



Mathematical Statistics
Stockholm University

**An epidemic model with exposure-
dependent severities**

Frank Ball and Tom Britton

Research Report 2005:7

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

<http://www.math.su.se/matstat>

An epidemic model with exposure-dependent severities

Frank Ball*, Tom Britton†

May 2005

Abstract

We consider a stochastic model for the spread of an SIR (susceptible \rightarrow infective \rightarrow removed) epidemic among a closed, finite population, in which there are two types of severity of infectious individuals, namely mild and severe. The type of severity depends on the amount of infectious exposure an individual receives, in that infectives are always initially mild but may become severe if additionally exposed. Large population properties of the model are derived. In particular, a coupling argument is used to provide a rigorous branching process approximation to the early stages of an epidemic, and an embedding argument is used to derive a strong law and associated central limit theorem for the final outcome of epidemics which take off. The basic reproduction number, which determines whether or not a major outbreak can occur given few initial infectives, depends only on parameters of the mild infectious state, whereas the final outcome in the event of a major outbreak depends also on parameters of the severe state. Moreover, the limiting final size proportions need not even be continuous in model parameters.

Keywords: Stochastic epidemic; final size; basic reproduction number; infections of varying severity; exposure; branching process; coupling; embedded process; weak convergence; central limit theorems.

AMS 2000 Subject Classification: Primary 92D30. Secondary 60F99; 60K99.

*School of Mathematical Sciences, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK. E-mail address: Frank.Ball@nottingham.ac.uk

†Postal address: Mathematical Statistics, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: tom.britton@math.su.se.

1 Introduction

This paper is concerned with an SIR (susceptible \rightarrow infective \rightarrow removed) epidemic model, describing the spread of an infectious disease in a closed finite community (see, for example, Lefèvre (1990) and Andersson and Britton (2000)). At a meeting in August 2003 in Mariefred, Sweden, comprising epidemiologists, mathematicians and statisticians, the need for epidemic models allowing different degrees of severity of disease was expressed by several epidemiologists. Examples of diseases for which this is relevant include measles (Aaby *et al.* (1998), Butler *et al.* (1994) and Morley and Aaby (1997)), varicella (Parang and Archana (2004)) and dengue fever (Mangada and Igarash (1998)). Typically, the degree of severity of a case depends on the amount of disease exposure an individual has received, and severe cases tend to spread the disease more than mild cases. In the present paper “severity” refers to how infectious an individual is and not on the degree of illness, though in reality the two are often correlated. It is the former that is relevant for the epidemic spread, which is the focus of the present paper, whereas the latter is of course important from both an individual and a national health perspective.

In this paper, a model which attempts to capture the feature of varying severity affecting the spread of a disease in a homogeneously mixing community is defined and studied. The model, which we call the exposure-dependent severity model (EDS for short), has two different severities, mild and severe, and is defined without any specific disease in mind. Individuals who become infected are initially mild cases but they may subsequently become severe cases if they are exposed further to the disease. (The extension to more than two different severities is discussed briefly at the conclusion of the paper.) The large population behaviour of the EDS model is analysed. The basic reproduction number R_0 , which determines whether a major outbreak is possible given few initial infectives, depends only on parameters governing the mild infectious state. However, the final size in the event of a major outbreak depends also on parameters governing the severe state. Further, the limiting final size as the population becomes infinite depends on the parameters in a surprising way, it being discontinuous in the initial proportion infected as well as in other model parameters.

The paper is organised as follows. The EDS model is defined in Section 2, where both a real-time description and a so-called Sellke construction of the epidemic (Sellke (1983)), that is used to analyse the model, are presented. A heuristic explanation of the final size equations in a large community is given in Section 3, which also contains a numerical example illustrating that the asymptotic final proportion infected of different types may be discontinuous in model parameters. The asymptotic behaviour of the model as the population

size tends to infinity is analysed rigorously in Section 4. In Section 4.1, an embedding argument is used to prove a strong law and associated central limit theorem for the final outcome of an epidemic initiated by a strictly positive asymptotic proportion of infectives. Section 4.2 treats the case when the number of initial infectives is held fixed as the population size tends to infinity. A coupling argument is used to justify a branching process approximation for the early stages of an epidemic and the final outcome of an epidemic that takes off is studied using embedding techniques. The paper concludes with a brief discussion in Section 5.

2 The exposure-dependent severity epidemic model

2.1 Definition of the model

Consider a closed homogeneously mixing population consisting initially (i.e. at time $t = 0$) of n susceptible individuals, m mildly infectious individuals, 0 severely infectious individuals and 0 removed individuals. A given mildly (severely) infectious individual makes contacts with any given individual at the points of a homogeneous Poisson process having rate λ_M/n (λ_S/n) throughout an infectious period $I^{(M)}$ ($I^{(S)}$) having distribution F_M (F_S). Each time a susceptible (mildly infected) individual is contacted it becomes mildly infectious (severely infectious) with probability p_M (p_S). Contacts with severely infectious individuals have no effect. All infectious periods, contact processes and outcomes of contacts are mutually independent. The epidemic continues until there is no mildly or severely infectious individual in the population.

It is worth emphasising that the mild and severe infectious periods of an individual may or may not overlap, and if they overlap the mild infectious period may end before or after the severe infectious period has ended. To be more specific, if a mildly infected individual is contacted successfully during its mild infectious period then it contacts any given individual at rate $(\lambda_M + \lambda_S)/n$ until one of its infectious periods ends. If the mild (severe) infectious period ends first then it contacts any given individual at rate λ_S/n (λ_M/n) until its severe (mild) infectious period ends, after which it is removed and plays no further role in the epidemic. Alternatively, if the first successful contact with a mildly infected individual occurs after its mild infectious period has ended, then it still becomes severely infected and again is removed at the end of its severe infectious period.

2.2 Sellke construction of the model

We now construct the model above using methods from Sellke (1983). Label the m initial mildly infectious individuals $-(m-1), -(m-2), \dots, 0$ and the initial susceptible individuals $1, 2, \dots, n$. For $i = -(m-1), -(m-2), \dots, n$, let $Q_i^{(M)}$ and $Q_i^{(S)}$ be exponentially distributed with intensities p_M and p_S , respectively, and let $I_i^{(M)}$ and $I_i^{(S)}$ be distributed according to F_M and F_S , respectively. All of these $4(m+n)$ random variables are mutually independent. The variable $Q_i^{(M)}$ is the *resistance* for individual i to become mildly infected and $Q_i^{(S)}$ the *extra* resistance for individual i to become severely infected once mildly infected. (Note that $Q_i^{(M)}$, $i = -(m-1), -(m-2), \dots, 0$, need not be defined as these individuals are mildly infected at the start of the epidemic.) The variables $I_i^{(M)}$ and $I_i^{(S)}$ denote individual i 's mild and severe infectious periods if they become so infected.

The epidemic is constructed as follows. At any time $t \geq 0$, individual i accumulates exposure to infection at rate $(\lambda_M Y_M(t) + \lambda_S Y_S(t)) / n$, where $Y_M(t)$ and $Y_S(t)$ denote respectively the number of mild and infectives at time t . An initial mild infective, j say, remains so for a period of length $I_j^{(M)}$. It becomes severely infectious if and when its accumulated exposure to infection reaches $Q_j^{(S)}$ and remains so for a period of length $I_j^{(M)}$. Similarly, an initial susceptible, i say, becomes mildly infectious if and when its accumulated exposure to infection reaches $Q_i^{(M)}$ and remains so for a period $I_i^{(M)}$. If and when individual i 's accumulated exposure to infection reaches $Q_i^{(M)} + Q_i^{(S)}$ it becomes severely infected and remains so for a period of length $I_i^{(S)}$. (As previously, note that a person can be both mildly and severely infectious at the same time, so ‘‘additionally infectious’’ might be a better name for the second state.) The epidemic ceases as soon as there is no mild or severe infective present in the population, i.e. when $Y_M(t) = Y_S(t) = 0$.

Observe that under the above construction it is possible for a mild infective to infect itself and hence become severely infected. If this possibility is also allowed in the model described in Section 2.1, then it is easily verified that the the Sellke construction yields a process that is probabilistically equivalent to that model. The infectious periods clearly follow the correct distributions and the lack-of-memory property of the exponential distribution ensures that the infection processes are equivalent.

The assumption that an individual can infect itself may appear unrealistic, although it can be thought of as permitting the possibility of a mild infective spontaneously becoming severely infectious. The constructions can be modified to exclude this possibility. However, the assumption does not affect the asymptotic behaviour of the model as $n \rightarrow \infty$, and the analysis is simplified

by retaining it.

The model can in fact be generalised to allow the contact rate to vary over time without affecting the final size distribution. The random quantities $\lambda_M I^{(M)}/n$ and $\lambda_S I^{(S)}/n$ are the accumulated infection forces exerted on a given individual by a given mild and severe infective, respectively. If the contact rate t time units after infection is $\lambda_M(t)$ and $\lambda_S(t)$ for mild and severe infectives, respectively, where the functions $\lambda_M(t)$ and $\lambda_S(t)$ may be deterministic or random, then it is clear from the Sellke construction that the final outcome of this extended epidemic model coincides with that of the previous model, provided $\lambda_M I^{(M)}$ and $\lambda_S I^{(S)}$ are chosen such that they have the same distribution as $\int_0^\infty \lambda_M(t) dt$ and $\int_0^\infty \lambda_S(t) dt$, respectively. Thus, for example, a latency period [$\lambda_i(t) = 0$ for $t \leq L_i$ ($i = M, S$)] does not affect the distribution of the final outcome.

3 Heuristics and a numerical example

3.1 Final size equations

Assume that the initial number of susceptibles n is large and let $\mu = m/n$. Let r_M denote the proportion of initial susceptibles who ultimately become mildly infected and note that some of these may also become severely infected. Let r_S denote the proportion of initial susceptibles who ultimately become severely infected, and let r_0 denote the proportion of initial mild infectives who become severely infected during the course of the epidemic. The total force of infection exerted on an individual during the entire epidemic is then given by

$$\tau = \lambda_M \iota_M (\mu + r_M) + \lambda_S \iota_S (\mu r_0 + r_S), \quad (3.1)$$

where ι_M and ι_S are the mean infectious periods for mild and severe infectives, respectively. This follows because $n(\mu + r_M)$ is the total number of mild infectives and each of them has contact with a given other individual with the average accumulated rate $\lambda_M \iota_M/n$, and similarly $n(\mu r_0 + r_S)$ is the total number of severe infectives, each of whom has accumulated contact rate $\lambda_S \iota_S/n$ with a specific other individual. A susceptible individual is infected by a single contact with probability p_M , so the probability that a given susceptible ultimately becomes mildly infected is $\exp[-p_M (\lambda_M \iota_M (\mu + r_M) + \lambda_S \iota_S (\mu r_0 + r_S))]$, which also gives the proportion of susceptibles who are eventually mildly infected, i.e. r_M . To derive an expression for r_S , the probability that a susceptible ultimately becomes severely infected, requires a bit more thought. For this event to happen, a susceptible individual must first become mildly infected and later severely infected. From the Sellke construction this happens if the sum of the mild and

severe resistances do not exceed the total force of infection. The mild resistance (denoted $Q_i^{(M)}$ in Section 2.2), i.e. the accumulated force necessary for a susceptible to become mildly infected, is exponentially distributed with parameter p_M . Similarly, the additional severe resistance $Q_i^{(S)}$ is exponentially distributed with parameter p_S . Thus $r_S = P[Q_i^{(M)} + Q_i^{(S)} < \lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S)]$, conditioning on $Q_i^{(M)}$ yields $r_S = 1 - (p_M e^{-p_S \tau} - p_S e^{-p_M \tau}) / (p_M - p_S)$ if $p_M \neq p_S$ and $r_S = 1 - (1 + p\tau)e^{-p\tau}$ if $p_M = p_S (= p \text{ say})$; recall that τ is defined in (3.1). Finally, an initial mild infective becomes severely infected if its severe resistance is less than the total force of infection, which happens with probability $r_0 = 1 - e^{-p_S \tau}$.

To summarise, we have the following set of balancing equations for the final proportions infected of different types r_M, r_S, r_0 . If $p_M \neq p_S$ then

$$r_M = 1 - e^{-p_M(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S))}, \quad (3.2)$$

$$r_S = 1 - \frac{p_M e^{-p_S(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S))} - p_S e^{-p_M(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S))}}{p_M - p_S}, \quad (3.3)$$

$$r_0 = 1 - e^{-p_S(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S))}. \quad (3.4)$$

If $p_M = p_S = p$ then (3.3) is replaced by

$$r_S = 1 - (1 + p(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S)))e^{-p(\lambda_M \iota_M(\mu + r_M) + \lambda_S \iota_S(\mu r_0 + r_S))}. \quad (3.3')$$

Suppose that $\mu > 0$. In Section 4.1 (Corollary 1), we prove that under mild regularity conditions, as $n \rightarrow \infty$, the proportions ultimately infected of the different types converge almost surely to the *smallest* positive solution of the balancing equations (3.2) to (3.4). We also prove an associated central limit theorem (Theorem 1).

Suppose that $\mu = 0$, which for example would be the case if m is held fixed as $n \rightarrow \infty$. Note that r_0 no longer enters the right hand sides of the equations (3.2) to (3.4), although (3.4) still gives the probability that a given initial mild case ultimately becomes severely infected. Equations (3.2) and (3.3) now admit the solution $r_M = r_S = 0$, corresponding to the case when the epidemic fails to take off. If the epidemic does take off then the asymptotic proportions ultimately infected of different types are given by the smallest strictly positive solution of (3.2) and (3.3) and satisfy a central limit theorem; see Theorem 3 in Section 4.2.

Suppose that m is held fixed as $n \rightarrow \infty$. Then during the early stages of the epidemic, the probability that a contact is with a previously infected individual is very small and the process of mild infectives can be approximated by a branching process. Moreover, as $n \rightarrow \infty$, severe cases arise only if the branching process does not go extinct. This approximation is made fully rigorous

in Theorem 2 of Section 4.2. The offspring distribution of the approximating branching process has mean $R_0 = \lambda_M \iota_M p_M$, since, as $n \rightarrow \infty$, a given mild infective makes contacts at rate λ_M throughout an infectious period having mean ι_M and each contact is successful with probability p_M . The quantity R_0 is known as the reproduction number (see, for example, Heesterbeek and Dietz (1996)) and is a threshold parameter for the epidemic, in that if m is small and n large, the epidemic can take off only if $R_0 > 1$. Note that R_0 depends only on parameters relating to *mild* infectives but if the epidemic takes off then its size depends also on parameters relating to severe infectives.

3.2 A numerical example

We now illustrate with an example what the deterministic solutions to the balancing equations (3.2) to (3.4) may look like. We let $\mu = 0.01$, $\lambda_M \iota_M = 1$, $\lambda_S \iota_S = 2$ and $p_M = p_S = p$, and examine the solutions (r_M, r_S, r_0) as a function of p . In Figure 1 the equations have been solved numerically and all solutions for r_M are plotted in the range $0.8 \leq p \leq 0.9$ (as might be guessed from the figure, there is only one solution for $p < 0.8$ and $p > 0.9$). As noted above, the

Figure 1: All solutions for r_M in the range $0.8 \leq p \leq 0.9$ for the equations (3.2) to (3.4) with $\mu = 0.01$, $\lambda_M \iota_M = 1$, $\lambda_S \iota_S = 2$ and $p_M = p_S = p$.

final proportion infected converges almost surely to the *smallest* positive solution to the set of equations (3.2) to (3.4). As a consequence, we see that the

example exhibits a discontinuity at around $p^* \approx 0.8705$. For p slightly smaller than this critical value the limiting proportion mildly infected is around 0.158 whereas just above p^* the limiting proportion mildly infected suddenly jumps up to around 0.735! For the same p^* there are similar discontinuities for r_S and r_0 (not shown in the figure). There are only discontinuities if the reproduction number R_0 is smaller than its threshold of 1 and, at the same time, the epidemic is started with a positive fraction initially infectious (i.e. $\mu > 0$). The heuristic argument for the discontinuity is that, even though the model is below threshold, the initial proportion mildly infected makes the epidemic “less subcritical” and as the parameters move towards criticality (p grows in our example) there is a critical point (below threshold!) at which the initial proportion mildly infectious suddenly makes the epidemic supercritical. A similar behaviour was observed by Scalia-Tomba (1985) for an SIR model in which a susceptible has to be successfully contacted twice before it becomes infected, i.e. for an extreme case of the EDS model, in which mild infectives make no infectious contacts ($\lambda_M \iota_M = 0$). Also, in several deterministic models for infectious diseases allowing endemic situations, as opposed to the transient nature of the present SIR model in a closed community, similar discontinuities have been observed, for example, by van den Driessche and Watmough (2002). This is known as (backward) bifurcation referring to the situation where equations have one solution up until a point where there suddenly appears more than one solution (cf. Figure 1).

We now compare the deterministic solutions discussed above with some simulations from the stochastic model in a finite population. Figure 2 shows the empirical distribution of the final outcome of the stochastic EDS epidemic, when $n = 1000, m = 10, \lambda_M = 1, \lambda_S = 2, p_M = p_S = p$, and $I^{(M)}$ and $I^{(S)}$ are each exponentially distributed with unit mean (i.e. same parameter values as in the deterministic example). The results are based on 10000 simulations. The top two figures, showing the number of mildly infected, can be compared with the deterministic solutions of Figure 1, evaluated at $p = 0.8$ and $p = 0.9$ respectively. Since the epidemic is started with 10 initially mild infectives it might very well die out quickly for both choices of p . This is confirmed by the simulations. In the deterministic solutions (assuming an outbreak occurs) of Figure 1 it is seen that the ultimate fraction mildly infected is about 0.045 when $p = 0.8$ and 0.795 when $p = 0.9$. In the simulations for $p = 0.8$ nearly all simulations result in small outbreaks, i.e. not too far from 0.045, but a fraction of about one percent, hardly visible in the figure, have outbreaks of more than 50% mildly infected. On the other hand, in the simulations for $p = 0.9$ a larger fraction of the simulations have large outbreaks close to 0.795, but there are still quite a few small outbreaks. Loosely speaking, the general conclusion in the stochastic setting seems to be that the outbreak size for a given p will

be close to the final size solution of the deterministic equation for a possibly slightly different \tilde{p} . As a consequence, for p close to the bifurcation point the outbreak may agree with the final size equation for \tilde{p} on the other side of the bifurcation point.

Figure 2: Final outcome of 10000 simulations of the stochastic EDS model when $p = 0.8$ and when $p = 0.9$. See text for further details.

4 Rigorous asymptotic analysis of the model

In this section we examine asymptotic properties of the model when n tends to infinity, so we equip all parameters depending on n with an n -index. We treat two different initial configurations. Let $\mu_n = n^{-1}m_n$. Then, either there is a positive proportion of initial mild infectives ($\mu_n \rightarrow \mu$ as $n \rightarrow \infty$, where $\mu > 0$) or else there is a fixed number of initial mild infectives, i.e. $m_n = m$. In the first case we prove a strong law and a central limit theorem for the ultimate proportion infected of the different types. In the second case we prove a threshold limit theorem indicating that the epidemic either never takes off and thus infects just a few individuals, or else it takes off and a more or less deterministic fraction of the population will ultimately become infected. A central limit theorem for the final outcome of an epidemic that takes off is also shown. Approximations of the initial stages of the epidemic are obtained by coupling the epidemic process with a suitable branching process and in case of

a major outbreak properties of its final outcome are obtained using embedding techniques.

4.1 The case $\mu_n \rightarrow \mu > 0$ as $n \rightarrow \infty$

We study the asymptotic final outcome of the model as $n \rightarrow \infty$ by adapting the embedding argument of Scalia-Tomba (1985), (1990). We use weak convergence in the space of bounded functions on $[0, \infty]$ equipped with the supremum metric. For any real-valued function $f = f(t)$ with domain $[0, \infty]$, let $\|f\| = \sup_{t \in [0, \infty]} |f(t)|$.

Consider the Sellke construction of Section 2.2. For $i = 1, 2, \dots, n$ and $t \in [0, \infty]$, let $R_i^{(M)}(t) = 1_{\{Q_i^{(M)} \leq t\}}$, $R_i^{(S)}(t) = 1_{\{Q_i^{(M)} + Q_i^{(S)} \leq t\}}$ and $A_i(t) = \lambda_M I_i^{(M)} R_i^{(M)}(t) + \lambda_S I_i^{(S)} R_i^{(S)}(t)$. Thus, $A_i(t)$ is the total force of infection exerted on the population by individual i if he or she is exposed to t units of infectious pressure. For $i = -(m_n - 1), -(m_n - 2), \dots, 0$ and $t \in [0, \infty]$, let $R_i^{(0)}(t) = 1_{\{Q_i^{(S)} \leq t\}}$ and $A_i(t) = \lambda_S I_i^{(S)} R_i^{(0)}(t)$, so $A_i(t)$ is the ‘severe’ infectious pressure exerted by initial mild infective i on the population if he or she is exposed to t units of infectious pressure. For $t \in [0, \infty]$, let $\tilde{A}_{\bullet n}(t) = \sum_{i=-(m_n-1)}^n A_i(t)$, $R_{\bullet n}^{(M)}(t) = \sum_{i=1}^n R_i^{(M)}(t)$, $R_{\bullet n}^{(S)}(t) = \sum_{i=1}^n R_i^{(S)}(t)$ and $R_{\bullet n}^{(0)}(t) = \sum_{i=-(m_n-1)}^0 R_i^{(0)}(t)$.

The final outcome of the epidemic can be obtained as follows. The m_n initial mildly infectious individuals exert in total $T_0^{(n)} = \sum_{i=-(m_n-1)}^0 \lambda_M I_i^{(M)}$ units of infectious pressure on the population from their mild infectious state. These $T_0^{(n)}$ units of infectious pressure will create $\tilde{A}_{\bullet n}(n^{-1}T_0^{(n)})$ further units of infectious pressure, which may in turn create further infectious pressure. For $k = 0, 1, \dots$, let

$$T_{k+1}^{(n)} = T_0^{(n)} + \tilde{A}_{\bullet n}(n^{-1}T_k^{(n)}). \quad (4.1)$$

Then $k_n^* = \min\{k: T_{k+1}^{(n)} = T_k^{(n)}\}$ is well defined since the population is finite. Let $T_\infty^{(n)} = T_{k_n^*}^{(n)}$ and $Z_n^{(k)} = R_{\bullet n}^{(k)}(n^{-1}T_\infty^{(n)})$ ($k = 0, M, S$). It is easily verified that $T_\infty^{(n)}$ is the total force of infection exerted during the course of the epidemic, also known as the total cost of the epidemic, $Z_n^{(0)}$ is the number of initial mild cases that eventually become severe cases, and $Z_n^{(M)}$ and $Z_n^{(S)}$ are the numbers of initially susceptible individuals that at some time become mild and severe cases, respectively. Note that $Z_n^{(M)}$ includes those individuals who eventually go on to become severe cases.

To study the asymptotic behaviour of $(Z_n^{(0)}, Z_n^{(M)}, Z_n^{(S)})$, it is convenient to assume that epidemics for different n are constructed using a common set of random variables $\{(Q_i^{(M)}, Q_i^{(S)}, I_i^{(M)}, I_i^{(S)}): i \in \mathbb{Z}\}$, defined on a probability

space (Ω, \mathcal{F}, P) and distributed as described in Section 2.2. For $t \in [0, \infty]$, let $r_j(t) = E[R_1^{(j)}(t)]$ ($j = M, S$), $a(t) = E[A_1(t)] = \lambda_M \iota_M r_M(t) + \lambda_S \iota_S r_S(t)$, $r_0(t) = \mu E[R_0^{(0)}(t)]$ and $a_0(t) = \mu E[A_0(t)] = \lambda_S \iota_S r_0(t)$, where $\iota_j = E[I_1^{(j)}]$ ($j = M, S$). It is easily verified that

$$\begin{aligned} r_0(t) &= \mu(1 - e^{-ps t}), \\ r_M(t) &= 1 - e^{-p_M t} \\ \text{and } r_S(t) &= \begin{cases} 1 - (p_M e^{-p_S t} - p_S e^{-p_M t}) / (p_M - p_S) & \text{if } p_M \neq p_S, \\ 1 - e^{-p t} (1 + p t) & \text{if } p_M = p_S = p. \end{cases} \end{aligned}$$

Lemma 1 *Suppose that $n^{-1} \mu_n \rightarrow \mu$ as $n \rightarrow \infty$, where $\mu > 0$, and that $\iota_j < \infty$ ($j = M, S$). Let $\tilde{a}(t) = a_0(t) + a(t)$. Then*

$$\|n^{-1} \tilde{A}_{\bullet n} - \tilde{a}\| \xrightarrow{a.s.} 0 \quad \text{as } n \rightarrow \infty. \quad (4.2)$$

Proof For $t \in [0, \infty]$, $n^{-1} \tilde{A}_{\bullet n}(t) = \frac{m_n}{n} \frac{1}{m_n} \sum_{i=(m_n-1)}^0 A_i(t) + \frac{1}{n} \sum_{i=1}^n A_i(t) \xrightarrow{a.s.} \tilde{a}(t)$ as $n \rightarrow \infty$, by the strong law of large numbers. Thus there exists $E \in \mathcal{F}$ such that $P(E) = 1$ and for all $\omega \in E$,

$$\lim_{n \rightarrow \infty} n^{-1} \tilde{A}_{\bullet n}(t, \omega) = \tilde{a}(t) \quad (t \in (\mathbb{Q} \cap [0, \infty]) \cup \{\infty\}). \quad (4.3)$$

Fix $\omega \in E$ and $\varepsilon > 0$. Now, $\tilde{a}(0) = 0$ and $\tilde{a}(\infty) < \infty$, so, since $\tilde{a}(t)$ is nondecreasing with t , there exists $q \in \mathbb{N}$ and $t_1, t_2, \dots, t_q \in \mathbb{Q}$ such that $0 = t_0 < t_1 < t_2 < \dots < t_q < t_{q+1} = \infty$ and

$$0 < \tilde{a}(t_i) - \tilde{a}(t_{i-1}) < \varepsilon/2 \quad (i = 1, 2, \dots, q+1). \quad (4.4)$$

From (4.3), there exists $N \in \mathbb{N}$ such that

$$|n^{-1} \tilde{A}_{\bullet n}(t_i, \omega) - \tilde{a}(t_i)| < \varepsilon/2 \quad (i = 0, 1, \dots, q+1; n \geq N). \quad (4.5)$$

Since $\tilde{A}_{\bullet n}(\cdot, \omega) = \tilde{A}_{\bullet n}(t, \omega)$ is also nondecreasing with t , it follows using (4.4) and (4.5) that $\|n^{-1} \tilde{A}_{\bullet n}(\cdot, \omega) - \tilde{a}(\cdot)\| < \varepsilon$ ($n \geq N$). Thus $\|n^{-1} \tilde{A}_{\bullet n}(\cdot, \omega) - \tilde{a}(\cdot)\| \rightarrow 0$ as $n \rightarrow \infty$, as $\varepsilon > 0$ is arbitrary, and the lemma follows, since $P(E) = 1$.

Corollary 1 *Suppose that the conditions of Lemma 1 are satisfied. Let $\tau = \tau(\mu) = \min\{t > 0: t = \lambda_M \mu_M + \tilde{a}(t)\}$ and suppose that $\tilde{a}'(\tau) < 1$, where ' denotes first derivative. Then $n^{-1} T_\infty^{(n)} \xrightarrow{a.s.} \tau$ and $n^{-1} Z_n^{(i)} \xrightarrow{a.s.} r_i(\tau)$ ($i = 0, M, S$) as $n \rightarrow \infty$.*

Proof Let $\bar{T}_k^{(n)} = n^{-1}T_k^{(n)}$ ($k = 0, 1, \dots, \infty$). Then it follows from (4.1) and the definition of $T_\infty^{(n)}$ that $\bar{T}_\infty^{(n)} = \min\{t \geq 0: t = \bar{T}_0^{(n)} + n^{-1}\tilde{A}_{\bullet n}(t)\}$. Since \tilde{a}' is continuous on $[0, \infty)$ and $\tilde{a}'(\tau) < 1$, there exists $\Delta > 0$ such that $t > \lambda_M\mu_M + \tilde{a}(t)$ for all $t \in (\tau, \tau + \Delta)$. Now $\bar{T}_0^{(n)} \xrightarrow{a.s.} \lambda_M\mu_M$ as $n \rightarrow \infty$, by the strong law of large numbers, which in conjunction with Lemma 1 implies that there exists $F \in \mathcal{F}$ such that $P(F) = 1$ and, for all $\omega \in F$,

$$\lim_{n \rightarrow \infty} \|n^{-1}\tilde{A}_{\bullet n}(\cdot, \omega) - \tilde{a}(\cdot)\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \bar{T}_0^{(n)} = \lambda_M\mu_M. \quad (4.6)$$

Fix $\omega \in F$ and $\epsilon \in (0, \Delta)$. Let $C_1 = \min_{t \in [0, \tau - \epsilon]} \{\lambda_M\mu_M + \tilde{a}(t) - t\}$ and $C_2 = \tau + \epsilon - \lambda_M\mu_M - \tilde{a}(\tau + \epsilon)$. Note that the definitions of τ and Δ imply respectively that $C_1 > 0$ and $C_2 > 0$. Further, by (4.6), there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $\inf_{t \in [0, \tau - \epsilon]} \{\bar{T}_0^{(n)}(\omega) + n^{-1}\tilde{A}_{\bullet n}(t, \omega) - t\} > \frac{1}{2}C_1$ and $\tau + \epsilon - \bar{T}_0^{(n)}(\omega) - n^{-1}\tilde{A}_{\bullet n}(\tau + \epsilon, \omega) > \frac{1}{2}C_2$, whence $\bar{T}_\infty^{(n)}(\omega) \in (\tau - \epsilon, \tau + \epsilon)$. Thus $\bar{T}_\infty^{(n)}(\omega) \rightarrow \tau$ as $n \rightarrow \infty$, as $\epsilon \in (0, \Delta)$ is arbitrary, so $\bar{T}_\infty^{(n)} \xrightarrow{a.s.} \tau$ as $n \rightarrow \infty$, since $P(F) = 1$. It is easily seen that equivalent results to (4.2) hold for $R_{\bullet n}^{(k)}$ ($k = 0, M, S$), from which it is immediate that $n^{-1}Z_n^{(k)} \xrightarrow{a.s.} r_k(\tau)$ ($k = 0, M, S$) as $n \rightarrow \infty$.

Remark 1 *It is easily verified that τ and $r_i(\tau)$ ($i = 0, M, S$) correspond to the smallest positive solution of the heuristic equations (3.1) to (3.4).*

We now derive a central limit theorem for the final outcome of the epidemic. Before stating the main result some more notation is required. For $t \in [0, \infty]$, let $\mathbf{X}_n(t) = (\tilde{A}_{\bullet n}(t), R_{\bullet n}^{(0)}(t), R_{\bullet n}^{(M)}(t), R_{\bullet n}^{(S)}(t))^\top$ and $\Sigma(t) = \lim_{n \rightarrow \infty} n^{-1}\text{var}(\mathbf{X}_n(t))$. Write $\Sigma(t)$ as

$$\Sigma(t) = \begin{bmatrix} \tilde{\sigma}_A^2(t) & \sigma_{A0}(t) & \sigma_{AM}(t) & \sigma_{AS}(t) \\ \sigma_{A0}(t) & \sigma_0^2(t) & \sigma_{0M}(t) & \sigma_{0S}(t) \\ \sigma_{AM}(t) & \sigma_{0M}(t) & \sigma_M^2(t) & \sigma_{MS}(t) \\ \sigma_{AS}(t) & \sigma_{0S}(t) & \sigma_{MS}(t) & \sigma_S^2(t) \end{bmatrix}. \quad (4.7)$$

Let $\sigma_i^2 = \text{var}(I_1^{(i)})$ ($i = M, S$). Elementary calculation yields that $\sigma_0^2(t) = r_0(t)(1 - \mu^{-1}r_0(t))$, $\sigma_i^2(t) = r_i(t)(1 - r_i(t))$ ($i = M, S$), $\sigma_{0M}(t) = \sigma_{0S}(t) = 0$, $\sigma_{MS}(t) = r_S(t)(1 - r_M(t))$, $\sigma_{A0}(t) = \lambda_S\iota_S\sigma_0^2(t)$, $\sigma_{AM}(t) = \lambda_M\iota_M\sigma_M^2(t) + \lambda_S\iota_S\sigma_{MS}(t)$, $\sigma_{AS}(t) = \lambda_M\iota_M\sigma_{MS}(t) + \lambda_S\iota_S\sigma_S^2(t)$ and $\tilde{\sigma}_A^2(t) = \lambda_S^2(\sigma_S^2r_0(t) + \iota_S^2\sigma_0^2(t)) + \lambda_M^2(r_M(t)\sigma_M^2 + \iota_M^2\sigma_M^2(t)) + \lambda_S^2(r_S(t)\sigma_S^2 + \iota_S^2\sigma_S^2(t)) + 2\lambda_M\lambda_S\iota_M\iota_S\sigma_{MS}(t)$. For $t \in [0, \infty]$, let $B(t) = [b_{ij}(t)]$ be the 4×4 matrix with elements given by $b_{11}(t) = (1 - \tilde{a}'(t))^{-1}\tilde{a}'(t)$, $b_{21}(t) = (1 - \tilde{a}'(t))^{-1}r_0'(t)$, $b_{31}(t) = (1 - \tilde{a}'(t))^{-1}r_M'(t)$,

$b_{41}(t) = (1 - \tilde{a}'(t))^{-1} r'_S(t)$ and $b_{ij}(t) = 0$ if $j > 1$. Let J be the 4×4 matrix with elements all equal to one. For $n = 1, 2, \dots$, let $\mathbf{Z}_n = (\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}), Z_n^{(0)}, Z_n^{(M)}, Z_n^{(S)})^\top$ and for $t \in [0, \infty]$, let $\boldsymbol{\mu}_Z(t) = (\tilde{a}(t), r_0(t), r_1(t), r_2(t))^\top$.

Theorem 1 *Suppose that the conditions of Lemma 1 and Corollary 1 are satisfied, that $\sigma_i^2 < \infty$ ($i = M, S$) and that $n^{1/2}(\mu_n - \mu) \rightarrow 0$ as $n \rightarrow \infty$. Then $n^{-1/2}(\mathbf{Z}_n - n\boldsymbol{\mu}_Z(\tau)) \xrightarrow{D} \mathbf{Z}$ as $n \rightarrow \infty$, where \mathbf{Z} is a 4-dimensional zero-mean normal random vector with variance matrix, Σ_Z say, given by*

$$\Sigma_Z = (I + B(\tau))\Sigma(\tau)(I + B(\tau)^\top) + \mu\lambda_M^2\sigma_M^2B(\tau)JB(\tau)^\top.$$

Proof For $t \in [0, \infty]$, let $A_{\bullet n}(t) = \sum_{i=1}^n A_i(t)$ and $A_{\bullet n}^{(0)}(t) = \sum_{i=-(m_n-1)}^0 A_i(t)$. The sample paths $t \mapsto A_1(t)$ are cadlag and nondecreasing, so by van der Vaart and Wellner (1996), Example 2.11.16,

$$n^{-1/2}(A_{\bullet n} - E[A_{\bullet n}]) \xrightarrow{w} X \quad \text{as } n \rightarrow \infty, \quad (4.8)$$

where X is a zero-mean Gaussian process and \xrightarrow{w} denotes weak convergence. An equivalent result to (4.8) holds for each of $A_{\bullet n}^{(0)}$ and $R_{\bullet n}^{(k)}$ ($k = 0, M, S$). For $t \in [0, \infty]$, let $\mathbf{X}_{\bullet n}(t) = (\tilde{A}_{\bullet n}(t), R_{\bullet n}^{(0)}(t), R_{\bullet n}^{(M)}(t), R_{\bullet n}^{(S)}(t))^\top$. Then

$$n^{-1/2}(\mathbf{X}_{\bullet n} - E[\mathbf{X}_{\bullet n}]) \xrightarrow{w} \mathbf{X} \quad \text{as } n \rightarrow \infty, \quad (4.9)$$

where \mathbf{X} is a 4-dimensional zero-mean Gaussian process whose covariance function satisfies $\text{var}(\mathbf{X}(t)) = \Sigma(t)$ ($t \in [0, \infty]$). (It is easily seen that the finite-dimensional distributions of $n^{-1/2}(\mathbf{X}_{\bullet n} - E[\mathbf{X}_{\bullet n}])$ converge to those of \mathbf{X} by using the Cramér–Wold device, and asymptotic tightness of $n^{-1/2}(\mathbf{X}_{\bullet n} - E[\mathbf{X}_{\bullet n}])$ follows using Lemma 1.4.3 and Theorem 1.5.4 of van der Vaart and Wellner (1996).) Corollary 1 implies that $\bar{T}_\infty^{(n)} \xrightarrow{P} \tau$ as $n \rightarrow \infty$, so by Slutsky’s lemma and the continuous mapping theorem (van der Vaart and Wellner (1996), Example 1.4.7 and Theorem 1.3.6, respectively) it follows from (4.9) that

$$n^{-1/2}(\mathbf{X}_{\bullet n}(\bar{T}_\infty^{(n)}) - E[\mathbf{X}_{\bullet n}](\bar{T}_\infty^{(n)})) \xrightarrow{D} \mathbf{X}(\tau) \quad \text{as } n \rightarrow \infty.$$

Now,

$$n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\tau)) = n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\bar{T}_\infty^{(n)})) + n^{1/2}(\tilde{a}(\bar{T}_\infty^{(n)}) - \tilde{a}(\tau)).$$

By the mean value theorem, $n^{1/2}(\tilde{a}(\bar{T}_\infty^{(n)}) - \tilde{a}(\tau)) = n^{1/2}(\bar{T}_\infty^{(n)} - \tau)\tilde{a}'(\eta_n)$ for some η_n lying between $\bar{T}_\infty^{(n)}$ and τ . Recall that $\bar{T}_\infty^{(n)} = \bar{T}_0^{(n)} + n^{-1}\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)})$ and $\tau = \lambda_M\mu_M + \tilde{a}(\tau)$. Thus

$$n^{1/2}(\bar{T}_\infty^{(n)} - \tau) = n^{1/2}(\bar{T}_0^{(n)} - \lambda_M\mu_M) + n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\tau)), \quad (4.10)$$

so

$$\begin{aligned} (1 - \tilde{a}'(\eta_n))n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\tau)) &= \tilde{a}'(\eta_n)n^{1/2}(\bar{T}_0^{(n)} - \lambda_M\mu\iota_M) \\ &\quad + n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\bar{T}_\infty^{(n)})). \end{aligned} \quad (4.11)$$

Now, $n^{-1/2}(E[\tilde{A}_{\bullet n}](\bar{T}_\infty^{(n)}) - n\tilde{a}(\bar{T}_\infty^{(n)})) = n^{1/2}(n^{-1}m_n - \mu)\mu^{-1}a_0(\bar{T}_\infty^{(n)}) \xrightarrow{P} 0$ as $n \rightarrow \infty$, as $n^{1/2}(n^{-1}m_n - \mu) \rightarrow 0$ and $\bar{T}_\infty^{(n)} \xrightarrow{P} \tau$ as $n \rightarrow \infty$. Further, by the central limit theorem, $n^{1/2}(\bar{T}_0^{(n)} - \lambda_M\mu\iota_M) \xrightarrow{D} W_0$ as $n \rightarrow \infty$, where W_0 is independent of $\mathbf{X}(\tau)$ and $W_0 \sim N(0, \mu\lambda_M^2\sigma_M^2)$. For $t \in [0, \infty]$, let $\mathbf{X}(t) = (X_{\tilde{A}}(t), X_R^{(0)}(t), X_R^{(M)}(t), X_R^{(S)}(t))^\top$. Now, $\eta_n \xrightarrow{P} \tau$ as $n \rightarrow \infty$ so, since $\tilde{a}(t)$ is continuously differentiable and $\tilde{a}'(\tau) < 1$, it follows from (4.11) and Slutsky's theorem that

$$n^{-1/2}(\tilde{A}_{\bullet n}(\bar{T}_\infty^{(n)}) - n\tilde{a}(\tau)) \xrightarrow{D} (1 - \tilde{a}'(\tau))^{-1}(\tilde{a}'(\tau)W_0 + X_A(\tau)) \quad \text{as } n \rightarrow \infty. \quad (4.12)$$

It then follows using (4.10) that

$$n^{1/2}(\bar{T}_\infty^{(n)} - \tau) \xrightarrow{D} (1 - \tilde{a}'(\tau))^{-1}(W_0 + X_A(\tau)) \quad \text{as } n \rightarrow \infty. \quad (4.13)$$

For $k = 0, M, S$,

$$n^{-1/2}(R_{\bullet n}^{(k)}(\bar{T}_\infty^{(n)}) - nr_k(\tau)) = n^{-1/2}(R_{\bullet n}^{(k)}(\bar{T}_\infty^{(n)}) - nr_k(\bar{T}_\infty^{(n)})) + n^{1/2}(r_k(\bar{T}_\infty^{(n)}) - r_k(\tau)). \quad (4.14)$$

Applying the mean value theorem to the final term in (4.14), using (4.13) and arguing as above yields that, as $n \rightarrow \infty$,

$$n^{-1/2}(R_{\bullet n}^{(k)}(\bar{T}_\infty^{(n)}) - nr_k(\tau)) \xrightarrow{D} X_R^{(k)}(\tau) + (1 - a'(\tau))^{-1}r'_k(\tau)(W_0 + X_A(\tau)). \quad (4.15)$$

Recall that $Z_n^{(k)} = R_{\bullet n}^{(k)}(\bar{T}_\infty^{(n)})$ ($k = 0, M, S$). It follows from (4.12) and (4.15) that $n^{-1/2}(\mathbf{Z}_n - n\boldsymbol{\mu}_Z(\tau)) \xrightarrow{D} \mathbf{Z}$ as $n \rightarrow \infty$, where $\mathbf{Z} = (I + B(\tau))\mathbf{X}(\tau) + B(\tau)\mathbf{W}$ and $\mathbf{W} = (W_0, W_0, W_0, W_0)^\top$. Theorem 1 follows on noting that $\mathbf{X}(\tau)$ and \mathbf{W} are independent.

4.2 The case $m_n = m$

We study the initial behaviour of the epidemic by adapting the coupling argument of Ball and Donnelly (1995) to the present model. Let (Ω, \mathcal{F}, P) be a probability space on which are defined the following independent sets of random quantities:

- (i) $\mathcal{H}_i = (I_i^{(M)}, I_i^{(S)}, \eta_i^{(M)}, \eta_i^{(S)}, \xi_i^{(M)}, \xi_i^{(S)})$ ($i = -(m-1), -(m-2), \dots$), independent and identically distributed according to $\mathcal{H} = (I^{(M)}, I^{(S)}, \eta^{(M)}, \eta^{(S)}, \xi^{(M)}, \xi^{(S)})$, where the components of \mathcal{H} are independent, $I^{(k)}$ is distributed according to F_k ($k = M, S$) and $\eta^{(M)}, \eta^{(S)}, \xi^{(M)}, \xi^{(S)}$ are homogeneous Poisson processes on $[0, \infty)$ with respective rates $\lambda_M, \lambda_S, \lambda_M, \lambda_S$;
- (ii) $\chi_i^{(n)}$ ($n = 1, 2, \dots; i = 1, 2, \dots$), where for each $n = 1, 2, \dots, \chi_1^{(n)}, \chi_2^{(n)}, \dots$ are independent and uniformly distributed on $\{1, 2, \dots, n\}$;
- (iii) $\tilde{\chi}_i$ ($i = 1, 2, \dots$), independent and uniformly distributed on $\{-(m-1), -(m-2), \dots, 0\}$;
- (iv) U_i ($i = 1, 2, \dots$), independent and uniformly distributed on $(0, 1)$.

Denote the epidemic described in Section 2.1 by E_n . A realisation of E_n can be constructed as follows. For $i = -(m-1), -(m-2), \dots, 0$, the initial mild infective i makes contacts with the initial susceptibles (mild infectives) at the points of $\eta_i^{(M)}$ ($nm^{-1}\xi_i^{(M)}$) during his or her mild infectious period $I_i^{(M)}$. (If ξ is a simple point process on $[0, \infty)$ with points at $t_1 < t_2 < \dots$ and $\alpha > 0$ then the point process with points at $\alpha t_1 < \alpha t_2 < \dots$ is denoted by $\alpha\xi$, which hence has changed rate by a factor $1/\alpha$.) For $k = 1, 2, \dots$, the k th contact made with initial susceptibles (mild infectives) is with individual $\chi_k^{(n)}$ ($\tilde{\chi}_k$). For $k = 1, 2, \dots$, if the k th contact occurring in E_n is with a susceptible then that individual becomes mildly infective if $U_k < p_M$ (otherwise the contact is ignored), if it is with a mild infective then that individual becomes a severe case if $U_k < p_S$ (otherwise the contact is ignored), if it is with an immune individual then nothing happens. The k th initial susceptible infected by the epidemic adopts the random quantity \mathcal{H}_k ; if t denotes this individual's time of infection then he or she makes contacts during $[t, t + I_k^{(M)}]$ at times given by $\{t + \eta_k^{(M)}\} \cup \{t + nm^{-1}\xi_k^{(M)}\}$. Suppose that a mild case (governed by \mathcal{H}_k , say) becomes a severe case at time t , then in addition to any contacts remaining from his or her mild infectious period, he or she makes contacts during $[t, t + I_k^{(S)}]$ at times given by $\{t + \eta_k^{(S)}\} \cup \{t + nm^{-1}\xi_k^{(S)}\}$. The epidemic stops when there is no mild or severe case in the population.

The above random quantities can also be used to define a realisation of a Crump–Mode–Jagers branching process, with m initial ancestors, labelled $-(m-1), -(m-2), \dots, 0$, as follows. For $i = -(m-1), -(m-2), \dots, 0$, initial ancestor i lives for time $I_i^{(M)}$ and has potential births at times given by $\eta_i^{(M)}$. For $k = 1, 2, \dots$, the k th potential birth in the branching process gives rise to a new individual if $U_k \leq p_M$. Suppose that the k th actual birth in the branching process occurs at time $t > 0$. Then that individual lives until time

$t + I_k^{(M)}$ and has potential births at times given by $\{t + \eta_k^{(M)}\} \cap [t, t + I_k^{(M)}]$. For $t \geq 0$, let $Y_n(t)$ denote the number of mildly infectious individuals in the epidemic E_n at time t and let $Y(t)$ denote the number of individuals alive in the branching process at time t . Observe that $Y_n(t)$ and $Y(t)$ coincide at least up until the first time a contact is made with an individual who is not susceptible.

For $n = 1, 2, \dots$, let $M_n = \min \{k \geq 2: \chi_k^{(n)} \in \{\chi_1^{(n)}, \chi_2^{(n)}, \dots, \chi_{k-1}^{(n)}\}\}$. Note that this is the birthday problem and it is well known (e.g. Aldous (1985), page 96) that $n^{-1/2}M_n \xrightarrow{D} M$ as $n \rightarrow \infty$, where M is a random variable with probability density function $f(x) = x \exp(-\frac{1}{2}x^2)$ ($x > 0$). As in the proof of Theorem 2.1 of Ball and Donnelly (1995), the Skorohod representation theorem implies that we may assume that $\chi_i^{(n)}$ ($n = 1, 2, \dots; i = 1, 2, \dots$) are constructed so that $n^{-1/2}M_n \xrightarrow{\text{a.s.}} M$ as $n \rightarrow \infty$, where M is now also defined on (Ω, \mathcal{F}, P) . Let $A_{\text{EXT}} \in \mathcal{F}$ denote the set on which the branching process $\{Y(t): t \geq 0\}$ becomes extinct and let Z denote the total progeny of the branching process, excluding the m initial ancestors. (Thus $Z(\omega) < \infty$ if and only if $\omega \in A_{\text{EXT}}$.)

Theorem 2 (a) For P -almost all $\omega \in A_{\text{EXT}}$, as $n \rightarrow \infty$,

$$(i) \quad Z_n^{(M)} \rightarrow Z \quad \text{and}$$

$$(ii) \quad Z_n^{(k)} \rightarrow 0 \quad (k = 0, S).$$

(b) For P -almost all $\omega \in \Omega \setminus A_{\text{EXT}}$,

$$Z_n^{(M)} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Proof Suppose that $\omega \in A_{\text{EXT}}$ and let $Z_p(\omega)$ denote the total number of potential births made in the branching process. Then for almost all $\omega \in A_{\text{EXT}}$, $Z_p(\omega) < \infty$ and $M_n(\omega) > \frac{1}{2}n^{1/2}M(\omega)$ for all sufficiently large n . Thus, for such ω , for all sufficiently large n every birth in the branching process corresponds to a new mildly infectious case in the epidemic E_n . For $i = -(m-1), -(m-2), \dots$, let $W_{n,i}$ be the time that the individual governed by \mathcal{H}_i has to wait in E_n from its initial infection until he or she contacts an initial infective. Note that $W_{n,i} \xrightarrow{\text{a.s.}} \infty$ as $n \rightarrow \infty$ ($i = -(m-1), -(m-2), \dots$). Hence, for P -almost $\omega \in A_{\text{EXT}}$, no contacts are made with initial infectives in E_n for all sufficiently large n , and part (a) of the theorem follows. Part (b) is proved by noting that, for $\ell \in \mathbb{N}$, if $Z \geq \ell$ then the above argument shows that $Z_n^{(M)} \geq \ell$ for almost all $\omega \in \Omega$.

Theorem 2 enables the distribution of the size of epidemics that do not take off to be approximated for large n . We now determine the asymptotic

distribution of those epidemics that do take off. Let p_{EXT} denote the extinction probability of the branching process $\{Y(t): t \geq 0\}$.

Lemma 2 *Let (b_n) be any sequence of real numbers such that $b_n \rightarrow \infty$ and $n^{-1}b_n \rightarrow 0$ as $n \rightarrow \infty$. Then*

$$\lim_{n \rightarrow \infty} P(Z_n^{(M)} \leq b_n) = p_{\text{EXT}}.$$

Proof For $k = 1, 2, \dots$,

$$\liminf_{n \rightarrow \infty} P(Z_n^{(M)} \leq b_n) \geq \liminf_{n \rightarrow \infty} P(Z_n^{(M)} \leq k) = P(Z \leq k), \quad (4.16)$$

by Theorem 2. Letting $k \rightarrow \infty$ in (4.16) yields

$$\liminf_{n \rightarrow \infty} P(Z_n^{(M)} \leq b_n) \geq p_{\text{EXT}}. \quad (4.17)$$

For $\varepsilon \in (0, 1)$, let $\{Y_\varepsilon(t): t \geq 0\}$ be the branching process constructed from $\{Y(t): t \geq 0\}$, by ignoring births in $\{Y(t): t \geq 0\}$ independently and with probability ε . Let $p_{\text{EXT}}(\varepsilon)$ denote the extinction probability of $\{Y_\varepsilon(t): t \geq 0\}$. Following Whittle (1955), note that if $Z_n^{(M)} \leq \varepsilon n$ then the branching process $\{Y_\varepsilon(t): t \geq 0\}$ is a lower bound for $\{Y_n(t): t \geq 0\}$, so if $Z(\varepsilon)$ denotes the total progeny (not including the initial ancestors) of $\{Y_\varepsilon(t): t \geq 0\}$ then

$$\limsup_{n \rightarrow \infty} P(Z_n^{(M)} \leq b_n) \leq \limsup_{n \rightarrow \infty} P(Z(n^{-1}b_n) \leq b_n) \leq \limsup_{n \rightarrow \infty} p_{\text{EXT}}(n^{-1}b_n).$$

Now $p_{\text{EXT}}(n^{-1}b_n) \rightarrow p_{\text{EXT}}$ as $n \rightarrow \infty$, so

$$\limsup_{n \rightarrow \infty} P(Z_n^{(M)} \leq b_n) \leq p_{\text{EXT}}. \quad (4.18)$$

The lemma follows from (4.17) and (4.18).

We now give a more precise definition of an epidemic taking off. For $n = 1, 2, \dots$, the epidemic E_n is said to take off if it infects at least $\log n$ initial susceptibles, i.e. if the event $G_n = \{Z_n^{(M)} \geq \log n\}$ occurs. It follows from Theorem 2 and its proof that $P(G_n) \rightarrow 1 - p_{\text{EXT}}$ as $n \rightarrow \infty$.

Let $R_0 = \lambda_M p_M t_M$. Then it follows from standard branching process theory that $p_{\text{EXT}} < 1$ if and only if $R_0 > 1$. Suppose that $R_0 > 1$ and let $\tau = \min\{t > 0: t = a(t)\}$. Recalling (4.7), for $t \geq 0$, let

$$\tilde{\Sigma}(t) = \begin{bmatrix} \sigma_A^2(t) & \sigma_{AM}(t) & \sigma_{AS}(t) \\ \sigma_{AM}(t) & \sigma_M^2(t) & \sigma_{MS}(t) \\ \sigma_{AS}(t) & \sigma_{MS}(t) & \sigma_S^2(t) \end{bmatrix},$$

where $\sigma_A^2(t)$ is obtained by letting $\mu \rightarrow 0$ in the expression for $\tilde{\sigma}_A^2(t)$, and let $\tilde{B}(t) = [\tilde{b}_{ij}(t)]$ be the 3×3 matrix with elements $\tilde{b}_{11}(t) = (1 - a'(t))^{-1} a'(t)$, $\tilde{b}_{21}(t) = (1 - a'(t))^{-1} r'_M(t)$, $\tilde{b}_{31}(t) = (1 - a'(t))^{-1} r'_S(t)$ and $\tilde{b}_{ij}(t) = 0$ if $j > 1$. For $n = 1, 2, \dots$, let $\tilde{\mathbf{Z}}_n = (A_{\bullet n}(\bar{T}_\infty^{(n)}), Z_n^{(M)}, Z_n^{(S)})^\top$ and for $t \in [0, \infty]$, let $\boldsymbol{\mu}_{\tilde{Z}}(t) = (a(t), r_M(t), r_S(t))^\top$.

Theorem 3 *Suppose that $R_0 > 1$, $\sigma_i^2 < \infty$ ($i = M, S$) and that $a'(\tau) < 1$. Then,*

$$n^{-1/2}(\tilde{\mathbf{Z}}_n - n\boldsymbol{\mu}_{\tilde{Z}}(\tau)) \mid G_n \xrightarrow{D} \tilde{\mathbf{Z}} \quad \text{as } n \rightarrow \infty,$$

where $\tilde{\mathbf{Z}}$ is a 3-dimensional zero-mean normal random vector with variance matrix, $\tilde{\Sigma}_Z$ say, given by

$$\tilde{\Sigma}_Z = (I + \tilde{B}(\tau))\tilde{\Sigma}(\tau)(I + \tilde{B}(\tau)).$$

Proof First note that it follows using Lemma 2 that $\lim_{n \rightarrow \infty} P(Z_n^{(M)} > b_n \mid G_n) = 1$ for any sequence (b_n) such that $b_n \rightarrow \infty$ and $n^{-1}b_n \rightarrow 0$ as $n \rightarrow \infty$. It then follows (see the proof of Theorem 3.12 of Ball and Neal (2003)) that there exists $b > 0$ such that $\lim_{n \rightarrow \infty} P(Z_n^{(M)} > bn \mid G_n) = 1$. Observe that Lemma 1 holds in the present setting, with $\mu = 0$, so $\tilde{a}(t) = a(t)$ ($t \in [0, \infty]$). The equation $a(t) = t$ has roots $t = 0$ and $t = \tau$, and arguing as in the proof of Corollary 1 shows that $\min\{\bar{T}_\infty^{(n)}, |\bar{T}_\infty^{(n)} - \tau|\} \xrightarrow{\text{a.s.}} 0$ as $n \rightarrow \infty$. Recall that $Z_n^{(M)} = R_{\bullet n}^{(M)}(\bar{T}_\infty^{(n)})$, so again arguing as in the proof of Corollary 1, $\bar{T}_\infty^{(n)} \mid G_n \xrightarrow{P} \tau$ and $Z_n^{(i)} \mid G_n \xrightarrow{P} r_i(\tau)$ ($i = M, S$) as $n \rightarrow \infty$. (The zero root of $a(t) = t$ is excluded since $\lim_{n \rightarrow \infty} P(Z_n^{(M)} > bn \mid G_n) = 1$.)

A realisation of $\tilde{\mathbf{Z}}_n \mid G_n$ can be constructed as follows. Construct the epidemic E_n in real time using the construction described at the start of this subsection. Stop this construction as soon as $\log n$ of the initial susceptibles have been infected, and let T_n^* denote the sum of the remaining mild infectious periods of all those individuals that are infectious at that time. Now use the embedding construction of Section 2.3.1, with $m + \lceil \log n \rceil + 1$ initial mild infectives, who exert $T_0^{(n)} = \lambda_M T_n^*$ units of infectious pressure on the population, and $n - \lceil \log n \rceil - 1$ initial susceptibles, where $\lceil \cdot \rceil$ denotes integer part. Note that $n^{-1/2}T_0^{(n)} \xrightarrow{\text{a.s.}} 0$ as $n \rightarrow \infty$. Theorem 3 now follows by a similar argument to the proof of Theorem 1, since $\bar{T}_\infty^{(n)} \mid G_n \xrightarrow{P} \tau$ as $n \rightarrow \infty$.

5 Discussion

The EDS model of the paper is aimed at capturing the possibility for a transmittable disease to have different severities, with these different severities also

affecting the amount of possible further transmission. The model does so by letting infected individuals first become mildly infectious and later, if exposed sufficiently further to the disease, become severely infectious. This exposure dependent severity is one way of modelling the phenomenon of interest. A perhaps less realistic feature of the model is that the severe infectious state can occur during and/or after the mild infectious state. As mentioned previously, “additionally infectious” might be a better name for the second state. Perhaps more realistic would be to allow for an increased infectivity during the infectious period if exposed to enough additional exposure, but that this increased infectivity stops when the (original) infectious period is over. However, such a model is less tractable from a mathematical point of view.

An alternative model trying to capture the same feature is where each individual can only become mildly or severely infected (and not both as in the EDS model), and where the chances that a contacted individual becomes mildly or severely infected depend on the state of the person transmitting the disease. An advantage with such an infector-dependent severity (IDS) model is perhaps that an infected individual is only ever in one of the two infectious states, but this might also be a disadvantage: once infected the severity might very well depend on possible extra exposure. The authors intend to study such an IDS model in a forthcoming paper.

The present model permits only two types of severity, mild and severe. It is mathematically straightforward to extend the present model to allow for several severity stages. The same qualitative features remain, i.e. that the reproduction number R_0 depends only on parameters of the mildest state, whereas the final size in case of a major outbreak depends on parameters of all states, and possibly in a discontinuous way. In fact, the model can be extended further to allow severity to depend continuously on exposure to disease, by defining the model directly in terms of the functions A_i of Section 4.1 and noting that (subject to a mild moment condition) the proofs still hold if the sample paths of A_i are cadlag and nondecreasing.

Important problems in preventing infectious disease outbreaks, beside modelling the epidemic, are statistical inference and studying the effects of vaccination. Neither of these problems has been addressed in the present paper. In particular, an interesting open problem is to model and analyse how vaccination affects susceptibility and potential severity, and how this in turn affects R_0 . The fact that the final size may be discontinuous in the model parameters indicate that interesting features may also be found in terms of vaccination. Indeed, the simulations described in Section 3.2 suggest that reducing R_0 to below its usual critical value of one may *not* be sufficient to prevent major outbreaks.

Acknowledgements

This research was supported by the Swedish Research Council. We thank the organisers and participants of the Mariefred Epidemic Meeting in 2003.

References

- AABY, P., BUKH, J., LISSE, I. M. AND DA SILVA, M. C. (1988). Further community studies on the role of overcrowding and intensive exposure on measles mortality. *Rev. Infect. Dis.* **10**, 474–477.
- ALDOUS, D. J. (1985). Exchangeability and related topics. In *Ecole d'Été de Probabilités de Saint-Flour XIII, 1993* (Lecture Notes in Mathematics **1117**), ed. Hennequin, P. L. Springer, Berlin.
- ANDERSSON, H. AND BRITTON, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*. (Lecture Notes in Statistics **151**). Springer, New York.
- BALL, F. G. AND DONNELLY, P. J. (1995). Strong approximations for epidemic models. *Stoch. Proc. Appl.* **55**, 1–21.
- BALL, F. G. AND NEAL, P. J. (2003). The great circle epidemic model. *Stoch. Proc. Appl.* **107**, 233–268.
- BUTLER, J. C., PROCTOR, M. E., FESSLER, K., HOPFENSBERGER, D. J., SOSIN, D. M. AND DAVIS, J. P. (1994). Household-acquisition of measles and illness severity in an urban community in the United States. *Epidemiol. Infect.* **112**, 569–577.
- HEESTERBEEK, J. A. P. AND DIETZ, K. (1996). The concept of R_0 in epidemic theory. *Statistica Neerlandica* **50**, 89–110.
- 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 in Biomath. **86**), eds. J.-P. Gabriel, C. Lefèvre and P. Picard. 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. *BMJ* **314**, 1692.
- PARANG, N. M. AND ARCHANA, C. (2004) Varicella. *eMedicine* September 15, 2004.
<http://www.emedicine.com/ped/topic2385.htm>
- SCALIA-TOMBA, G.-P. (1985). Asymptotic final size distribution for some chain-binomial models. *Adv. Appl. Prob.* **17**, 477–495.

SCALIA-TOMBA, G.-P. (1990). On the asymptotic final size distribution of epidemics in heterogeneous populations. In *Stochastic Processes in Epidemic Theory* (Lecture Notes in Biomath. **86**), eds. J.-P. Gabriel, C. Lefèvre and P. Picard. Springer, Berlin, pp. 189–196.

SELLKE, T. (1983). On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Prob.* **20**, 390–394.

VAN DEN DRIESSCHE, P. AND WATMOUGH, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29-48.

VAN DER VAART, A. W. AND WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes: with Applications to Statistics*. Springer, New York.