



Epidemics on Networks with Large Initial Conditions or Changing Structure

Joel C. Miller*

School of Mathematical Sciences, School of Biological Sciences, and Monash Academy for Cross & Interdisciplinary Mathematics, Monash University, Melbourne, Victoria, Australia

Abstract

In this paper we extend previous work deriving dynamic equations governing infectious disease spread on networks. The previous work has implicitly assumed that the disease is initialized by an infinitesimally small proportion of the population. Our modifications allow us to account for an arbitrarily large initial proportion infected. This helps resolve an apparent paradox in earlier work whereby the number of susceptible individuals could increase if too many individuals were initially infected. It also helps explain an apparent small deviation that has been observed between simulation and theory. An advantage of this modification is that it allows us to account for changes in the structure or behavior of the population during the epidemic.

Citation: Miller JC (2014) Epidemics on Networks with Large Initial Conditions or Changing Structure. PLoS ONE 9(7): e101421. doi:10.1371/journal.pone.0101421

Editor: Caroline Colijn, Imperial College London, United Kingdom

Received: March 31, 2014; **Accepted:** June 6, 2014; **Published:** July 8, 2014

Copyright: © 2014 Joel C. Miller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: JCM was supported in part by 1) the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, Department of Homeland Security and the Fogarty International Center, National Institutes of Health, and 2) the Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard School of Public Health under Award Number U54GM088558 from the National Institute Of General Medical Sciences. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The author has declared that no competing interests exist.

* Email: joel.c.miller.research@gmail.com

Introduction

The mathematical study of infectious disease spread has contributed significantly to our ability to design effective interventions to reduce disease spread. Most of the earliest models were based on the assumption that disease transmission occurs as a Poisson process and each transmission reaches an individual chosen randomly from the population. This implicitly assumes that partnership duration is very brief. These models have been modified to account for a number of different effects, such as demographic groups [1].

More recently, attempts have been made to incorporate the “network” structure of the population (see, *e.g.*, [2]). Typically these focus on trying to understand the role played by “high-degree” individuals (those individuals with many contacts), and they come in one of two flavors: they either continue the assumption of fleeting partnerships (the timescale of individual transmissions is long compared to the timescale of individual partnerships) [1,3–6], or they take the opposite limit in which the partnership network is static (the timescale of the epidemic is short compared to the timescale of individual partnerships) [7–14]. These two approaches do not address the intermediate regimes.

Recent work has shown that for susceptible-infectious-recovered (SIR) models, it is possible to unify these two approaches with an “edge-based compartmental model” (EBCM) that allows partnership duration to range continuously from zero to infinite [15–17] [for susceptible-infectious-susceptible (SIS) models, the picture is more complicated, see for example [18]]. The resulting models are

low-dimensional and contain many standard models as special cases [17]. Unfortunately, these models are derived under the assumption that the initial proportion infected is infinitesimally small (while the absolute number infected is sufficiently large that the dynamics are deterministic). It is assumed that by the time the equations are used, any early transients have died away. A consequence of this assumption is that the models break down if $\mathcal{R}_0 < 1$ (that is, if the average number of infections caused by an infected individual early in the epidemic is less than 1) or if the initial proportion infected is not negligible.

The failure if the initial proportion infected is not negligible was observed by [14]. This paper used an early (static network) version of the equations of [15] from [7] and compared them with simulation. A small discrepancy in final sizes was noted. This discrepancy was not present for equations of [19], a system requiring $\mathcal{O}(M)$ equations where M is the maximum degree, or for another system introduced in [14] which required $\mathcal{O}(M^2)$ equations.

In the remainder of this paper, we generalize the EBCM equations for the spread of infectious disease assuming a finite proportion of the population is initially infected. We test the resulting equations against simulations, analyze their predictions for different disease scenarios, and investigate the cause of the discrepancies found in previous work. For simplicity we focus on the static network limit. The method we introduce is straightforward to adapt to dynamic networks.

Analysis

We modify the approach of [15] which assumed an infinitesimal initial proportion infected. We adapt the approach to consider a wide range of possible initial conditions. We assume that the dynamics of the epidemic may be treated as deterministic, which means we assume the population is very large and the initial number infected is large enough for the epidemic to behave deterministically. The assumption that behavior is deterministic may be understood qualitatively as equivalent to the claim that no individual has a large enough effect to alter the dynamics of the disease at the population-scale: the time (or even if) a single given individual becomes infected has a negligible impact on the proportion infected. If stochastic effects are still important but $R_0 > 1$, then these equations may become accurate at a later time once sufficient numbers are infected.

We assume the population consists of $N \gg 1$ individuals. Each is assigned a degree k (independently of degrees of other individuals) with probability $P(k)$ where P defines a probability distribution on the non-negative integers. The network is wired together using the ‘‘Configuration Model’’ (or ‘‘Molloy-Reed’’) approach [11,20]: each individual is assigned a number of stubs (or half-edges) equal to its degree. Pairs of stubs are then wired together to form edges/partnerships. It is likely that this algorithm produces a handful of self-loops or repeated edges, but although they may be present, their density (*i.e.*, the probability a given individual is involved in a self-loop or repeated edge) goes to zero like $1/N$.

We define a *test individual* u to be a random individual chosen at time 0. Because we assume that the spread is deterministic, this means that the probability u is in a given state is equal to the proportion of the population in that state. So we focus on calculating the probability u is susceptible, infected, or recovered. We modify u so that it does not transmit to any of its partners if ever infected. This assumption does not affect the probability u is in any given state, but it does prevent a correlation between the statuses of different partners which would be caused by infection traveling through u . This allows us to treat the partners of u as independent and so each partner of u may independently transmit to u . It is important to note that this assumption has no impact on the probability u is in any given state and therefore, it does not affect our calculation of the proportion of the population in each state. Further discussion of the test individual is in [21].

Variables and Parameters

We introduce our variables, our parameters, and their definitions in table 2. The starting point is the test individual u . The remaining variables and parameters can be broadly divided into four groups.

- S , I , and R denote the proportion of the population in each state, or equivalently the probability that the test individual u is in each state.
- θ , ϕ_S , ϕ_I , and ϕ_R give the probability a partner of u has a given status and the probability the partner has transmitted to u : θ is the probability the partner has not transmitted to u and ϕ_S , ϕ_I , and ϕ_R give the probability the partner has not transmitted to u and is susceptible, infected, or recovered respectively.
- $P(k)$ tells us the probability a random individual has degree k , while $S(k,0)$ tells us the probability a random individual of degree k is initially susceptible. The function $\psi(x) = \sum_k S(k,0) P(k)x^k$ encodes P and S . We define $\langle K \rangle = \sum_k kP(k)$ to be the average degree.
- We have two disease parameters to consider: β , the transmission rate, and γ , the recovery rate.

Given our definitions, $\psi(\theta(t))$ is the probability u is susceptible at time t . By noting that $\psi(\theta(t)) = S(t)$, we will be able to close our system of equations.

The main distinction between this approach and the previous EBCM approach [15] is that we use just the initially susceptible individuals to define ψ while the earlier work took $\psi(x) = \sum_k P(k)x^k$, the probability generating function for the degree distribution. The earlier work then assumed the disease had already been spreading prior to time 0 and defined θ to be the probability that a partner has never transmitted [so $\theta(0) < 1$] whereas here we take θ to be the probability that a partner has not transmitted given that it had not prior to time 0 [so $\theta(0) = 1$].

Equation Derivation

We will find a closed system of equations based on these variables. We begin by looking at $S(t)$. If the test individual u has degree k and is susceptible at $t=0$, then the probability it is susceptible at some later time is $\theta(t)^k$. If we do not know k or whether u is susceptible at $t=0$, then the probability u is susceptible at time t is the sum over all k of the product of the probability u is initially susceptible $S(k,0)$ with the probability u is still susceptible θ^k . We have $S(t) = \sum_k S(k,0)P(k)\theta(t)^k = \psi(\theta(t))$. Thus we conclude

$$S(t) = \psi(\theta(t))$$

We know that R solves $\dot{R} = \gamma I$. We also have a conservation rule that $S + I + R = 1$, so $I = 1 - S - R$. Thus our equations are

$$S = \psi(\theta)$$

$$I = 1 - S - R$$

$$\dot{R} = \gamma I$$

Assuming $\theta(t)$ is known, then this system with an initial condition for R completely defines S , I , and R . This is shown in the flow diagram in figure 1. Other formulations are possible, for example $S = \psi(\theta)$, $\dot{I} = -\psi'(\theta)\dot{\theta} - \gamma I$, $\dot{R} = \gamma I$. However we find our system to be preferable because it minimizes the number of differential equations.

In order to close this system of equations we need an equation giving θ . Recall that $\theta(t)$ is the probability a partner v of u that had not yet transmitted to u by time 0 has still not transmitted by time t . This is broken into three disjoint sub-compartments

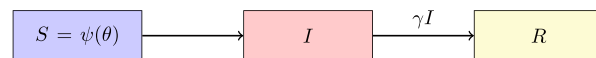


Figure 1. Flow diagram showing the flux of individuals between the different compartments. Because we have an explicit expression for S , if we know θ we do not need to explicitly determine the flux from S to I .

doi:10.1371/journal.pone.0101421.g001

$\theta = \phi_S + \phi_I + \phi_R$ based on whether v has not transmitted and is susceptible, infected, or recovered. Because ϕ_I is the probability v has not transmitted and is infected, it is straightforward to see that $\dot{\theta} = -\beta\phi_I$. So if we can find ϕ_I in terms of θ , then we arrive at a single equation for θ , which can be used to provide θ for the S , I , and R equations.

To do this, we use the fact that $\phi_I = \theta - \phi_S - \phi_R$ and find ϕ_S and ϕ_R in terms of θ . We turn to figure 2. The recovery rate is γ and the transmission rate is β , so we have $\dot{\phi}_R = -\gamma\dot{\theta}/\beta$. We can integrate this, and using the fact that $\theta(0)=1$ we find $\phi_R = \gamma(1-\theta)/\beta + \phi_R(0)$. To find ϕ_S in terms of θ , we note that the probability u has an edge to a node that is susceptible at time $t=0$ is $\phi_S(0)$. The probability the susceptible partner has degree k is $kP(k)S(k,0)/\sum_k kP(k)S(k,0)$, so the probability an initially susceptible partner is susceptible at some later time is $\sum_k kP(k)S(k,0)\theta^{k-1}/\sum_k kP(k)S(k,0) = \psi'(\theta)/\psi'(1)$. Thus $\dot{\phi}_S(t) = \phi_S(0)\psi'(\theta(t))/\psi'(1)$. We arrive at

$$\dot{\phi}_I = \theta - \phi_S(0) \frac{\psi'(\theta)}{\psi'(1)} - \frac{\gamma}{\beta}(1-\theta) - \phi_R(0)$$

and $\dot{\theta} = -\beta\phi_I$ becomes

$$\dot{\theta} = -\beta\theta + \beta\phi_S(0) \frac{\psi'(\theta)}{\psi'(1)} + \gamma(1-\theta) + \beta\phi_R(0)$$

with $\theta(0)=1$. This completes our system.

Our final closed system of equations is

$$\dot{\theta} = -\beta\theta + \beta\phi_S(0) \frac{\psi'(\theta)}{\psi'(1)} + \gamma(1-\theta) + \beta\phi_R(0) \tag{1}$$

$$\dot{R} = \gamma I, \quad S = \psi(\theta), \quad I = 1 - S - R \tag{2}$$

where $\theta(0)=1$, and $R(0)$ is given by the initial conditions. These equations lead to earlier equations of [7,15,22] if $\mathcal{R}_0 > 1$ and $1 - \phi_S(0)$, $\phi_I(0)$, $\phi_R(0)$ and $1 - \theta(0)$ are all infinitesimally small. If $\mathcal{R}_0 < 1$, the error caused by the approximations $\phi_R(0)=0$ and $\phi_S(0)=1$ is comparable to the actual number infected.

Final size relation. The final size relation assuming small initial condition is well-known [11]. The final size relation for

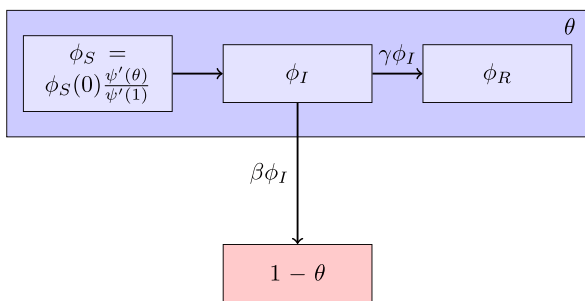


Figure 2. Flow diagram for the flux of partners through different states. The top three boxes ϕ_S , ϕ_I , and ϕ_R make up θ and represent the different states the partner can be in if it has not transmitted. The lower box $1 - \theta$ is the probability the partner has transmitted.

doi:10.1371/journal.pone.0101421.g002

larger initial conditions has recently been found [21] in a more general case not assuming constant transmission and recovery rates. It can be derived easily for this model by setting $\dot{\theta}=0$. We find

$$\theta(\infty) = \frac{\beta \left[\phi_S(0) \frac{\psi'(\theta(\infty))}{\psi'(1)} + \phi_R(0) \right] + \gamma}{\beta + \gamma} \tag{3}$$

$$R(\infty) = 1 - \psi(\theta(\infty)) \tag{4}$$

Model Validation

We now compare our model with simulations for populations that satisfy the Configuration Model/Molloy-Reed model assumptions. Although an earlier version of these equations was found to have minor discrepancies [14], we show that once we appropriately account for the initial condition, the calculation becomes correct.

Final Size Comparison. To show that our new equations accurately calculate the impact of the initial conditions, we first consider epidemic spread in networks with the same degree distribution as in [14] (table 1), but with varying numbers initially infected and varying population sizes. We then consider the impact of selecting high or low degree nodes as the earliest infected individuals, using networks whose degree distributions more clearly show the impact of biased selection of the initial individuals.

We run a large number of simulations for each number of initial infections. For each simulation we generate a new network. Our simulation technique is similar to those recently described by [8,19,23]. In the Configuration Model framework, each node is assigned a degree, nodes are given stubs (or half-edges), and then stubs are randomly paired together. In the simulations we use, each node is assigned a degree, nodes are given stubs, and then the disease begins to spread in the network before stubs are paired. Each time the disease transmits along a stub that stub is joined to a randomly selected as-yet-unpaired stub. If the partner is susceptible, then it becomes infected. If not, nothing happens. Once stubs are paired they remain in their edge. This approach is equivalent to constructing the network in advance and then following the disease, but it is more efficient computationally because it only constructs those parts of the network the disease traces.

Randomly selected initial infections. We first consider varying numbers of randomly chosen initial infected individuals. In figure 3 we take the degree distribution from table 1.

We randomly select a proportion ρ of the population to initially infect. We have $S(k,0)=1-\rho$ for all k , so $\psi(x)=(1-\rho)\sum_k P(k)x^k$. Similarly we have $\phi_S(0)=1-\rho$. Because the epidemic begins with no recovered individuals, we take $\phi_R(0)=0$. We take $\beta=0.1$ and $\gamma=0.2$ (though all that matters for the final size is their ratio).

We take populations of 100, 1000, and 10000 and perform many simulations. We compare the final sizes observed with the final size relation of equations (3) and (4). The equations are derived in the infinite population limit, but in figure 3 we see that even with populations of only 100 they give a good prediction of the observed behavior. As the population size increases, the

Table 1. The degree distribution used in simulations in [14].

k	$P(k)$
1	18.118×10^{-3}
2	72.536×10^{-3}
3	145.222×10^{-3}
4	194.589×10^{-3}
5	195.962×10^{-3}
6	156.857×10^{-3}
7	105.280×10^{-3}
8	59.713×10^{-3}
9	30.066×10^{-3}
10	21.657×10^{-3}

doi:10.1371/journal.pone.0101421.t001

distribution becomes narrower and the simulations collapse more tightly around the prediction.

Biased initial infections. To show that the approach we have derived can also be applied to cases where the initial infected individuals are selectively chosen based on their degree, we use a different degree distribution that helps highlight the effect. We take $P(1)=P(9)=1/2$. We consider two options. In the first approach, individuals with higher degree are preferentially selected. To do the selection, we choose an individual with probability proportional to the square of its degree, and infect it. We repeat this until a proportion ρ of the population is infected. In the second approach individuals are chosen with probability proportional to the square of their inverse degree until a proportion ρ is infected. We take $\beta=0.1$ and $\gamma=0.6$.

Using these rules, we clearly see that $S(k,0)$ is not uniform. Instead, for the case where individuals are selected with probability proportional to their squared degree, we find that $S(k,0)=\alpha^{k^2}$ where α solves $\sum_k P(k)\alpha^{k^2}=1-\rho$. We find $\phi_S(0)=\sum_k kS(k,0)P(k)\alpha^{k^2}/\sum_k kP(k)$. In the case where individuals are selected with probability inversely proportional to their squared degree, we find that $S(k,0)=\alpha^{1/k^2}$ where α solves $\sum_k P(k)\alpha^{1/k^2}=1-\rho$, and $\phi_S(0)=\sum_k kS(k,0)P(k)\alpha^{1/k^2}/\sum_k kP(k)$.

We compare predictions and simulations in populations of 1000 individuals in figure 4. In the limit of a negligible initial proportion infected, the final size of epidemics in these networks is about 4%. As we increase the number of initially infected individuals, we increase the final size because of these individuals and because of the additional infections they lead to. At small amounts, increasing the initial number of high degree nodes has a much larger impact on the final size because they cause more additional infections. However, as the amount of infection initially present is increased this effect becomes less important: the high degree individuals would become infected anyway. So the largest gain in final size comes from infecting low degree individuals who would not receive an infection from their partners. The ‘‘kinks’’ that occur just above 50% initially infected are because effectively all individuals of high (left) or low (right) degree are initially infected.

Dynamic Calculation. We now look at the performance of the dynamic equations. The dynamic prediction is more easily affected by noise than the final size prediction, so we use larger population sizes. We again take the degree distribution of table 1. We begin with 5% infected, either randomly chosen, or chosen as before proportional to the square of the degree. A comparison of individual simulations with theoretical predictions is in figure 5. The theory accurately predicts the dynamics of epidemics in large populations, but there is significant stochasticity in smaller

Table 2. The variables we need to calculate the epidemic dynamics. In all of these u is a test individual: randomly chosen from the population and modified so that it cannot infect others, although it can become infected.

Variable/parameter	Definition
Test Individual u	A randomly member of the population chosen at time $t=0$ who is prevented from transmitting to its partners.
$S(t)$	The proportion of the entire population that is susceptible.
$I(t)$	The proportion of the entire population that is infected.
$R(t)$	The proportion of the entire population that is recovered.
$\theta(t)$	The probability a random partner v of u that did not transmit to u by time 0 has not transmitted to u by time t .
$\phi_S(t)$	The probability a random partner v that did not transmit to u by time 0 is susceptible at time t .
$\phi_I(t)$	The probability a random partner v that did not transmit to u by time 0 is infected at time t but has not transmitted to u .
$\phi_R(t)$	The probability a random partner v that did not transmit to u by time 0 is recovered at time t and never transmitted to u .
$P(k)$	The probability an individual has degree k .
$\langle K \rangle = \sum_k kP(k)$	The average degree.
$S(k,0)$	The probability an individual with degree k is initially susceptible.
$\psi(\theta(t)) = \sum_k S(k,0)P(k)\theta(t)^k$	The probability that the test individual u is susceptible at time t . In a large population this should equal $S(t)$.
β	the per-edge transmission rate
γ	the per-individual recovery rate

doi:10.1371/journal.pone.0101421.t002

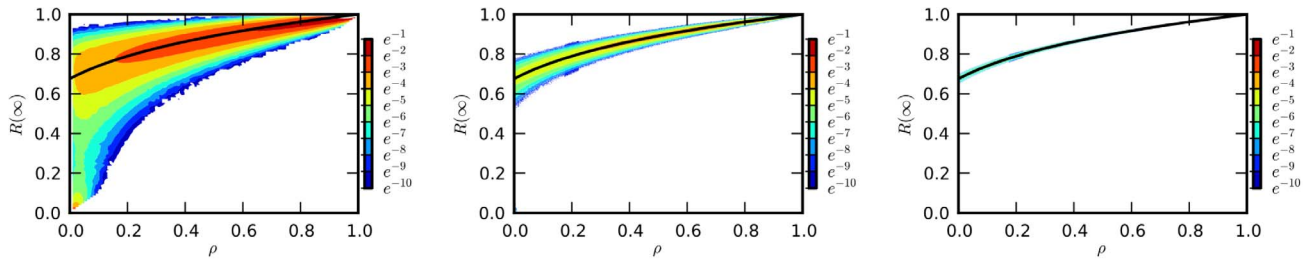


Figure 3. Results of simulations for $N = 100, 1000,$ and 10000 individuals. The solid curve gives our prediction for the final sizes of epidemic in a large population. Colors are log scale giving probability of that particular epidemic size. Each simulation is for a new network generated using the $P(k)$ from table 1, with $\beta=0.1$ and $\gamma=0.2$. We randomly select a proportion ρ of the population to initially infect and compare final size with the prediction of theory. The number of simulations for each ρ for $N=100, 1000,$ and 10000 was 50000, 11500, and 3000 respectively. To show that this is sufficient to resolve the distribution, for $.175 < \rho < .225$ there were 2000000, 50000 and 10000 simulations performed respectively for each N . This only slightly improves the tails of the distribution. Note that when the initial number (not proportion) of infections is small, a large fraction of simulations end without an epidemic. doi:10.1371/journal.pone.0101421.g003

populations. The theory is able to capture the impact of the biased selection of initial infections.

Intervention Impact

We can use our equations to compare the impact of several interventions. We consider an epidemic spreading in the population, and at some intermediate time we introduce a change in the disease or population. Because the system changes at a time with a non-negligible amount of infection in the population, the equations derived assuming a negligible proportion infected fail.

Consider a population in which $P(4)=P(5)=P(6)=1/3$. Assume we initially infect a small, randomly chosen proportion of the population, ρ at $t=0$. Thus we have $\psi(x)=(1-\rho)(x^4+x^5+x^6)/3$, $\phi_S(0)=1-\rho$, $\phi_I(0)=\rho$, and $\phi_R(0)=0$. We take $\beta=1$ and $\gamma=1/2$.

We consider three interventions that may be introduced at time t_1 . All are aimed at “halving” the transmission rate, but they do this in different ways. In mass-action based models, these would all have the same effect. We can clearly identify differences using our approach.

- An intervention that reduces β by a factor of 2.
- An intervention that reduces β so that per-contact transmission probability $\beta/(\beta+\gamma)$ is reduced by a factor of 2.

- An intervention that eliminates half of the partnerships randomly.

The distinction between the first two comes from the fact that partnerships have duration, so reducing β by a factor of 2 does not reduce the infection probability by a factor of 2. The expected number of transmissions an individual sends to a given partner is β/γ , but once the partner is infected, the subsequent transmissions are irrelevant. If we use mass action assumptions however, each transmission is to a replacement partner. In each case, halving β halves the total number of transmissions, but when partnerships have nonzero duration, a larger proportion of transmissions have no effect. The probability of transmitting at least once along a static partnership is $\beta/(\beta+\gamma)$. So to reduce infection probability by a given factor requires a larger reduction to β . Note that the work of [24] suggests that in Configuration Model networks the final size of our second and third intervention will be the same, (but that in clustered networks it will be different).

We will demonstrate our approach in all three cases, restarting the calculations when the intervention is put into place. In all cases, this allows us to use the conditions at t_1 to predict the final size. We take $\psi_0(x)=\psi(x)$, θ_0 , $\phi_{S,0}$, $\phi_{I,0}$, and $\phi_{R,0}$ to correspond to time less than t_1 . We use a subscript of 1 for times after t_1 . We

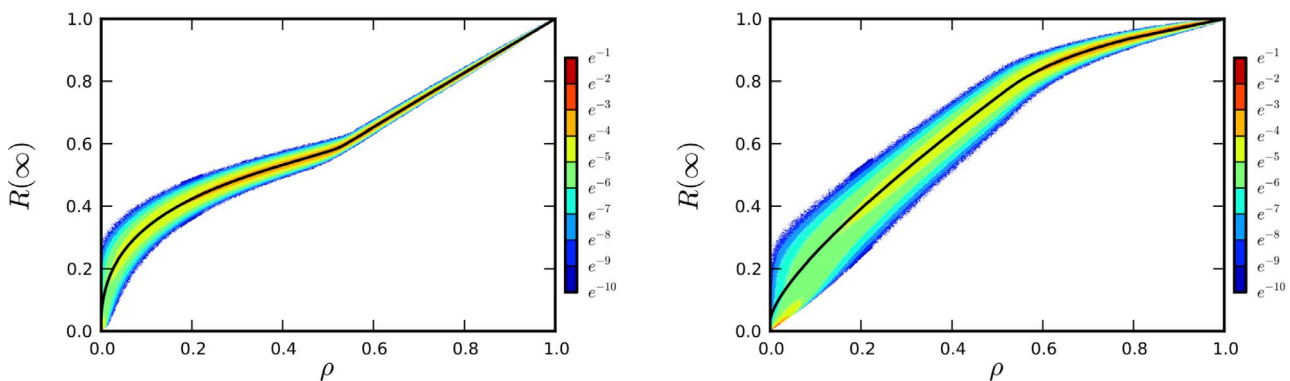


Figure 4. Epidemic final sizes in population of 1000 individuals with half having degree 9 and half with degree 1. The disease parameters are $\beta=0.1$, $\gamma=0.6$. Results of simulations having initial infections chosen with probability proportional to square of degree (left) or inverse square of degree (right). For each initial number of infections, 22500 simulations were performed, each with a different network. For the range $0.175 < \rho < 0.225$, 200000 simulations were performed to give insight into how well resolved the distribution is. Note that for small numbers of initial infections, epidemics are less likely when the lower degree individuals are chosen. doi:10.1371/journal.pone.0101421.g004

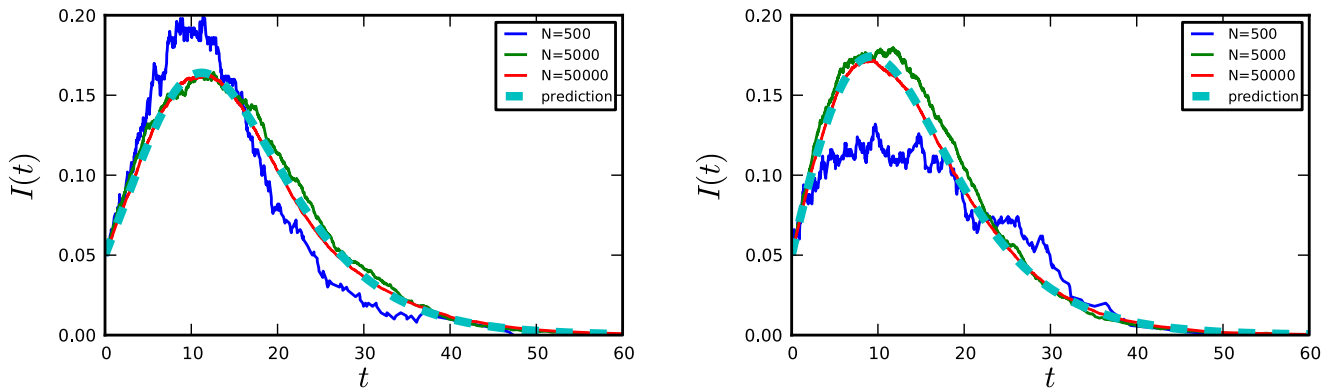


Figure 5. A comparison of the observed and predicted number of infections from simulations. Left: 5% initially infected, chosen randomly from the population. Right: 5% initially infected, chosen with probability proportional to squared degree. Each simulation curve represents a single simulation of the given size. As population size increases, the results converge to the theoretical prediction. doi:10.1371/journal.pone.0101421.g005

solve the original equations, and then use the results to initialize the second set of variables.

Case 1. We begin by reducing β by a factor of 2 at time t_1 . Until time t_1 , we are solving the original equations. By solving the original system until t_1 we have $\theta_0(t_1)$. The probability an individual of degree k is susceptible at time t_1 is $S(k, t_1) = S(k, 0)\theta_0(t_1)^k$. So our new $\psi(x)$ is $\psi_1(x) = \sum_k P(k) S(k, 0)\theta_0(t_1)^k x^k = \psi(\theta_0(t_1)x)$. We take our new θ_1 to have $\theta_1(t_1) = 1$. The intervention we are doing has no impact on the probability a partner is in any given state. $\phi_S, \phi_I,$ and ϕ_R keep the same proportion, but are scaled up to sum to 1 so each is scaled by $\theta_0(t_1)$. For example, $\phi_{S,1}(t_1) = \phi_{S,0}(t_1)/\theta_0(t_1) = \phi_{S,0}(0)\psi'_0(\theta_0(t_1))/\psi'_0(1)\theta_0(t_1)$.

We restart the solutions with these new values.

Case 2. The total probability of transmitting to a partner is $\beta/(\beta + \gamma)$. For this intervention we change β so that $\beta/(\beta + \gamma)$ is reduced by a factor of 2 at time t_1 . This proceeds exactly as above except that the new value of β must be smaller.

Case 3. When we delete half the edges at random, we do not affect the probability that a random partner is in any given state. So the ϕ variables rescale in the same way as for changing β in the previous cases. However, ψ undergoes a more significant change. As a starting point, consider $P_1(k_1)$, the probability an individual has degree k_1 after edges are deleted. This depends on $P_0(k_0)$, the probability of having k_0 edges prior to deletion. The relation is

$$P_1(k_1) = \sum_{k_0} P_0(k_0) \binom{k_0}{k_1} \left(\frac{1}{2}\right)^{k_0}$$

The probability the individual has degree k_1 and is susceptible at time t_1 is

$$Q(k_1) = \sum_{k_0} P_0(k_0) S(k_0, 0) \binom{k_0}{k_1} \left(\frac{\theta_0(t_1)}{2}\right)^{k_0}$$

So if we restart the calculations at $t = t_1$ we have $S(k_1, t_1) = Q(k_1)/P_1(k_1)$ and

$$\begin{aligned} \psi_1(x) &= \sum_{k_1} S(k_1, t_1) P_1(k_1) x^{k_1} \\ &= \sum_{k_1} Q(k_1) x^{k_1} \\ &= \sum_{k_1} \sum_{k_0} P_0(k_0) S(k_0, 0) \binom{k_0}{k_1} \left(\frac{\theta_0(t_1)}{2}\right)^{k_0} x^{k_1} \\ &= \sum_{k_0} P_0(k_0) S(k_0, 0) \theta_0(t_1)^{k_0} \sum_{k_1} \binom{k_0}{k_1} \left(\frac{1}{2}\right)^{k_0 - k_1} \left(\frac{x}{2}\right)^{k_1} \\ &= \sum_{k_0} P_0(k_0) S(k_0, 0) \theta_0(t_1)^{k_0} \left(\frac{x+1}{2}\right)^{k_0} \\ &= \psi_0\left(\theta_0(t_1) \left[\frac{1+x}{2}\right]\right) \end{aligned}$$

(in general if we delete edges with probability p and keep with probability $q = 1 - p$, then the new function is $\psi_1(x) = \psi_0([\theta_0(t_1)][p + qx])$). Using this new $\psi_1(x)$, the same system of equations holds.

Figure 6 compares these strategies. As anticipated, the final sizes resulting from cases 2 and 3 are identical, regardless of the time of intervention (indeed this is straightforward to show from the final size relation). However, we see that the dynamics are significantly different.

There are other ways we could capture these interventions mathematically. We note that in both case 1 and case 2, we could also capture the intervention by simply using a step change in β . Case 3 could be captured by reducing ϕ_S, ϕ_I and ϕ_R each by half and placing the other half into a new inactive compartment ϕ_X for the deleted edges. However, each of these is *ad hoc* and dependent on the precise details of the case. Using the approach presented above, we have a standard approach that will apply across a wide range of interventions.

Bifurcation analysis

We try to gain a better understanding of the epidemic threshold and what happens to the final size as the initial proportion infected is increased. Consider the final size relation found from

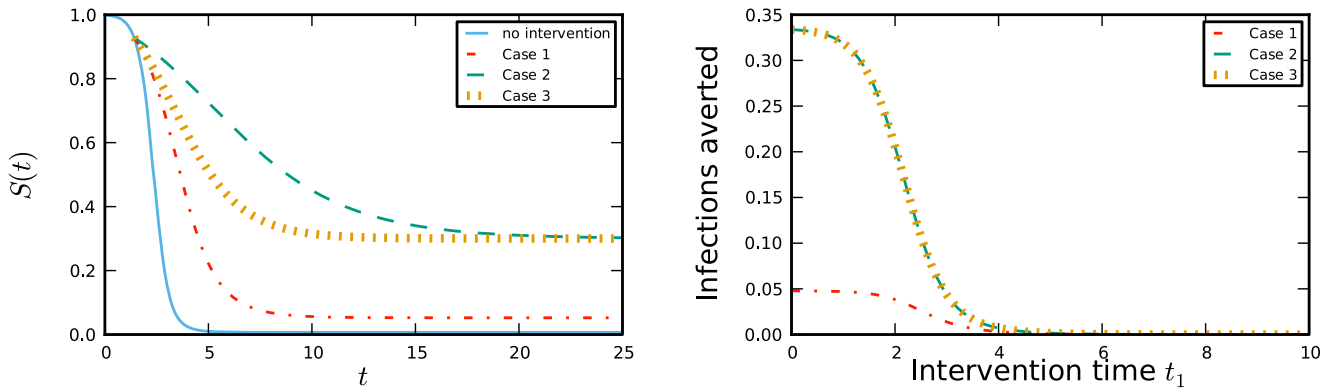


Figure 6. The impact of interventions. Epidemics begin at $t=0$ with 0.001 of the population infected. Left: epidemic curve without interventions, and with each intervention introduced at time $t_1=1.5$. Right: horizontal axis is t_1 , showing final effectiveness if interventions introduced at different times
doi:10.1371/journal.pone.0101421.g006

$$\theta = \frac{\beta \left[\phi_S(0) \frac{\psi'(\theta)}{\psi'(1)} + \phi_R(0) \right] + \gamma}{\beta + \gamma} \quad \mu = 0 \quad \text{or} \quad \mu \approx \frac{2}{\psi'''(1)} \left[\frac{\psi'(1)(\beta + \gamma)}{\beta} - \psi''(1) \right] \quad (5)$$

If $\phi_S(0) + \phi_R(0) = 1$, then we find that $\theta = 1$ is a solution to these equations. Biologically this states that if there is no infection initially [$\phi_I(0) = 0$] there will be no infection later. To study what happens when $\phi_I(0) > 0$ (but possibly arbitrarily small), we begin by first analyzing the structure of the dynamical equation for θ under the assumption that $\phi_S(0) = 1$ and $\phi_R(0) = 0$, taking $\theta < 1$. These assumptions contradict our initial conditions, but understanding the bifurcation in this system first will lead to an easier understanding of the full system with $\phi_I(0) > 0$.

If $\phi_S(0) = 1$ and $\phi_R(0) = 0$, the differential equation for θ becomes

$$\dot{\theta} = -\beta\theta + \beta \frac{\psi'(\theta)}{\psi'(1)} + \gamma(1 - \theta)$$

which has $\dot{\theta} = 0$ whenever

$$\theta = \frac{\beta \frac{\psi'(\theta)}{\psi'(1)} + \gamma}{\beta + \gamma}$$

Clearly $\theta = 1$ is an equilibrium. Close to $\theta = 1$, we write $\theta = 1 + \mu$, so $\psi'(\theta)/\psi'(1) = 1 + \mu\psi''(1)/\psi'(1) + \mu^2\psi'''(1)/2\psi'(1) + \mathcal{O}(\mu^3)$. The value of $\psi''(1)/\psi'(1)$ plays a key role in this analysis. Its biological interpretation comes from looking at a random infected individual's partner v early in the epidemic. If v is infected by that random infected individual, then $\psi''(1)/\psi'(1)$ is the expected number of additional partners v has (its *excess degree*).

Substituting into the equation for the equilibrium we have

$$1 + \mu = 1 + \mu\beta \frac{\psi''(1) + \mu\psi'''(1)/2 + \mathcal{O}(\mu^2)}{\psi'(1)(\beta + \gamma)}$$

which yields

So there is a bifurcation as the bracketed term passes through zero, when $(\beta + \gamma)/\beta = \psi''(1)/\psi'(1)$. This is the well-known epidemic threshold [11]. The bifurcation is transcritical and corresponds to \mathcal{R}_0 increasing through 1. If $\psi''(1)/\psi'(1) < (\beta + \gamma)/\beta$ (that is $\mathcal{R}_0 < 1$) then μ is positive and the corresponding equilibrium has $\theta > 1$ and is unstable, while the equilibrium at $\theta = 1$ is stable. In our case, we will not observe $\theta > 1$ because θ is a probability. If however $\psi''(1)/\psi'(1) > (\beta + \gamma)/\beta$ (that is $\mathcal{R}_0 > 1$), then the corresponding equilibrium has $\theta < 1$ and is stable while the equilibrium at $\theta = 1$ is unstable.

Previous studies found these approximate fixed points for θ , and used a slightly different definition for θ such that $\theta(0)$ was slightly less than 1. So the biologically implausible prediction is that for $\mathcal{R}_0 < 1$ the value of θ increases towards the stable fixed point at 1. For $\mathcal{R}_0 > 1$ the prediction is more meaningful: θ decreases away from the unstable fixed point at 1 to the lower, stable fixed point. When we do a more careful analysis, we now have $\theta(0) = 1$ and $\phi_I(0) > 0$. The stable fixed point that was at 1 for $\mathcal{R}_0 < 1$ will be slightly decreased to $\theta_0 < 1 = \theta(0)$. So θ will decrease to θ_0 when $\mathcal{R}_0 < 1$. For $\mathcal{R}_0 > 1$ the unstable fixed point that was at $\theta_0 = 1$ is slightly increased, so $\theta_0 > 1 = \theta(0)$. So θ will decrease away from θ_0 . When the number of infections decays immediately ($\mathcal{R}_0 < 1$) we care about a small error in the prediction caused by the fact that the initial proportion infected is not exactly zero, while if the number of infections grows ($\mathcal{R}_0 > 1$), the error caused by treating the initial proportion infected as asymptotically 0 is insignificant.

We are now able to consider the effect of realistic initial conditions. We keep $\phi_R(0) = 0$, but take $\phi_I(0)$ to be a small positive number with $\theta(0) = 1$ and $\phi_S(0) = 1 - \phi_I(0)$. The bifurcation diagram changes slightly. Compared to the equations assuming $\phi_S(0) = 1$, this has the effect of decreasing $\dot{\theta}$ slightly, so the equilibrium values shift as shown in figure 7. In fact, strictly speaking, there is no bifurcation for $\phi_I(0) > 0$.

Below the bifurcation value of μ , the equilibrium $\theta = 1$ is slightly reduced to a $\theta_0 < 1$, but remains stable. The solution with initial condition $\theta = 1$ converges to this equilibrium. Above the bifurcation, the $\theta = 1$ equilibrium is slightly increased to a $\theta_0 > 1$ and is unstable. The other equilibrium with smaller θ is stable and

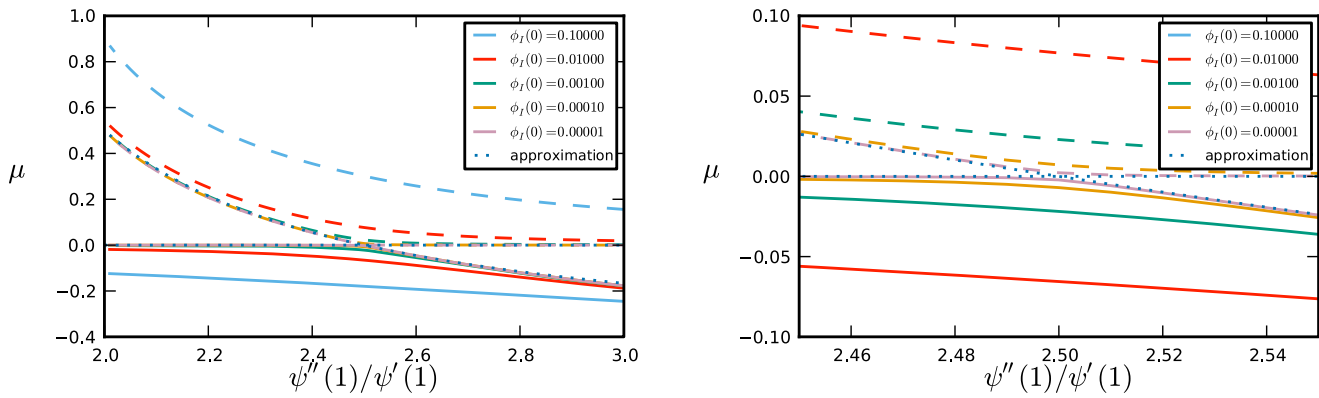


Figure 7. Bifurcation diagram with $\phi_R(0)=0$, $\phi_S(0)=1-\phi_I(0)$, and $\phi_I(0)$ as given. The figure on right zooms in on the bifurcation point. Disease parameters are $\beta=1$ and $\gamma=1.5$. In each all members of the population have degree either 3 or 4, with the proportions chosen so that $\psi''(1)/\psi'(1)$ takes the values on the horizontal axis. The dashed curves denote unstable equilibria and the solid curves stable equilibria. Approximate curves (dotted) come from equation (5). Only the equilibria with $\mu < 0$ are biologically meaningful. doi:10.1371/journal.pone.0101421.g007

its location is decreased slightly. Thus the solution with initial condition $\theta=1$ decreases and converges to the stable solution.

Results and Discussion

In this paper we have extended previous mathematical methods for epidemic spread in networks to allow for simple models that can capture the impact of a non-negligible initial condition.

Our system of equations (1) and (2) are mathematically simple and can be solved numerically with standard tools. Changes in the population’s degree distribution affect ψ , but do not otherwise alter the structure of the equations, and in particular, the population can have arbitrarily large maximum degree without requiring any increase in the number of equations.

Our modeling approach accurately predicts the size and dynamics of simulated epidemics with arbitrary sized initial conditions. The approach allows us to compare interventions introduced during the epidemic. We further see that the discrepancy found by [14] apparently results from the fact that the initial proportion infected was nonnegligible.

We are able to investigate the dynamical structure of the equilibria found in the equations, showing how the epidemic threshold is modified when a non-negligible proportion of the population is infected. This helps resolve the apparent contradiction in earlier work in which θ , a measure of the probability of having escaped transmission, could increase in time. On biological grounds θ must be monotonically decreasing. The mathematical inconsistency resulted from the assumption of a small initial condition. For $\mathcal{R}_0 < 1$, in the small initial condition limit, there is an attracting equilibrium value where the cumulative amount of infection is very small. For a reduced initial infection, the equilibrium moves, going to no infection in the limit. The previous analyses effectively froze this equilibrium at its asymptotic limit and then considered a small, but nonzero initial amount of infection. So the initial condition was on the wrong side of the equilibrium and the mathematical model attempts to return to its frozen equilibrium of no cumulative infection rather than approaching the true equilibrium with a small number of recovered individuals. When $\mathcal{R}_0 > 1$, this effect does not matter because the frozen equilibrium is repelling and still on the correct side of the initial condition.

One of the most obvious applications of these results is to the understanding of the impact of an intervention that begins after a

disease has established itself. This has been a weakness of network models for some time: the earliest low-dimensional models could only calculate static quantities such as the final size of epidemics assuming no intervention, while more recent approaches that calculate the dynamics [7,15,22] have been restricted to the assumption of asymptotically small initial conditions, again with no change in the population behavior. Because we now have a low-dimensional model that can account for large initial conditions, we can use this to restart our calculations when an intervention is to be implemented, or we can use the final size relation to quickly compare intervention effectiveness.

We have analyzed the bifurcation structure of the final size relation, and used this to explain an apparent discrepancy in earlier work if $\mathcal{R}_0 < 1$. The previous models that assumed small initial condition also implicitly assume that $\mathcal{R}_0 > 1$. This resulted in a disturbing prediction for $\mathcal{R}_0 < 1$ that transmissions could be reversed as time progresses, and infected individuals are uninfected. Once we correctly account for the initial condition this apparent discrepancy disappears.

If variance is large enough, then there may be a small number of very high degree individuals who have a macroscopic effect on the dynamics. Usually, increasing the population size will “drown out” the effects of individuals. However, when the variance is sufficiently large increasing the population size results in a small number of much higher degree individuals. These again have a macroscopic effect on the dynamics. Deterministic predictions will not be accurate: for example, how long the highest degree individual remains infected will influence the final size. The work of [8] rigorously studied the equations using small initial conditions, and showed that if all moments up to the fifth moment were finite, then these equations are accurate in the limit of a large network. More recently [25] has shown that the equations remain valid with much weaker assumptions. The equations will fail if the degree distribution is too broad exactly because these very rare individuals are able to have a population-scale impact on the dynamics, so the specific time and duration of their infections matter.

We expect to be able to recover from this challenge however. After the epidemic has run for a short period of time in a Configuration Model network, all of these high degree individuals have been infected and recovered. The remaining population will have significantly reduced moments. At this point, the stochastic effects are “frozen in”: the dynamics are now deterministic. We

can use the observed conditions at this time to initialize our new system of equations.

There are some assumptions implicit in our derivation that deserve further attention. The model fails if $\theta(0)$, $\phi_S(0)$, $\phi_I(0)$, or $\phi_R(0)$ depend on degree of u . So if for example, we select high degree individuals and then infect their partners (leaving the high degree individuals uninfected), the model will not account for the fact that higher degree individuals are more likely to have infected partners at $t=0$. The approach will fail. This is discussed in more detail in [26].

To be clear, the model does not fail if the initial individuals infected have higher (or lower) degree. This simply affects the initial conditions. Indeed, we expect that if the infection is initially spreading stochastically in the population, and we set $t=0$ to be

when enough cases are infected to have deterministic behavior, we will see that at $t=0$ a disproportionate number of higher degree individuals have been infected. So ϕ_I and ϕ_R may initially be larger than I and R .

We finally note that in [15,16], a number of generalizations of the EBCM equations were considered. The basic approach we have used here can be applied to any of those generalizations.

Author Contributions

Conceived and designed the experiments: JCM. Performed the experiments: JCM. Analyzed the data: JCM. Contributed reagents/materials/analysis tools: JCM. Contributed to the writing of the manuscript: JCM.

References

1. Anderson RM, May RM (1991) *Infectious Diseases of Humans*. Oxford: Oxford University Press.
2. Bansal S, Grenfell B, Meyers L (2007) When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface* 4: 879–891.
3. May RM, Anderson RM (1988) The transmission dynamics of human immunodeficiency virus (HIV). *Philosophical Transactions of the Royal Society London B* 321: 565–607.
4. May RM, Lloyd AL (2001) Infection dynamics on scale-free networks. *Physical Review E* 64: 066112.
5. Moreno Y, Pastor-Satorras R, Vespignani A (2002) Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems* 26: 521–529.
6. Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. *Physical Review Letters* 86: 3200–3203.
7. Volz EM (2008) SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology* 56: 293–310.
8. Decreusefond L, Dhersin JS, Moyal P, Tran VC (2012) Large graph limit for an SIR process in random network with heterogeneous connectivity. *The Annals of Applied Probability* 22: 541–575.
9. House T, Keeling M (2011) Insights from unifying modern approximations to infections on networks. *Journal of The Royal Society Interface* 8: 67–73.
10. Eames K, Keeling M (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences* 99: 13330–13335.
11. Newman MEJ (2002) Spread of epidemic disease on networks. *Physical Review E* 66: 016128.
12. Meyers LA, Pourbohloul B, Newman MEJ, Skowronski DM, Brunham RC (2005) Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology* 232: 71–81.
13. Meyers LA, Newman M, Pourbohloul B (2006) Predicting epidemics on directed contact networks. *Journal of Theoretical Biology* 240: 400–418.
14. Lindquist J, Ma J, van den Driessche P, Willeboordse F (2011) Effective degree network disease models. *Journal of Mathematical Biology* 62: 143–164.
15. Miller JC, Slim AC, Volz EM (2012) Edge-based compartmental modelling for infectious disease spread. *Journal of the Royal Society Interface* 9: 890–906.
16. Miller JC, Volz EM (2013) Incorporating disease and population structure into models of SIR disease in contact networks. *PLoS One* 8: e69162.
17. Miller JC, Volz EM (2013) Model hierarchies in edge-based compartmental modeling for infectious disease spread. *Journal of Mathematical Biology* 67: 869–899.
18. Kiss I, Berthouze L, Taylor T, Simon P (2012) Modelling approaches for simple dynamic networks and applications to disease transmission models. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*.
19. Ball F, Neal P (2008) Network epidemic models with two levels of mixing. *Mathematical Biosciences* 212: 69–87.
20. Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms* 6: 161–179.
21. Miller JC (2012) A note on the derivation of epidemic final sizes. *Bulletin of Mathematical Biology* 74: 2125–2141.
22. Miller JC (2011) A note on a paper by Erik Volz: SIR dynamics in random networks. *Journal of Mathematical Biology* 62: 349–358.
23. Noël PA, Allard A, Hébert-Dufresne L, Marceau V, Dubé IJ (2012) Propagation on networks: An exact alternative perspective. *Physical Review E* 85: 03118.
24. Miller JC (2008) Bounding the size and probability of epidemics on networks. *Journal of Applied Probability* 45: 498–512.
25. Janson S, Luczak M, Windridge P (2013) Law of large numbers for the SIR epidemic on a random graph with given degrees. arXiv preprint arXiv:13085493.
26. Miller JC, Kiss IZ (2014) Epidemic spread in networks: Existing methods and current challenges. *Mathematical Modelling of Natural Phenomena*.