View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Nottingham ePrints

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Ball, Frank and Sirl, David J. (2017) Evaluation of vaccination strategies for SIR epidemics on random networks incorporating household structure. Journal of Mathematical Biology . pp. 1-48. ISSN 1432-1416

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/42915/8/article-1.%20Sirl.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see:
http://creativecommons.org/licenses/by/2.5/

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

J. Math. Biol.
DOI 10.1007/s00285-017-1139-0

Mathematical Biology

Evaluation of vaccination strategies for SIR epidemics on random networks incorporating household structure

Frank Ball[1] · **David Sirl**[1]

Received: 1 November 2016 / Revised: 12 April 2017
© The Author(s) 2017. This article is an open access publication

Abstract This paper is concerned with the analysis of vaccination strategies in a stochastic susceptible → infected → removed model for the spread of an epidemic amongst a population of individuals with a random network of social contacts that is also partitioned into households. Under various vaccine action models, we consider both household-based vaccination schemes, in which the way in which individuals are chosen for vaccination depends on the size of the households in which they reside, and acquaintance vaccination, which targets individuals of high degree in the social network. For both types of vaccination scheme, assuming a large population with few initial infectives, we derive a threshold parameter which determines whether or not a large outbreak can occur and also the probability of a large outbreak and the fraction of the population infected by a large outbreak. The performance of these schemes is studied numerically, focusing on the influence of the household size distribution and the degree distribution of the social network. We find that acquaintance vaccination can significantly outperform the best household-based scheme if the degree distribution of the social network is heavy-tailed. For household-based schemes, when the vaccine coverage is insufficient to prevent a major outbreak and the vaccine is imperfect, we find situations in which both the probability and size of a major outbreak under the scheme which minimises the threshold parameter are *larger* than in the scheme which maximises the threshold parameter.

Keywords Branching process · Configuration model · Epidemic process · Final size · Random graph · Threshold behaviour · Vaccination

✉ David Sirl
david.sirl@nottingham.ac.uk

Frank Ball
frank.ball@nottingham.ac.uk

1 School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

Published online: 20 June 2017

Mathematics Subject Classification Primary 92D30 (Epidemiology); Secondary 60J85 (Applications of BPs) · 05C80 (Random graphs)

1 Introduction and description of results

Mathematical models for the spread of infectious disease have much to offer in terms of understanding past outbreaks, predicting likely behaviours of future outbreaks and predicting the effect of interventions or mitigating strategies. In the last decade or two there has been considerable interest and work on network epidemic models. These involve supplanting the traditional assumption of homogeneous mixing of homogeneous individuals with some random graph structure, with specific interest in being able to control the degree distribution, reflecting the varying numbers of people with which different individuals tend to interact. Other structures, for example including households and stratification of populations, have been studied for longer (Bartoszyński 1972; Ball et al. 1997; Watson 1972; Scalia-Tomba 1986); but structures with a 'social network' type of interpretation start around the turn of the millenium with the works of Andersson (1997, 1998), Diekmann et al. (1998) and Newman (2002). In this and most other papers in the field we typically have in mind an infection spreading through a human population. However, much the same ideas apply to mathematical models of a variety of other motivating applications, such as the spread of rumours or information through human populations, infection or information spread through a population of other animals and virus spread through a network of computers.

In this paper we build on the model of Ball et al. (2009, 2010) which includes household and network structure to include vaccination, with some emphasis on so-called acquaintance vaccination (Cohen et al. 2003; Britton et al. 2007) as elucidated in Ball and Sirl (2013) in a model without household structure. In the model of Ball et al. (2010), a population of fixed size is given social network structure via the configuration model random graph (see e.g. Bollobás 1980; Newman et al. 2001; Newman 2002) and the population is also partitioned into households (see e.g. Ball et al. 1997). A stochastic Susceptible → Infective → Removed (SIR) epidemic model is then defined on this population structure. A first quantity of interest in this model is the final size, which is the (random) number of initial susceptibles that are infected at some point during the epidemic. In line with much of modern stochastic epidemic theory, one can use branching process approximations to prove a threshold theorem (valid in the large population limit) which determines whether the infection will necessarily die out relatively quickly, resulting in a small final size, or whether it is possible for the epidemic to take off and infect a substantial fraction of the population. These methods also yield approximations for the probability that a supercritical epidemic will take off and using closely related methods one can also study final size properties of such a large outbreak.

This paper provides tools for studying the effect of introducing vaccination into this model. *Households-based* vaccination schemes are those that can be described in terms of the distribution of the number of vaccinated individuals in households of size n, for every household size n in the population. This includes as special cases the situation when we vaccinate individuals who are chosen uniformly at random from

the population (the distributions are binomial) and vaccinating households at random (the distributions are concentrated at 0 and n). Optimal schemes in this context often resemble the equalising strategy (Ball et al. 1997, Sect. 5.2) where one vaccinates preferentially in larger households, there being more of a herd immunity effect available to exploit in those larger households on top of the direct protection of vaccinated individuals. *Acquaintance vaccination* schemes exploit the heterogeneity of individuals' connectivities in a network to preferentially target better-connected individuals for vaccination. Instead of vaccinating individuals chosen from the population in some way, one samples individuals and then vaccinates their friends—their acquaintances in the network. This exploits the so-called friendship paradox: the observation that, for most people, their friends on average have more friends than they do (Feld 1991).

Our main theoretical results are the calculation of asymptotic final size quantities, via appropriate branching process approximations, when we include vaccination in Ball et al.'s (2010) household-network model. This extends Becker and Starczak's (1997) and Ball and Lyne's (2002, 2006) results on the standard households model to have network-based (rather than homogeneous mixing) casual contacts. It also extends Ball and Sirl's (2013) results on acquaintance vaccination in a population with network (but not household) structure. We also explore the model numerically and find that there can be substantial differences in the performance of the different vaccine allocation strategies.

Results like those in this paper can inform about which mechanism of spreading (household or network) can be most beneficially targeted for reducing the impact of outbreaks of disease. This seems likely to be relevant when considering, for example, how best to reduce the impact of any future Ebola outbreak in West Africa (Adams 2016) using vaccines that are currently under development. Other possible applications include pandemic influenza (Halloran et al. 2008) and smallpox (Halloran et al. 2002). For all of these diseases a salient feature is that they have markedly enhanced transmission within small social groups such as households. In each of these possible applications some further refinements of the underlying model may be valuable.

The remainder of the paper is structured as follows. In Sect. 2 we specify our models for the population structure and evolution of the epidemic, then outline the analysis of the final outcome of the epidemic and lastly introduce the model we use for the action of a vaccine on individuals who receive it. In Sect. 3 we consider the effect of households-based vaccination, including optimal households-based strategies, and in Sect. 4 we consider the effect of acquaintance vaccination. The analysis in these sections is of the same final outcome measures as in the basic model in Sect. 2. Some exploration of the behaviour of the model (mainly numerical) is presented in Sect. 5. Lastly we offer some concluding remarks in Sect. 6. Details of many of the calculations relating to Sect. 4 are given in Appendix A, while in Appendix B some known exact results for the final outcome and susceptibility sets of multitype stochastic SIR epidemics (households in our setting) are stated and applied to the asymptotic analyses of Sects. 2, 3 and 4.

2 Model, threshold behaviour and vaccination

2.1 Model

The model we consider is one for the spread of an SIR epidemic on a finite random network incorporating household structure (Ball et al. 2010). We assume that the population consists of N individuals and is partitioned into m households, of which m_n are of size n ($n = 1, 2, \ldots$). Thus $m = \sum_{n=1}^{\infty} m_n$ and $N = \sum_{n=1}^{\infty} n m_n$. The network of possible global (i.e. between-household) contacts is constructed using the configuration model (with random, rather than specified, degree sequence). Thus each individual is assigned a number of 'half-edges' independently, according to an arbitrary but specified discrete random variable D having mass function $P(D = k) = p_k$ ($k = 0, 1, \ldots$). Then all such half-edges are paired up uniformly at random to form the edges in the graph describing the global network. If the total number of half-edges is odd, we ignore the single leftover half-edge.

Our analysis is asymptotic as the number of households $m \to \infty$. We require that, as $m \to \infty$, $m_n/m \to \rho_n$ ($n = 1, 2, \ldots$), where $(\rho_1, \rho_2, \ldots)$ is a proper probability distribution (i.e. $\sum_{n=1}^{\infty} \rho_n = 1$) having finite mean $\mu_H = \sum_{n=1}^{\infty} n\rho_n$. Thus μ_H is the mean household size in the limiting population. We also require that $\mu_D = E[D]$ is finite. These assumptions are sufficient for our analysis. If we make the stronger assumptions that $\sigma_D^2 = \mathrm{var}(D)$ and $\sum_{n=1}^{\infty} n^2 \rho_n$ are both finite, then parallel edges and self-loops, between either individuals or households, become sparse in the global network as $n \to \infty$ (Ball et al. 2010, Sect. 1.2).

The epidemic is initiated by a single individual, chosen uniformly at random from the population, becoming infected, with the other individuals in the population all assumed to be susceptible. The infectious periods of different infectives are each distributed according to a random variable I, having an arbitrary but specified distribution, which is most conveniently specified in terms of its moment generating function $\phi_I(\theta) = E[\mathrm{e}^{-\theta I}]$ ($\theta \geq 0$). Throughout its infectious period, a given infective makes infectious contact with any given member of its household at the points of a Poisson process having rate λ_L and with any given global neighbour at the points of a Poisson process with rate λ_G. (Note that λ_L and λ_G are both individual to individual contact rates.) If an individual so contacted is susceptible then it becomes infected, otherwise the contact has no effect. Contacted susceptibles are immediately able to infect other individuals. An infective individual becomes removed at the end of its infectious period and plays no further role in the epidemic. All infectious periods, global degrees and Poisson processes are assumed to be mutually independent. The epidemic ceases as soon as there is no infective in the population.

For ease of exposition we assume throughout that there is no latent period and that the epidemic is started by a single infective chosen uniformly at random from the population. As explained by Ball et al. (2010), these assumptions may be relaxed without compromising mathematical tractability. In particular, our results are related to the final outcome of the epidemic, the distribution of which is invariant to very general assumptions concerning a latent period (see e.g. Pellis et al. 2008).

2.2 Threshold behaviour

2.2.1 Early stages of epidemic

Recall that the global network is formed by pairing up the half-edges uniformly at random. It follows that in the early stages of an epidemic the probability that a global contact is with an individual residing in a previously infected household is small, indeed it is zero in the limit as $m \to \infty$. Thus, in the early stages of an epidemic, the process of infected households can be approximated by a branching process. The individuals in this branching process correspond to infected households in the epidemic process, and the offspring of a given individual in the branching process are all households that are contacted globally by members of the local (within-household) epidemic in the parent household.

The offspring distribution for this branching process is usually different in the initial generation than in all subsequent generations. The number of global neighbours that the initial infective in the population may infect is distributed according to D, as that individual is chosen uniformly at random from the population. The number of global neighbours that the initial infective in any subsequent infected household (in the branching process approximation) may infect is distributed according to $\tilde{D} - 1$, where $\tilde{D}$ is the degree of a typical neighbour of a typical individual in the network. The -1 arises because such an infective has been infected through the global network, so one of its neighbours (i.e. its infector) is not available for further infection. Note that a given half-edge is k times as likely to be paired with a half-edge emanating from an individual with degree k than with one emanating from an individual with degree 1, so $P(\tilde{D} = k) = \mu_D^{-1} k p_k$ $(k = 1, 2, \ldots)$. The distributions of D and $\tilde{D} - 1$ are equal if and only if D has a Poisson distribution. For any non-negative integer valued random variable X, we denote its probability generating function (PGF) by f_X, so $f_X(s) = E[s^X]$ $(0 \le s \le 1)$. We note for future reference that $f_{\tilde{D}-1}(s) = f_D'(s)/\mu_D$.

The approximation of the early stages of the epidemic process by the above branching process is made mathematically fully rigorous by Ball et al. (2009) and Ball and Sirl (2012). The latter shows that a sequence of epidemic processes, indexed by m, and the approximating branching process can be constructed on the same probability space so that, as $m \to \infty$, the number of households ultimately infected in the epidemic process converges almost surely to the total progeny of the branching process. Thus, provided m is large, whether or not the epidemic can become established and lead to a major outbreak is determined by whether or not the branching process is supercritical.

Let C and $\tilde{C}$ be random variables describing the number of offspring of the initial and a typical subsequent individual, respectively, in the branching process. Then standard branching process theory (e.g. Haccou et al. 2005, Theorem 5.2), gives that the extinction probability of the branching process is strictly less than one if and only if $R_* = E[\tilde{C}] > 1$. Thus R_* serves as a threshold parameter for the epidemic model. We now outline the calculation of R_*. Further details are given by Ball et al. (2010).

First note that, since the degree and household size of an individual are independent, the probability that a typical globally infected individual resides in a household of size n is given by $\tilde{\rho}_n = \mu_H^{-1} n \rho_n$ $(n = 1, 2, \ldots)$. (An individual chosen uniformly at random

from the population is n times as likely to reside in a given household of size n than in a given household of size 1.) Thus,

$$E[\tilde{C}] = \sum_{n=1}^{\infty} \tilde{\rho}_n E\left[\tilde{C}^{(n)}\right], \tag{2.1}$$

where $\tilde{C}^{(n)}$ is the number of global infections emanating from a typical size-n single-household epidemic initiated by a single infective who is infected through the global network. Consider such a size-n single-household epidemic. Label the household members $0, 1, \ldots, n-1$, where 0 is the initial infective, and write

$$\tilde{C}^{(n)} = C_0 + \sum_{i=1}^{n-1} \chi_i C_i, \tag{2.2}$$

where $\chi_i = 1$ if individual i is infected by the single-household epidemic, otherwise $\chi_i = 0$, and C_i is the number of global infections made by individual i (assuming it is infected). (Throughout the paper we adopt the usual convention that empty sums are equal to 0, so, for example, $\tilde{C}^{(1)} = C_0$.) Let $T^{(n)} = \sum_{i=1}^{n-1} \chi_i$ be the final size of the single-household epidemic, not including the initial infective, and $\mu^{(n)}(\lambda_L) = E[T^{(n)}]$. (A formula for $\mu^{(n)}(\lambda_L)$ is given in Eq. (B.25) in Appendix B.6.) Whether or not a given individual, i say, is infected by the single-household epidemic is independent of its infectious period, so χ_i and C_i are independent. Thus, taking expectations of (2.2) and exploiting symmetries yields

$$E\left[\tilde{C}^{(n)}\right] = E[C_0] + \mu^{(n)}(\lambda_L) E[C_1]. \tag{2.3}$$

Since the probability that a Poisson process of rate λ_G has no points in a time of length I is $\mathrm{e}^{-\lambda_G I}$ we have that the probability an infected individual infects a given global neighbour is $p_G = 1 - E[\mathrm{e}^{-\lambda_G I}] = 1 - \phi_I(\lambda_G)$. We also have that the number of uninfected global neighbours of individual 0 is distributed according to $\tilde{D} - 1$. Thus $E[C_0] = p_G \mu_{\tilde{D}-1}$, where $\mu_{\tilde{D}-1} = E[\tilde{D} - 1] = \mu_D + \frac{\sigma_D^2}{\mu_D} - 1$. Similarly, $E[C_1] = p_G \mu_D$, since the number of uninfected global neighbours of individual 1 is distributed according to D. Substituting these results into (2.3) and then into (2.1) yields

$$R_* = p_G \left[\mu_{\tilde{D}-1} + \mu_D \sum_{n=1}^{\infty} \tilde{\rho}_n \mu^{(n)}(\lambda_L)\right]. \tag{2.4}$$

Let p_{maj} be the probability that a major outbreak occurs. Then standard branching process theory [e.g. Haccou et al. (2005, Theorem 5.2)], shows that $p_{\mathrm{maj}} = 1 - f_C(\sigma)$, where σ is the smallest solution of $f_{\tilde{C}}(s) = s$ in [0, 1]. Note that, analagous to (2.1), $f_C(s) = \sum_{i=1}^{\infty} \tilde{\rho}_n f_{C^{(n)}}(s)$ and $f_{\tilde{C}}(s) = \sum_{n=1}^{\infty} \tilde{\rho}_n f_{\tilde{C}^{(n)}}(s)$. The PGFs $f_{C^{(n)}}$ and $f_{\tilde{C}^{(n)}}$ can be calculated using the methods described in Appendices B.1 and B.2. As noted by Ball et al. (2010, Sect. 3.2), these calculations are much simpler in the case where

the infectious period is constant, since in this case an infected individual infects its global neighbours independently.

2.2.2 *Final outcome of major outbreak*

We now consider the fraction of the population that are ultimately infected by a major outbreak. The key tool we use is the *susceptibility set* (Ball 2000; Ball and Lyne 2001; Ball and Neal 2002), which we now describe. Let $\mathcal{N} = \{1, 2, \ldots, N\}$ denote the entire population of N individuals. For each $i \in \mathcal{N}$, by sampling from the infectious period distribution and then the relevant Poisson processes for local and global contacts, draw up a (random) list of individuals i would make infectious contact with if it was to become infected. Then construct a random directed graph on $\mathcal{N}$, in which for any pair, (i, j) say, of individuals there is an arc from i to j if and only if j is in i's list. The susceptibility set of a given individual, i say, consists of those individuals from which there is a chain of arcs to i in the graph (including i itself). Note that any given individual is ultimately infected by the epidemic if and only if the initial infective belongs to its susceptibility set.

Similarly to the early stages of an epidemic, we can approximate the susceptibility set of an individual, i say, chosen uniformly at random from the population, by a households-based branching process. We first consider i's *local* susceptibility set, i.e. the susceptibility set obtained when only local (within-household) contacts are considered. Suppose that this local susceptibility set has size $M^{(n)}+1$, where n denotes the size of i's household; so $M^{(n)}$ is the size of the local susceptibility set excluding i. (The probability mass function of $M^{(n)}$ is given by Eq. (B.23) in Appendix B.5.) Let B be the number of individuals, who are not in i's household, that in the random directed graph have an edge leading directly *to* one of the $M^{(n)} + 1$ individuals in i's local susceptibility set. Each individual in i's local susceptibility set has global degree distributed independently according to D, and, as $m \to \infty$, each global neighbour of i's local susceptibility set enters i's susceptibility set independently with probability p_G. Thus, taking expectations with respect to i's household size,

$$f_B(s) = \sum_{n=1}^{\infty} \tilde{\rho}_n f_{B^{(n)}}(s), \tag{2.5}$$

where

$$f_{B^{(n)}}(s) = f_D(1 - p_G + p_G s) f_{M^{(n)}}(f_D(1 - p_G + p_G s)). \tag{2.6}$$

The above B individuals form the first generation of our approximating branching process. We next consider each of these B individuals in turn, construct the local susceptibility sets in their respective households (which are distinct with probability tending to one as $m \to \infty$) and then examine the global neighbours of these local susceptibility sets to obtain the second generation of the approximating branching process, and so on. Note that the initial individual in each of these B local susceptibility sets has degree distributed according to $\tilde{D}$ and one of its global neighbours is already in i's susceptibility set. Thus, as above, the offspring distribution for the branching process is different for the initial individual than for all subsequent individuals. If we

let $\tilde{B}$ denote the offspring random variable of a typical non-initial individual, then arguing as in the derivation of (2.5),

$$f_{\tilde{B}}(s) = \sum_{n=1}^{\infty} \tilde{\rho}_n f_{\tilde{B}^{(n)}}(s),$$

where

$$f_{\tilde{B}^{(n)}}(s) = f_{\tilde{D}-1}(1 - p_G + p_G s) f_{M^{(n)}}(f_D(1 - p_G + p_G s)). \tag{2.7}$$

The probability that the approximating branching process survives (i.e. does not go extinct) is given by $z = 1 - f_B(\xi)$, where ξ is the smallest solution of $f_{\tilde{B}}(s) = s$ in $[0, 1]$. It is straightforward to show that $E[\tilde{B}] = E[\tilde{C}]$, so $z > 0$ if and only if $R_* > 1$. Moreover, if $R_* > 1$ then z is the expected proportion of the population that is ultimately infected by a major outbreak in the limit as $m \to \infty$; see Ball et al. (2009) for a formal proof when all the households have the same size. Furthermore, the proof of Ball et al. (2014, Theorem 3.4) can be adapted to show that, as $m \to \infty$, the proportion of the population that is ultimately infected by a major outbreak converges in probability to z. Thus we refer to z as the relative final size of a major outbreak.

2.3 Vaccination

In modelling vaccination there are two distinct aspects that must be modelled: who gets vaccinated and what happens to those who are vaccinated. Vaccine *allocation* models (addressing the former issue) are our focus in this paper. We now outline the vaccine *action* models (addressing the latter aspect) that we allow for in our analysis.

We use a model for vaccine action, proposed by Becker and Starczak (1998), in which the vaccine response of an individual who is vaccinated is described by a random vector (A, B) which takes values in $[0, \infty)^2$. Here A denotes the relative susceptibility (compared to an unvaccinated individual) and B the relative infectivity if the vaccinated individual becomes infected. Thus all Poisson processes concerning potential infection of the individual have their rates multiplied by A and the Poisson processes governing the contacts the individual makes, if infected, have their rates multiplied by B. The vaccine responses of distinct vaccinees are assumed to be mutually independent. Within this framework we consider two special cases, the *all-or-nothing* and the *non-random* vaccine responses, where (A, B) is supported on 2 and 1 points respectively. Our methods extend to vaccine action models where (A, B) is supported on finitely many points; the extension is straightforward analytically but quickly becomes numerically cumbersome if there are more than a few possible outcomes of (A, B).

The all-or-nothing model (see e.g. Halloran et al. 1992) is obtained by setting $P((A, B) = (0, 0)) = 1 - P((A, B) = (1, 1)) = \varepsilon$, so vaccinated individuals are rendered completely immune with probability ε, otherwise the vaccine has no effect. The non-random model assumes that $P((A, B) = (a, b)) = 1$, for some (a, b), so all vaccinated individuals respond identically (see e.g. Ball and Lyne 2006). An

important special case is the *leaky* model (see e.g. Halloran et al. 1992), when $b = 1$, so vaccination does not affect an individual's ability to transmit the disease if they become infected. Setting $\varepsilon = 1$ in the all-or-nothing model or $a = b = 0$ in the non-random model yields a *perfect* vaccine response; that is one in which all vaccinated individuals are rendered completely immune.

3 Households based vaccination

3.1 Introduction

In this section we consider vaccine allocation strategies based on household size and determine their impact on R_*, p_{maj} and z under the all-or-nothing and non-random vaccine action models. For $n = 1, 2, \ldots$ and $v = 0, 1, \ldots, n$, let x_{nv} denote the proportion of households of size n that have v members vaccinated. Let p_V denote the proportion of the population that is vaccinated, i.e. the *vaccination coverage*. Then p_V is also the probability that an individual chosen uniformly at random from the population is vaccinated. Conditioning on the size of such an individual's household yields

$$p_V = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v=0}^{n} \frac{v}{n} x_{nv}. \tag{3.1}$$

We derive results for an arbitrary but specified vaccine allocation however in the numerical studies we consider four allocation schemes: uniformly chosen households, uniformly chosen individuals, 'best' and 'worst'. In the uniformly chosen households scheme, households are chosen uniformly at random and all of their members are vaccinated. Thus, if the vaccination coverage is p_V, $x_{nv} = p_V \delta_{vn} + (1 - p_V)\delta_{v0}$ ($n = 1, 2, \ldots$; $v = 0, 1, \ldots, n$), where $\delta_{vk} = 1$ if $v = k$ and $\delta_{vk} = 0$ if $v \neq k$. In the uniformly chosen individuals scheme, individuals are chosen uniformly at random and vaccinated, so $x_{nv} = \binom{n}{v} p_V^v (1 - p_V)^{n-v}$ ($n = 1, 2, \ldots$; $v = 0, 1, \ldots, n$). The best and worst schemes are the allocations that make R_* respectively as small as possible and as large as possible, for a given vaccination coverage.

3.2 All-or-nothing vaccine action

To analyse the consequences of a vaccination scheme using an all-or-nothing vaccine action model it is convenient to use the concept of a potential infectious global contact. Consider a given infected individual, i say, and a given global neighbour j of i. Then j is a potential infectious global contact of i if it is in i's list of individuals it makes infectious contact with (see the start of Sect. 2.2.2). The potential infectious global contact becomes an actual infectious global contact if j is unvaccinated or is vaccinated but the vaccination fails.

The early stages of an epidemic with vaccination are approximated by a branching process of (potentially) infected households, and the offspring of a given individual in the branching process are all households with which members of the single-household epidemic in the parent household make a potential global infectious contact. As in the

case without vaccination, the offspring distribution of this branching process is usually different in the initial generation from all subsequent generations. Let $\tilde{C}'$ denote the offspring random variable for a non-initial individual. Conditioning first on the size of the corresponding household and then on the number of people vaccinated in that household yields, in an obvious notation,

$$f_{\tilde{C}'}(s) = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v=0}^{n} x_{nv} f_{\tilde{C}'_{n,v}}(s). \tag{3.2}$$

To determine $f_{\tilde{C}'_{n,v}}(s)$, consider a household in state (n, v), i.e. of size n having v members vaccinated. For $k = 0, 1, \ldots, v$, the probability that k vaccinations are successful is $\binom{v}{k}\varepsilon^k(1-\varepsilon)^{v-k}$, and given that k vaccinations are successful, the probability that the initial potentially contacted individual in that household is susceptible and thus triggers a local epidemic is $\frac{n-k}{n}$. Moreover, if such a local epidemic is triggered, the number of potential infectious global contacts emanating from the local epidemic is distributed as $\tilde{C}^{(n-k)}$, where $\tilde{C}^{(n)}$ is as in Eq. (2.1). Thus,

$$f_{\tilde{C}'_{n,v}}(s) = \sum_{k=0}^{v} \binom{v}{k} \varepsilon^k (1-\varepsilon)^{v-k} \left[\frac{(n-k)}{n} f_{\tilde{C}^{(n-k)}}(s) + \frac{k}{n} \right]. \tag{3.3}$$

The distribution of the offspring random variable, C' say, for the initial generation depends on how the initial infective is chosen. For $n = 1, 2, \ldots$ and $v = 0, 1, \ldots, n$, let $p^V_{n,v}$ be the probability that a vaccinated individual chosen uniformly at random resides in a household in state (n, v) and let $p^U_{n,v}$ be the corresponding probability for an unvaccinated individual. Then

$$p^V_{n,v} = \frac{\tilde{\rho}_n x_{nv} \frac{v}{n}}{p_V} \quad \text{and} \quad p^U_{n,v} = \frac{\tilde{\rho}_n x_{nv} \left(1 - \frac{v}{n}\right)}{1 - p_V}. \tag{3.4}$$

Thus, if the epidemic is started by an individual chosen uniformly at random from all unvaccinated individuals being infected, then C' is distributed as C'_U, where

$$f_{C'_U}(s) = \sum_{n=1}^{\infty} \sum_{v=0}^{n-1} p^U_{n,v} \sum_{k=0}^{v} \binom{v}{k} \varepsilon^k (1-\varepsilon)^{v-k} f_{C^{(n-k)}}(s).$$

Alternatively, if the epidemic is started by choosing a vaccinated individual uniformly at random, who triggers an outbreak only if its vaccination fails, then C' is distributed as C'_V, where

$$f_{C'_V}(s) = \varepsilon + (1-\varepsilon) \sum_{n=1}^{\infty} \sum_{v=1}^{n} p^V_{n,v} \sum_{k=0}^{v-1} \binom{v-1}{k} \varepsilon^k (1-\varepsilon)^{v-1-k} f_{C^{(n-k)}}(s).$$

The probability of a major outbreak may now be calculated as at the end of Sect. 2.2.1. Again, the formulae simplify appreciably if the infectious period is constant.

A post-vaccination threshold parameter is given by $R_v = E[\tilde{C}'] = f'_{\tilde{C}'}(1)$. Differentiating (3.2) and (3.3), or a direct calculation, yields

$$R_v = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v=0}^{n} x_{nv} \mu_{n,v}, \tag{3.5}$$

where $\mu_{n,v} = E[\tilde{C}'_{n,v}]$ is given by

$$\mu_{n,v} = \sum_{k=0}^{v} \binom{v}{k} \varepsilon^k (1-\varepsilon)^{v-k} \left(\frac{n-k}{k}\right) E[\tilde{C}^{(n-k)}].$$

As with the case of no vaccination, we can determine the relative final size of a major outbreak by considering a households-based branching process that approximates the susceptibility set of a typical individual. As with the forward process, it is convenient to consider potential global neighbours when constructing this branching process. Thus we start with an individual, i^* say, chosen uniformly at random from the population, construct its local susceptibility set, taking the vaccine status of individuals in the household into account, then determine which global neighbours of individuals in this local susceptibility set would enter the susceptibility set of i^* if they were susceptible (i.e. unvaccinated or unsuccessfully vaccinated). These individuals correspond to the first generation of the approximating backward branching process. Suppose that there are B' such individuals. Next we take each of these B' individuals in turn, first determine whether they really do enter the susceptibility set of i^* (this happens with probability 1 if the individual is unvaccinated and with probability $1-\varepsilon$ if it is vaccinated, independently for distinct individuals) and if they do enter the susceptibility set of i^*, determine the number of potential global neighbours of its local susceptibility set to obtain its offspring in the branching process, and so on.

Let $\tilde{B}'$ be the offspring random variable for any non-initial individual in this backward branching process. Conditioning first on the state (n, v) of that individual's household and then on whether it joins the susceptibility set of i^*, we obtain

$$f_{\tilde{B}'}(s) = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v=0}^{n} x_{nv} f_{\tilde{B}'_{n,v}}(s),$$

where

$$f_{\tilde{B}'_{n,v}}(s) = \sum_{k=0}^{v} \binom{v}{k} \varepsilon^k (1-\varepsilon)^{v-k} \left[\frac{(n-k)}{n} f_{\tilde{B}^{(n-k)}}(s) + \frac{k}{n}\right]$$

and $f_{\tilde{B}^{(n)}}$ is as in Eq. (2.7).

The distribution of B' depends on how the initial individual i^* is chosen. If i^* is chosen uniformly at random from all unvaccinated individuals, then B' is distributed as B'_U, say, and conditioning on the state (n, v) of i^*'s household yields

$$f_{B'_U}(s) = \sum_{n=1}^{\infty}\sum_{v=0}^{n-1} p_{n,v}^U \sum_{k=0}^{v}\binom{v}{k}\varepsilon^k(1-\varepsilon)^{v-k} f_{B^{(n-k)}}(s),$$

where $f_{B^{(n)}}$ is as in Eq. (2.6). If i^* is chosen uniformly at random from all vaccinated individuals, then B' is distributed as B'_V, where

$$f_{B'_V}(s) = \varepsilon + (1-\varepsilon)\sum_{n=1}^{\infty}\sum_{v=1}^{n} p_{n,v}^V \sum_{k=0}^{v-1}\binom{v-1}{k}\varepsilon^k(1-\varepsilon)^{v-1-k} f_{B^{(n-k)}}(s).$$

Let ξ^V be the smallest solution of $f_{\tilde{B}'}(s) = s$ in $[0, 1]$. Then the proportion of unvaccinated individuals that are ultimately infected by a major outbreak is $z^U = 1 - f_{B'_U}(\xi^V)$ and the corresponding proportion for vaccinated individuals is $z^V = 1 - f_{B'_V}(\xi^V)$. The overall proportion of the population infected by a major outbreak is $z = p_V z^V + (1 - p_V)z^U$.

3.3 Non-random vaccine action

Analysing the consequences of a vaccination scheme using a non-random vaccine action model is more difficult than with an all-or-nothing vaccine action model since disease spread is now genuinely two type. Thus we now consider two types of individual: type-U, unvaccinated, and type-V, vaccinated. The early stages of the epidemic can again be approximated by a branching process of infected households. This is now a two-type branching process, with the type of an infected household being given by the type of the initial case in that household. The offspring of a given individual in the branching process correspond to the households that are contacted globally by members of that individual's corresponding single-household epidemic in the epidemic process.

Let $\boldsymbol{C}^U = (C_{UU}, C_{UV})$ denote the offspring random variable for the initial individual in the above branching process, given that individual is of type U, and let $\boldsymbol{C}^V = (C_{VU}, C_{VV})$ be the corresponding offspring random variable when the initial individual has type V. Thus, for example, C_{UU} and C_{UV} are respectively the number of unvaccinated and vaccinated individuals globally infected by the initially infected household, given that the first infective in that household is unvaccinated. Define $\tilde{\boldsymbol{C}}^U = (\tilde{C}_{UU}, \tilde{C}_{UV})$ and $\tilde{\boldsymbol{C}}^V = (\tilde{C}_{VU}, \tilde{C}_{VV})$ similarly for subsequent individuals in the branching process. For $\boldsymbol{s} = (s_U, s_V) \in [0, 1]^2$, let $f_{\boldsymbol{C}^U}(\boldsymbol{s}) = E[s_U^{C_{UU}} s_V^{C_{UV}}]$ and define $f_{\boldsymbol{C}^V}(\boldsymbol{s})$, $f_{\tilde{\boldsymbol{C}}^U}(\boldsymbol{s})$ and $f_{\tilde{\boldsymbol{C}}^V}(\boldsymbol{s})$ similarly. (Here and henceforth, for any discrete random vector $\boldsymbol{X}$ we denote its joint PGF by $f_{\boldsymbol{X}}$.) Recall that when the network of global contacts is formed half-edges are paired uniformly at random. It follows that, for $A \in \{U, V\}$,

$$f_{\tilde{\boldsymbol{C}}^A}(\boldsymbol{s}) = \sum_{n=1}^{\infty}\sum_{v=0}^{n} p_{n,v}^A f_{\tilde{\boldsymbol{C}}_{n,v}^A}(\boldsymbol{s}),$$

where $\tilde{\boldsymbol{C}}^A_{n,v}$ is a random vector giving the numbers of unvaccinated and vaccinated global infections that emanate from a non-initial infected household of size n, having v members vaccinated, whose primary case is of type A. The distributions of $\boldsymbol{C}^U$ and $\boldsymbol{C}^V$ depend on how the initial infective is chosen. If it is chosen uniformly at random from all individuals of the appropriate type in the population, then, for $A \in \{U, V\}$,

$$f_{\boldsymbol{C}^A}(\boldsymbol{s}) = \sum_{n=1}^{\infty} \sum_{v=0}^{n} p^A_{n,v} f_{\boldsymbol{C}^A_{n,v}}(\boldsymbol{s}),$$

where $\boldsymbol{C}^A_{n,v}$ is defined analogously to $\tilde{\boldsymbol{C}}^A_{n,v}$ but for the initial infected household.

Let

$$\tilde{M} = \begin{bmatrix} \tilde{m}_{UU} & \tilde{m}_{UV} \\ \tilde{m}_{VU} & \tilde{m}_{VV} \end{bmatrix},$$

where, for example, $\tilde{m}_{UU} = E[\tilde{C}_{UU}]$. The post-vaccination threshold parameter R_v is given by the dominant eigenvalue (a real, positive eigenvalue of maximum modulus) of $\tilde{M}$. It is well known (e.g. Haccou et al. 2005, p. 123) that, provided $\tilde{m}_{UV}\tilde{m}_{VU} \neq 0$, the two-type branching process has non-zero probability of surviving if and only if $R_v > 1$. Moreover, if $\tilde{m}_{UV}\tilde{m}_{VU} \neq 0$ and $R_v > 1$, then the survival probability (and hence the probability of a major outbreak) can be determined as follows. Let $\boldsymbol{\sigma} = (\sigma_U, \sigma_V)$ be the unique solution in $[0, 1)^2$ of the equations

$$\sigma_U = f_{\tilde{C}^U}(\sigma_U, \sigma_V), \quad \sigma_V = f_{\tilde{C}^V}(\sigma_U, \sigma_V). \tag{3.6}$$

Then, if the epidemic is started by an unvaccinated individual chosen uniformly at random from the population becoming infected, the probability of a major outbreak is $p^U_{\text{maj}} = 1 - f_{C^U}(\boldsymbol{\sigma})$. The corresponding probability when the initial infective is a vaccinated individual is $p^V_{\text{maj}} = 1 - f_{C^V}(\boldsymbol{\sigma})$. Calculation of the PGFs $f_{\boldsymbol{C}^U_{n,v}}$, $f_{\boldsymbol{C}^V_{n,v}}$, $f_{\tilde{\boldsymbol{C}}^U_{n,v}}$ and $f_{\tilde{\boldsymbol{C}}^V_{n,v}}$, which is rather involved unless the infectious period is constant, requires the methods described in Appendices B.1 and B.3. Calculation of $\tilde{M}$, which is sufficient for determining optimal vaccination strategies, is simpler and we now outline it.

Consider first a local (single-household) epidemic in a household in state (n, v), initiated by one of the household members becoming infected. Let $T^{(n,v)}_U$ and $T^{(n,v)}_V$ denote the number of unvaccinated and vaccinated individuals ultimately infected by this local epidemic, not including the initial case. For $A, A' \in \{U, V\}$, let $\mu^{(n,v)}(A, A')$ be the mean of $T^{(n,v)}_{A'}$ given that the initial case is of type A. (Calculation of $\mu^{(n,v)}(A, A')$ is described in Appendix B.6.) Also, define the vaccine-status-dependent marginal transmission probabilities $p^{\text{NR}}_G(U, U)$, $p^{\text{NR}}_G(U, V)$, $p^{\text{NR}}_G(V, U)$ and $p^{\text{NR}}_G(V, V)$ between global neighbours, where, for example, $p^{\text{NR}}_G(U, V)$ is the probability that an unvaccinated infective infects a given vaccinated global neighbour. Then

$$P_G^{\mathrm{NR}} = \begin{bmatrix} p_G^{\mathrm{NR}}(U,U) & p_G^{\mathrm{NR}}(U,V) \\ p_G^{\mathrm{NR}}(U,V) & p_G^{\mathrm{NR}}(V,V) \end{bmatrix} = \begin{bmatrix} 1-\phi_I(\lambda_G) & 1-\phi_I(a\lambda_G) \\ 1-\phi_I(b\lambda_G) & 1-\phi_I(ab\lambda_G) \end{bmatrix}, \quad (3.7)$$

recalling that ϕ_I is the moment generating function of the generic infectious period random variable I.

Conditioning on the state (n, v) of the infected household shows that, for $A, A' \in \{U, V\}$,

$$\tilde{m}_{AA'} = \sum_{n=1}^{\infty} \sum_{v=0}^{n} p_{n,v}^{A} \tilde{m}_{AA'}^{(n,v)}, \quad (3.8)$$

where $\tilde{m}_{AA'}^{(n,v)}$ is defined analogously to $\tilde{m}_{AA'}$ but for a household in state (n, v). Further, arguing as in the derivations of (2.3) and (2.4), yields

$$\tilde{m}_{AA'}^{(nv)} = \left[\mu_{\tilde{D}-1} p_G^{\mathrm{NR}}(A, A') + \mu_D \left(\mu^{(n,v)}(A, U) p_G^{\mathrm{NR}}(U, A') + \mu^{(n,v)}(A, V) p_G^{\mathrm{NR}}(V, A')\right)\right] p_{A'}, \quad (3.9)$$

where $p_U = 1 - p_V$. Hence, if we let

$$F = \begin{bmatrix} F_{UU} & F_{UV} \\ F_{VU} & F_{VV} \end{bmatrix} \quad \text{and} \quad D_V = \begin{bmatrix} 1-p_V & 0 \\ 0 & p_V \end{bmatrix},$$

where, for $A, A' \in \{U, V\}$,

$$F_{AA'} = \sum_{n=1}^{\infty} \sum_{v=0}^{n} p_{n,v}^{A} \mu^{(n,v)}(A, A'),$$

then (3.8) and (3.9) yield

$$\tilde{M} = (\mu_{\tilde{D}-1} I + \mu_D F) P_G^{\mathrm{NR}} D_V. \quad (3.10)$$

Turning to the relative final size of a major outbreak, we consider a households-based branching process that approximates the susceptibility set of a given individual. This is now a two-type branching process, with type (U or V) corresponding to the type of the primary member of the corresponding local susceptibility set. Define the offspring random variables $\boldsymbol{B}^U = (B_{UU}, B_{UV})$, $\boldsymbol{B}^V = (B_{VU}, B_{VV})$, $\tilde{\boldsymbol{B}}^U = (\tilde{B}_{UU}, \tilde{B}_{UV})$ and $\tilde{\boldsymbol{B}}^V = (\tilde{B}_{VU}, \tilde{B}_{VV})$ for this branching process in the obvious fashion (cf. the forward process offspring random variables $\boldsymbol{C}^U$, $\boldsymbol{C}^V$, $\tilde{\boldsymbol{C}}^U$ and $\tilde{\boldsymbol{C}}^V$ and the notation used in Sect. 3.2). We determine first the PGFs $f_{\tilde{\boldsymbol{B}}^U}(\boldsymbol{s})$ and $f_{\tilde{\boldsymbol{B}}^V}(\boldsymbol{s})$.

First note that, for $A \in \{U, V\}$, conditioning on the state (n, v) of a household yields, in obvious notation,

$$f_{\tilde{\boldsymbol{B}}^A}(\boldsymbol{s}) = \sum_{n=1}^{\infty} \sum_{v=0}^{n} p_{n,v}^{A} f_{\tilde{\boldsymbol{B}}_{n,v}^A}(\boldsymbol{s}).$$

Fix $A \in \{U, V\}$ and (n, v), and let $\boldsymbol{M}_A^{(n,v)} = (M_{AU}^{(n,v)}, M_{AV}^{(n,v)})$, where $M_{AU}^{(n,v)}$ and $M_{AV}^{(n,v)}$ are respectively the numbers of unvaccinated and vaccinated individuals, not including i^* itself, in the local susceptibility set of a typical type-A individual, i^* say, that resides in a household in state (n, v). Then $\tilde{\boldsymbol{B}}_{n,v}^A$ admits the decomposition

$$\tilde{\boldsymbol{B}}_{n,v}^A = \tilde{\boldsymbol{B}}_{n,v}^{A0} + \sum_{i=1}^{M_{AU}^{(n,v)}} \tilde{\boldsymbol{B}}_{n,v}^{AU}(i) + \sum_{j=1}^{M_{AV}^{(n,v)}} \tilde{\boldsymbol{B}}_{n,v}^{AV}(j), \tag{3.11}$$

where $\tilde{\boldsymbol{B}}_{n,v}^{A0}$, $\tilde{\boldsymbol{B}}_{n,v}^{AU}(i)$ and $\tilde{\boldsymbol{B}}_{n,v}^{AV}(j)$ are the contributions to $\tilde{\boldsymbol{B}}_{n,v}^A$ from the primary individual i^*, the ith secondary unvaccinated member and the jth secondary vaccinated member of i^*'s local susceptibility set, respectively. Let D_{i^*} denote the number of neighbours i^* has in the global network, so $D_{i^*} \sim \tilde{D}$, and recall that one of these global neighbours is used when i^* joins the susceptibility set process. Thus,

$$\tilde{\boldsymbol{B}}_{n,v}^{A0} = \sum_{k=1}^{D_{i^*}-1} \boldsymbol{\chi}_k^A,$$

where $\boldsymbol{\chi}_k^A = (1, 0)$ if the kth global neighbour of i^* is unvaccinated and joins the susceptibility set process, $\boldsymbol{\chi}_k^A = (0, 1)$ if this neighbour is vaccinated and joins the susceptibility set process, and $\boldsymbol{\chi}_k^A = (0, 0)$ otherwise. Note that, independently, each such global neighbour is vaccinated with probability p_V, and it joins the susceptibility set process with probability $p_G^{\mathrm{NR}}(U, A)$ if it is unvaccinated and with probability $p_G^{\mathrm{NR}}(V, A)$ if it is vaccinated. Thus

$$f_{\tilde{\boldsymbol{B}}_{n,v}^{A0}}(\boldsymbol{s}) = f_{\tilde{D}-1}(p^A(\boldsymbol{s})),$$

where

$$p^A(\boldsymbol{s}) = f_{\boldsymbol{\chi}_1^A}(\boldsymbol{s}) = 1 - (1 - p_V)p_G^{\mathrm{NR}}(U, A)(1 - s_U) - p_V p_G^{\mathrm{NR}}(V, A)(1 - s_V).$$

A similar argument shows that $f_{\tilde{\boldsymbol{B}}_{n,v}^{AU}(i)}(\boldsymbol{s}) = f_D(p^U(\boldsymbol{s}))$ and $f_{\tilde{\boldsymbol{B}}_{n,v}^{AV}(j)}(\boldsymbol{s}) = f_D(p^V(\boldsymbol{s}))$. Exploiting the mutual independence of all the random quantities in (3.11) except the components of $\boldsymbol{M}_A^{(n,v)}$ then yields

$$f_{\tilde{\boldsymbol{B}}_{n,v}^A}(\boldsymbol{s}) = f_{\tilde{D}-1}(p^A(\boldsymbol{s}))f_{\boldsymbol{M}_A^{(n,v)}}(f_D(p^U(\boldsymbol{s})), f_D(p^V(\boldsymbol{s}))).$$

Calculation of the mass function of $\boldsymbol{M}_A^{(n,v)}$ is described in Appendix B.5.

The distributions of $\boldsymbol{B}^U$ and $\boldsymbol{B}^V$ depend on how the initial individual for the susceptibility set process is chosen. For $A \in \{U, V\}$, if this initial individual is chosen uniformly at random from all type-A individuals in the population, then

$$f_{\boldsymbol{B}^A}(\boldsymbol{s}) = \sum_{n=1}^{\infty}\sum_{v=0}^{n} p_{n,v}^A f_{\boldsymbol{B}_{n,v}^A}(\boldsymbol{s}),$$

where

$$f_{\boldsymbol{B}_{n,v}^A}(\boldsymbol{s}) = f_D(p^A(\boldsymbol{s})) f_{\boldsymbol{M}_A^{(n,v)}}(f_D(p^U(\boldsymbol{s})), f_D(p^V(\boldsymbol{s}))).$$

Suppose that $\tilde{m}_{UV}\tilde{m}_{VU} \neq 0$ and $R_v > 1$. Let $\boldsymbol{\xi} = (\xi_U, \xi_V)$ be the unique solution in $[0, 1)^2$ of the equations

$$\xi_U = f_{\tilde{\boldsymbol{B}}^U}(\xi_U, \xi_V), \quad \xi_V = f_{\tilde{\boldsymbol{B}}^V}(\xi_U, \xi_V).$$

Then the proportions of unvaccinated and vaccinated individuals that are infected by a major epidemic are given by $z^U = 1 - f_{\boldsymbol{B}^U}(\xi_U, \xi_V)$ and $z^V = 1 - f_{\boldsymbol{B}^V}(\xi_U, \xi_V)$, respectively. Thus the overall proportion of individuals infected by a major epidemic is $z = p_V z^V + (1 - p_V)z^U$.

3.4 Optimal vaccination strategies

A main aim of a vaccination scheme is to reduce the threshold parameter R_* to below one, i.e. to make $R_v \leq 1$, and thus prevent a major outbreak occurring. The vaccine response may be such that $R_v > 1$ even if the entire population is vaccinated, in which case vaccination by itself is insufficient to be sure of preventing a major outbreak. However, if $R_* > 1$ and it is possible to make $R_v \leq 1$ then it is of interest to determine the allocation of vaccines that reduces R_v to 1 with the minimum vaccination coverage p_V.

Suppose that the population has a maximum household size $n_{\max} < \infty$. Then p_V is a linear function of x_{nv} ($n = 1, 2, \ldots, n_{\max}$; $v = 0, 1, \ldots, n$) (recall (3.1)), as is R_v when the vaccine action is all-or-nothing (recall (3.5)). Thus in this case determining the allocation of vaccines that (i) minimises p_V subject to $R_v \leq 1$ or (ii) minimises R_v subject to an upper bound on p_V are both linear programming problems. Moreover, as we outline below, the method of Ball and Lyne (2002, 2006) can be used to construct explicitly the solutions of these linear programming problems. The situation is in general more complicated if the vaccine action is non-random, since then R_v is the dominant eigenvalue of the matrix $\tilde{M}$, and the corresponding optimisation problems are non-linear. However, the problem is linear if $\text{rank}(\tilde{M}) = 1$, a sufficient condition for which is $\text{rank}(P_G^{\text{NR}}) = 1$, i.e. $p_G^{\text{NR}}(U, U)p_G^{\text{NR}}(V, V) = p_G^{\text{NR}}(U, V)p_G^{\text{NR}}(V, U)$. Note that $\text{rank}(P_G^{\text{NR}}) = 1$ if either $a = 1$ or $b = 1$, so a leaky vaccine response satisfies this constraint.

Consider the non-random vaccine response and suppose that $\text{rank}(\tilde{M}) = 1$. Then $R_v = \text{trace}(\tilde{M})$ and, recalling (3.10), it follows using (3.4), (3.8) and (3.9) that

$$R_v = \sum_{n=1}^{\infty}\sum_{v=0}^{n} \tilde{\rho}_n x_{nv} \mu_{n,v}^{\text{NR}}, \tag{3.12}$$

where

$$\mu_{n,v}^{\mathrm{NR}} = \left(1 - \frac{v}{n}\right)\left\{\mu_{\tilde{D}-1} p_G^{\mathrm{NR}}(U,U) + \mu_D\left[\mu^{(n,v)}(U,U)p_G^{\mathrm{NR}}(U,U) + \mu^{(n,v)}(U,V)p_G^{\mathrm{NR}}(V,U)\right]\right\}$$
$$+ \frac{v}{n}\left\{\mu_{\tilde{D}-1} p_G^{\mathrm{NR}}(V,V) + \mu_D\left[\mu^{(n,v)}(V,U)p_G^{\mathrm{NR}}(U,V) + \mu^{(n,v)}(V,V)p_G^{\mathrm{NR}}(V,V)\right]\right\}.$$

Observe that, when $\mathrm{rank}(\tilde{M}) = 1$, R_v takes the same form as for the all-or-nothing vaccine response; compare (3.12) and (3.5).

To characterise the optimal vaccination schemes in these cases, it is convenient to consider a finite population of m households, with maximum household size $n_{\max}$. Let $m_n = m\rho_n$ be the number of households of size n and let $h_{nv} = m_n x_{nv}$ be the number of households in state (n, v). Then $\tilde{\rho}_n = nm_n/N$, where N is the total population size, and, writing $\mu_{n,v}$ for $\mu_{n,v}^{\mathrm{AoN}}$ or $\mu_{n,v}^{\mathrm{NR}}$, as appropriate, (3.5) or (3.12) implies that

$$R_v = \sum_{n=1}^{n_{\max}} \sum_{v=0}^{n} h_{nv} M_{n,v}, \tag{3.13}$$

where $M_{n,v} = n\mu_{n,v}/N$, and (3.1) yields

$$p_V = \frac{1}{N} \sum_{n=1}^{n_{\max}} \sum_{v=0}^{n} v h_{nv}.$$

Observe that (3.13) implies that R_v is obtained by summing $M_{n,v}$ over all households in the population. For $n = 1, 2, \ldots, n_{\max}$ and $v = 0, 1, \ldots, n-1$, let $G_{n,v} = M_{n,v} - M_{n,v+1}$ be the reduction in R_v obtained by vaccinating one further individual in a household in state (n, v). If $G_{n,v}$ is decreasing in v for each fixed n (so successive vaccinations in the same household have diminishing returns), then it is straightforward to determine optimal vaccination schemes (Ball and Lyne 2002, 2006). One simply orders the states (n, v) according to decreasing $G_{n,v}$, and then uses this ordering to determine the order in which individuals in the population are vaccinated, stopping the process when either the vaccination coverage reaches the desired level or when $R_v \leq 1$, depending on the optimisation problem under consideration. (The 'worst' scheme is obtained by vaccinating whole households in increasing order of $M_{n,0} - M_{n,n}$.) If for some n, say $n = n'$, $G_{n',v}$ is not decreasing in v, then only those states, (n', v') say, on the lower edge of the convex hull of $\{(v, G_{n',v}) \colon v = 0, 1, \ldots, n'-1\}$ can be part of an optimal vaccination scheme. It is still possible to give explicit solutions of associated optimisation problems (cf. Ball et al. 2004), and of 'worst' schemes, but the details are more involved.

4 Acquaintance vaccination

4.1 Introduction

The acquaintance vaccine allocation model proposed by Cohen et al. (2003) and further analysed by Britton et al. (2007), both in the setting of a standard network model

(i.e. a population modelled by a configuration model, without household structure), is as follows. Each individual in the population is sampled independently a Poisson distributed number of times, with mean $\kappa > 0$, and each time an individual is sampled it chooses one of its neighbours uniformly at random, with replacement, and that neighbour is vaccinated. If a sampled individual has no neighbours then that sampling is ignored. Individuals are vaccinated at most once, even if they are chosen to be vaccinated more than once.

If the vaccine is perfect then the early stages and final outcome of an epidemic with this acquaintance vaccination model can be analysed relatively easily using branching process approximations. Essentially this is because the epidemic involves only unvaccinated individuals, and the degrees of the neighbours of an unvaccinated individual are mutually independent. However, if the vaccine is imperfect, then the epidemic may also involve vaccinated individuals and the degrees of the neighbours of a vaccinated individual are dependent. (A low-degree neighbour of a given individual, A say, is more likely to nominate A for vaccination than a high-degree neighbour but, if A is vaccinated, at least one of A's neighbours nominated A for vaccination.) It follows that the independence property required for a branching process approximation breaks down. This dependence can be overcome if individuals are also typed by their degree but, unless the support of D is small, the calculations become computationally prohibitively expensive. Indeed, infinite-type branching processes are required if the support of D is countably infinite. For these reasons, Ball and Sirl (2013) introduced an alternative acquaintance vaccine allocation model and analysed it, in the setting of a standard network model (without households). We now extend this analysis to the present network and households model.

4.2 Model and results

4.2.1 Acquaintance vaccination model

We assume that each individual is sampled independently with probability p_S and then each sampled individual nominates each of its global neighbours independently with probability p_N. All individuals that are nominated at least once are then vaccinated. Thus individuals are sampled only once and it is easily seen that the degrees of the neighbours of both vaccinated and unvaccinated individuals are mutually independent, thus facilitating branching process approximations which do not involve typing by degree. Note that the probability that an individual is not named by a given neighbour is $1 - p_S p_N$, so the probability it is vaccinated is

$$p_V = 1 - \sum_{d=0}^{\infty} p_d (1 - p_S p_N)^d = 1 - f_D(1 - p_S p_N), \tag{4.1}$$

which, of course, also gives the vaccination coverage.

We now describe the methods used to find threshold parameter, the major outbreak probability and the relative final size of a major outbreak. All details of the calculations are deferred to Appendix A.

Table 1 Numerical indices of the 6 types of household involved in non-initial generations of the branching process approximations

1	2	3	4	5	6
(N, S)	(V, S)	(U, S)	(N, S^c)	(V, S^c)	(U, S^c)

4.2.2 Early stages of the epidemic

We approximate the early stages of the epidemic, with vaccination, by a multitype (forward) branching process of infected households, in which households are typed by the type of their primary (globally contacted) case. Such primary cases are typed by (i) whether they are named (N), vaccinated (V) or unvaccinated (U) and (ii) whether or not they are sampled and thus might name their neighbours for vaccination (S and S^c). Here N means that a primary case was named by its global infector and therefore is vaccinated, V means that it is not named by its global infector but it is nevertheless vaccinated (i.e. it is named by another neighbour) and U means that it is unvaccinated (i.e. not named by any of its neighbours). Thus there are 6 types of infected households, which for notational convenience we give numerical indices as shown in Table 1. Secondary infected cases in a household, and also the primary case in the initially infected household, need only to be typed V or U, according to whether or not they are vaccinated. This forward branching process is in the same spirit as that described in Sect. 3.3 for household-based vaccination with a non-random vaccine action model, except to maintain the independence required for a branching process there are now 6 rather than 2 types (and an additional 2 types that appear only in the initial generation).

The offspring random variables for this forward branching process are as follows. For $i = 1, 2, \ldots, 6$, let $\tilde{\boldsymbol{C}}_i = (\tilde{C}_{i1}, \tilde{C}_{i2}, \ldots, \tilde{C}_{i6})$ denote the offspring random variable for a type-i non-initial individual in the forward branching process. Thus $\tilde{C}_{ij}$ is the number of type-j primary household cases emanating from a typical single household epidemic that is initiated by a single type-i primary case. Similarly, for $A \in \{U, V\}$, let $\boldsymbol{C}_A = (C_{A1}, C_{A2}, \ldots, C_{A6})$ denote the offspring random variable for a type-A individual in the initial generation of the forward branching process. In contrast to similar analyses of previous models in this paper, when analysing the all-or-nothing vaccine response, we consider actual, rather than potential, global infections. This facilitates simultaneous analysis of the two vaccine action models to a much larger extent than is otherwise possible, but has the consequence that the formulae for p_{maj} and z below are slightly different for the different models of vaccine action.

A threshold parameter is then given by R_v, the dominant eigenvalue of the matrix $\tilde{M} = [\tilde{m}_{ij}]$, where $\tilde{m}_{ij} = E[\tilde{C}_{ij}]$. We assume that $\tilde{M}$ is positively regular, i.e. that for some $n > 0$ the matrix $\tilde{M}^n$ has all of its entries strictly positive (e.g. Mode 1971, Sect. 1.6), which will be satisfied in all practical situations. If $R_v > 1$ then a major outbreak occurs with strictly positive probability; otherwise the probability of a major outbreak is zero.

The probability of a major outbreak can be written in terms of the joint PGFs $f_{\tilde{\boldsymbol{C}}_i}$ and $f_{\boldsymbol{C}_A}$ of the random variables $\tilde{\boldsymbol{C}}_i$ ($i = 1, 2, \ldots, 6$) and $\boldsymbol{C}_A$ ($A \in \{U, V\}$). Let

p^U_{maj} be the probability of a major outbreak if the initial infective is unvaccinated and p^V_{maj} be the corresponding probability when the initial infective is vaccinated. With an all-or-nothing vaccine this latter probability also assumes that the vaccination of the initial case is unsuccessful, since the branching process considers actual global infections. Conditioning on a single initial infective of type $A \in \{U, V\}$, we have $p^A_{\text{maj}} = 1 - f_{\boldsymbol{C}_A}(\boldsymbol{\sigma})$, where $\boldsymbol{\sigma} = (\sigma_i, i = 1, 2, \ldots, 6)$ is the unique solution of $f_{\tilde{\boldsymbol{C}}_i}(\boldsymbol{\sigma}) = \sigma_i$ $(i = 1, 2, \ldots, 6)$ in $[0, 1)^6$ (cf. (3.6)). The probability of a major outbreak with an initial infective chosen uniformly from the whole population is therefore

$$p_{\text{maj}} = \begin{cases} p_V p^V_{\text{maj}} + (1 - p_V) p^U_{\text{maj}}, & \text{non-random vaccine,} \\ p_V (1 - \varepsilon) p^V_{\text{maj}} + (1 - p_V) p^U_{\text{maj}}, & \text{all-or-nothing vaccine.} \end{cases}$$

(With an all-or-nothing vaccine, if the initial infective is vaccinated then a major outbreak can occur only if that vaccination is unsuccessful, which occurs with probability $1 - \varepsilon$.)

4.2.3 Final outcome of a major outbreak

We approximate the susceptibility set of a given individual, i^* say, by a households-based (backward) multitype branching process, where, as in Sect. 3.3, the type of a household is given by the type of the primary member of the corresponding local susceptibility set. As previously, the fraction of individuals that are ultimately infected by a major outbreak is given by the survival probability of this branching process. In the initial generation there are two types, V and U, depending on whether or not i^* is vaccinated. In subsequent generations, there are 6 types, numbered 1–6 as above, where now, for example, N means that the primary case was named by the individual that it contacts globally to join i^*'s susceptibility set. Similar to before, secondary members of a local susceptibility set need only to be typed V or U. Let $\boldsymbol{B}_U = (B_{U1}, B_{U2}, \ldots, B_{U6})$ and $\boldsymbol{B}_V = (B_{V1}, B_{V2}, \ldots, B_{V6})$ denote the offspring random variables for the initial generation of this branching process, and let $\tilde{\boldsymbol{B}}_i = (\tilde{B}_{i1}, \tilde{B}_{i2}, \ldots, \tilde{B}_{i6})$ $(i = 1, 2, \ldots, 6)$ be the offspring random variables for all subsequent generations.

Suppose that $R_v > 1$, so a major outbreak is possible. The asymptotic proportions of unvaccinated and vaccinated individuals that are infected by a major outbreak are given by $z^U = 1 - f_{\boldsymbol{B}_U}(\tilde{\boldsymbol{\xi}})$ and $z^V = 1 - f_{\boldsymbol{B}_V}(\tilde{\boldsymbol{\xi}})$, respectively, where $\tilde{\boldsymbol{\xi}} = (\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_6)$ is the unique solution in $[0, 1)^6$ of the equations $f_{\tilde{\boldsymbol{B}}_i}(\boldsymbol{s}) = s_i$ $(i = 1, 2, \ldots, 6)$. The overall proportion of the population infected by a major outbreak is therefore

$$z = \begin{cases} p_V z^V + (1 - p_V) z^U, & \text{non-random vaccine,} \\ p_V (1 - \varepsilon) z^V + (1 - p_V) z^U, & \text{all-or-nothing vaccine.} \end{cases}$$

5 Model exploration

In this section we explore the behaviour of the model and its dependence on some key parameters. The main focus is on numerical comparison of the various vaccine

allocation regimes for a variety of household size and network degree distributions. First, however, we comment briefly on the relationship between outcomes in simulated finite populations and our asymptotic analytical results, and also give some discussion of the choice of the parameters p_S and p_N in acquaintance vaccination.

5.1 Outcomes in finite populations

One can compare the analytical, asymptotic, quantities of interest (p_{maj} and z) to simulation-based estimates of the corresponding quantities on finite populations. We find [consistent with Ball et al. (2009, Sect. 5)] that the agreement becomes reasonable with the number of households m in the low hundreds and very close indeed for m over about 1000, though the convergence is a little slower for heavier tailed degree distributions.

5.2 Acquaintance vaccination dependence on p_S and p_N

The vaccine coverage $c = p_V = 1 - f_D(1 - p_S p_N)$ in our acquaintance vaccination model is determined by the product $p_S p_N$ [see Eq. (4.1)]. Here we investigate the dependence of the model outcomes (R_v, p_{maj} and z) on the precise values of p_N and p_S, for a fixed coverage.

In the case of a perfect vaccine some analytical progress is possible. It can be shown (see Appendix A.2.3 for details) that in this case R_v has the form

$$R_v = \eta(p')(1 - 2p' + p' p_N),$$

where the function η depends on p_S and p_N only through their product $p' = p_S p_N$. It is immediate that R_v is increasing in p_N for fixed p'. Thus we have that the best effect is achieved by taking $(p_S, p_N) = (1, p')$ and the worst with $(p_S, p_N) = (p', 1)$. The corresponding values of R_v are $R_v^b = \eta(p')(1 - p')^2$ and $R_v^w = \eta(p')(1 - p')$, so we have $R_v^b = (1 - p')R_v^w$. Precisely as in the network-only model without households, we note that this difference is generally quite small and is at its largest when p' is large, so there is high coverage and the epidemic is likely to be subcritical in any event.

When the vaccine action is not perfect we cannot make analytical progress in this direction, but extensive numerical calculations suggest that R_v, p_{maj} and z are usually monotonic in p_N when $p_S p_N$ is fixed and that in any case the difference in outcomes between the best and worst choices are very small. We do not explicitly demonstrate this here, but we note that (as might be expected) this is the same as Ball and Sirl (2013) found for the model without households. It can be seen to some extent in all of the plots in Sect. 5.3, where the two acquaintance vaccination plots are barely distinguishable in most cases.

In the numerical studies below we use the terms 'best' and 'worst' to refer to the acquaintance vaccination schemes with $(p_S, p_N) = (1, p')$ and $(p_S, p_N) = (p', 1)$ even though the names are not necessarily correct. As noted above, the precise choice of p_S and p_N seems to have only a very weak influence on the final outcome of the epidemic model and these two cases seem to give bounds for the quantities of interest.

5.3 Effect of different vaccine allocation regimes

We now explore the relative effectiveness of the different vaccine allocation regimes discussed in this paper, across a small but representative variety of household size and degree distributions. In the following figures we plot an outcome (R_v or z) against the vaccine coverage c for several vaccine allocation strategies; with all other parameters held fixed. Using different plots within the same figure we vary some of these other parameters. The vaccine allocation methods we consider are (i) uniformly chosen individuals (Ind UAR), (ii) uniformly chosen households (HH UAR), (iii & iv) best and worst household-based allocation (HH Best and HH Worst), (v & vi) 'best' and 'worst' acquaintance vaccination (Acq Best and Acq Worst); see the end of Sect. 3.1 for descriptions of the household-based vaccination schemes.

In this numerical section we use household size distributions $\rho_{\mathrm{UK}} = (31, 32, 16, 14, 5, 2)/100$ and $\rho_{\mathrm{Pak}} = (9, 45, 77, 118, 141, 146, 123, 104, 237)/1000$, which have respective means 2.4 and 6.8, as realistic household size distributions from the UK (Office for National Statistics 2011, Table 3.1) and Pakistan (National Institute of Population Studies 2013, Table 2.9), respectively. The degree distributions we use are the standard Poisson distribution and a power law distribution with exponential cutoff (see, for example, Amaral et al. 2000). For this latter distribution we use the notation $D \sim \mathrm{PowC}(a, \kappa)$ to denote that $p_k \propto k^{-a}\mathrm{e}^{-k/\kappa}$ $(k = 1, 2, \dots)$, i.e. a power law with index a and exponential cutoff at about κ. In particular, we use the degree distribution $D \sim \mathrm{PowC}(2, 120)$, which has mean $\mu_D \approx 3.001$ and $\sigma_D^2 \approx 66$. We also use the notation $\mathrm{Gam}(k, r)$ to denote a Gamma distribution with shape parameter k and scale parameter r (and thus mean kr and variance kr^2).

First we look at the relative effectiveness of the various allocation schemes, and how this changes with the household size distribution and the network degree distribution. Figure 1 shows plots of the post-vaccination threshold parameter R_v as a function of the vaccine coverage c for the 6 vaccine allocation schemes with a perfect vaccine. The different plots use different combinations of household size and network degree distribution, with all other parameters kept fixed (full details are in the figure caption). Firstly we note that the patterns of the household-based allocation schemes are consistent with those of Ball and Lyne (2006, Sect. 4) for the standard households model. We see that when the network degree distribution is not very variable (Poisson) good household-based schemes perform similarly to acquaintance vaccination. On the other hand, when the network degree distribution is much more variable (cut-off power law) we see that acquaintance vaccination significantly outperforms the household-based schemes, though to a slightly lesser extent when the household size distribution is also more variable.

Figures 2 and 3 are similar to Fig. 1, but we consider imperfect vaccine action models (with the same *efficacy* $1-E[AB] = 0.7$) and we plot both the post-vaccination threshold parameter R_v and expected final size of a large outbreak z. In Fig. 2 we use an all-or-nothing vaccine action model with success probability $\varepsilon = 0.7$ and consider two network degree distributions, with fixed (less variable) household size distribution. In Fig. 3 we use a non-random vaccine action model with relative susceptibility and infectivity $a = 0.5$ and $b = 0.6$, respectively, and vary the household size distibution, with a fixed (more variable) network degree distribution. Note that for this choice of a

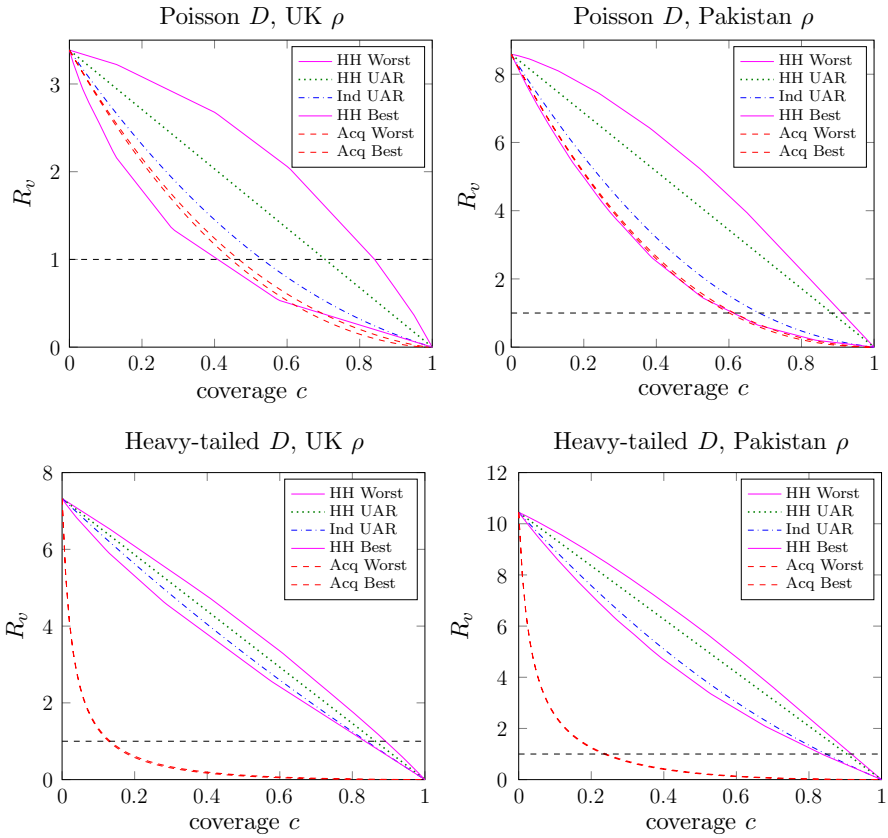

Fig. 1 Plots of R_v versus coverage (c) for various kinds of allocation of a perfect vaccine. We use two different household size distributions: *left plots* $\rho = \rho_{\text{UK}}$, *right plots* $\rho = \rho_{\text{Pak}}$; and two different degree distributions: *upper plots* $D \sim \text{Poi}(5)$, *lower plots* $D \sim \text{PowC}(2, 120)$. Other parameters are $I \sim \text{Gam}(5, 1/5)$, $\lambda_L = 1$, $\lambda_G = 0.3$

and b, determining the 'best' and 'worst' household-based vaccine allocation schemes are not linear programming problems (see Sect. 3.4). In our numerical routine we use the MATLAB constrained optimisation solver `fmincon`.

We see in the upper plots (of R_v) in Figs. 2 and 3 broadly similar patterns to those in Fig. 1. The lower plots in Fig. 2 show the ordering of z for the different allocation regimes being the same as the ordering of R_v. These two lower plots look qualitatively quite different, but it is important to note that in one the vaccine can and in the other the vaccine cannot bring the epidemic below threshold. These plots suggest that, as might be expected, when the network degree distribution is not so variable (e.g. the Poisson case) then vaccine allocation should be focussed on households-based methods, whilst when the network degree distribution is more variable (e.g. the cutoff power law) then targeting vaccination effort based on the network might give better results. Precisely which method is preferable will of course depend heavily on the other parameters of the model, but we have demonstrated that either allocation method, household-based or acquaintance-based, can be preferable to the other.

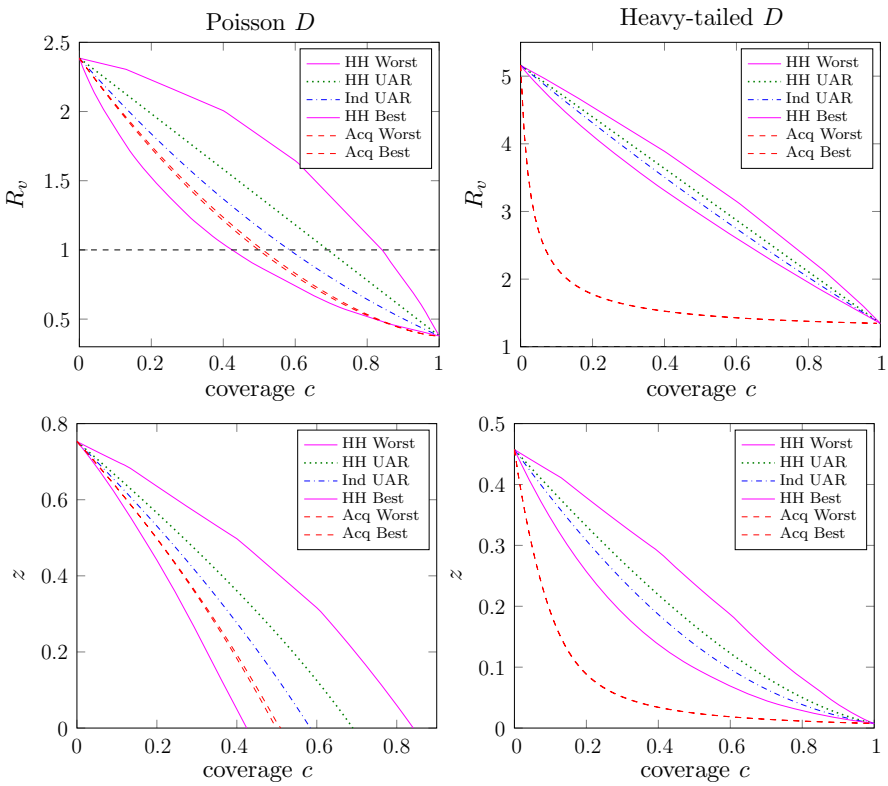

Fig. 2 Plots of R_v and z versus coverage (c) for various kinds of allocation of an all-or-nothing vaccine with success probability $\varepsilon = 0.7$. We use two different degree distributions: *left plots* $D \sim \text{Poi}(5)$, *right plots* $D \sim \text{PowC}(2, 120)$. Other parameters are $\rho = \rho_{\text{UK}}$, $I \sim \text{Gam}(5, 1/5)$, $\lambda_L = 1$, $\lambda_G = 0.2$

The lower plots of Fig. 3 follow similar patterns in that we are considering the case of a quite variable network degree distribution so acquaintance vaccination outperforms the household-based methods. There are however some unexpected patterns in the lower-right plot of Fig. 3, in that the ordering of the various household-based allocation regimes are not the same as in the corresponding plot of R_v immediately above it. In particular, for lower coverages the 'worst' households based allocation outperforms, in terms of z, the 'best'! This demonstrates that when the epidemic is well above threshold, optimising vaccine allocation based on R_* does not necessarily result in an expected final size that is as low as possible; cf. Keeling and Ross (2015), who observe a similar phenomenon in the standard households model with a perfect vaccine. The threshold parameter R_v measures household-to-household transmission, but does not directly take into account the size of the within-household outbreaks. This could perhaps be resolved by optimising an individual-based threshold parameter instead (see Pellis et al. 2012; Ball et al. 2016), but the optimisation problem would be more difficult than the one we have considered. We also note that this phenomenon appears to arise only when the epidemic is well above criticality.

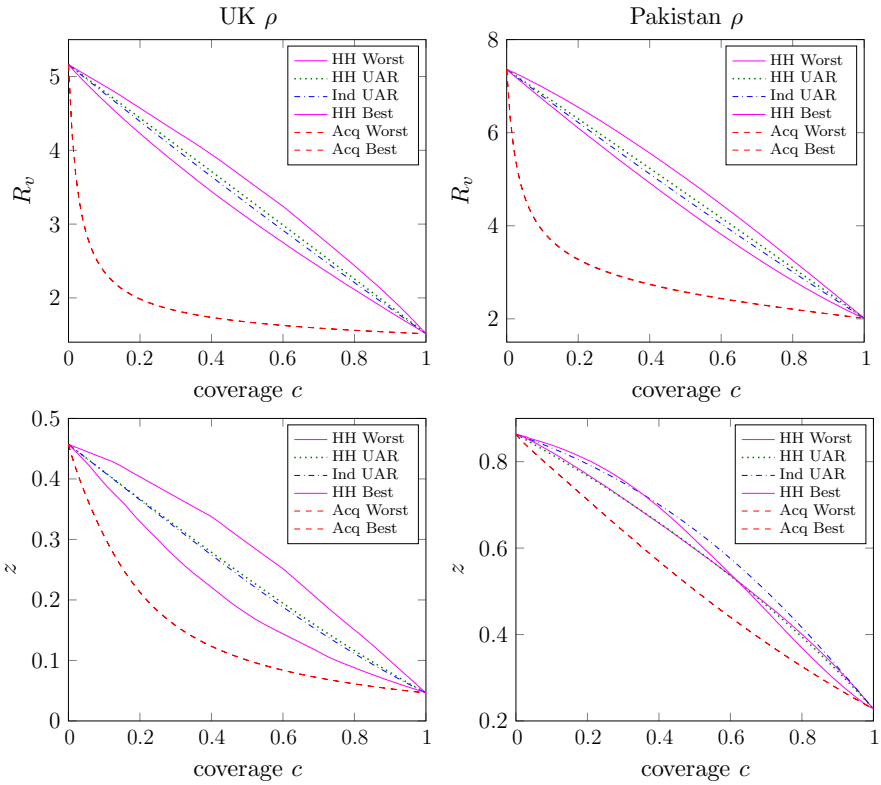

Fig. 3 Plots of R_v and z versus coverage (c) for various kinds of allocation of a non-random vaccine with relative susceptibility and infectivity $a = 0.5$ and $b = 0.6$ respectively. We use two different household size distributions: *left plots* $\rho = \rho_{\text{UK}}$, *right plots* $\rho = \rho_{\text{Pak}}$. Other parameters are $D \sim \text{PowC}(2, 120)$, $I \sim \text{Gam}(5, 1/5)$, $\lambda_L = 1$, $\lambda_G = 0.2$

Lastly we note that the behaviour of p_{maj} is broadly similar to that of z. We do not present any plots, but they have shapes and patterns similar to those in the z plots that are shown, including the unexpected ordering observed in the lower-right plot of Fig. 3.

6 Concluding remarks

In this paper we have analysed vaccine allocation strategies in stochastic SIR epidemic models upon populations with household and random network (of configuration model type) structure. By exploiting branching process approximations we derive asymptotic results describing the threshold behaviour of this epidemic model when there are few initial infectives and the final outcome in the event of a major outbreak. Particular attention has been paid to the analysis of acquaintance vaccination, which aims to target vaccination at individuals who are highly connected in the network. We find that acquaintance vaccination potentially offers substantial benefits over other (households-based) vaccine allocation regimes.

Whilst we have shown that acquaintance vaccination is potentially useful, it is clearly not feasible to implement in practice in human populations, so investigation of a more ethically acceptable allocation regime that preserves the targeting of well-connected individuals is a clear direction for future work. It is also likely that the effectiveness of acquaintance vaccination will depend on the amount of clustering in the network in an interesting way; we have touched on this through the use of different household size distributions, but clearly there is scope for considerably more work in this direction. Further issues that warrant more investigation include (i) comparison to the optimal network-based vaccine allocation, assuming knowledge of the degree of every individual in the network (Ball and Sirl 2012, Sect. 5 and Appendix B), (ii) the determination of optimal or near-optimal vaccination strategies based on household and network information and (iii) extending acquaintance vaccination to include the possiblity of naming individuals who are in the same household.

A particularly striking feature of the numerical study is the fact that, when vaccine coverage is insufficient to prevent a major outbreak, the ordering of the performance of the household-based allocation strategies can be different depending on whether R_v (the post-vaccination R_*) or z is used as the measure of performance. We emphasise that this does not imply that minimising R_v is not worthwhile. When using household-based vaccine allocation with a limited quantity of vaccine (i.e. a maximum vaccination coverage, c say), a natural approach is first to minimise R_v among all strategies having coverage c. If the minimised value of R_v is no greater than 1 then major outbreaks can be prevented and R_v can be used to determine the vaccine allocation that prevents a major outbreak with minimum vaccination coverage. If the minimised value of R_v is not less than 1 then households-based vaccination cannot prevent major outbreaks with the available quantity of vaccine. It then makes sense to minimise the expected cost of an outbreak; measured for example by p_{maj}, z or some combination of these such as their product, which gives the probability that a typical individual is infected by the epidemic. The numerical examples in Sect. 5.3 suggest that in these circumstances minimising R_v (which is a much simpler optimisation problem than, for example, minimising z) still produces a near-optimal vaccine allocation unless, post-vaccination, the epidemic is still well above threshold. However, this topic is not straightforward and clearly warrants further research. It will be investigated in more detail in a subsequent paper.

We also note in closing that there are some numerical challenges involved in implementing the methods we have presented. These are particularly relevant for the calculations relating to the forward process (i.e. calculation of p_{maj}), but do also apply to the backward process (calculation of z). The main issue that arises is that of slowly converging infinite sums and the resulting possibilities for numerical overflow and underflow. Writing doubly or triply-infinite sums as one infinite and one or two finite sums (for example $\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{i,j} = \sum_{k=0}^{\infty}\sum_{l=0}^{k} a_{k-l,l}$) helps in some regards (e.g. faster computing since there is only one truncation to have to control the error of) but hinders in others (e.g. slower computing since methods to avoid underflow and overflow errors become more involved).

Acknowledgements This work was supported by the Engineering and Physical Sciences Research Council [grant numbers EP/K032208/1, EP/E038670/1]. This work was partially supported by a grant from the

Simons Foundation. The authors thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Theoretical Foundations for Statistical Network Analysis, where work on this paper was undertaken. The authors also thank the Associate Editor and reviewers for their constructive comments which have significantly improved the presentation of the paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Acquaintance vaccination branching process offspring distributions

In this appendix we consider the detail of calculating the offspring PGFs and means of the approximating branching processes for the acquaintance vaccination model described in Sect. 4.

In our analysis of the approximating branching processes (both forward and backward) it is convenient to think of the process of an individual in a household producing offspring (i.e. infecting network neighbours and thus contributing to the offspring of its household) in two phases. In the forward process the first phase involves consideration of the individual's network neighbours, excluding its infector if it is a primary case. We consider how many such neighbours there are and what type they (and therefore their household) would be if they were to be infected. We call these individuals *next-generation neighbours*. In the second phase we consider which of these next-generation neighbours are actually infected and thus become primary cases in the offspring households. The above description of the first phase applies *mutatis mutandis* to the backward process; here the second phase is different in that the possible infectious contacts originate from outside rather than within the household under consideration. Since the first phase is the same in both the forward and backward branching processes we first consider it separately.

In Sect. A.1 we study the next-generation neighbour distributions. In Sect. A.2 we determine the mean offspring matrix for the forward process, and hence the post-vaccination threshold parameter R_v; in Sect. A.3 we consider the offspring distribution PGFs for the forward process, and hence the probability of a major outbreak; and in Sect. A.4 we determine the offspring PGFs for the backward process, and hence the relative final size of a major outbreak. The offspring distribution PGFs for the forward process are generally more complicated and some further detail of their calculation is described in Appendix B.4.

A.1 Degree and next-generation neighbour distributions

As outlined above, the first phase of individuals producing offspring (which this subsection describes) applies equally to the forward and backward processes. To simplify the exposition we use only the language of the forward process of epidemic spread. We also use 'household' and 'individual' in the population sense (as opposed to the

branching process sense where an individual corresponds to a household in the epidemic model).

Recall the typing described in Sect. 4.2, in which infectious households are characterised by the type $(1, 2, \ldots, 6)$ of their primary case (see Table 1). Secondary cases in a household and the primary case in the initially infected household are typed (V or U) according to whether they are vaccinated or unvaccinated. Recall also that next-generation neighbours of an infected individual are the individuals (in neighbouring households) that it may infect. In this subsection we first calculate the degree distribution of the various secondary and primary cases conditional on their type. Then we consider the next-generation neighbour distributions: the number of network neighbours, excluding its infector, of each secondary type that a typical type-i individual has (for $i \in \{1, 2, \ldots, 6, U, V\}$).

A.1.1 Degree distributions

Let D_V and D_U respectively denote the degree of a typical vaccinated and unvaccinated secondary case. Then, recalling from Eq. (4.1) that $p_V = 1 - f_D(1 - p_S p_N)$,

$$P(D_U = d) = \frac{P(D = d)P(U|D = d)}{P(U)} = \frac{p_d(1 - p_S p_N)^d}{1 - p_V} \quad (d = 0, 1, \ldots) \tag{A.1}$$

and, similarly,

$$P(D_V = d) = \frac{p_d(1 - (1 - p_S p_N)^d)}{p_V} \quad (d = 1, 2, \ldots).$$

These follow because an unvaccinated individual has avoided being named by any of its neighbours and a vaccinated individual has not done so.

Now consider an individual of one of the primary types $1, 2, \ldots, 6$. Note that whether or not it is sampled is independent of its degree. Thus primary cases of types 1 and 4 have the same degree distribution, as do primary cases of types 2 and 5, and primary cases of types 3 and 6. Let $\tilde{D}_N$, $\tilde{D}_V$ and $\tilde{D}_U$ denote generic random variables having these respective distributions. First note that $\tilde{D}_N \stackrel{D}{=} \tilde{D}$, where $\stackrel{D}{=}$ denotes equality in distribution. Second, consider a typical unvaccinated primary case. It has unconditional degree distribution $\tilde{D}$ but, in addition to not being named by its infector, we also know that it is not named by any of its other global neighbours. Thus

$$P(\tilde{D}_U = d) = \frac{P(\tilde{D} = d)P(U|\tilde{D} = d)}{1 - \tilde{p}_V} = \frac{\tilde{p}_d(1 - p_N p_S)^{d-1}}{1 - \tilde{p}_V} \quad (d = 1, 2, \ldots), \tag{A.2}$$

where

$$\tilde{p}_V = \sum_{d=1}^{\infty} \tilde{p}_d(1 - (1 - p_N p_S)^{d-1}) = 1 - f_{\tilde{D}-1}(1 - p_N p_S) \tag{A.3}$$

is the probability that a typical unnamed neighbour of an infector is vaccinated. Similarly,

$$P(\tilde{D}_V = d) = \frac{\tilde{p}_d(1-(1-p_N p_S)^{d-1})}{\tilde{p}_V} \quad (d = 2, 3, \ldots). \tag{A.4}$$

A.1.2 Next-generation neighbour distributions

For $i = 1, 2, \ldots, 6$ we write $\tilde{\boldsymbol{X}}_i = (\tilde{X}_{i1}, \tilde{X}_{i2}, \ldots, \tilde{X}_{i6})$ for the number of next-generation neighbours of a type-i primary case; i.e. $\tilde{X}_{ij}$ is the number of neighbours of a type-i individual who will become type-j individuals if they are infected by the type-i primary case. For $A = U, V$ we similarly define $\boldsymbol{X}_A = (X_{A1}, X_{A2}, \ldots, X_{A6})$ to be the corresponding quantities for a secondary case or the primary case in the initial household. To calculate the threshold parameter we need expressions for the means $\hat{\mu}_{ij}$ and $\hat{\mu}_{Aj}$ of the elements of these random vectors and to calculate the offspring distribution PGFs we require the joint PGFs $f_{\tilde{\boldsymbol{X}}_i}$ and $f_{\boldsymbol{X}_A}$.

Means The mean $E[\tilde{X}_{ij}] = \hat{\mu}_{ij}$ can be written as $\hat{\mu}_{ij} = \tilde{\mu}_i^G \hat{p}_{ij}$, where $\tilde{\mu}_i^G$ is the mean number of network neighbours a typical type-i primary case has in addition to its infector and $\hat{p}_{ij}$ is the probability that a given such global neighbour is of type j (or, more accurately, the probability it will be of type j if it is infected). Similarly we write $E[X_{Aj}] = \hat{\mu}_{Aj} = \mu_A^G \hat{p}_{Aj}$ for the secondary cases and primary case in the initial household.

Consider a type-1 (i.e. (N, S)) primary case. Its global degree is distributed as $\tilde{D}_N \stackrel{D}{=} \tilde{D}$, so $\tilde{\mu}_1^G = \mu_{\tilde{D}-1}$. A given neighbour is named with probability p_N, and, if that neighbour is unnamed, it is vaccinated with probability $\tilde{p}_V$ and otherwise unvaccinated. Further, independently, that neighbour is sampled with probability p_S. So, $\hat{p}_{11} = p_N p_S$, $\hat{p}_{12} = (1-p_N)\tilde{p}_V p_S$, $\hat{p}_{13} = (1-p_N)(1-\tilde{p}_V)p_S$, $\hat{p}_{14} = p_N(1-p_S)$, $\hat{p}_{15} = (1-p_N)\tilde{p}_V(1-p_S)$ and $\hat{p}_{16} = (1-p_N)(1-\tilde{p}_V)(1-p_S)$. The situation is similar for a type-4 (i.e. (N, S^c)) individual, except a type-4 individual is not sampled and hence cannot name its global neighbours. Hence $\tilde{\mu}_4^G = \mu_{\tilde{D}-1}$, $\hat{p}_{41} = \hat{p}_{44} = 0$, $\hat{p}_{42} = \tilde{p}_V p_S$, $\hat{p}_{43} = (1-\tilde{p}_V)p_S$, $\hat{p}_{45} = \tilde{p}_V(1-p_S)$ and $\hat{p}_{46} = (1-\tilde{p}_V)(1-p_S)$.

Next consider a type-2 (i.e. (V, S)) primary case, i^* say. Its global degree is distributed as $\tilde{D}_V$, so $\tilde{\mu}_2^G = \mu_{\tilde{D}_V-1}$ and, using (A.4),

$$\mu_{\tilde{D}_V-1} = \frac{\mu_{\tilde{D}-1} - (1-p_N p_S) f'_{\tilde{D}-1}(1-p_N p_S)}{\tilde{p}_V}.$$

Let $\tilde{N}_S$ be the number of next-generation neighbours of i^* that are sampled. Then $\tilde{N}_S \geq 1$, since i^* is vaccinated but not named by its global infector, and, for $k = 1, 2, \ldots, d-1$,

$$P(\tilde{N}_S = k \mid \tilde{D}_V = d) = \frac{\binom{d-1}{k} p_S^k (1-p_S)^{d-1-k}(1-(1-p_N)^k)}{1-(1-p_N p_S)^{d-1}}, \tag{A.5}$$

whence, using (A.4),

$$\mu_{\tilde{N}_S} = \frac{p_S\left(\mu_{\tilde{D}-1} - (1-p_N)f'_{\tilde{D}-1}(1-p_N p_S)\right)}{\tilde{p}_V}. \tag{A.6}$$

It follows that the probability that a given next-generation neighbour of i^* is sampled is $p_S^{\tilde{V}} = \mu_{\tilde{N}_S}/\tilde{\mu}_2^G$ and that, for $j = 1, 2, \ldots, 6$, $\hat{p}_{2j}$ is given by $\hat{p}_{1j}$ with p_S replaced by $p_S^{\tilde{V}}$. Similar arguments for a type-5 (i.e. (U, S^c)) primary case establish that $\tilde{\mu}_5^G = \mu_{\tilde{D}_V-1}$ and, for $j = 1, 2, \ldots, 6$, $\hat{p}_{5j}$ is given by $\hat{p}_{4j}$ with p_S replaced by $p_S^{\tilde{V}}$.

Now consider a type-3 (i.e. (U, S)) primary case, j^* say. Its global degree is distributed as $\tilde{D}_U$, so $\tilde{\mu}_3^G = \mu_{\tilde{D}_U-1}$ and, using (A.2),

$$\mu_{\tilde{D}_U-1} = \frac{(1-p_N p_S)f'_{\tilde{D}-1}(1-p_N p_S)}{1-\tilde{p}_V}. \tag{A.7}$$

Since j^* is not vaccinated, none of its next-generation neighbours name j^*, so the probability that a given such neighbour, k^* say, is sampled is given by

$$p_S^U = P(k^* \text{ sampled} \mid k^* \text{ does not name } j^*) = \frac{p_S(1-p_N)}{1-p_S p_N}.$$

It follows that, for $j = 1, 2, \ldots, 6$, $\hat{p}_{3j}$ is given by $\hat{p}_{1j}$ with p_S replaced by p_S^U. Similarly, $\tilde{\mu}_6^G = \mu_{\tilde{D}_U-1}$ and, for $j = 1, 2, \ldots, 6$, $\hat{p}_{6j}$ is given by $\hat{p}_{4j}$ with p_S replaced by p_S^U.

Now consider an individual, i^* say, of type U (i.e. chosen uniformly from the population and unvaccinated). Its global degree is distributed as D_U so, using (A.1), its expected number of next-generation neighbours is

$$\mu_U^G = \mu_{D_U} = \frac{(1-p_S p_N)f'_D(1-p_S p_N)}{1-p_V}. \tag{A.8}$$

Further, i^* is sampled with probability p_S, and the type of a given neighbour of i^* is distributed according to $\hat{p}_{3j}$ if i^* is sampled and according to $\hat{p}_{6j}$ if i^* is not sampled. Thus $\hat{p}_{Uj} = p_S\hat{p}_{3j} + (1-p_S)\hat{p}_{6j}$.

Finally, consider an individual, j^* say, of type V (i.e. chosen uniformly from the population and vaccinated). Let N_S and N_{S^c} denote the number of next-generation neighbours of j^* that are sampled and not sampled, respectively. Then arguing as in the derivation of (A.6) shows that

$$\mu_{N_S} = \frac{p_S\left(\mu_D - (1-p_N)f'_D(1-p_S p_N)\right)}{p_V}$$

and

$$\mu_{N_{S^c}} = \frac{(1-p_S)\left(\mu_D - f'_D(1-p_S p_N)\right)}{p_V}.$$

Further, j^* is sampled with probability p_S. If j^* is sampled then a given neighbour is named, vaccinated or unvaccinated with respective probabilities p_N, $(1-p_N)\tilde{p}_V$ and $(1-p_N)(1-\tilde{p}_V)$; whilst if j^* is not sampled these probabilities are 0, $\tilde{p}_V$ and $1-\tilde{p}_V$. We therefore have $\hat{p}_{Vj} = p_S\hat{p}_{1j} + (1-p_S)\hat{p}_{4j}$.

Joint PGFs We determine next the PGFs $f_{\tilde{\boldsymbol{X}}_i}$ $(i = 1, 2, \ldots, 6)$. For $i = 1, 2, \ldots, 6$, let $\tilde{D}(i)$ denote the degree of a typical type-i primary individual, who thus has $\tilde{D}(i)-1$ next-generation neighbours. We therefore have $\tilde{D}(i)-1 \stackrel{D}{=} \sum_{j=1}^{6} \tilde{X}_{ij}$. For $i = 1, 3, 4$ and 6, the types of these $\tilde{D}(i)-1$ next-generation neighbours are chosen independently, according to $(\hat{p}_{ij},\ j = 1, 2, \ldots, 6)$, so

$$f_{\tilde{\boldsymbol{X}}_i}(\boldsymbol{s}) = f_{\tilde{D}(i)-1}(\hat{g}_i(\boldsymbol{s})) \quad (i = 1, 3, 4, 6),$$

where $\hat{g}_i(\boldsymbol{s}) = \sum_{j=1}^{6} \hat{p}_{ij}s_j$. Further $\tilde{D}(1) \stackrel{D}{=} \tilde{D}(4) \sim \tilde{D}$ and $\tilde{D}(3) \stackrel{D}{=} \tilde{D}(6) \sim \tilde{D}_U$ [see (A.2)].

Now consider a typical type-2 primary individual. The situation is more complex here (and for type-5 individuals) because we know that at least one of its neighbours named it for vaccination and thus must be sampled. Let $\tilde{N}_S$ and $\tilde{N}_{S^c}$ be the number of its $\tilde{D}(2)-1$ next-generation neighbours that are sampled and unsampled, respectively, so $\tilde{N}_S + \tilde{N}_{S^c} = \tilde{D}(2)-1$. Each of these $\tilde{D}(2)-1$ neighbours is named independently with probability p_N and any neighbour not so named is vaccinated with probability $\tilde{p}_V$. Thus

$$f_{\tilde{\boldsymbol{X}}_2}(\boldsymbol{s}) = E\left[(\hat{g}_{21}(\boldsymbol{s}))^{\tilde{N}_S}(\hat{g}_{22}(\boldsymbol{s}))^{\tilde{N}_{S^c}}\right],$$

where $\hat{g}_{21}(\boldsymbol{s}) = p_N s_1 + (1-p_N)\tilde{p}_V s_2 + (1-p_N)(1-\tilde{p}_V)s_3$ and $\hat{g}_{22}(\boldsymbol{s}) = p_N s_4 + (1-p_N)\tilde{p}_V s_5 + (1-p_N)(1-\tilde{p}_V)s_6$. Furthermore, $\tilde{D}(2) \sim \tilde{D}_V$, whence using (A.4) and (A.5) we obtain

$$f_{\tilde{\boldsymbol{X}}_2}(\boldsymbol{s}) = \tilde{p}_V^{-1}\left[f_{\tilde{D}-1}\left(\hat{g}_2(\boldsymbol{s},0)\right) - f_{\tilde{D}-1}\left(\hat{g}_2(\boldsymbol{s},p_N)\right)\right], \tag{A.9}$$

where $\hat{g}_2(\boldsymbol{s},x) = p_S(1-x)\hat{g}_{21}(\boldsymbol{s}) + (1-p_S)\hat{g}_{22}(\boldsymbol{s})$. A similar argument shows that

$$f_{\tilde{\boldsymbol{X}}_5}(\boldsymbol{s}) = \tilde{p}_V^{-1}\left[f_{\tilde{D}-1}\left(\hat{g}_5(\boldsymbol{s},0)\right) - f_{\tilde{D}-1}\left(\hat{g}_5(\boldsymbol{s},p_N)\right)\right], \tag{A.10}$$

where $\hat{g}_5(\boldsymbol{s},x) = p_S(1-x)\hat{g}_{51}(\boldsymbol{s}) + (1-p_S)\hat{g}_{52}(\boldsymbol{s})$, with $\hat{g}_{51}(\boldsymbol{s}) = \tilde{p}_V s_2 + (1-\tilde{p}_V)s_3$ and $\hat{g}_{52}(\boldsymbol{s}) = \tilde{p}_V s_5 + (1-\tilde{p}_V)s_6$.

A typical type-U secondary individual, j_U^* say, has degree distributed according to D_U [recall (A.1)], so its number of next-generation neighbours is also distributed as

D_U. Further, j_U^* is sampled with probability p_S, in which case, apart from its degree, j_U^* behaves similarly to a type-3 primary individual, otherwise j_U^* is not sampled and behaves similarly to a type-6 primary individual. Thus,

$$f_{\boldsymbol{X}_U}(\boldsymbol{s}) = p_S f_{D_U}(\hat{g}_3(\boldsymbol{s})) + (1 - p_S) f_{D_U}(\hat{g}_6(\boldsymbol{s})).$$

A typical type-V secondary individual, j_V^* say, is also sampled with probability p_S. A similar argument to the derivation of (A.9) and (A.10), using appropriate modifications as j_V^* has degree distributed according to D_V, yields

$$\begin{aligned} f_{\boldsymbol{X}_V}(\boldsymbol{s}) = p_V^{-1} \big\{ & p_S \left[f_D\left(\hat{g}_2(\boldsymbol{s}, 0)\right) - f_D\left(\hat{g}_2(\boldsymbol{s}, p_N)\right) \right] \\ & + (1 - p_S) \left[f_D\left(\hat{g}_5(\boldsymbol{s}, 0)\right) - f_D\left(\hat{g}_5(\boldsymbol{s}, p_N)\right) \right] \big\}. \end{aligned}$$

A.2 Threshold parameter

As far as is possible, we treat the all-or-nothing and non-random models simultaneously. As explained in the second paragraph of Sect. 4.2.2, in contrast to Sect. 3.2, when analysing the all-or-nothing vaccine response we consider actual, rather than potential, global infections. The forward branching process for both models is then similar to that used for the non-random model in Sect. 3.3, except now it is a 6-type (rather than 2-type) branching process. As always, the offspring distributions are different in the initial generation from subsequent generations, but only those for the latter are required to determine the threshold parameter. Recall that, for $i = 1, 2, \ldots, 6$, $\tilde{\boldsymbol{C}}_i = (\tilde{C}_{i1}, \tilde{C}_{i2}, \ldots, \tilde{C}_{i6})$ denotes the offspring random variable for a type-i non-initial individual in the forward branching process. Thus $\tilde{C}_{ij}$ is the number of type-j primary household cases emanating from a typical single-household epidemic that is initiated by a single type-i primary case. Since the post-vaccination threshold parameter R_v is the dominant eigenvalue of $\tilde{M} = [\tilde{m}_{ij}]$, our goal is to determine these matrix entries $\tilde{m}_{ij} = E[\tilde{C}_{ij}]$.

To calculate $\tilde{m}_{ij}$ it is convenient to decompose it into $\tilde{m}_{ij} = \tilde{m}_{ij}^P + \tilde{m}_{ij}^S$, where $\tilde{m}_{ij}^P$ is the mean number of global infections of type-j individuals emanating from the primary type-i case and $\tilde{m}_{ij}^S$ is the mean total number of global infections of type-j individuals emanating from all secondary cases in the single-household epidemic.

A.2.1 Infections emamating from primary cases

Define the matrix P_G of marginal global transmission probabilities, so $P_G = P_G^{\mathrm{NR}}$ [recall (3.7)] if the vaccine action is non-random and $P_G = P_G^{\mathrm{AoN}}$ if the vaccine action is all-or-nothing, where

$$P_G^{\mathrm{AoN}} = \begin{bmatrix} p_G^{\mathrm{AoN}}(U, U) & p_G^{\mathrm{AoN}}(U, V) \\ p_G^{\mathrm{AoN}}(V, U) & p_G^{\mathrm{AoN}}(V, V) \end{bmatrix} = (1 - \phi_I(\lambda_G)) \begin{bmatrix} 1 & 1 - \varepsilon \\ 1 & 1 - \varepsilon \end{bmatrix}.$$

The marginal global transmission probabilities, for the types $i, j = 1, 2, \ldots, 6$, are then given by $p_{ij}^G = p_G(\alpha(i), \alpha(j))$, where $\alpha(k) = U$ if $k = 3, 6$ and $\alpha(k) = V$ if $k = 1, 2, 4, 5$. Since each type-j next-generation neighbour of a type-i infective is infected with probability p_{ij}^G it is immediate that $\tilde{m}_{ij}^P = \hat{\mu}_{ij} p_{ij}^G$, where $\hat{\mu}_{ij}$ is the mean number of next-generation neighbours calculated in Sect. A.1.2.

A.2.2 Infections emamating from secondary cases

We first note that, for $j = 1, 2, \ldots, 6$, $\tilde{m}_{1j}^S = \tilde{m}_{2j}^S = \tilde{m}_{4j}^S = \tilde{m}_{5j}^S$ ($= \tilde{m}_{Vj}^S$ say) and $\tilde{m}_{3j}^S = \tilde{m}_{6j}^S$ ($= \tilde{m}_{Uj}^S$ say). Further, for $A \in \{U, V\}$ and $j = 1, 2, \ldots, 6$,

$$\tilde{m}_{Aj}^S = \mu_S(A, U)\hat{m}_{Uj} + \mu_S(A, V)\hat{m}_{Vj},$$

where, for $A, A' \in \{U, V\}$, $\mu_S(A, A')$ is the mean number of type-A' secondary cases in a typical single-household epidemic initiated by a primary case of type A and, for $A \in \{U, V\}$, $\hat{m}_{Aj}$ is the mean number of type-j primary cases emanating from a typical type-A secondary case.

Now observe that $\hat{m}_{Aj} = \hat{\mu}_{Aj} p_G(A, \alpha(j))$, where $\alpha(j)$ is as in Sect. A.2.1. Conditioning on the size of a typical globally contacted household, we obtain in an obvious notation that

$$\mu_S(A, A') = \sum_{n=1}^{\infty} \tilde{\rho}_n \mu_S^{(n)}(A, A') \quad (A, A' \in \{U, V\}).$$

Note that in the limit as the number of households $m \to \infty$, secondary individuals in a typical household have no common global neighbour with probability one, so they are vaccinated independently, each with probability p_V. Thus, again in an obvious notation,

$$\mu_S^{(n)}(A, A') = \sum_{k=0}^{n-1} \binom{n-1}{k} p_V^k (1 - p_V)^{n-1-k} \mu_S^{(n,k)}(A, A').$$

If the vaccine action is non-random then, recalling the notation in Sect. 3.3, $\mu_S^{(n,k)}(A, A') = \mu^{(n,k+\delta_{A,V})}(A, A')$, where $\delta_{A,V} = 1$ if $A = V$ and 0 otherwise. If the vaccine action is all-or-nothing, we condition on the number of successfully vaccinated secondary individuals in the household to obtain, for $A \in \{U, V\}$,

$$\mu_S^{(n,k)}(A, V) = \sum_{l=0}^{\min(k,n-2)} \binom{k}{l} \varepsilon^l (1 - \varepsilon)^{k-l} \mu^{(n-l)}(\lambda_L) \left(\frac{k-l}{n-l-1} \right)$$

and

$$\mu_S^{(n,k)}(A, U) = \sum_{l=0}^{\min(k,n-2)} \binom{k}{l} \varepsilon^l (1 - \varepsilon)^{k-l} \mu^{(n-l)}(\lambda_L) \left(\frac{n-k-1}{n-l-1} \right).$$

(Note that $\mu^{(n-l)}(\lambda_L) = 0$ when $l = n - 1$.)

A.2.3 Perfect vaccine

In the situation where the vaccine is perfect, we have only households of types 3 and 6 (i.e. (U, S) and (U, S^c)); and since these two types are sampled independently with the same probability, the forward process reduces to single-type. The mean number of next-generation neighbours of an infected individual is $\mu_{\tilde{D}_U-1} + \mu_S(U, U)\mu_{D_U}$. Each infected individual is sampled with probability p_S^U, so fails to name each next-generation neighbour with probability $p_S^U(1-p_N)+1-p_S^U$. Then each such unnamed individual avoids vaccination (by other neighbours in the network) with probability $1 - \tilde{p}_V$. Thus

$$R_v = (\mu_{\tilde{D}_U-1} + \mu_{D_U}\mu_S(U, U))(p_S^U(1 - p_N) + 1 - p_S^U)(1 - \tilde{p}_V)p_G.$$

Using (A.3), (A.7) and (A.8) this can be written as

$$\begin{aligned} R_v &= f_{\tilde{D}-1}(1 - p_S p_N)p_G \left(\frac{f'_{\tilde{D}-1}(1 - p_S p_N)}{f_{\tilde{D}-1}(1 - p_S p_N)} + \mu_S(U, U)\frac{f'_D(1 - p_S p_N)}{f_D(1 - p_S p_N)} \right) \\ &\quad \times (1 - 2p_S p_N + p_S p_N^2) \\ &= \eta(p')(1 - 2p' + p' p_N), \end{aligned}$$

where $p' = p_S p_N$ and η depends on p_S and p_N only through p'.

A.3 Probability of a major outbreak

To determine the probability of a major outbreak we need to derive the PGFs $f_{\tilde{\boldsymbol{C}}_i}(\boldsymbol{s})$ and $f_{\boldsymbol{C}_A}(\boldsymbol{s})$, where $\boldsymbol{s} = (s_1, s_2, \ldots, s_6) \in [0, 1]^6$, for $i = 1, 2, \ldots, 6$ and $A = U, V$; i.e. the offspring PGFs for the forward process.

Recall that secondary individuals in a household are vaccinated independently, each with probability p_V. Thus, conditioning first on household size and then on the number of secondary members that are vaccinated yields, for $i = 1, 2, \ldots, 6$,

$$f_{\tilde{\boldsymbol{C}}_i}(\boldsymbol{s}) = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v_s=0}^{n-1} \binom{n-1}{v_s} p_V^{v_s}(1 - p_V)^{n-1-v_s} f_{\tilde{\boldsymbol{C}}_i^{(n,v_s)}}(\boldsymbol{s}),$$

where $\tilde{\boldsymbol{C}}_i^{(n,v_s)}$ is the offspring random variable for a type-i non-initial individual, given that the corresponding household is of size n and has v_s secondary members vaccinated. Similarly, and in an obvious notation, for $A \in \{U, V\}$,

$$f_{\boldsymbol{C}_A}(\boldsymbol{s}) = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v_s=0}^{n-1} \binom{n-1}{v_s} p_V^{v_s}(1 - p_V)^{n-1-v_s} f_{\boldsymbol{C}_A^{(n,v_s)}}(\boldsymbol{s}).$$

With an all-or-nothing vaccine action, we may condition also on the number of vaccinations of secondary members that are unsuccessful to obtain, for $i = 1, 2, \ldots, 6$,

$$f_{\tilde{\boldsymbol{C}}_i^{(n,v_S)}}(\boldsymbol{s}) = \sum_{u_S=0}^{v_S} \binom{v_S}{u_S} (1-\varepsilon)^{u_S} \varepsilon^{v_S-u_S} f_{\hat{\boldsymbol{C}}_i^{(n-v_S+u_S,u_S)}}(\boldsymbol{s}),$$

where $\hat{\boldsymbol{C}}_i^{(n',v')}$ is the offspring random variable for a type-i non-initial individual, given that the corresponding household contains $n'-1$ other susceptibles, of which v' are unsuccessfully vaccinated. (Note that, unlike with household-based vaccination, the vaccine status of these susceptibles is important as it affects their degree distributions.) Similarly, and in an obvious notation, for $A \in \{U, V\}$,

$$f_{\boldsymbol{C}_A^{(n,v_S)}}(\boldsymbol{s}) = \sum_{u_S=0}^{v_S} \binom{v_S}{u_S} (1-\varepsilon)^{u_S} \varepsilon^{v_S-u_S} f_{\hat{\boldsymbol{C}}_A^{(n-v_S+u_S,u_S)}}(\boldsymbol{s}).$$

The PGFs $f_{\tilde{\boldsymbol{C}}_i^{(n,v_S)}}(\boldsymbol{s})$ and $f_{\boldsymbol{C}_A^{(n,v_S)}}(\boldsymbol{s})$ when the vaccine action is non-random and $f_{\hat{\boldsymbol{C}}_i^{(n',v')}}(\boldsymbol{s})$ and $f_{\hat{\boldsymbol{C}}_A^{(n',v')}}(\boldsymbol{s})$ when the vaccine action is all-or-nothing can be calculated using the methods described in Appendices B.1 and B.4. As with other calculations of PGFs relating to forward branching process approximations, these simplify appreciably if the infectious period is constant.

A.4 Final outcome of major outbreak

To calculate the asymptotic final size of a major outbreak we need the joint PGFs $f_{\tilde{\boldsymbol{B}}_i}(\boldsymbol{s})$, for $i = 1, 2, \ldots, 6$, and $f_{\boldsymbol{B}_A}(\boldsymbol{s})$, for $A = U, V$; where $\boldsymbol{s} = (s_1, s_2, \ldots, s_6) \in [0, 1]^6$. We determine first the former joint PGFs. For $i = 1, 2, \ldots, 6$, let $\boldsymbol{M}_i = (M_{iU}, M_{iV})$, where M_{iU} and M_{iV} are respectively the number of unvaccinated and vaccinated secondary individuals in the local susceptibility set of a typical type-i primary individual. Then $\tilde{\boldsymbol{B}}_i$ admits the decomposition

$$\tilde{\boldsymbol{B}}_i = \hat{\boldsymbol{B}}_i^P + \sum_{k=1}^{M_{iU}} \hat{\boldsymbol{B}}_U^S(k) + \sum_{l=1}^{M_{iV}} \hat{\boldsymbol{B}}_V^S(l), \tag{A.11}$$

where $\hat{\boldsymbol{B}}_i^P$, $\hat{\boldsymbol{B}}_U^S(k)$ and $\hat{\boldsymbol{B}}_V^S(l)$ are the respective contributions to $\tilde{\boldsymbol{B}}_i$ from the primary individual, the kth unvaccinated secondary individual and the lth vaccinated secondary individual in the local susceptibility set. All summands on the right hand side of (A.11) are mutually independent and also independent of $\boldsymbol{M}_i$, so, in an obvious notation,

$$f_{\tilde{\boldsymbol{B}}_i}(\boldsymbol{s}) = f_{\hat{\boldsymbol{B}}_i^P}(\boldsymbol{s}) f_{\boldsymbol{M}_i}(f_{\hat{\boldsymbol{B}}_U^S}(\boldsymbol{s}), f_{\hat{\boldsymbol{B}}_V^S}(\boldsymbol{s})). \tag{A.12}$$

To derive the PGF $f_{\boldsymbol{M}_i}$, note that $f_{\boldsymbol{M}_1} = f_{\boldsymbol{M}_2} = f_{\boldsymbol{M}_4} = f_{\boldsymbol{M}_5}$ ($= f_{\boldsymbol{M}_V}$ say) and $f_{\boldsymbol{M}_3} = f_{\boldsymbol{M}_6}$ ($= f_{\boldsymbol{M}_U}$ say). Further, conditioning on the size of the household and

recalling that secondary individuals in the household are vaccinated independently, each with probability p_V, yields

$$f_{\boldsymbol{M}_A}(\boldsymbol{s}) = \sum_{n=1}^{\infty} \tilde{\rho}_n \sum_{v=0}^{n-1} \binom{n-1}{v} p_V^v (1-p_V)^{n-1-v} f_{\boldsymbol{M}_A^{(n,v+\delta_{A,V})}}(\boldsymbol{s}) \quad (\boldsymbol{s} \in [0,1]^2),$$

for $A \in \{U, V\}$, where $\boldsymbol{M}_A^{(n,v)}$ is as in Sect. 3.3. Calculation of $f_{\boldsymbol{M}_A^{(n,v)}}(\boldsymbol{s})$ for the two models of vaccine action is addressed in Appendix B.5.

We determine next the PGFs $f_{\hat{\boldsymbol{B}}_i^P}$ $(i = 1, 2, \ldots, 6)$. We do so by conditioning on $\tilde{\boldsymbol{X}}_i$, the numbers of next-generation neighbours of different types of a typical type-i primary case, i^* say, and then considering how many of the $\tilde{\boldsymbol{X}}_i$ individuals of the various types actually join the susceptibility set. Each of i^*'s $\tilde{D}(i) - 1 = \sum_{j=1}^{6} \tilde{X}_{ij}$ next-generation neighbours in the construction of the backward process enters the susceptibility set independently, with probability depending on the types of the two individuals involved. The probability that a type-j neighbour of a type-i individual joins the susceptibility set is

$$p_{ij}^B = \begin{cases} p_{ji}^G & \text{non-random vaccine,} \\ p_{ij}^G & \text{all-or-nothing vaccine.} \end{cases}$$

(The different formulae arise from the fact that in the all-or-nothing case, in addition to requiring a contact from the neighbour to the type-i individual of interest, a vaccinated neighbour is able to join the susceptibility set only if its vaccination fails.) Hence, for $i = 1, 2, \ldots, 6$,

$$f_{\hat{\boldsymbol{B}}_i^P}(\boldsymbol{s}) = f_{\tilde{\boldsymbol{X}}_i}(\boldsymbol{h}_i^B(\boldsymbol{s})),$$

where $\boldsymbol{h}_i^B(\boldsymbol{s}) = (h_{i1}^B(s_1), h_{i2}^B(s_2), \ldots, h_{i6}^B(s_6))$, with $h_{ij}^B(s) = 1 - p_{ij}^B + s p_{ij}^B$ $(j = 1, 2, \ldots, 6)$. The PGFs $f_{\hat{\boldsymbol{B}}_U^S}$ and $f_{\hat{\boldsymbol{B}}_V^S}$ follow using exactly the same arguments. In an obvious notation,

$$f_{\hat{\boldsymbol{B}}_A^S}(\boldsymbol{s}) = f_{\boldsymbol{X}_A}(\boldsymbol{h}_A^B(\boldsymbol{s})) \quad (A \in \{U, V\}),$$

where $\boldsymbol{h}_U^B = \boldsymbol{h}_3^B$ and $\boldsymbol{h}_V^B = \boldsymbol{h}_1^B$.

Finally, we determine the offspring PGFs for the initial generation, $f_{\boldsymbol{B}_U}$ and $f_{\boldsymbol{B}_V}$. Observe that, for $A \in \{U, V\}$, in the initial generation, a primary individual of type A behaves according to the same probability law as a typical secondary type-A individual (in any generation), so (A.12) becomes

$$f_{\boldsymbol{B}_A}(\boldsymbol{s}) = f_{\hat{\boldsymbol{B}}_A^S}(\boldsymbol{s}) f_{\boldsymbol{M}_A}(f_{\hat{\boldsymbol{B}}_U^S}(\boldsymbol{s}), f_{\hat{\boldsymbol{B}}_V^S}(\boldsymbol{s})) \quad (A \in \{U, V\})$$

and $f_{\boldsymbol{B}_A}(\boldsymbol{s})$ can be evaluated using results given above.

Appendix B: Properties of single-household epidemics

In this appendix we first outline the theory of final state random variables as studied by Ball and O'Neill (1999). We then apply this (in Sects. B.2, B.3, B.4) to calculating the PGFs of offspring distributions of the various forward branching process approximations in the paper. In Sects. B.5 and B.6 we give results pertaining to the size of local susceptibility sets and the expected size of local epidemics.

We use the following notation. The symbols $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote the non-negative integers and reals, respectively. For vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ in $\mathbb{R}^m$, we define $\boldsymbol{x}\boldsymbol{y} = \prod_{i=1}^m x_i y_i$ and $\boldsymbol{x}^{\boldsymbol{y}} = \prod_{i=1}^m x_i^{y_i}$. We write $\boldsymbol{x} \leq \boldsymbol{y}$ if $x_i \leq y_i$ $(i = 1, 2, \ldots, m)$ and $\boldsymbol{x} < \boldsymbol{y}$ if, in addition, $x_i < y_i$ for at least one i. For $n, k \in \mathbb{Z}_+$, $n_{[k]}$ denotes the falling factorial $n(n-1)\ldots(n-k+1)$, with $n_{[0]} = 1$. For $\boldsymbol{n}, \boldsymbol{k} \in \mathbb{Z}_+^m$, we define $\boldsymbol{n}_{[\boldsymbol{k}]} = \prod_{i=1}^m n_{i[k_i]}$. For $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{Z}_+^m$, we write $\sum_{\boldsymbol{k}=\boldsymbol{i}}^{\boldsymbol{j}}$ for $\sum_{k_1=i_1}^{j_1} \sum_{k_2=i_2}^{j_2} \cdots \sum_{k_m=i_m}^{j_m}$. We take $\mathbf{0}$ and $\mathbf{1}$ to be vectors with all entries 0 and 1 respectively, with the dimensions being clear from the context. Finally, we let $f^{(r)}$ denote the r-th order derivative of the function f.

B.1 Multitype single-household epidemic and key result

Consider a single-household SIR epidemic model with m types of individuals, labelled $1, 2, \ldots, m$. Suppose that, for $i = 1, 2, \ldots, m$, there are initially a_i infectives and n_i susceptibles of type i, and let $\boldsymbol{a} = (a_1, a_2, \ldots, a_m)$ and $\boldsymbol{n} = (n_1, n_2, \ldots, n_m)$. For $i = 1, 2, \ldots, m$, the infectious periods of type-i infectives are each distributed according to a random variable $I^{(i)}$. For $i, j = 1, 2, \ldots, m$, the individual-to-individual infection rate from a given type-i infective to a given type-j susceptible is λ_{ij}. As in Sect. 2.1, such infections are governed by Poisson processes, and all Poisson processes and infectious periods are mutually independent.

To each type-i infective we attach a vector of p non-negative integer-valued random attributes distributed according to a vector random variable $\boldsymbol{A}^{(i)} = (A_1^{(i)}, A_2^{(i)}, \ldots, A_p^{(i)})$. In our applications $\boldsymbol{A}^{(i)}$ will describe the numbers of global infections of different types of individuals made by an infective in the single-household epidemic. The realisations of the random variables $(I^{(i)}, \boldsymbol{A}^{(i)})$ are independent for distinct infectives and identically distributed for infectives of the same type. Note that $I^{(i)}$ and $\boldsymbol{A}^{(i)}$ may be dependent. For $i = 1, 2, \ldots, m$, let T_i be the number of initial susceptibles of type i that are ultimately infected by the epidemic and let $\boldsymbol{A}^{(i)}(T_i)$ be the sum of the attribute vectors over all $a_i + T_i$ infectives of type i. Further, let $\boldsymbol{T} = (T_1, T_2, \ldots, T_m)$ and let $\boldsymbol{A}(\boldsymbol{T}) = \sum_{i=1}^m \boldsymbol{A}^{(i)}(T_i)$ be the sum of the attribute vectors over all $\sum_{i=1}^m (a_i + T_i)$ infectives in the epidemic. Thus, in applications, $\boldsymbol{A}(\boldsymbol{T})$ will give the offspring random variable for the forward branching process. For $\boldsymbol{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$ and $\boldsymbol{s} = (s_1, s_2, \ldots, s_p) \in [0, 1]^p$, let

$$\phi(\boldsymbol{x}, \boldsymbol{s}) = E[\boldsymbol{x}^{\boldsymbol{n}-\boldsymbol{T}} \boldsymbol{s}^{\boldsymbol{A}(\boldsymbol{T})}].$$

We give below an expression for $\phi(\boldsymbol{x}, s)$ in terms of multivariate Gontcharoff polynomials, first studied by Lefèvre and Picard (1990), which we now define. Let $\boldsymbol{U} =$

$(\boldsymbol{u}_{\boldsymbol{j}} \in \mathbb{R}^m : \boldsymbol{j} \in \mathbb{Z}_+^m)$ be a collection of real numbers. The Gontcharoff polynomials associated with $\boldsymbol{U}$, denoted $(G_{\boldsymbol{k}}(\boldsymbol{x}|\boldsymbol{U}),\ \boldsymbol{k} \in \mathbb{Z}_+^m, \boldsymbol{x} \in \mathbb{R}^m)$, are defined recursively by

$$\sum_{\boldsymbol{j}=\boldsymbol{0}}^{\boldsymbol{k}} \boldsymbol{k}_{[\boldsymbol{j}]} \boldsymbol{u}_{\boldsymbol{j}}^{\boldsymbol{k}-\boldsymbol{j}} G_{\boldsymbol{j}}(\boldsymbol{x}|\boldsymbol{U}) = \boldsymbol{x}^{\boldsymbol{k}} \quad (\boldsymbol{k} \geq \boldsymbol{0}).$$

Note that $G_{\boldsymbol{0}}(\boldsymbol{x}|\boldsymbol{U}) \equiv 1$ and that $G_{\boldsymbol{k}}(\boldsymbol{x}|\boldsymbol{U})$ is a polynomial of degree $k_1, k_2, \ldots, k_m$ in $x_1, x_2, \ldots, x_m$, respectively, depending only on $(\boldsymbol{u}_{\boldsymbol{j}} : \boldsymbol{j} < \boldsymbol{k})$. The following result is derived easily from Ball and O'Neill (1999, Theorem 5.1), so its proof is omitted.

Theorem 1 *For $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{s} \in [0, 1]^p$,*

$$\phi(\boldsymbol{x}, \boldsymbol{s}) = \sum_{\boldsymbol{j}=\boldsymbol{0}}^{\boldsymbol{n}} \boldsymbol{n}_{[\boldsymbol{j}]} (\boldsymbol{\psi}(\boldsymbol{s}, \boldsymbol{j}))^{\boldsymbol{n}+\boldsymbol{a}-\boldsymbol{j}} G_{\boldsymbol{j}}(\boldsymbol{x}|\boldsymbol{U}), \tag{B.1}$$

where $\boldsymbol{\psi}(\boldsymbol{s}, \boldsymbol{j}) = (\psi_1(\boldsymbol{s}, \boldsymbol{j}), \psi_2(\boldsymbol{s}, \boldsymbol{j}), \ldots, \psi_m(\boldsymbol{s}, \boldsymbol{j}))$, with

$$\psi_i(\boldsymbol{s}, \boldsymbol{j}) = E\left[\exp\left(-I^{(i)} \sum_{k=1}^{m} \lambda_{ik} j_k\right) \boldsymbol{s}^{\boldsymbol{A}^{(i)}}\right] \quad (i = 1, 2, \ldots, m),$$

and $\boldsymbol{U} = (\boldsymbol{u}_{\boldsymbol{j}} : \boldsymbol{j} \in \mathbb{Z}_+^m)$ has components $\boldsymbol{u}_{\boldsymbol{j}} = \boldsymbol{\psi}(\boldsymbol{s}, \boldsymbol{j})$.

We now describe how Theorem 1 can be used to determine offspring PGFs for the forward branching processes in the earlier parts of the paper. Recall that primary and secondary infectives in a household typically have distinct global degree distributions. Hence, in addition to being typed according to their vaccination status, individuals also need to be typed as primary or secondary. Thus we may write $m = m_P + m_S$, where types $1, 2, \ldots, m_P$ correspond to primary individuals and types $m_P+1, m_P+2, \ldots, m$ to secondary individuals. Write $\boldsymbol{a} = (\boldsymbol{a}_P, \boldsymbol{a}_S)$, $\boldsymbol{n} = (\boldsymbol{n}_P, \boldsymbol{n}_S)$, $\boldsymbol{x} = (\boldsymbol{x}_P, \boldsymbol{x}_S)$, $\boldsymbol{T} = (\boldsymbol{T}_P, \boldsymbol{T}_S)$ and $\boldsymbol{\psi}(\boldsymbol{s}, \boldsymbol{j}) = (\boldsymbol{\psi}_P(\boldsymbol{s}, \boldsymbol{j}), \boldsymbol{\psi}_S(\boldsymbol{s}, \boldsymbol{j}))$ in the obvious fashion. Note that all susceptibles are secondary individuals, so $\boldsymbol{n}_P = \boldsymbol{0}$, and all initial infectives are primary individuals, so $\boldsymbol{a}_S = \boldsymbol{0}$. It follows that $\boldsymbol{T}_P = \boldsymbol{0}$ and the index $\boldsymbol{j}$ of the summation in (B.1) takes the form $(\boldsymbol{0}, \boldsymbol{j}_S)$. Let

$$\tilde{\phi}(\boldsymbol{x}_S, \boldsymbol{s}) = E\left[\boldsymbol{x}_S^{\boldsymbol{n}_S - \boldsymbol{T}_S} \boldsymbol{s}^{\boldsymbol{A}(\boldsymbol{T})}\right] \quad (\boldsymbol{x}_S \in \mathbb{R}^{m_S},\ \boldsymbol{s} \in [0, 1]^p).$$

Then the following corollary follows easily from Theorem 1. For $i = 1, 2, \ldots, m$, let $\boldsymbol{\lambda}_S^{(i)} = (\lambda_{i,m_P+1}, \lambda_{i,m_P+2}, \ldots, \lambda_{i,m})^\top$, where $\top$ denotes transpose.

Corollary 2 *For $\boldsymbol{x}_S \in \mathbb{R}^{m_S}$ and $\boldsymbol{s} \in [0, 1]^p$,*

$$\tilde{\phi}(\boldsymbol{x}_S, \boldsymbol{s}) = \sum_{\boldsymbol{j}_S=\boldsymbol{0}}^{\boldsymbol{n}_S} \boldsymbol{n}_{S_{[\boldsymbol{j}_S]}} (\tilde{\boldsymbol{\psi}}_P(\boldsymbol{s}, \boldsymbol{j}_S))^{\boldsymbol{a}_P} (\tilde{\boldsymbol{\psi}}_S(\boldsymbol{s}, \boldsymbol{j}_S))^{\boldsymbol{n}_S - \boldsymbol{j}_S} G_{\boldsymbol{j}_S}(\boldsymbol{x}_S|\boldsymbol{U}_S), \tag{B.2}$$

where $\tilde{\boldsymbol{\psi}}_P(\boldsymbol{s}, \boldsymbol{j}_S) = \left(\tilde{\psi}_1(\boldsymbol{s}, \boldsymbol{j}_S), \tilde{\psi}_2(\boldsymbol{s}, \boldsymbol{j}_S), \ldots, \tilde{\psi}_{m_P}(\boldsymbol{s}, \boldsymbol{j}_S)\right)$
and $\tilde{\boldsymbol{\psi}}_S(\boldsymbol{s}, \boldsymbol{j}_S) = \left(\tilde{\psi}_{m_P+1}(\boldsymbol{s}, \boldsymbol{j}_S), \tilde{\psi}_{m_P+2}(\boldsymbol{s}, \boldsymbol{j}_S), \ldots, \tilde{\psi}_m(\boldsymbol{s}, \boldsymbol{j}_S)\right)$, *with*

$$\tilde{\psi}_i(\boldsymbol{s}, \boldsymbol{j}_S) = E\left[\exp(-I^{(i)}\boldsymbol{j}_S\boldsymbol{\lambda}_S^{(i)})\boldsymbol{s}^{\boldsymbol{A}^{(i)}}\right] \quad (i = 1, 2, \ldots, m) \tag{B.3}$$

and $\boldsymbol{U}_S = (\boldsymbol{u}_{\boldsymbol{j}_S}^S : \boldsymbol{j}_S \in \mathbb{Z}_+^{m_S})$ *has components* $\boldsymbol{u}_{\boldsymbol{j}_S}^S = \tilde{\boldsymbol{\psi}}_S(\boldsymbol{s}, \boldsymbol{j}_S)$.

We are primarily concerned with determining the PGF of $\boldsymbol{A}(\boldsymbol{T})$, which of course is obtained by setting $\boldsymbol{x}_S = \mathbf{1}$ in (B.2). Thus, all that remains is to determine $\tilde{\psi}_i(\boldsymbol{s}, \boldsymbol{j}_S)$, which is application dependent. Note from (B.3) that it is sufficient to determine, for $\theta \in \mathbb{R}_+$ and $\boldsymbol{s} \in [0, 1]^p$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = E\left[\exp(-\theta I^{(i)})\boldsymbol{s}^{\boldsymbol{A}^{(i)}}\right] \quad (i = 1, 2, \ldots, m),$$

which we now do for the various models in the paper.

B.2 No vaccination

This model is studied by Ball et al. (2010), so we just state the result. Note that there are two types of individual, primary (P) and secondary (S), with their number of uninfected global neighbours distributed according to $D_P \sim \tilde{D} - 1$ and $D_S \sim D$, respectively. Further $p = 1$, so $\boldsymbol{s}$ is a scalar, s say. For a degree distribution D and real numbers c_1, c_2, θ, define the function $F(D, c_1, c_2, \theta)$ by

$$F(D, c_1, c_2, \theta) = \sum_{r=0}^{\infty} \frac{c_1^r}{r!}\phi_I(\theta + r\lambda_G) f_D^{(r)}(c_2). \tag{B.4}$$

Then, for $A \in \{P, S\}$,

$$\hat{\psi}_A(\theta, s) = F(D_A, 1 - s, s, \theta),$$

cf. Ball et al. (2010, Theorem 1). Equation (B.4), in conjunction with Corollary 2, enables the PGFs $f_{C^{(n)}}$ and $f_{\tilde{C}^{(n)}}$, defined in Sect. 2.2.1, to be evaluated.

B.3 Households based vaccination

Recall from Sect. 3 that with an all-or-nothing vaccine action all required PGFs and means can be expressed in terms of $f_{C^{(n)}}$, $f_{\tilde{C}^{(n)}}$ and $\mu_n(\lambda_L)$, so here we need consider only the non-random vaccine action model. This model has four types of individuals: primary-unvaccinated (PU), primary-vaccinated (PV), secondary-unvaccinated (SU) and secondary-vaccinated (SV), their number of uninfected global neighbours being distributed as D_{PU}, D_{PV}, D_{SU} and D_{SV}, respectively. Recall from Sect. 3.3 that $D_{PU} \stackrel{D}{=} D_{PV} \sim \tilde{D} - 1$ and $D_{SU} \stackrel{D}{=} D_{SV} \sim D$. For $F \in \{PU, PV, SU, SV\}$, the random attribute of interest is $\boldsymbol{A}^{(F)} = (A_U^{(F)}, A_V^{(F)})$, where $A_U^{(F)}$ and $A_V^{(F)}$ are the

number of unvaccinated and vaccinated global infections made by a type-F infective in the single household epidemic. Thus $m_S = 2$ and we determine $\hat{\psi}_F(\theta, \boldsymbol{s})$, where $\boldsymbol{s} = (s_U, s_V)$.

Consider first $\hat{\psi}_{SU}(\theta, \boldsymbol{s})$. Conditioning on the degree D_{SU} and infectious period I of a typical type-SU infective, i^* say, yields

$$\hat{\psi}_{SU}(\theta, \boldsymbol{s}) = E_{D_{SU},I}\left[\mathrm{e}^{-\theta I} E\left[\boldsymbol{s}^{\boldsymbol{A}^{(SU)}} \mid D_{SU}, I\right]\right]. \tag{B.5}$$

Let N_V denote the number of i^*'s uninfected global neighbours that are vaccinated. Then $\left(A_U^{(SU)} \mid N_V, D_{SU}, I\right) \sim \mathrm{Bin}\left(D_{SU} - N_V, 1 - \mathrm{e}^{-\lambda_G I}\right)$ and independently $\left(A_V^{(SU)} \mid N_V, D_{SU}, I\right) \sim \mathrm{Bin}\left(N_V, 1 - \mathrm{e}^{-\lambda_G a I}\right)$, so

$$\begin{aligned} &E\left[\boldsymbol{s}^{\boldsymbol{A}^{(SU)}} \mid N_V, D_{SU}, I\right] \\ &\quad = \left(\mathrm{e}^{-\lambda_G I} + (1 - \mathrm{e}^{-\lambda_G I}) s_U\right)^{D_{SU} - N_V} \left(\mathrm{e}^{-\lambda_G a I} + (1 - \mathrm{e}^{-\lambda_G a I}) s_V\right)^{N_V}, \end{aligned} \tag{B.6}$$

whence, since $(N_V \mid D_{SU}) \sim \mathrm{Bin}(D_{SU}, p_V)$,

$$E\left[\boldsymbol{s}^{\boldsymbol{A}^{(SU)}} \mid D_{SU}, I\right] = \left((1 - p_V)\left[\mathrm{e}^{-\lambda_G I} + (1 - \mathrm{e}^{-\lambda_G I}) s_U\right] + p_V\left[\mathrm{e}^{-\lambda_G a I} + (1 - \mathrm{e}^{-\lambda_G a I}) s_V\right]\right)^{D_{SU}}. \tag{B.7}$$

(As usual, $\mathrm{Bin}(n, p)$ denotes a binomial distribution with parameters $n \in \mathbb{Z}_+$ and $p \in [0, 1]$.) Let $\hat{a}(\boldsymbol{s}) = p_V s_V + (1 - p_V) s_U$, $\hat{b}(\boldsymbol{s}) = p_V(1 - s_V)$ and $\hat{c}(\boldsymbol{s}) = (1 - p_V)(1 - s_U)$. Then, using (B.5) and (B.7),

$$\begin{aligned} \hat{\psi}_{SU}(\theta, \boldsymbol{s}) &= E_{D_{SU},I}\left[\mathrm{e}^{-\theta I}\left(\hat{a}(\boldsymbol{s}) + \hat{b}(\boldsymbol{s})\mathrm{e}^{-\lambda_G a I} + \hat{c}(\boldsymbol{s})\mathrm{e}^{-\lambda_G I}\right)^{D_{SU}}\right] \\ &= \sum_{k=0}^{\infty} P(D_{SU} = k) E_I\left[\mathrm{e}^{-\theta I} \sum_{r=0}^{k} \sum_{l=0}^{k-r} \frac{k!}{r!l!(k-r-l)!} (\hat{a}(\boldsymbol{s}))^{k-l-r} (\hat{b}(\boldsymbol{s})\mathrm{e}^{-\lambda_G a I})^r (\hat{c}(\boldsymbol{s})\mathrm{e}^{-\lambda_G I})^l\right] \\ &= \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \sum_{k=r+l}^{\infty} P(D_{SU} = k) \frac{k!}{r!l!(k-r-l)!} (\hat{a}(\boldsymbol{s}))^{k-l-r} (\hat{b}(\boldsymbol{s}))^r (\hat{c}(\boldsymbol{s}))^l E\left[\mathrm{e}^{-(\theta + (ar+l)\lambda_G) I}\right]. \end{aligned} \tag{B.8}$$

For a degree distribution D and real numbers $c_1, c_2, c_3, \theta, a, b$, define the function $G(D, c_1, c_2, c_3, \theta, a, b)$ by

$$G(D, c_1, c_2, c_3, \theta, a, b) = \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{c_2^r}{r!} \frac{c_3^l}{l!} \phi_I(\theta + \lambda_G b(ar + l)) f_D^{(r+l)}(c_1). \tag{B.9}$$

Then (B.8) implies that

$$\hat{\psi}_{SU}(\theta, \boldsymbol{s}) = G(D_{SU}, \hat{a}(\boldsymbol{s}), \hat{b}(\boldsymbol{s}), \hat{c}(\boldsymbol{s}), \theta, a, 1). \tag{B.10}$$

A similar argument shows that

$$\hat{\psi}_{SV}(\theta, \boldsymbol{s}) = G(D_{SV}, \hat{a}(\boldsymbol{s}), \hat{b}(\boldsymbol{s}), \hat{c}(\boldsymbol{s}), \theta, a, b), \tag{B.11}$$

and that $\hat{\psi}_{PU}(\theta, \boldsymbol{s})$ is given by (B.10), with D_{SU} replaced by D_{PU}, and $\hat{\psi}_{PV}(\theta, \boldsymbol{s})$ is given by (B.11), with D_{SV} replaced by D_{PV}.

If a closed-form expression for f_D is available then $G(D, c_1, c_2, c_3, \theta, a, b)$ may be computed by using a finite truncation of (B.9). If a closed-form expression for f_D is unavailable then $G(D, c_1, c_2, c_3, \theta, a, b)$ may be computed by using a finite truncation of a corresponding triple sum (cf. (B.8)).

Note that if the infectious period is constant, say $I \equiv \iota$, then (B.5) to (B.7) imply that

$$\hat{\psi}_{SU}(\theta, \boldsymbol{s}) = \mathrm{e}^{-\theta\iota} f_{D_{SU}} \left(\hat{a}(\boldsymbol{s}) + \mathrm{e}^{-\lambda_G a\iota}\hat{b}(\boldsymbol{s}) + \mathrm{e}^{-\lambda_G \iota}\hat{c}(\boldsymbol{s})\right)$$

and corresponding expressions for $\hat{\psi}_{SV}(\theta, \boldsymbol{s})$, $\hat{\psi}_{PU}(\theta, \boldsymbol{s})$ and $\hat{\psi}_{PV}(\theta, \boldsymbol{s})$ are easily derived.

B.4 Acquaintance vaccination

This model has eight types of individuals, six primary and two secondary, which, as in Sect. 4.2, we label $1, 2, \ldots, 6$ and U, V, respectively. For each type i, the random attribute of interest $\boldsymbol{A}^{(i)} = (A_1^{(i)}, A_2^{(i)}, \ldots, A_6^{(i)})$, where $A_j^{(i)}$ is the number of global infections of a type j individual made by an infective type i individual. Thus we determine, in an obvious notation, $\hat{\psi}_i(\theta, \boldsymbol{s})$ $(i = 1, 2, \ldots, 6)$, $\hat{\psi}_U(\theta, \boldsymbol{s})$ and $\hat{\psi}_V(\theta, \boldsymbol{s})$; for $\theta \in \mathbb{R}_+$ and $\boldsymbol{s} = (s_1, s_2, \ldots, s_6) \in [0, 1]^6$. It is convenient to treat the all-or-nothing and non-random vaccine action models separately.

B.4.1 All-or-nothing vaccine action

For $i = 1, 2, \ldots, 6$, recall that $\tilde{D}(i)$ is the global degree of a typical type-i primary individual, i^* say, and also that $\tilde{\boldsymbol{X}}_i = (\tilde{X}_{i1}, \tilde{X}_{i2}, \ldots, \tilde{X}_{i6})$, where $\tilde{X}_{ij}$ is the number of i^*'s next-generation neighbours that have type j. Recall that types 3 and 6 have not been vaccinated and types 1, 2, 4 and 5 have been vaccinated. Thus

$$\left(A_j^{(i)} | \tilde{X}_{ij}, I\right) \sim \begin{cases} \mathrm{Bin}\left(\tilde{X}_{ij}, 1 - \mathrm{e}^{-\lambda_G I}\right) & \text{if } j = 3, 6, \\ \mathrm{Bin}\left(\tilde{X}_{ij}, (1-\varepsilon)(1 - \mathrm{e}^{-\lambda_G I})\right) & \text{if } j = 1, 2, 4, 5, \end{cases}$$

and

$$E\left[\boldsymbol{s}^{\boldsymbol{A}^{(i)}} | \tilde{\boldsymbol{X}}_i, I\right] = \prod_{j=3,6} \left(\mathrm{e}^{-\lambda_G I} + (1 - \mathrm{e}^{-\lambda_G I}) s_j\right)^{\tilde{X}_{ij}} \times \prod_{j=1,2,4,5} \left(\varepsilon + (1-\varepsilon)\mathrm{e}^{-\lambda_G I} + (1-\varepsilon)(1 - \mathrm{e}^{-\lambda_G I}) s_j\right)^{\tilde{X}_{ij}}. \tag{B.12}$$

Further, for $i = 1, 3, 4, 6$, the types of the $\tilde{D}(i) - 1$ of i^*'s next-generation neighbours are chosen independently according to $\hat{p}_{ij}$, defined in Sect. A.1.2, so

$$E\left[\boldsymbol{s}^{\boldsymbol{A}^{(i)}} | \tilde{D}(i), I\right] = \left(\hat{p}_i(\boldsymbol{s}, \varepsilon) + (1 - \hat{p}_i(\boldsymbol{s}, \varepsilon))\mathrm{e}^{-\lambda_G I}\right)^{\tilde{D}(i)-1}, \tag{B.13}$$

where

$$\hat{p}_i(\boldsymbol{s}, \varepsilon) = \sum_{j=3,6} \hat{p}_{ij} s_j + \sum_{j=1,2,4,5} \hat{p}_{ij}(\varepsilon + (1 - \varepsilon)s_j).$$

Hence,

$$\begin{aligned}\hat{\psi}_i(\theta, \boldsymbol{s}) &= E\left[\mathrm{e}^{-\theta I}\boldsymbol{s}^{\boldsymbol{A}^{(i)}}\right] \\ &= E_I\left[\mathrm{e}^{-\theta I} E_{\tilde{D}(i)}\left[E\left[\boldsymbol{s}^{\boldsymbol{A}^{(i)}} | \tilde{D}(i), I\right]\right]\right] \\ &= E_I\left[\mathrm{e}^{-\theta I} E_{\tilde{D}(i)}\left[(\hat{p}_i(\boldsymbol{s}, \varepsilon) + (1 - \hat{p}_i(\boldsymbol{s}, \varepsilon))\mathrm{e}^{-\lambda_G I})^{\tilde{D}(i)-1}\right]\right].\end{aligned} \tag{B.14}$$

A similar argument to the derivation of (B.8), using the binomial theorem rather than the multinomial theorem, then yields

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = F(\tilde{D}(i) - 1, 1 - \hat{p}_i(\boldsymbol{s}, \varepsilon), \hat{p}_i(\boldsymbol{s}, \varepsilon), \theta).$$

For $i = 2$ and $i = 5$, (B.12) still holds but $f_{\tilde{\boldsymbol{X}}_i}(\boldsymbol{s})$ is given by (A.9) and (A.10), respectively. Omitting the details, similar arguments to the above show that, for $i = 2, 5$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1}\left[F\left(\tilde{D} - 1, \tilde{p}_i(\boldsymbol{s}, 0), \tilde{q}_i(\boldsymbol{s}, 0), \theta\right) - F\left(\tilde{D} - 1, \tilde{p}_i(\boldsymbol{s}, p_N), \tilde{q}_i(\boldsymbol{s}, p_N), \theta\right)\right], \tag{B.15}$$

where $\tilde{p}_i(\boldsymbol{s}, x) = p_S(1-x)\tilde{p}_{i1}(\boldsymbol{s}) + (1-p_S)\tilde{p}_{i2}(\boldsymbol{s})$ and $\tilde{q}_i(\boldsymbol{s}, x) = 1 - p_S x - \tilde{p}_i(\boldsymbol{s}, x)$, with

$$\begin{aligned}
\tilde{p}_{21}(\boldsymbol{s}) &= p_N(1 - \varepsilon)(1 - s_1) + (1 - p_N)[\tilde{p}_V(1 - \varepsilon)(1 - s_2) + (1 - \tilde{p}_V)(1 - s_3)], \\
\tilde{p}_{22}(\boldsymbol{s}) &= p_N(1 - \varepsilon)(1 - s_4) + (1 - p_N)[\tilde{p}_V(1 - \varepsilon)(1 - s_5) + (1 - \tilde{p}_V)(1 - s_6)], \\
\tilde{p}_{51}(\boldsymbol{s}) &= \tilde{p}_V(1 - \varepsilon)(1 - s_2) + (1 - \tilde{p}_V)(1 - s_3) \quad \text{and} \\
\tilde{p}_{52}(\boldsymbol{s}) &= \tilde{p}_V(1 - \varepsilon)(1 - s_5) + (1 - \tilde{p}_V)(1 - s_6).
\end{aligned}$$

To simplify the exposition we write (B.15) as

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1}\Delta_{0,p_N} F\left(\tilde{D} - 1, \tilde{p}_i(\boldsymbol{s}, x), \tilde{q}_i(\boldsymbol{s}, x), \theta\right).$$

In the sequel we use, without comment, a similar notation for differences of other functions evaluated at 0 and p_N.

Turning now to secondary individuals, note that a type-U individual has global degree and therefore number of next-generation neighbours distributed as D_U. Such an individual is sampled with probability p_S, in which case the type of each next-generation neighbour is distributed according to $\hat{p}_{3j}$, otherwise it is unsampled and the type of each next-generation neighbour is distributed according to $\hat{p}_{6j}$. Thus

$$\hat{\psi}_U(\theta, \boldsymbol{s}) = p_S F\left(D_U, 1 - \hat{p}_3(\boldsymbol{s}, \varepsilon), \hat{p}_3(\boldsymbol{s}, \varepsilon), \theta\right) + (1 - p_S) F\left(D_U, 1 - \hat{p}_6(\boldsymbol{s}, \varepsilon), \hat{p}_6(\boldsymbol{s}, \varepsilon), \theta\right). \tag{B.16}$$

Similarly,

$$\hat{\psi}_V(\theta, \boldsymbol{s}) = p_V^{-1}\left[p_S \Delta_{0,p_N} F\left(D_V, \tilde{p}_2(\boldsymbol{s}, x), \tilde{q}_2(\boldsymbol{s}, x), \theta\right) + (1 - p_S)\Delta_{0,p_N} F\left(D_V, \tilde{p}_5(\boldsymbol{s}, x), \tilde{q}_5(\boldsymbol{s}, x), \theta\right)\right]. \tag{B.17}$$

To simplify the exposition we write (B.16) and (B.17) as

$$\hat{\psi}_U(\theta, \boldsymbol{s}) = p_S^{3,6} F(D_U, 1 - \hat{p}(\boldsymbol{s}, \varepsilon), \hat{p}(\boldsymbol{s}, \varepsilon), \theta)$$

and

$$\hat{\psi}_V(\theta, \boldsymbol{s}) = p_V^{-1} p_S^{2,5} \Delta_{0,p_N} F(D_V, \tilde{p}(\boldsymbol{s}, x), \tilde{q}(\boldsymbol{s}, x), \theta).$$

In the sequel, we use a similar notation without comment.

The above expressions simplify appreciably if the infectious period is constant, say $I \equiv \iota$. In that case, Eq. (B.14) implies that

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \mathrm{e}^{-\theta\iota} f_{\tilde{D}(i)-1}\left(\hat{p}_i(\boldsymbol{s}, \varepsilon) + (1 - \hat{p}_i(\boldsymbol{s}, \varepsilon))\mathrm{e}^{-\lambda_G \iota}\right) \quad (i = 1, 3, 4, 6)$$

and exploiting the PGFs for $\tilde{\boldsymbol{X}}_2$ and $\tilde{\boldsymbol{X}}_5$ [see (A.9) and (A.10)] yields

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1} \mathrm{e}^{-\theta\iota} \Delta_{0,p_N} f_{\tilde{D}-1}\left(\tilde{q}_i(\boldsymbol{s}, x) + \tilde{p}_i(\boldsymbol{s}, x)\mathrm{e}^{-\lambda_G \iota}\right) \quad (i = 2, 5).$$

Similar arguments show that

$$\hat{\psi}_U(\theta, \boldsymbol{s}) = \mathrm{e}^{-\theta\iota} p_S^{3,6} f_{D_U}\left(\hat{p}(\boldsymbol{s}, \varepsilon) + (1 - \hat{p}(\boldsymbol{s}, \varepsilon))\mathrm{e}^{-\lambda_G \iota}\right)$$

and

$$\hat{\psi}_V(\theta, \boldsymbol{s}) = p_V^{-1} \mathrm{e}^{-\theta\iota} p_S^{2,5} \Delta_{0,p_N} f_{D_V}\left(\tilde{q}(\boldsymbol{s}, x) + \tilde{p}(\boldsymbol{s}, x)\mathrm{e}^{-\lambda_G \iota}\right).$$

B.4.2 Non-random vaccine

Note that now

$$(A_j^{(i)} | \tilde{X}_{ij}, I) \sim \begin{cases} \mathrm{Bin}(\tilde{X}_{ij}, 1 - \mathrm{e}^{-\lambda_G I}) & \text{if } i, j \in \{3, 6\}, \\ \mathrm{Bin}(\tilde{X}_{ij}, 1 - \mathrm{e}^{-a\lambda_G I}) & \text{if } i \in \{3, 6\},\ j \in \{1, 2, 4, 5\}, \\ \mathrm{Bin}(\tilde{X}_{ij}, 1 - \mathrm{e}^{-b\lambda_G I}) & \text{if } i \in \{1, 2, 4, 5\},\ j \in \{3, 6\}, \\ \mathrm{Bin}(\tilde{X}_{ij}, 1 - \mathrm{e}^{-ab\lambda_G I}) & \text{if } i, j \in \{1, 2, 4, 5\}. \end{cases} \tag{B.18}$$

For $i \in \{1, 3, 4, 6\}$, let

$$\hat{a}_i(\boldsymbol{s}) = \sum_{j=1}^{6} \hat{p}_{ij} s_j, \quad \hat{b}_i(\boldsymbol{s}) = \sum_{j=1,2,4,5} \hat{p}_{ij}(1 - s_j) \quad \text{and} \quad \hat{c}_i(\boldsymbol{s}) = \sum_{j=3,6} \hat{p}_{ij}(1 - s_j).$$

Suppose that $i \in \{3, 6\}$. Then, arguing as in the derivation of (B.13),

$$E\left[\boldsymbol{s}^{\boldsymbol{A}^{(i)}} | \tilde{D}(i), I\right] = \left(\hat{a}_i(\boldsymbol{s}) + \hat{b}_i(\boldsymbol{s})\mathrm{e}^{-a\lambda_G I} + \hat{c}_i(\boldsymbol{s})\mathrm{e}^{-\lambda_G I}\right)^{\tilde{D}(i)-1},$$

whence

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = E_I\left[\mathrm{e}^{-\theta I} E_{\tilde{D}(i)}\left[\left(\hat{a}_i(\boldsymbol{s}) + \hat{b}_i(\boldsymbol{s})\mathrm{e}^{-a\lambda_G I} + \hat{c}_i(\boldsymbol{s})\mathrm{e}^{-\lambda_G I}\right)^{\tilde{D}(i)-1}\right]\right]. \tag{B.19}$$

Arguing as in the derivation of (B.8) then yields that

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = G\left(\tilde{D}(i) - 1, \hat{a}_i(\boldsymbol{s}), \hat{b}_i(\boldsymbol{s}), \hat{c}_i(\boldsymbol{s}), \theta, a, 1\right).$$

Similarly, for $i \in \{1, 4\}$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = G\left(\tilde{D}(i) - 1, \hat{a}_i(\boldsymbol{s}), \hat{b}_i(\boldsymbol{s}), \hat{c}_i(\boldsymbol{s}), \theta, a, b\right).$$

Suppose that $i = 5$. Then, using (B.18),

$$E\left[\boldsymbol{s}^{\boldsymbol{A}^{(5)}} | \tilde{\boldsymbol{X}}_5, I\right] = \prod_{j=3,6}\left(\mathrm{e}^{-b\lambda_G I} + (1 - \mathrm{e}^{-b\lambda_G I})s_j\right)^{\tilde{X}_{5j}} \times \prod_{j=1,2,4,5}\left(\mathrm{e}^{-ab\lambda_G I} + (1 - \mathrm{e}^{-ab\lambda_G I})s_j\right)^{\tilde{X}_{5j}}.$$

Invoking (A.10) now gives

$$\hat{\psi}_5(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1} E\left[\mathrm{e}^{-\theta I}\Delta_{0,p_N} f_{\tilde{D}-1}\left(\hat{g}_5(\boldsymbol{h}_5^F(\boldsymbol{s}, I), x)\right)\right], \tag{B.20}$$

where $\boldsymbol{h}_5^F(\boldsymbol{s}, I) = (h_{51}^F(s_1, I), h_{52}^F(s_2, I), \ldots, h_{56}^F(s_6, I))$, with $h_{5j}^F(s, I) = \mathrm{e}^{-b\lambda_G I} + (1 - \mathrm{e}^{-b\lambda_G I})s$ if $j = 3, 6$ and $h_{5j}^F(s, I) = \mathrm{e}^{-ab\lambda_G I} + (1 - \mathrm{e}^{-ab\lambda_G I})s$ if $j = 1, 2, 4, 5$; and $\hat{g}_5(\boldsymbol{s}, x)$ is as defined immediately following (A.10) in Sect. A.1.2. A simple calculation now shows that

$$\hat{g}_5\left(\boldsymbol{h}_5^F(\boldsymbol{s}, I), x\right) = \hat{a}_5(\boldsymbol{s}, x) + \hat{b}_5(\boldsymbol{s}, x)\mathrm{e}^{-ab\lambda_G I} + \hat{c}_5(\boldsymbol{s}, x)\mathrm{e}^{-b\lambda_G I}, \tag{B.21}$$

where $\hat{a}_5(\boldsymbol{s}, x) = p_S(1 - x)[\tilde{p}_V s_2 + (1 - \tilde{p}_V)s_3] + (1 - p_S)[\tilde{p}_V s_5 + (1 - \tilde{p}_V)s_6]$, $\hat{b}_5(\boldsymbol{s}, x) = p_S(1 - x)\tilde{p}_V(1 - s_2) + (1 - p_S)\tilde{p}_V(1 - s_5)$ and lastly

$\hat{c}_5(\boldsymbol{s}, x) = p_S(1-x)(1-\tilde{p}_V)(1-s_3) + (1-p_S)(1-\tilde{p}_V)(1-s_6)$. Arguing as in the derivation of (B.8) now gives, for $i = 5$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1} \Delta_{0,p_N} G\left(\tilde{D}-1, \hat{a}_i(\boldsymbol{s}, x), \hat{b}_i(\boldsymbol{s}, x), \hat{c}_i(\boldsymbol{s}, x), \theta, a, b\right). \tag{B.22}$$

A similar argument shows that (B.22) holds also for $i = 2$, with

$$\begin{aligned}
\hat{a}_2(\boldsymbol{s}, x) &= p_S(1-x)\{p_N s_1 + (1-p_N)[\tilde{p}_V s_2 + (1-\tilde{p}_V)s_3]\} \\
&\quad + (1-p_S)\{p_N s_4 + (1-p_N)[\tilde{p}_V s_5 + (1-\tilde{p}_V)s_6]\}, \\
\hat{b}_2(\boldsymbol{s}, x) &= p_S(1-x)[p_N(1-s_1) + (1-p_N)\tilde{p}_V(1-s_2)] \\
&\quad + (1-p_S)[p_N(1-s_4) + (1-p_N)\tilde{p}_V(1-s_5)] \text{ and} \\
\hat{c}_2(\boldsymbol{s}, x) &= (1-p_N)(1-\tilde{p}_V)[p_S(1-x)(1-s_3) + (1-p_S)(1-s_6)].
\end{aligned}$$

Expressions for $\hat{\psi}_U(\theta, \boldsymbol{s})$ and $\hat{\psi}_V(\theta, \boldsymbol{s})$ are derived in exactly the same way as with an all-or-nothing vaccine, yielding

$$\hat{\psi}_U(\theta, \boldsymbol{s}) = p_S^{3,6} G\left(D_U, \hat{a}(\boldsymbol{s}), \hat{b}(\boldsymbol{s}), \hat{c}(\boldsymbol{s}), \theta, a, 1\right)$$

and

$$\hat{\psi}_V(\theta, \boldsymbol{s}) = p_V^{-1} p_S^{2,5} \Delta_{0,p_N} G\left(D_V, \hat{a}(\boldsymbol{s}, x), \hat{b}(\boldsymbol{s}, x), \hat{c}(\boldsymbol{s}, x), \theta, a, b\right).$$

As usual, the above expressions simplify if $I \equiv \iota$. Eq. (B.19) then implies that, for $i = 3, 6$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \mathrm{e}^{-\theta\iota} f_{\tilde{D}(i)}\left(\hat{a}_i(\boldsymbol{s}) + \hat{b}_i(\boldsymbol{s})\mathrm{e}^{-a\lambda_G \iota} + \hat{c}_i(\boldsymbol{s})\mathrm{e}^{-\lambda_G \iota}\right),$$

which also holds for $i = 1, 4$, provided λ_G is replaced by $b\lambda_G$. Further, exploiting (B.20), (B.21) and similar equations for $\hat{\psi}_2(\theta, \boldsymbol{s})$, gives, for $i = 2, 5$,

$$\hat{\psi}_i(\theta, \boldsymbol{s}) = \tilde{p}_V^{-1}\mathrm{e}^{-\theta\iota} \Delta_{0,p_N} f_{\tilde{D}-1}\left(\hat{a}_i(\boldsymbol{s}, x) + \hat{b}_i(\boldsymbol{s}, x)\mathrm{e}^{-ab\lambda_G \iota} + \hat{c}_i(\boldsymbol{s}, x)\mathrm{e}^{-b\lambda_G \iota}\right).$$

Lastly, we also find that

$$\hat{\psi}_U(\theta, \boldsymbol{s}) = \mathrm{e}^{-\theta\iota} p_S^{3,6} f_{D_U}\left(\hat{a}(\boldsymbol{s}) + \hat{b}(\boldsymbol{s})\mathrm{e}^{-a\lambda_G \iota} + \hat{c}(\boldsymbol{s})\mathrm{e}^{-\lambda_G \iota}\right)$$

and

$$\hat{\psi}_V(\theta, s) = p_V^{-1}\mathrm{e}^{-\theta\iota} p_S^{2,5} \Delta_{0,p_N} f_{D_V}\left(\hat{a}(\boldsymbol{s}, x) + \hat{b}(\boldsymbol{s}, x)\mathrm{e}^{-ab\lambda_G \iota} + \hat{c}(\boldsymbol{s}, x)\mathrm{e}^{-b\lambda_G \iota}\right).$$

B.5 Local susceptibility set size

In this section we give formulae for the probability mass functions of the size of local susceptibility sets. The corresponding PGFs are then simple to calculate numerically. Consider first the model without vaccination and recall the definition of $M^{(n)}$ in Sect. 2.2.2. Then it follows directly from Ball (2000, Lemma 3.1) (see also Ball and Neal (2002, Lemma 3.1), which gives the same result but not in terms of Gontcharoff polynomials) that

$$P(M^{(n)} = k) = (n-1)_{[k]} q_{k+1}^{n-1-k} G_{k+1}(1|V) \quad (k = 0, 1, \ldots, n-1), \tag{B.23}$$

where $q_i = \phi_I(\lambda_L i)$ and $V = (q_{i+1} : i = 0, 1, \ldots)$.

Consider next the model with a non-random vaccine action and recall the definition of $\boldsymbol{M}_A^{(n,v)} = (M_{AU}^{(n,v)}, M_{AV}^{(n,v)})$ $(A \in \{U, V\})$ in Sect. 3.3. It is convenient to now use a different notation. For $\boldsymbol{n} = (n_U, n_V) \in \mathbb{Z}_+^2$, let $\hat{\boldsymbol{M}}_U^{(\boldsymbol{n})} = (\hat{M}_{UU}^{(\boldsymbol{n})}, \hat{M}_{UV}^{(\boldsymbol{n})})$, where $\hat{M}_{UU}^{(\boldsymbol{n})}$ and $\hat{M}_{UV}^{(\boldsymbol{n})}$ are the numbers of unvaccinated and vaccinated individuals, respectively, in the local susceptibility set of an unvaccinated individual, i^* say, who resides in a household containing n_U other unvaccinated individuals and n_V vaccinated individuals, so the household has size $n_U + n_V + 1$. Note that $\hat{M}_{UU}^{(\boldsymbol{n})}$ does *not* include the individual i^*. Define $\hat{\boldsymbol{M}}_V^{(\boldsymbol{n})} = (\hat{M}_{VU}^{(\boldsymbol{n})}, \hat{M}_{VV}^{(\boldsymbol{n})})$ similarly, but for a vaccinated individual. Hence, to connect with the notation in Sect. 3.3, $\boldsymbol{M}_U^{(n,v)} = \hat{\boldsymbol{M}}_U^{(n-1-v,v)}$ and $\boldsymbol{M}_V^{(n,v)} = \hat{\boldsymbol{M}}_V^{(n-v,v-1)}$. For $\boldsymbol{n} \geq \boldsymbol{0}$ and $A \in \{U, V\}$, the probability mass function of $\hat{\boldsymbol{M}}_A^{(\boldsymbol{n})}$ can be derived using the same argument as in the proof of Lemma 3.1 in Ball (2000). For brevity, we omit the details and just state the result.

For $\boldsymbol{j} = (j_U, j_V) \geq \boldsymbol{0}$, let $\boldsymbol{q}_{\boldsymbol{j}} = (q_{\boldsymbol{j}}^U, q_{\boldsymbol{j}}^V)$, where $q_{\boldsymbol{j}}^U = \phi_I(\lambda_L(j_U + bj_V))$ and $q_{\boldsymbol{j}}^V = \phi_I(a\lambda_L(j_U + bj_V))$. Let $\boldsymbol{1}_U = (1, 0)$ and $\boldsymbol{1}_V = (0, 1)$. Then, for $\boldsymbol{n} \geq \boldsymbol{0}$ and $A \in \{U, V\}$,

$$P\left(\hat{\boldsymbol{M}}_A^{(\boldsymbol{n})} = \boldsymbol{k}\right) = \boldsymbol{n}_{[\boldsymbol{k}]} \boldsymbol{q}_{\boldsymbol{k}+\boldsymbol{1}_A}^{\boldsymbol{n}-\boldsymbol{k}} G_{\boldsymbol{k}}(\boldsymbol{1}|\boldsymbol{V}^A) \quad (\boldsymbol{0} \leq \boldsymbol{k} \leq \boldsymbol{n} - \boldsymbol{1}_A), \tag{B.24}$$

where $\boldsymbol{V}^A = (\boldsymbol{v}_{\boldsymbol{j}}^A : \boldsymbol{j} \geq \boldsymbol{0})$ has components $\boldsymbol{v}_{\boldsymbol{j}}^A = \boldsymbol{q}_{\boldsymbol{j}+\boldsymbol{1}_A}^A$.

In the model with an all-or-nothing vaccine action we can calculate the mass function of $\boldsymbol{M}_A^{(n,v)}$ by conditioning on the number, v_S say, of the v vaccinations that are successful. The mass function of $\boldsymbol{M}_A^{(n-v_S,v-v_S)}$ can then be calculated as above, but with $q_{\boldsymbol{j}}^U = q_{\boldsymbol{j}}^V = \phi_I(\lambda_L(j_U + j_V))$.

B.6 Mean local epidemic size

The mean single-household epidemic final size $\mu_n(\lambda_L)$, defined in Sect. 2.2.1, is given in terms of Gontcharoff polynomials by

$$\mu_n(\lambda_L) = n - 1 - \sum_{i=1}^{n-1} (n-1)_{[i]} q_i^{n-i} G_{i-1}(1|V), \tag{B.25}$$

where $q_i = \phi_I(i\lambda_L)$ and $V = (q_{i+1} : i = 0, 1, \ldots)$; cf. Lefévre and Picard (1990, Corollary 3.3).

Finally, we consider the means $\mu^{(n,v)}(A, A')$ $(A, A' \in \{U, V\})$ defined in Sect. 3.3. Again it is convenient to use a different notation. For $\boldsymbol{n} = (n_U, n_V) \geq \boldsymbol{0}$, let $\hat{\mu}^{(\boldsymbol{n})}(U, U)$ and $\hat{\mu}^{(\boldsymbol{n})}(U, V)$ denote respectively the mean number of unvaccinated and vaccinated susceptibles that are ultimately infected in a single-household epidemic with 1 initial infective, who is unvaccinated, n_U unvaccinated susceptibles and n_V vaccinated susceptibles. Define $\hat{\mu}^{(\boldsymbol{n})}(V, U)$ and $\hat{\mu}^{(\boldsymbol{n})}(V, V)$ similarly, for when the initial infective is vaccinated. Thus, to connect with the notation in Sect. 3.3, for $A \in \{U, V\}$, $\mu^{(n,v)}(U, A) = \hat{\mu}^{(n-1-v,v)}(U, A)$ and $\mu^{(n,v)}(V, A) = \hat{\mu}^{(n-v,v-1)}(V, A)$. Suitable differentiation applied to Theorem 3.5 of Ball (1986), or Corollary 4.4 of Picard and Lefèvre (1990), yields that, for $\boldsymbol{n} \geq 0$ and $A, A' \in \{U, V\}$,

$$\hat{\mu}^{(\boldsymbol{n})}(A, A') = n_{A'} - \sum_{\boldsymbol{k}=\boldsymbol{1}_{A'}}^{\boldsymbol{n}} \boldsymbol{n}_{[\boldsymbol{k}]} \boldsymbol{q}_{\boldsymbol{k}}^{\boldsymbol{n}+\boldsymbol{1}_A-\boldsymbol{k}} G_{\boldsymbol{k}-\boldsymbol{1}_{A'}}(\boldsymbol{1}|\boldsymbol{V}^{A'}),$$

where $\boldsymbol{q}_{\boldsymbol{k}}$ $(\boldsymbol{k} \geq \boldsymbol{0})$ and $\boldsymbol{V}^A$ $(A \in \{U, V\})$ are as in (B.24).

References

Adams B (2016) Household demographic determinants of Ebola epidemic risk. J Theoret Biol 392:99–106

Amaral LAN, Scala A, Barthélémy M, Stanley H (2000) Classes of small-world networks. Proc Nat Acad Sci USA 97:11149–11152

Andersson H (1997) Epidemics in a population with social structures. Math Biosci 140:79–84

Andersson H (1998) Limit theorems for a random graph epidemic model. Ann Appl Probab 8(4):1331–1349

Ball FG (1986) A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. Adv Appl Probab 18(2):289–310

Ball FG (2000) Susceptibility sets and the final outcome of stochastic SIR epidemic models (Research Report No. 00-09). Division of Statistics, School of Mathematical Sciences, University of Nottingham

Ball FG, O'Neill PD (1999) The distribution of general final state random variables for stochastic epidemic models. J Appl Probab 36(2):473–491

Ball FG, Lyne OD (2001) Stochastic multitype SIR epidemics among a population partitioned into households. Adv Appl Probab 33(1):99–123

Ball FG, Lyne OD (2002) Optimal vaccination policies for stochastic epidemics among a population of households. Math Biosci 177–178:333–354

Ball FG, Neal PJ (2002) A general model for stochastic SIR epidemics with two levels of mixing. Math Biosci 180:73–102

Ball FG, Lyne OD (2006) Optimal vaccination schemes for epidemics among a population of households, with application to variola minor in Brazil. Stat Methods Med Res 15(5):481–497

Ball FG, Sirl DJ (2012) An SIR epidemic model on a population with random network and household structure and several types of individuals. Adv Appl Probab 44:63–86

Ball FG, Sirl DJ (2013) Acquaintance vaccination in an epidemic on a random graph with specified degree distribution. J Appl Probab 50:1147–1168 (*Correction (2015), ibid. 52:908.*)

Ball FG, Mollison D, Scalia-Tomba G (1997) Epidemics with two levels of mixing. Ann Appl Probab 7(1):46–89

Ball FG, Britton T, Lyne OD (2004) Stochastic multitype epidemics in a community of households: estimation and form of optimal vaccination schemes. Math Biosci 191(1):19–40
Ball FG, Sirl DJ, Trapman P (2009) Threshold behaviour and final outcome of an epidemic on a random network with household structure. Adv Appl Probab 41(3):765–796
Ball FG, Sirl DJ, Trapman P (2010) Analysis of a stochastic SIR epidemic on a random network incorporating household structure. Math Biosci 224(2):53–73 (*Erratum (2010), ibid*. 225:81.)
Ball FG, Sirl DJ, Trapman P (2014) Epidemics on random intersection graphs. Ann Appl Probab 24:1081–1128
Ball FG, Pellis L, Trapman P (2016) Reproduction numbers for epidemic models with households and other social structures II: comparisons and implications for vaccination. Math Biosci 274:108–139
Bartoszyński R (1972) On a certain model of an epidemic. Zastos Mat 13:139–151
Becker NG, Starczak DN (1997) Optimal vaccination strategies for a community of households. Math Biosci 139:117–132
Becker NG, Starczak DN (1998) The effect of random vaccine response on the vaccination coverage required to prevent epidemics. Math Biosci 154:117–135
Bollobás B (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. Euro J Comb 1(4):311–316
Britton T, Janson S, Martin-Löf A (2007) Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. Adv Appl Probab 39(4):922–948
Cohen R, Havlin S, ben Avraham D (2003) Efficient immunization strategies for computer networks and populations. Phys Rev Lett 91(24):247901
Diekmann O, de Jong MCM, Metz JAJ (1998) A deterministic epidemic model taking account of repeated contacts between the same individuals. J Appl Probab 35(2):448–462
Feld SL (1991) Why your friends have more friends than you do. Am J Sociol 96:1464–1477
Haccou P, Jagers P, Vatutin V (2005) Branching processes: variation, growth, and extinction of populations. Cambridge University Press, Cambridge
Halloran ME, Haber M, Longini I (1992) Interpretation and estimation of vaccine efficacy under heterogeneity. Am J Epidemiol 136:328–343
Halloran ME, Longini I, Nizam A, Yang Y (2002) Containing bioterrorist smallpox. Science 298:1428–1432
Halloran ME, Ferguson N, Eubank S, Longini I, Cummings D, Lewis B, Cooley P (2008) Containing bioterrorist smallpox. Proc Natl Acad Sci USA 105:4639–4644
Keeling MJ, Ross JV (2015) Optimal prophylactic vaccination in segregated populations: when can we improve on the equalising strategy? Epidemics 11:7–13
Lefèvre C, Picard P (1990) A nonstandard family of polynomials and the final size distribution of Reed-Frost epidemic processes. Adv Appl Probab 22(1):25–48
Mode CJ (1971) Multitype branching processes, theory and applications. American Elsevier Publishing Co. Inc, New York
National Institute of Population Studies [Pakistan] and ICF International (2013) Pakistan demographic and health survey 2012–13. http://www.dhsprogram.com/pubs/pdf/FR290/FR290.pdf. Accessed 24 June 2015
Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66(1):016128
Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. Phys Rev E 64(2):026118
Office for National Statistics [UK] (2011) General lifestyle survey. http://www.ons.gov.uk/ons/rel/ghs/general-lifestyle-survey/2011/index.html. Accessed 24 June 2015
Pellis L, Ferguson NM, Fraser C (2008) The relationship between real-time and discrete-generation models of epidemic spread. Math Biosci 216(1):63–70
Pellis L, Ball F, Trapman P (2012) Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R_0. Math Biosci 235(1):85–97
Picard P, Lefèvre C (1990) A unified analysis of the final size and severity distribution in collective Reed-Frost epidemic processes. Adv Appl Probab 22(2):269–294
Scalia-Tomba G (1986) Asymptotic final size distribution of the multitype Reed-Frost process. J Appl Probab 23(3):563–584
Watson RK (1972) On an epidemic in a stratified population. J Appl Probab 9:659–666