_R^{(n)}(t) \le 2[\tilde{y}_M(t)\mu_T^{(M)}(\tilde{x}(t_0)) + \tilde{y}_S(t)\mu_T^{(S)}(\tilde{x}(t_0))]) = 1. \tag{4.9}$$

Let $\varepsilon > 0$ be given. Since $\tilde{z}(t) \uparrow \tilde{z}(\tilde{\tau})$ and $(\tilde{y}_M(t), \tilde{y}_S(t)) \to (0, 0)$ as $t \to \tilde{\tau}-$, there exists $t_1 \in (t_0, \tilde{\tau})$ such that

$$\tilde{z}_M(\tilde{\tau}) - \tilde{z}_M(\tilde{t_1}) + \tilde{z}_S(\tilde{\tau}) - \tilde{z}_S(t_1) < \frac{\varepsilon}{3} \quad \text{and} \quad \tilde{y}_M(t)\mu_T^{(M)}(\tilde{x}(t_0)) + \tilde{y}_S(t)\mu_T^{(S)}(\tilde{x}(t_0)) < \frac{\varepsilon}{6}.$$

Then $\lim_{n \to \infty} \mathrm{P}(|n^{-1}(\tilde{\boldsymbol{Z}}(\tilde{\tau}^{(n)}) - \tilde{\boldsymbol{Z}}(t_1))| < \varepsilon/3) = 1$, using (4.9), and $|\tilde{z}(t_1) - \tilde{z}(\tilde{\tau})| < \varepsilon/3$. Furthermore, $\lim_{n \to \infty} \mathrm{P}(|n^{-1}\tilde{\boldsymbol{Z}}^{(n)}(t_1) - \tilde{z}(t_1)| < \varepsilon/3) = 1$, using Theorem 4.2. The last three facts imply (4.5), as required.

We now seek a central limit theorem (CLT) for the final outcome of the IDS epidemic. Let $\partial \tilde{F}(\tilde{\boldsymbol{x}}) = [\partial_j \tilde{F}_i(\tilde{\boldsymbol{x}})]$ denote the matrix of first partial derivatives of $\tilde{F}(\tilde{\boldsymbol{x}})$ and, for $0 \le s \le t \le \tilde{\tau}$, let $\tilde{\Phi}(t, s)$ be the solution of the matrix differential equation

$$\frac{\partial}{\partial t}\tilde{\Phi}(t, s) = \partial \tilde{F}(\tilde{\boldsymbol{x}}(t))\tilde{\Phi}(t, s), \quad \tilde{\Phi}(s, s) = \boldsymbol{I}, \tag{4.10}$$

where $\boldsymbol{I}$ denotes the 4×4 identity matrix. Let $\tilde{G}(\tilde{\boldsymbol{x}}) = \sum_{l \in \boldsymbol{\Delta}} \boldsymbol{l}\boldsymbol{l}^\top \tilde{\beta}_l(\tilde{\boldsymbol{x}})$. The following theorem is an immediate consequence of Theorem 11.2.3 of Ethier and Kurtz (1986) since $\sum_{l \in \boldsymbol{\Delta}} |\boldsymbol{l}|^2 \sup_{\tilde{\boldsymbol{x}} \in S} \tilde{\beta}_l(\tilde{\boldsymbol{x}}) < \infty$.

Theorem 4.4. *For $t \in [0, \tilde{\tau}]$, let $\tilde{\boldsymbol{V}}^{(n)}(t) = \sqrt{n}(n^{-1}\tilde{\boldsymbol{X}}^{(n)}(t) - \tilde{\boldsymbol{x}}(t))$, and suppose that*

$$\lim_{n \to \infty} \sqrt{n}(n^{-1}m_M^{(n)} - \mu_M) = 0 \quad \text{and} \quad \lim_{n \to \infty} \sqrt{n}(n^{-1}m_S^{(n)} - \mu_S) = 0.$$

Then, for any $t_0 < \tilde{\tau}$,

$$\{\tilde{\boldsymbol{V}}^{(n)}(t) : 0 \le t \le t_0\} \xrightarrow{\mathrm{w}} \{\tilde{\boldsymbol{V}}(t) : 0 \le t \le t_0\},$$

where '$\xrightarrow{\mathrm{w}}$' denotes weak convergence in the space of right-continuous functions from $[0, t_0] \to \mathbb{R}^4$ with left limits, endowed with the Skorohod topology, and $\{\tilde{\boldsymbol{V}}(t) : 0 \le t \le t_0\}$ is a zero-mean Gaussian process with $\tilde{\boldsymbol{V}}(0) = \boldsymbol{0}$ and covariance function given by

$$\mathrm{cov}(\tilde{\boldsymbol{V}}(t), \tilde{\boldsymbol{V}}(s)) = \int_0^{\min(t,s)} \tilde{\Phi}(t, u)\tilde{G}(\tilde{\boldsymbol{x}}(u))(\tilde{\Phi}(s, u))^\top \, \mathrm{d}u. \tag{4.11}$$

The final outcome of the nth epidemic is $\tilde{\boldsymbol{X}}^{(n)}(\tilde{\tau}^{(n)})$ and, in the notation of Section 2.2.2, $R_M^{(n)} = \tilde{Z}_M^{(n)}(\tilde{\tau}^{(n)}) - m_M^{(n)}$ and $R_S^{(n)} = n + m_M^{(n)} - \tilde{X}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{Z}_M^{(n)}(\tilde{\tau}^{(n)})$. Theorem 11.4.1 of Ethier and Kurtz (1986) would yield a CLT for $\tilde{\boldsymbol{X}}^{(n)}(\tilde{\tau}^{(n)})$, but to apply that theorem as stated requires the processes $\tilde{\boldsymbol{X}}^{(n)}$ and $\tilde{\boldsymbol{x}}$ to be defined on the interval $[0, t_1]$ for some $t_1 > \tilde{\tau}$, and that Theorems 4.2 and 4.4 hold on $[0, t_1]$.

Suppose that $\theta = \lim_{t \to \tilde{\tau}-}[\tilde{y}_M(t)/\tilde{y}_S(t)]$ exists. Then, using L'Hôpital's rule,

$$\begin{aligned}
\theta &= \lim_{t \to \tilde{\tau}-} \frac{\tilde{y}_M'(t)}{\tilde{y}_S'(t)} = \lim_{t \to \tilde{\tau}-}\left[\frac{\tilde{x}(t)(\lambda_M p_M^{(M)}\tilde{y}_M(t) + \lambda_S p_M^{(S)}\tilde{y}_S(t)) - \gamma_M\tilde{y}_M(t)}{\tilde{x}(t)(\lambda_M p_S^{(M)}\tilde{y}_M(t) + \lambda_S p_S^{(S)}\tilde{y}_S(t)) - \gamma_S\tilde{y}_S(t)}\right] \\
&= \frac{\tilde{x}(\tilde{\tau})(\lambda_M p_M^{(M)}\theta + \lambda_S p_M^{(S)}) - \gamma_M\theta}{\tilde{x}(\tilde{\tau})(\lambda_M p_S^{(M)}\theta + \lambda_S p_S^{(S)}) - \gamma_S},
\end{aligned}$$

provided that $\tilde{x}(\tilde{\tau})(\lambda_M p_S^{(M)}\theta + \lambda_S p_S^{(S)}) - \gamma_S \neq 0$. If this is the case then θ satisfies

$$\tilde{x}(\tilde{\tau})\lambda_M p_S^{(M)}\theta^2 + (x(\tilde{\tau})(\lambda_S p_S^{(S)} - \lambda_M p_M^{(M)}) + \gamma_M - \gamma_S)\theta - \tilde{x}(\tilde{\tau})\lambda_S p_M^{(S)} = 0. \quad (4.12)$$

Now $p_S^{(M)} > 0$ and $p_M^{(S)} > 0$, so the roots of (4.12) have opposite signs and θ is given by the positive root as it clearly cannot be negative. We do not have a proof that the limit θ exists or that $\tilde{x}(\tilde{\tau})(\lambda_M p_S^{(M)}\theta + \lambda_S p_S^{(S)}) - \gamma_S \neq 0$. However, the latter is easily checked numerically for any given case and our numerical studies support the former. We assume that these results both hold. It then follows, from (4.3a)–(4.3d), that $\tilde{F}(\tilde{x}(\tilde{\tau}))$ is well defined and, thus, $\tilde{x}$ can be extended to an interval of the form $[0, t_1]$ above. The jump intensities (4.1a)–(4.1d) can be extended similarly, allowing $\tilde{X}^{(n)}$ to be defined on $[0, t_1]$, but the corresponding drift function $\tilde{F}$ is not Lipschitz-continuous at $\tilde{x}(\tilde{\tau})$. We proceed on the basis that Theorem 11.4.1 of Ethier and Kurtz (1986) still applies in this setting.

Let $\phi(\tilde{x}, \tilde{y}_M, \tilde{y}_S, \tilde{z}_M) = \tilde{y}_M + \tilde{y}_S$, so $\tilde{\tau}^{(n)} = \inf\{t > 0 : \phi(\tilde{X}^{(n)}(t)) \leq 0\}$ and $\tilde{\tau} = \inf\{t > 0 : \phi(\tilde{x}(t)) \leq 0\}$. Now $\nabla\phi(\tilde{x}) = (0, 1, 1, 0)$ and, using (4.1a)–(4.1d),

$$\nabla\phi(\tilde{x}(\tilde{\tau}))\tilde{F}(\tilde{x}(\tilde{\tau}))$$
$$= \frac{[\tilde{x}(\tilde{\tau})(\lambda_M p_M^{(M)}\theta + \lambda_S p_M^{(S)}) - \gamma_M\theta] + [\tilde{x}(\tilde{\tau})(\lambda_M p_S^{(M)}\theta + \lambda_S p_S^{(S)}) - \gamma_S]}{\lambda_M\theta + \lambda_S}. \quad (4.13)$$

Since $\theta > 0$ and $\tilde{x}(\tilde{\tau})(\lambda_M p_S^{(M)}\theta + \lambda_S p_S^{(S)}) - \gamma_S \neq 0$, the two terms in square brackets in (4.13) are nonzero and have the same sign, which must be negative from the definition of $\tilde{\tau}$. Thus, $\nabla\phi(\tilde{x}(\tilde{\tau})) < 0$; whence Theorem 11.4.1 of Ethier and Kurtz (1986) yields

$$\sqrt{n}(n^{-1}\tilde{X}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{x}(\tilde{\tau})) \overset{D}{\to} \tilde{V}(\tilde{\tau}) - \frac{(0, 1, 1, 0)\tilde{V}(\tilde{\tau})}{(0, 1, 1, 0)\tilde{F}(\tilde{x}(\tilde{\tau}))}\tilde{F}(\tilde{x}(\tilde{\tau})) \quad \text{as } n \to \infty. \quad (4.14)$$

Let $\tilde{\Sigma}(t) = \text{cov}(\tilde{V}(t), \tilde{V}(t))$ denote the variance-covariance matrix of $\tilde{V}(t)$, and define matrices B and C by

$$B = I - \frac{\tilde{F}(\tilde{x}(\tilde{\tau}))(0, 1, 1, 0)}{(0, 1, 1, 0)\tilde{F}(\tilde{x}(\tilde{\tau}))} \quad \text{and} \quad C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 \end{bmatrix}. \quad (4.15)$$

Then (4.14) implies that

$$\sqrt{n}(n^{-1}\tilde{X}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{x}(\tilde{\tau})) \overset{D}{\to} N(\mathbf{0}, B\tilde{\Sigma}(\tilde{\tau})B^\top) \quad \text{as } n \to \infty. \quad (4.16)$$

Hence, the limiting variance-covariance matrix $\Sigma(\mu_M, \mu_S)$ in (2.7) is given by $\Sigma(\mu_M, \mu_S) = CB\tilde{\Sigma}(\tilde{\tau})B^\top C^\top$.

Note that it follows, from (4.10) and (4.11), that $\tilde{\Sigma}(t)$ satisfies the differential equation

$$\frac{d\tilde{\Sigma}}{dt} = \tilde{G}(\tilde{x}) + \partial\tilde{F}(\tilde{x})\tilde{\Sigma} + \tilde{\Sigma}[\partial\tilde{F}(\tilde{x})]^\top, \quad (4.17)$$

with initial condition $\tilde{\Sigma}(0) = \mathbf{0}$. Thus, $\tilde{\Sigma}(\tilde{\tau})$ can be computed by numerically solving the differential equations (4.3a)–(4.3d) and (4.17) simultaneously.

4.3. Few initial infectives: heuristics for a CLT in the case of a major outbreak

We continue to assume that the infectious periods follow exponential distributions, with rate γ_M for mild infectives and rate γ_S for severe infectives, but, as in Section 4.1, we now consider a sequence of epidemics $\{E^{(n)}: n \geq 1\}$, indexed by the initial number of susceptibles n, with each epidemic having m_M initial mild infectives and m_S initial severe infectives. Let $Y = \{(Y_M(t), Y_S(t)): t \geq 0\}$ be the branching process defined in Section 4.1, but with the above exponential lifetime distributions.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, on which is defined the branching process Y and, independently, a sequence $U_1, U_2, \ldots$ of independent and identically distributed random variables that are uniformly distributed on $(0, 1)$. For $\varepsilon \in (0, 1)$, let Y_ε denote the branching process obtained from Y by, for $i = 1, 2, \ldots$, deleting the ith birth in Y and all of its descendants if $U_i \leq \varepsilon$. For $n = 1, 2, \ldots$, a realisation of the epidemic $E^{(n)}$ is obtained from Y, by associating births with infections and, for $i = 1, 2, \ldots$, deleting the ith birth in Y and all of its descendants if $U_i < 1 - n^{-1} X^{(n)}(s-)$, where s is the time of the ith birth in Y. For $t \geq 0$, let $T(t)$, $Z_M(t)$, and $Z_S(t)$ respectively denote the total number of births, the total number of mild deaths, and the total number of severe deaths during $(0, t]$ in Y. Define $Y_{M\varepsilon}(t)$, $Y_{S\varepsilon}(t)$, and $T_\varepsilon(t)$ for Y_ε in the obvious fashion. Observe that if $\tau_\varepsilon^{(n)} = \inf\{t \geq 0: X^{(n)}(t) < (1 - \varepsilon)n\}$, where $\varepsilon \in (0, 1)$, then

$$n - T(t) \leq X^{(n)}(t) \leq n - T_\varepsilon(t), \qquad t \in [0, \tau_\varepsilon^{(n)}]. \tag{4.18}$$

Consider the branching process Y. For $t \geq 0$, let $\bar{m}_{ij}(t) = E[Y_j(t)]$ given that initially there is one individual, whose type is i, and let $\bar{M}(t) = [\bar{m}_{ij}(t)]$. The infinitesimal generator of the semigroup $\{\bar{M}(t): t \geq 0\}$ is

$$\bar{A} = \begin{bmatrix} \lambda_M p_M^{(M)} - \gamma_M & \lambda_M p_S^{(M)} \\ \lambda_S p_M^{(S)} & \lambda_S p_S^{(S)} - \gamma_S \end{bmatrix}.$$

Let α denote the maximal eigenvalue of $\bar{A}$ (i.e. the Malthusian parameter of Y), and let $v = (v_M, v_S)$ be the corresponding left eigenvector, normalized so that $v_M + v_S = 1$. The branching process Y is positive regular (as $p_S^{(M)} > 0$ and $p_M^{(S)} > 0$) and nonsingular, so it follows, using Theorem 2 of Athreya and Ney (1972, p. 206) and the theory of asymptotic growth and stabilisation of general multitype branching processes (see, e.g. Jagers (1991)), that there exists a random variable $W \geq 0$ such that, almost surely,

$$\lim_{t \to \infty} e^{-\alpha t} Y_M(t) = v_M W, \qquad \lim_{t \to \infty} e^{-\alpha t} Y_S(t) = v_S W, \tag{4.19}$$

and

$$\lim_{t \to \infty} e^{-\alpha t} T(t) = \frac{\lambda_M v_M + \lambda_S v_S}{\alpha} W. \tag{4.20}$$

Moreover, $W(\omega) > 0$ if and only if $\omega \in A_E^c$, i.e. if and only if Y does not go extinct.

Lemma 4.1. *Suppose that $R_0 > 1$. Then, there exists $\varepsilon_0 > 0$ such that*

$$\lim_{n \to \infty} P(X^{(n)}(\infty) \leq (1 - \varepsilon_0)n \mid A_E^c) = 1. \tag{4.21}$$

Proof. First note that (4.20) implies that, for P-almost all $\omega \in A_E^c$, $t_n = \inf\{t: T(t) > \log n\}$ is well defined and finite for $n = 1, 2, \ldots$. Also, by construction, the first $[\log n] + 1$ births in Y

correspond to infections in $E^{(n)}$ if and only if $U_i \geq 1 - (i-1)/n$ for $i = 1, 2, \ldots, [\log n] + 1$. A simple calculation then shows that

$$\lim_{n \to \infty} P(X^{(n)}(t_n) = n - T(t_n), \ Y_M^{(n)}(t_n) = Y_M(t_n), \ Y_S^{(n)}(t_n) = Y_S(t_n) \mid A_E^c) = 1. \quad (4.22)$$

It then follows, using (4.19) and (4.20), that

$$\lim_{n \to \infty} P\left(Y_M^{(n)}(t_n) > \frac{c_M}{2} \log n, \ Y_S^{(n)}(t_n) > \frac{c_S}{2} \log n \ \middle| \ A_E^c \right) = 1, \quad (4.23)$$

where

$$c_M = \lim_{n \to \infty} \frac{Y_M(t_n)}{T(t_n)} = \frac{\alpha v_M}{(\lambda_M v_M + \lambda_S v_S)} \quad \text{and} \quad c_S = \lim_{n \to \infty} \frac{Y_S(t_n)}{T(t_n)} = \frac{\alpha v_S}{(\lambda_M v_M + \lambda_S v_S)}.$$

The branching process Y_ε has basic reproduction number $R_0(\varepsilon) = (1-\varepsilon)R_0$. For $i = M, S$, let $\pi_i(\varepsilon)$ denote the extinction probability of Y_ε given that initially there is one individual, whose type is i. Since $R_0 > 1$, there exists $\varepsilon_0 > 0$ such that $R_0(\varepsilon_0) > 1$, which implies that $\pi_M(\varepsilon_0) < 1$ and $\pi_S(\varepsilon_0) < 1$.

For each $n = 1, 2, \ldots$, the construction described above (4.18) can be used to define a realisation of $E^{(n)}$, $\hat{E}^{(n)}$ say, and a realisation of Y_{ε_0}, with, in obvious notation,

$$\hat{Y}_M^{(n)}(0) = Y_{M\varepsilon_0}^{(n)}(0) = Y_M(t_n), \quad \hat{Y}_S^{(n)}(0) = Y_{S\varepsilon_0}^{(n)}(0) = Y_S(t_n), \quad \text{and} \quad \hat{X}^{(n)}(0) = n - T(t_n),$$

that are coupled so that $\hat{X}^{(n)}(t) \leq n - T_{\varepsilon_0}(t)$, provided that $\hat{X}^{(n)}(t) \geq (1 - \varepsilon_0)n$. Hence, recalling (4.22) and (4.23),

$$\lim_{n \to \infty} P(X^{(n)}(\infty) \leq (1 - \varepsilon_0)n \mid A_E^c) = \lim_{n \to \infty} P(\hat{X}^{(n)}(\infty) \leq (1 - \varepsilon_0)n)$$

$$\geq \lim_{n \to \infty} P(T_{\varepsilon_0}(\infty) \geq n\varepsilon_0)$$

$$\geq \lim_{n \to \infty} P(T_{\varepsilon_0}(\infty) = \infty)$$

$$\geq \lim_{n \to \infty} [1 - (\pi_M(\varepsilon_0))^{(c_M/2)\log n} (\pi_S(\varepsilon_0))^{(c_S/2)\log n}]$$

$$= 1,$$

and (4.21) follows.

It is convenient to return now to the construction of $(E^{(n)})$ and Y given in Section 4.1. As above, let $t_n = \inf\{t : T(t) > \log n\}$. Then, since $n^{-1/2}\eta^{(n)} \to \eta$ almost surely as $n \to \infty$, for P-almost all $\omega \in A_E^c$, the process of infectives in $E^{(n)}$ and the branching process Y coincide over $[0, t_n]$ for all sufficiently large n. Thus, for such n,

$$X^{(n)}(t_n) = n - T(t_n), \quad Y_M^{(n)}(t_n) = Y_M(t_n), \quad Y_S^{(n)}(t_n) = Y_S(t_n), \quad \text{and} \quad Z_M^{(n)}(t_n) = Z_M(t_n).$$

It then follows, using (4.19), (4.20), and a similar result for $Z_M(t)$, that, for P-almost all $\omega \in A_E^c$,

$$\lim_{n \to \infty} \left[\frac{Y_M^{(n)}(t_n)}{\log n} \right] = c_M, \quad \lim_{n \to \infty} \left[\frac{Y_S^{(n)}(t_n)}{\log n} \right] = c_S,$$

$$\text{and} \quad \lim_{n \to \infty} \left[\frac{Z_M^{(n)}(t_n)}{\log n} \right] = c_{ZM},$$

where $c_{ZM} = \gamma_M v_M/(\lambda_M v_M + \lambda_S v_S)$. Thus, as $n \to \infty$ and letting $g_n = n^{-1} \log n$,

$$\frac{1}{n}(X^{(n)}(t_n), Y_M^{(n)}(t_n), Y_S^{(n)}(t_n), Z_M^{(n)}(t_n)) \sim (1 - g_n, c_M g_n, c_S g_n, c_{ZM} g_n). \qquad (4.24)$$

Let $\hat{X}^{(n)}$ denote the time-changed process described in Section 4.2, but with $\hat{X}^{(n)}(0) = X^{(n)}(t_n)$. Recall that $\tau^{(n)} = \inf\{t > 0 \colon Y_M^{(n)}(t) + Y_S^{(n)}(t) = 0\}$, and let

$$\hat{\tau}^{(n)} = \inf\{t > 0 \colon \hat{Y}_M^{(n)}(t) + \hat{Y}_S^{(n)}(t) = 0\}.$$

In view of (4.24), let $\hat{x}^{(n)}$ denote the solution of the deterministic model (4.2) with

$$\hat{x}^{(n)}(0) = (1 - g_n, c_M g_n, c_S g_n, c_{ZM} g_n),$$

and let $\hat{x} = \lim_{n\to\infty} \hat{x}^{(n)}$. Then it is plausible that

$$n^{-1} \hat{X}^{(n)} \xrightarrow{W} x \quad \text{as } n \to \infty,$$

and that

$$n^{-1} \hat{X}^{(n)}(\hat{\tau}^{(n)}) \xrightarrow{P} \hat{x}(\hat{\tau}) \quad \text{as } n \to \infty,$$

where $\hat{\tau} = \inf\{t > 0 \colon \hat{y}_M(t) + \hat{y}_S(t) \le 0\}$.

Recall that $A^{(n)} = \int_0^\infty n^{-1}(\lambda_M Y_M^{(n)}(u) + \lambda_S^{(n)}(u))\, du$. Note that, since there are few initial infectives, it is possible that $A^{(n)} \to 0$ as $n \to \infty$, in which case the time-changed process $\tilde{X}^{(n)}$, defined in Section 4.2, is defined only for $t = 0$ in the limit as $n \to \infty$. Indeed, this is what happens if $\omega \in A_E$. However, Lemma 4.1 implies that there exists $\varepsilon_1 > 0$ such that

$$P(A^{(n)} \ge \varepsilon_1 \mid A_E^c) \to 1 \quad \text{as } n \to \infty,$$

so, conditional on the epidemic becoming established, the limiting processes $\lim_{n\to\infty} n^{-1} \tilde{X}^{(n)}$ and $\lim_{n\to\infty} n^{-1} \hat{X}^{(n)}$ do not get stuck at their initial state $(1, 0, 0, 0)$. Furthermore, since $\lim_{n\to\infty} n^{-1/2} \log n = 0$, it is also plausible that, conditional upon the outbreak becoming established, Theorem 4.4 and the asymptotic distribution of the final outcome given by (4.16) also hold with obvious modifications. In particular, the asymptotic mean and variance-covariance matrix of $\tilde{X}^{(n)}(\tilde{\tau}^{(n)})$, given A_E^c occurs, is obtained by setting $\tilde{x}(0) = (1 - \varepsilon, \varepsilon c_M, \varepsilon c_S, \varepsilon c_{ZM})^\top$ in (4.3a)–(4.3d), (4.15), and (4.16), and letting $\varepsilon \downarrow 0$.

5. Discussion

The aim of the paper is to incorporate the fact that many infectious diseases have varying severity of the disease, and this severity often affects an individual's degree of further spreading the disease. The IDS model allows for two different severities, mild and severe. It is quite straightforward to generalise the model to allow for an arbitrary number of severities, and the same proof techniques may be used. Another important step to making the model more realistic is to incorporate households and allow for the possiblility of a higher transmission rate and a greater proportion of severe contacts within households. In principle, a household model, e.g. Ball *et al.* (1997), could be combined with the present model for different severities, and this will be the subject of further research.

It would be of interest to study vaccination strategies particularly targeted at reducing the number of severe cases. In particular, if the basic epidemic model was multitype, e.g. incorporating age cohorts, it might be of special interest to reduce the number of severe cases in vulnerable groups, such as the very young and the very old.

The IDS model resembles a two-type epidemic model (e.g. Andersson and Britton (2000, Chapter 6)) in certain ways, with mild and severe specifying the two types. The difference between the IDS model and a two-type epidemic model is that an individual's type is not decided in advance, but only upon infection, and the probability that an infective becomes a specific type depends on who he or she was infected by. Nevertheless, in its initial stages, the IDS model behaves like a two-type epidemic model, in that given any IDS model the contact rates in the two-type epidemic can be chosen so that the two processes have the same approximating branching process. However, once an epidemic becomes established, the behaviours of the IDS and the two-type epidemic models no longer resemble each other.

In the IDS model the infectious state of an individual depends on the type of his or her infector. Ball and Britton (2005) treated a different model for varying severity in which each infected individual was initially mild upon infection, but later may become severely infected if additionally exposed to the disease. Which of the models is more realistic of course depends on the disease of interest, but probably a combination of the two would be most realistic. That is, where the severity of a given infective depends not only on who he or she was infected by but also on whether or not that infective has been exposed several times.

Acknowledgements

References

ALDOUS, D. J. (1985). *Exchangeability and Related Topics* (Ecole d'Été de Probabilités de Saint-Flour XIII—1983; Lecture Notes Math. **1117**), ed. P. L. Hennequin, Springer, Berlin.

ANDERSSON, H. AND BRITTON, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*. (Lecture Notes Statist. **151**). Springer, New York.

ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

BALL, F. G. AND BECKER, N. G. (2006). Control of transmission with two types of infection. *Math. Biosci.* **200,** 170–187.

BALL, F. G. AND BRITTON, T. (2005). An epidemic model with exposure-dependent severities. *J. Appl. Prob.* **42,** 932–949.

BALL, F. G. AND CLANCY, D. (1995). The final outcome of an epidemic with several different types of infective in a large population. *J. Appl. Prob.* **32,** 579–590.

BALL, F. G. AND DONNELLY, P. J. (1995). Strong approximations for epidemic models. *Stoch. Process. Appl.* **55,** 1–21.

BALL, F. G., MOLLISON, D. AND SCALIA-TOMBA, G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Prob.* **7,** 46–89.

BECKER, N. G. AND STARCZAK, D. N. (1998). The effect of random vaccine response on the vaccination coverage required to prevent epidemics. *Math. Biosci.* **154,** 117–135.

BECKER, N. G., BRITTON, T. AND O'NEILL, P. D. (2006). Estimating vaccine effects from studies of outbreaks in household pairs. *Statist. Medicine* **25,** 1079–1093.

ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.

HACCOU, P., JAGERS, P. AND VATUTIN, V. A. (2005). *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press.

HALLORAN, M. E., HABER, M. AND LONGINI, I. M. (1992). Interpretation and estimation of vaccine efficacy under heterogeneity. *Amer. J. Epidemiol.* **136,** 328–343.

HEESTERBEEK, J. A. P. AND DIETZ, K. (1996). The concept of R_0 in epidemic theory. *Statist. Neerlandica* **50,** 89–110.

JAGERS, P. (1991). The growth and stabilization of populations. *Statist. Sci.* **6,** 269–283.

KENDALL, W. S. AND SAUNDERS, I. W. (1983). Epidemics in competition. II. The general epidemic. *J. R. Statist. Soc. B* **45,** 238–244.

LEFÈVRE, C. (1990). Stochastic epidemic models for S-I-R infectious diseases: a brief survey of the recent theory. In *Stochastic Processes in Epidemic Theory* (Lecture Notes Biomath. **86**), eds J.-P. Gabriel *et al.*, Springer, Berlin, pp. 1–12.

MANGADA, M. N. M. AND IGARASHI, A. (1998). Molecular and in vitro analysis of eight dengue type 2 viruses isolated from patients exhibiting different disease severities. *Virology* **244,** 458–466.

MORLEY, D. C. AND AABY, P. (1997). Managing measles: size of infecting dose may be important. *British Medical J.* **314,** 1692.

PARANG, N. M. AND ARCHANA, C. (2004). Varicella. *eMedicine* 15 September 2004. Available at http://www.emedicine. com/ped/topic2385.htm.

PICARD, P. AND LEFÈVRE, C. (1990). A unified analysis of the final size and severity distribution in collective Reed–Frost epidemic processes. *Adv. Appl. Prob.* **22,** 269–294.

SCALIA-TOMBA, G.-P. (2005). Epidemic models, multi-strain. In *Encylopedia of Biostatistics*, 2nd edn, eds P. Armitage and N. Colton, John Wiley, Chichester, pp. 1672–1674.

SCALIA-TOMBA, G.-P. AND SVENSSON, Å. (2001). Competing epidemics in closed populations. Res. Rep. 2001:8, Department of Mathematical Statistics, Stockholm University.