



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Estimating finite-population reproductive numbers in heterogeneous populations

Lindsay T. Keegan<sup>a,\*</sup>, Jonathan Dushoff<sup>b</sup>

<sup>a</sup> 615 North Wolfe St, The Johns Hopkins University, Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>b</sup> 1280 Main Street West, McMaster University, Hamilton, Ontario, Canada



## HIGHLIGHTS

- Outline a framework for discussing the different types of heterogeneity.
- Found simple expressions for each of the four different types of heterogeneity and heterogeneity in intrinsic mixing.
- Showed heterogeneity in finite populations is more complicated than previously thought.
- Showed that heterogeneity in a finite population makes control easier than predicted by  $\mathcal{R}_0$  and the homogeneous finite-population reproductive number.

## ARTICLE INFO

### Article history:

Received 21 July 2015

Received in revised form

13 January 2016

Accepted 16 January 2016

Available online 15 February 2016

### Keywords:

Basic reproductive number

Heterogeneity

Super-spreaders

## ABSTRACT

The basic reproductive number,  $\mathcal{R}_0$ , is one of the most important epidemiological quantities.  $\mathcal{R}_0$  provides a threshold for elimination and determines when a disease can spread or when a disease will die out. Classically,  $\mathcal{R}_0$  is calculated assuming an infinite population of identical hosts. Previous work has shown that heterogeneity in the host mixing rate increases  $\mathcal{R}_0$  in an infinite population. However, it has been suggested that in a finite population, heterogeneity in the mixing rate may actually *decrease* the finite-population reproductive numbers. Here, we outline a framework for discussing different types of heterogeneity in disease parameters, and how these affect disease spread and control. We calculate “finite-population reproductive numbers” with different types of heterogeneity, and show that in a finite population, heterogeneity has complicated effects on the reproductive number. We find that simple heterogeneity *decreases* the finite-population reproductive number, whereas heterogeneity in the intrinsic mixing rate (which affects both infectiousness and susceptibility) *increases* the finite-population reproductive number when  $\mathcal{R}_0$  is small relative to the size of the population and *decreases* the finite-population reproductive number when  $\mathcal{R}_0$  is large relative to the size of the population. Although heterogeneity has complicated effects on the finite-population reproductive numbers, its implications for control are straightforward: when  $\mathcal{R}_0$  is large relative to the size of the population, heterogeneity decreases the finite-population reproductive numbers, making disease control or elimination easier than predicted by  $\mathcal{R}_0$ .

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

The effect of heterogeneity on disease dynamics is a foundational question in infectious disease modeling. Recently, renewed attention has been given to the impact of heterogeneity on the basic reproductive number,  $\mathcal{R}_0$  (Lloyd-Smith et al., 2005; Smith et al., 2007; Alex Perkins et al., 2013).  $\mathcal{R}_0$  is the average number of secondary infections from a single infectious individual in an otherwise totally

susceptible population. Kermack and McKendrick (1927) formulated  $\mathcal{R}_0$  assuming a disease spreading in a large, homogeneous population; this construction of  $\mathcal{R}_0$  has since dominated epidemic theory.

Despite the common, convenient assumption that diseases are spread in well mixed homogeneous populations, there is a great deal of evidence that population heterogeneity is an important determinant of disease spread. Woolhouse et al. (1997) argued that heterogeneity in disease spread is pervasive, and often characterized by the ‘20/80 rule’ – i.e., 20% of infected individuals cause 80% of cases. Perhaps the most famous example of disease heterogeneity is typhoid Mary, who is estimated to have caused over 50 new cases, despite typhoid having an  $\mathcal{R}_0$  of 2.8 (Pitzer et al., 2014;

\* Corresponding author.

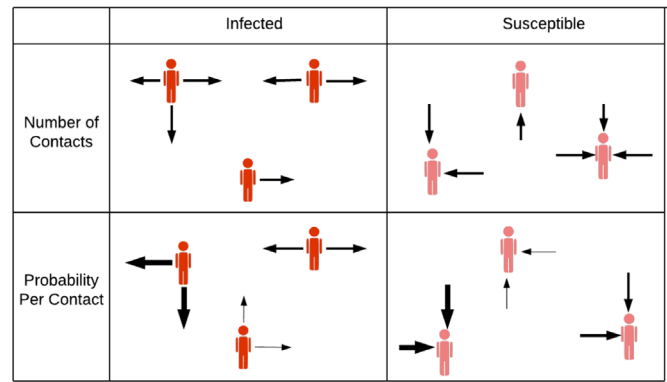
E-mail address: [keeganlt@mcmaster.ca](mailto:keeganlt@mcmaster.ca) (L.T. Keegan).

Brooks, 1996). Other examples include: vector-borne diseases, including malaria (Scott et al., 2006), dengue fever (De Benedictis et al., 2003) and West Nile Virus (Kilpatrick et al., 2006); sexually transmitted infections (STI), including HIV (Galvani and May, 2005) and gonorrhoea (Nold, 1980; Hethcote et al., 1982); and other directly transmitted diseases such as severe acute respiratory syndrome (SARS) (Lloyd-Smith et al., 2005), tuberculosis (TB) (Kline et al., 1995), and small pox (Lloyd-Smith et al., 2005). Individual variation has important implications for disease dynamics, including emergence, spread, and control.

The effect of population heterogeneity on the basic reproductive number,  $\mathcal{R}_0$ , has been well studied (Dye and Hasibeder, 1986; Lloyd-Smith et al., 2005; Alex Perkins et al., 2013; Smith et al., 2007; Diekmann et al., 1990; May and Anderson, 1984). May and Anderson explored the effects of heterogeneity on  $\mathcal{R}_0$  for directly transmitted diseases and showed that heterogeneity increases  $\mathcal{R}_0$ , making eradication harder under a homogeneously applied immunization program (May and Anderson, 1984). Diekmann et al. (1990) defined and calculated  $\mathcal{R}_0$  for heterogeneous populations of directly-transmitted diseases. They calculated  $\mathcal{R}_0$  in terms of the next-generation operator, which maps generations of infected individuals to each other. Like other calculations of  $\mathcal{R}_0$ , it assumes one infected individual in an otherwise very large population of susceptible hosts. Dye and Hasibeder (1986) showed that for vector-borne diseases, heterogeneity in the mosquito biting rate increases  $\mathcal{R}_0$ , which suggests that heterogeneity makes invasion likely and elimination more difficult than would be predicted by a standard calculation of  $\mathcal{R}_0$  based on parameters. Lloyd-Smith et al. (2005) showed that for directly transmitted diseases with heterogeneous transmission, like SARS, both the probability that an epidemic will occur and the subsequent size of the epidemic will be affected: heterogeneity made extinction more likely than predicted by the standard calculation of  $\mathcal{R}_0$ , but if an epidemic did occur, it was more likely to be explosive.

While heterogeneity increases  $\mathcal{R}_0$  in infinite populations, Smith et al. (2007) suggest that heterogeneity actually decreases  $\mathcal{R}_0$  in finite populations. They introduced the idea of calculating the expected number of secondary infections from a single infectious individual in a finite population of susceptible hosts. Motivated by malaria, a disease with a large  $\mathcal{R}_0$ , where  $\mathcal{R}_0$  can easily approach or exceed the size of the population, they simulated these finite-population reproductive numbers and show that in a finite population, heterogeneity actually decreases  $\mathcal{R}_0$ ; this is because in a finite population, individuals who are more susceptible are more likely to get infected multiple times, absorbing some possible infections. Keegan and Dushoff (2014) calculated these finite-population reproductive numbers for both vector-borne and directly-transmitted diseases assuming a well mixed host population.

Building on this previous work we calculate finite-population reproductive numbers for directly transmitted diseases under different assumptions of heterogeneity in transmission. We also discuss a framework for discussing different “types of heterogeneity” (Fig. 1) and their importance in terms of disease control and intervention. Like classical calculations of  $\mathcal{R}_0$ , we are only interested in the initial spread of infection and our calculations ignore the longer term depletion of susceptibles. Previous work (Smith et al., 2007; Keegan and Dushoff, 2014) also took this approach in calculating finite-population reproductive numbers. In a study of homogeneous finite population reproductive numbers, Ross (2011) did account for the depletion of susceptibles in their estimates of the finite-population reproductive numbers with homogeneous transmission and showed that this further reduced the finite-population reproductive number.



**Fig. 1.** Schematic of the different types of heterogeneity in individual level parameters. The red people are infected and the pink people are susceptible. Each arrow represents a contact and the weight of the arrow represents the probability that contact will result in infection. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

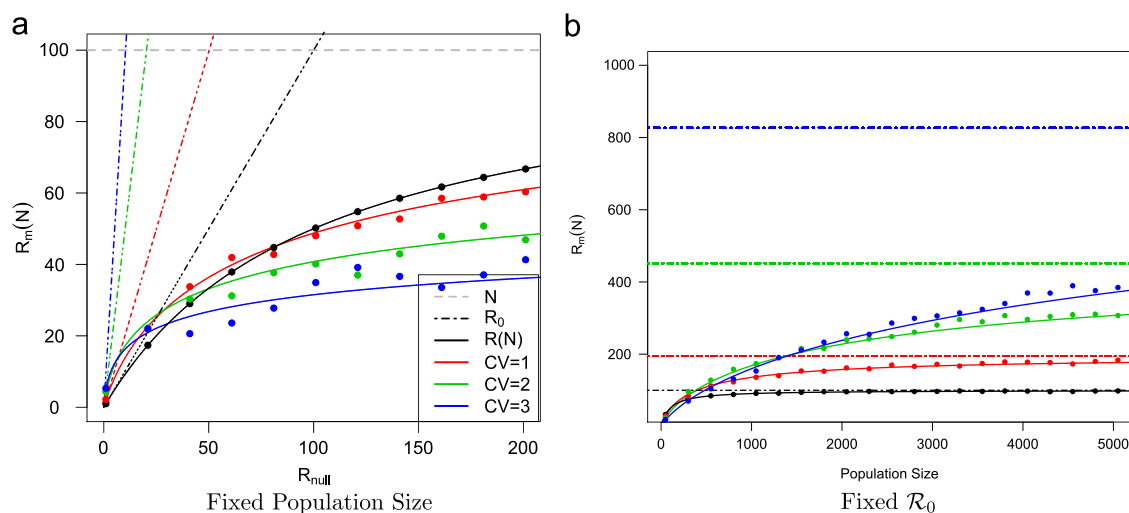
### 1.1. Heterogeneity framework

In general, heterogeneity in transmission can be broken into two categories: structural heterogeneity where individuals are separated into groups, either by age or by spatial structure; and heterogeneity in individual level parameters in which individuals exhibit different disease-related behaviors (Lloyd-Smith et al., 2005; Dushoff, 1999). Here, we outline the different types of heterogeneity in individual level parameters.

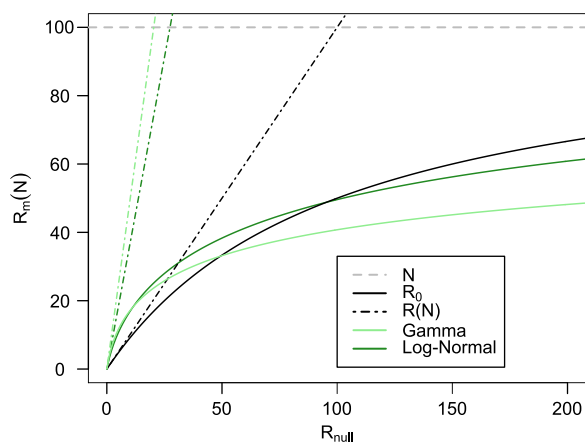
Although a great deal of work has been done in understanding how heterogeneity affects the spread and control of infectious diseases (e.g. Lloyd-Smith et al., 2005; Dushoff, 1999; Diekmann et al., 1990), we have found little detailed discussion of different types of heterogeneity in transmission and their effects. Often, heterogeneity is discussed in terms of presence/absence: a population is assumed to be homogeneous or it is not. Less attention is given to the type or types of heterogeneity in disease spread. When heterogeneity is discussed in more detail, it tends to be discussed in terms of heterogeneity in infected individuals, ie “super-spreaders” (Galvani and May, 2005; Lloyd-Smith et al., 2005; Stein, 2011) and “super-shedders” (Caroline Breese Hall, 2007; Stephens et al., 2009; Chase-Topping et al., 2008), likely because heterogeneity in susceptible individuals is harder to measure. However, clearly identifying and understanding the different types of heterogeneity and how they affect disease dynamics provides new opportunities for control. Here, we outline a framework for discussing heterogeneity in individual-level parameters.

#### 1.1.1. Mixing rate

The mixing rate, also called “contact rate” describes the number of contacts that an individual has that could result in an infection. Mixing rates vary by modes of transmission and by disease. The contact rate for a vector-borne disease depends on hosts being bitten by vectors and consequently, it is dependent both on host-related and vector-related factors. The mixing rate for an STI is the number of potentially infectious sexual contacts an individual has. This can be affected by a multitude of factors including condom use, etc. For other directly transmitted diseases, the mixing rate may be harder to quantify and depend on specific disease-related factors such as how long infectious particles remain in the air or stay alive on surfaces, and environmental factors such as humidity (Caroline Breese Hall, 2007; Bean et al., 1982; Karim et al., 1985; Miller and Artenstein, 1967; Kao and Yang, 2006).



**Fig. 2.** The finite-population reproductive number,  $\mathcal{R}_m(N)$ , for gamma distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=100$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 100$  (dot-dashed line). The points represent the average of 500,000 simulations of each.



**Fig. 3.** The finite-population reproductive numbers,  $\mathcal{R}_m(N)$ , versus the homogeneous basic reproductive number,  $\mathcal{R}_0$  for gamma distributed (light green) and log-normally distributed (forest green) heterogeneity for CV=2. The population size is  $N=100$ , the homogeneous  $\mathcal{R}_0$  is the solid black line and the heterogeneous reproductive numbers are the solid green lines, the homogeneous finite-population reproductive number,  $\mathcal{R}(N)$ , is the dotted black line, and the dot-dashed lines represent heterogeneous finite-population reproductive numbers with different coefficients of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Heterogeneity affects mixing rates in a multitude of ways. For a vector-borne disease, such as malaria, heterogeneity in the mixing rate is often a result of host-related factors, such as differential attractiveness to mosquitoes (e.g. Port et al., 1980; Shirai et al., 2004; Lindsay et al., 2000) or the use of bed nets (Lengeler et al., 2004). Heterogeneity in the mixing rate for sexually transmitted infections can be affected by including rate of sexual partner change, sexual practices (Joseph Hotz Avner Ahituv and Philipson, 1996), and access to condoms (Chimbindi et al., 2010; MacPhail et al., 2009; Bassett and Mhloyi, 1991; Pettifor et al., 2004). For directly transmitted diseases, contacts are often harder to define, and heterogeneity arises from a mixture of host, pathogen, and environmental factors. Heterogeneity in the mixing rate can be further broken down into the mixing rate of *infected* individuals and the mixing rate of *susceptible* individuals. Here, we assume that individuals have an intrinsic mixing rate that does not change with infection status.

**Heterogeneity in infected mixing:** The mixing rate of infected individuals is the number of contacts that an infected individual has that could result in an infection, during the course of infection. This type of heterogeneity is the most widely discussed and is often discussed in terms of infected individuals with large numbers of contacts, or super-spreaders (Lloyd-Smith et al., 2005; Galvani and May, 2005).

Malaria is a well known example of a disease with heterogeneity in the mixing rate: some hosts are more attractive to mosquitoes for a variety of reasons, including body size (Port et al., 1980), blood type (Shirai et al., 2004), pregnancy (Lindsay et al., 2000; Ansell et al., 2002), and alcohol consumption (Shirai et al., 2002; Lefèvre et al., 2010), among others. An example of a directly transmitted disease with heterogeneity in the mixing rate of infected individuals is SARS. During the 2002–2003 SARS epidemic, renewed attention was given to the role of super-spreading events in the epidemic; particularly in the context of control. In a model of SARS transmission, Lloyd-Smith et al. (2005) showed that targeted control was up to three times more effective than random control.

**Heterogeneity in susceptible mixing:** The mixing rate of susceptible individuals is the number of contacts that a susceptible individual has that could result in an infection. The susceptible mixing rate may or may not be closely related to the infected mixing rate, depending on whether contacts are symmetric, and whether behavior is changed by the disease. For example, vector-borne and sexually transmitted diseases involve symmetric contact (both the susceptible and infected individual need to be bitten by a vector, or to have sexual intercourse to transmit disease), whereas food-borne illnesses often involve asymmetric contact (food workers infect food consumers, but not the other way around). Ebola virus disease is an example where behavior is changed by disease: effective mixing rates of well people depend on how likely they are to be involved in care-giving, and likely vary less than the effective mixing rates of infectious people. Sick individuals may also voluntarily attempt to reduce risk (Ijumba et al., 2004; Funk et al., 2009).

### 1.1.2. Probability per contact

The probability of infection per contact describes the probability of successful transmission per potentially infectious contact. A “potentially infectious contact” is defined as one which would succeed in transmitting if both the probability of transmitting and

the probability of contracting an infection are 1. The probability of successful transmission is the product of an infectiousness probability and a susceptibility probability.

**Heterogeneity in infectiousness:** The infectiousness of an individual is the probability of transmitting an infection per potentially infectious contact, assuming the susceptibility probability is 1. This type of heterogeneity is often discussed in terms of individuals who shed a lot of virus, or super-shedders (Caroline Breese Hall, 2007; Stephens et al., 2009; Chase-Topping et al., 2008).

Influenza and tuberculosis are examples of directly transmitted diseases with heterogeneity in infectiousness. For influenza, viral shedding influences the per contact infectiousness, there is a large variation in the number aerosolized respiratory secretions, making some individuals more infectious than others (Caroline Breese Hall, 2007). Additionally, external factors such as taking antipyretics may also increase the per-contact probability of transmission by increasing both the rate and duration of viral shedding (Earn et al., 2014). For TB, access to health care interventions and proper nutrition may reduce the probability of transmitting per contact, by reducing the risk of pulmonary TB (Kline et al., 1995).

HIV is an example of an STI with heterogeneity in infectiousness. For HIV, the infectiousness per contact varies for a variety of reasons: individuals who are co-infected with another STI have been shown to be more infectious per contact (Galvin and Cohen, 2004; Abu-Raddad et al., 2008), while individuals on antiretroviral treatment are less infectious per contact (Wawer et al., 2005).

**Heterogeneity in susceptibility:** The susceptibility of an individual is the probability of contracting an infection per potentially infectious contact, assuming the infectiousness probability is 1. This is the least talked about as it is likely the hardest to measure and probably depends on the interaction between host, pathogen, and environment.

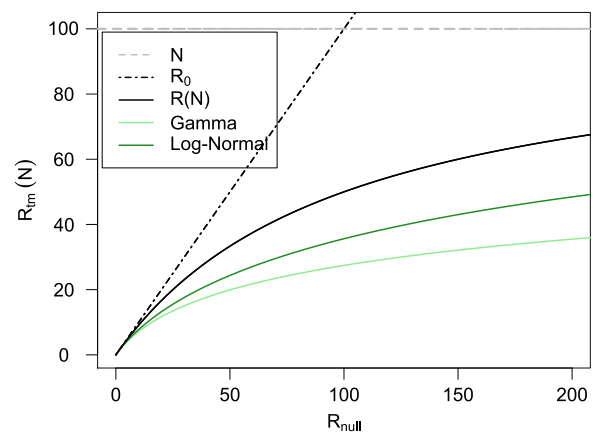
An example of a disease with heterogeneity in susceptibility is HIV. There are a multitude of factors that can cause susceptibility to vary including gender and circumcision. It has been suggested that women are physiologically more susceptible to HIV than men (Glynn et al., 2001; Quinn and Overbaugh, 2005) and that male circumcision has a protective effect against the per contact transmission of HIV (Williams et al., 2006; Bailey et al., 2007).

## 2. Methods

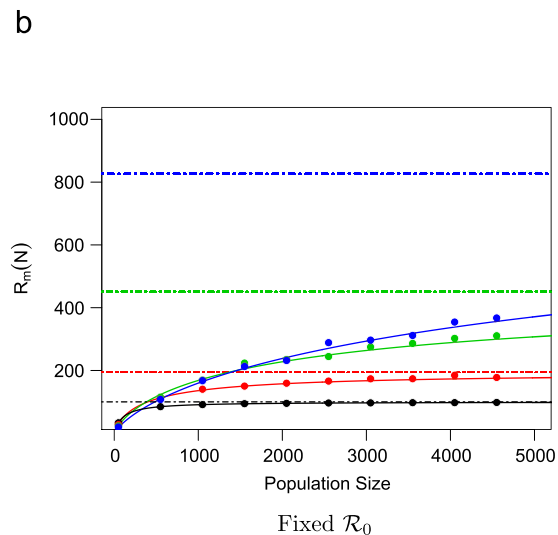
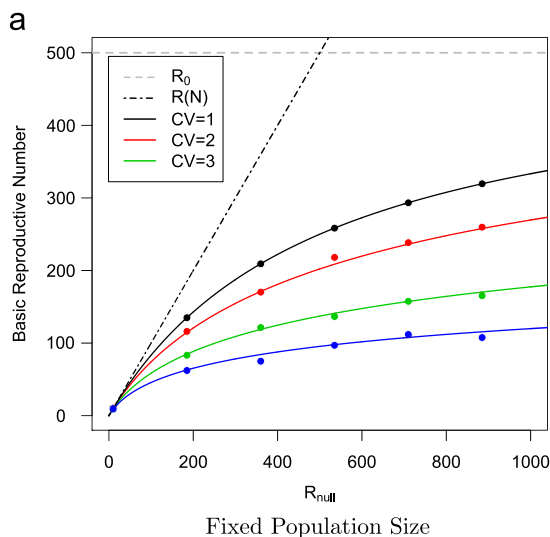
We calculate finite-population reproductive numbers for directly transmitted diseases under three heterogeneity assumptions: heterogeneity in intrinsic mixing rate (assumed to be independent of disease status),  $\mathcal{R}_m(N)$ ; heterogeneity in mixing rate when infected,  $\mathcal{R}_{tm}(N)$ ; and heterogeneity in the probability of contracting an infection per contact,  $\mathcal{R}_{sp}(N)$ . Two other assumptions: heterogeneity in the probability of transmitting an infection per contact,  $\mathcal{R}_{tp}(N)$  and heterogeneity in mixing rate when susceptible,  $\mathcal{R}_{sm}(N)$ , are shown in the Appendix.

We also talk about the different infinite population reproductive numbers. The basic reproductive number in a homogeneous population, which we refer to as “ $\mathcal{R}_{null}$ ” and the infinite population reproductive numbers with heterogeneity,  $\mathcal{R}_0$ .

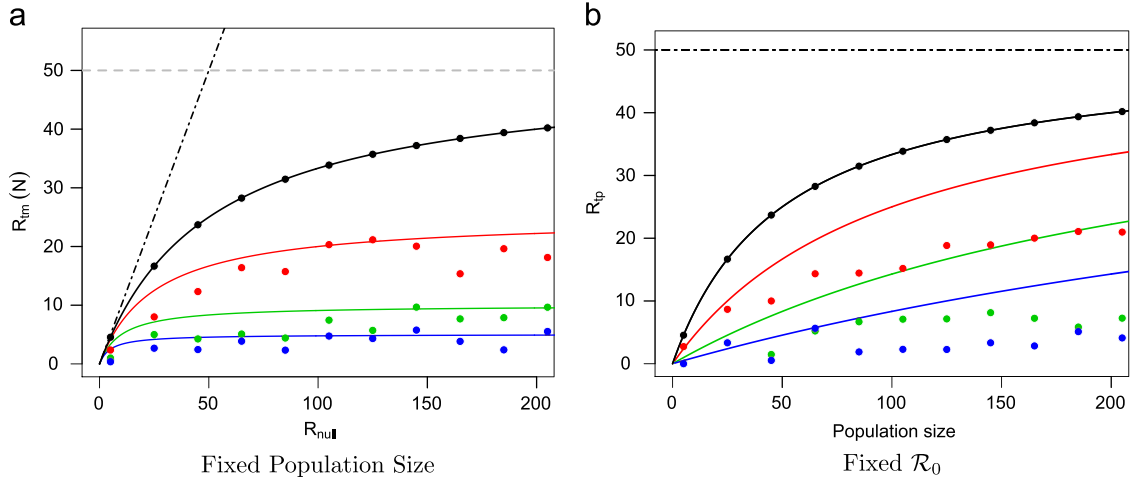
We numerically calculate each of the five finite-population reproductive numbers using two distributions of the underlying heterogeneity. We use gamma distributed and log-normally



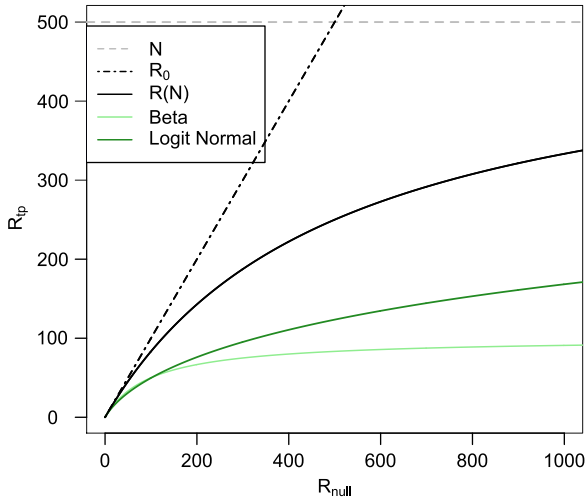
**Fig. 5.** The finite-population reproductive numbers,  $\mathcal{R}_{tm}(N)$ , versus the basic reproductive number,  $\mathcal{R}_0$  for gamma distributed (light green) and log-normally distributed (forest green) heterogeneity for  $CV=2$ . The population size is  $N=100$ ,  $\mathcal{R}_0$  is the solid black line, the homogeneous finite-population reproductive number,  $\mathcal{R}(N)$ , is the dotted black line, and the dot-dashed lines represent heterogeneous finite-population reproductive numbers with different coefficients of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)



**Fig. 4.** The finite-population reproductive number,  $\mathcal{R}_{tm}(N)$ , for gamma distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=50$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 50$  (dot-dashed line). The points represent the average of 500,000 simulations of each.



**Fig. 6.** The finite-population reproductive number,  $\mathcal{R}_{t_p}(N)$ , for beta distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=50$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0=50$  (dot-dashed line). The points represent the average of 500,000 simulations of each.



**Fig. 7.** The finite-population reproductive numbers,  $\mathcal{R}_{t_p}(N)$ , versus the basic reproductive number,  $\mathcal{R}_0$  for beta distributed (light green) and logit-normally distributed (forest green) heterogeneity for  $CV=2$ . The population size is  $N=500$ ,  $\mathcal{R}_0$  is the solid black line, the homogeneous finite-population reproductive number,  $\mathcal{R}(N)$ , is the dotted black line, and the dot-dashed lines represent heterogeneous finite-population reproductive numbers with different coefficients of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

distributed heterogeneity for  $\mathcal{R}_m(N)$ ,  $\mathcal{R}_{t_m}(N)$ , and  $\mathcal{R}_{s_m}(N)$  and beta distributed and logit-normally distributed heterogeneity for  $\mathcal{R}_{t_p}(N)$  and  $\mathcal{R}_{s_p}(N)$ .

2.1. Assumptions

We assume a finite population of size  $N$ . We talk about transmission,  $t = t_m t_p$ , as the product of the mixing rate of infected individuals,  $t_m$ , and the probability of transmitting an infection per infectious contact,  $t_p$ . Similarly, we talk about susceptibility,  $s = s_m s_p$ , as the product of the mixing rate of susceptible individuals,  $s_m$ , and the probability of contracting an infection per contact,  $s_p$ .

Here, we calculate all five of the finite-population reproductive numbers twice, once with each of the distributions of  $t_m$ ,  $s_m$  (gamma and log-normal distributions) and  $t_p$ ,  $s_p$  (beta and logit-normal distributions). We talk about the mean mixing rates  $\bar{t}_m$  and

$\bar{s}_m$  for infected and susceptible individuals and the mean probability of infection,  $\bar{t}_p$  and  $\bar{s}_p$ . Further, when all susceptible individuals are the same, we talk about the mean susceptibility,  $\bar{s}$  and when all infected individuals are the same, we talk about the mean infectiousness  $\bar{t}$ .

We assume that an infected individual of type  $y$  produces a geometric distribution of new contacts with mean of  $t(y)$ . This is equivalent to assuming that the infection and recovery processes are Markovian (Keegan and Dushoff, 2014). We assume that all hosts behave independently and we ignore the longer-term depletion of susceptibles. Since the host population is finite, some of the possible infections may fall on the same susceptible host, so the average number of realized infections, in general, will be smaller.

2.2. Calculation framework

We start with a known infectious individual of type  $y$ . We know that this infected individual produces a geometric distribution of potentially successful challenges (contacts) with mean  $t(y)$  (where  $t(y) = t_m(y)t_p(y)$ ). For each challenge, the risk to a particular susceptible individual of type  $x$ , is  $\frac{s(x)}{\bar{s}_m N}$ , where  $\bar{s}_m$  is the mean susceptible mixing rate.

The probability of escaping  $a$  challenges from an infected individual is  $\left(1 - \frac{s(x)}{\bar{s}_m N}\right)^a$ . The risk of being infected by at least one of those contacts is  $\left(1 - \left(1 - \frac{s(x)}{\bar{s}_m N}\right)^a\right)$ . Using the generating function method detailed in Keegan and Dushoff (2014), we find the risk to a susceptible individual of type  $x$  from a single infected individual of type  $y$  is:

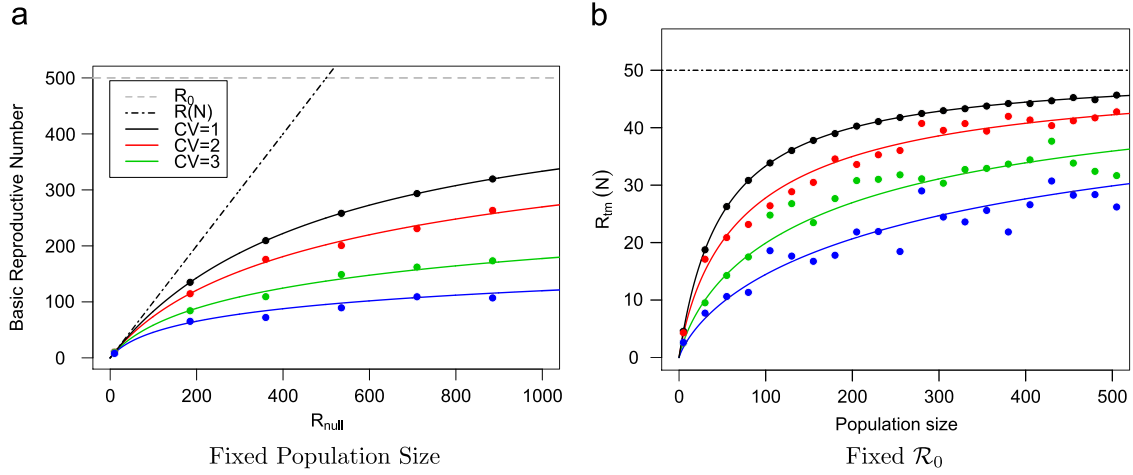
$$\frac{s(x)t(y)}{s(x)t(y) + \bar{s}_m N} \tag{1}$$

The expected number of infections from the known infectious individual is

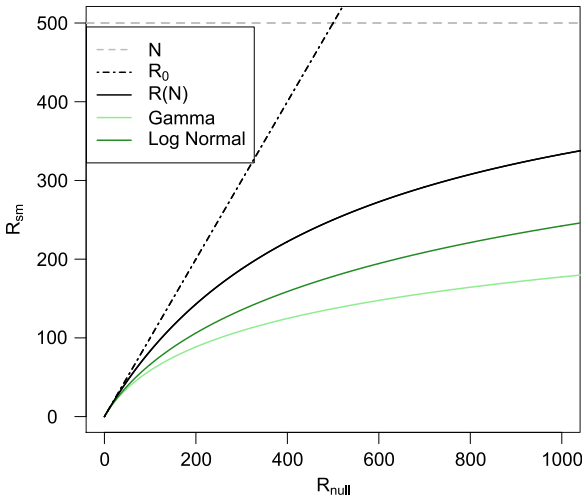
$$\left\langle \frac{s(x)t(y)}{s(x)t(y) + \bar{s}_m N} \right\rangle_x N \tag{2}$$

We then average over the distribution of “typical” individuals (Diekmann et al., 1990),  $s(y)/\bar{s}$ , and find:

$$R(N) = \left\langle \frac{s(y)}{\bar{s}} \left\langle \frac{s(x)t(y)N}{s(x)t(y) + \bar{s}_m N} \right\rangle_x \right\rangle_y \tag{3}$$



**Fig. 8.** The finite-population reproductive number,  $\mathcal{R}_{sm}(N)$ , for gamma distributed heterogeneity. The solid lines are the finite-population reproductive numbers with corresponding coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=50$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 50$  (dot-dashed line). The points represent the average of 500,000 simulations of each.



**Fig. 9.** The finite-population reproductive numbers,  $\mathcal{R}_{sm}(N)$ , versus the basic reproductive number,  $\mathcal{R}_0$  for gamma distributed (light green) and log-normally distributed (forest green) heterogeneity for  $CV=2$ . The population size is  $N=500$ ,  $\mathcal{R}_0$  is the solid black line, the homogeneous finite-population reproductive number,  $\mathcal{R}(N)$ , is the dotted black line, and the dot-dashed lines represent heterogeneous finite-population reproductive numbers with different coefficients of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Which is equivalent to

$$R(N) = \int \frac{s(y)}{\bar{s}} \int \frac{s(x)t(y)N}{s(x)t(y) + \bar{s}_m N} f(x)f(y) dx dy \quad (4)$$

where  $f(x)$  is the distribution of susceptible individuals and  $f(y)$  is the distribution of infected individuals. Here, we always choose the same distribution for  $f(x)$  and  $f(y)$  however, we allow the mean of the distributions to vary ( $s$ ,  $t$ , and  $m$ ).

### 2.2.1. Heterogeneity in mixing rate, $\mathcal{R}_m(N)$

We calculate the finite-population reproductive number for the case where we allow only heterogeneity in the mixing rates of infected and susceptible individuals.

To account for individuals contacting each other we need the rate at which susceptible individuals contact infected individuals  $s$  to be proportional to the rate at which infected individuals contact susceptibles  $t$ , so we write  $t = \mu$  and  $s = \mu/\bar{\mu}$ .

We find the expected number of new infections from our known infectious individual is

$$\left\langle \frac{\mu(x)\mu(y)N}{\mu(x)\mu(y) + \bar{\mu}N} \right\rangle_x \quad (5)$$

We then account for the distribution of typical individuals,  $\mu(y)/\bar{\mu}$  and find

$$\mathcal{R}_m(N) = \left\langle \frac{\mu(y)}{\bar{\mu}} \left\langle \frac{\mu(x)\mu(y)N}{\mu(x)\mu(y) + \bar{\mu}N} \right\rangle_x \right\rangle_y \quad (6)$$

## 3. Results

We calculate the finite population reproductive number for each of the four types of heterogeneity and for heterogeneity in intrinsic mixing. We find:

*Heterogeneity in intrinsic mixing,  $\mathcal{R}_m(N)$*

$$\mathcal{R}_m(N) = \left\langle \frac{\mu(y)}{\bar{\mu}} \left\langle \frac{\mu(x)\mu(y)N}{\mu(x)\mu(y) + \bar{\mu}N} \right\rangle_x \right\rangle_y \quad (7)$$

We solve  $\mathcal{R}_m(N)$  numerically, once for each of two distributions of mixing rates, gamma distributed heterogeneity in mixing (Fig. 2) and log-normally distributed heterogeneity in mixing (Fig. B3); we compare the results of the two distributions for a fixed coefficient of variation (Fig. 3).

*Heterogeneity in transmission mixing  $\mathcal{R}_{tm}(N)$*

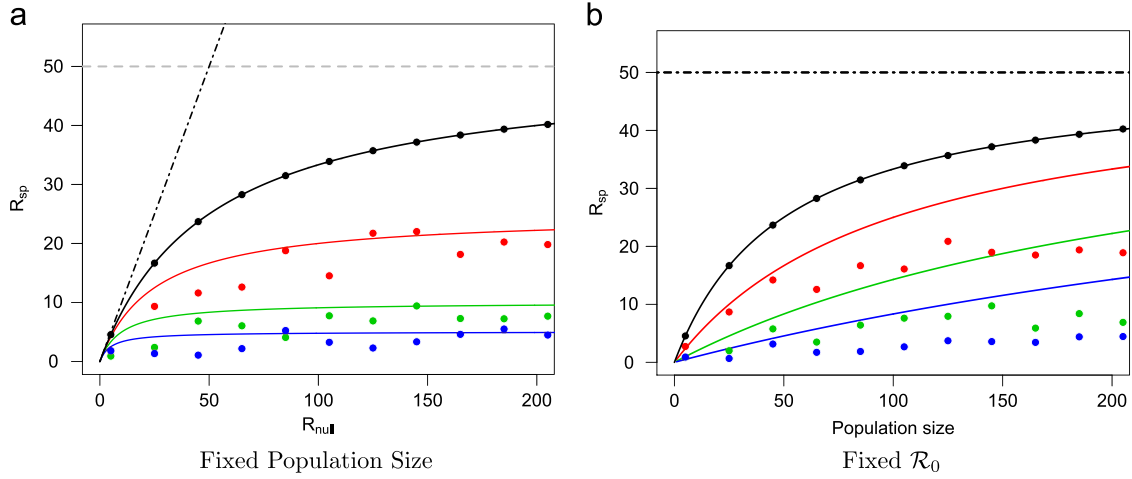
$$\mathcal{R}_{tm}(N) = \left\langle \frac{\bar{s}t(y)N}{\bar{s}t(y) + \bar{s}_m N} \right\rangle_y \quad (8)$$

We solve  $\mathcal{R}_{tm}(N)$  numerically, once for each of the two distributions of transmission mixing rates, gamma distributed heterogeneity in transmission mixing (Fig. 4) and log-normally distributed heterogeneity in transmission mixing (Fig. B2); we compare the results of the two distributions for a fixed coefficient of variation (Fig. 5).

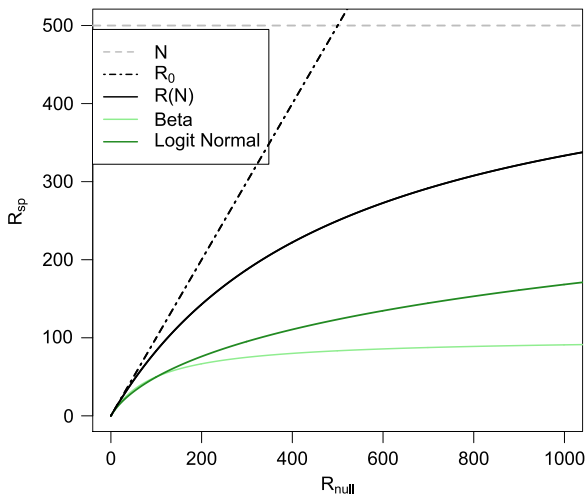
*Heterogeneity in transmission probability  $\mathcal{R}_{tp}(N)$*

$$\mathcal{R}_{tp}(N) = \left\langle \frac{t(y)\bar{s}N}{t(y)\bar{s} + \bar{s}_m N} \right\rangle_y \quad (9)$$

We solve for  $\mathcal{R}_{tp}(N)$  numerically, once for each of the two distributions of mixing rates, beta-distributed heterogeneity (Fig. 6) and logit-normally distributed heterogeneity (Fig. B2); we



**Fig. 10.** The finite-population reproductive number,  $\mathcal{R}_{sp}(N)$ , for beta distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=500$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 100$  (dot-dashed line). The points represent the average of 500,000 simulations of each.



**Fig. 11.** The finite-population reproductive numbers,  $\mathcal{R}_{sp}(N)$ , versus the basic reproductive number,  $\mathcal{R}_0$  for beta distributed (light green) and log-normally distributed (forest green) heterogeneity for  $CV=2$ . The population size is  $N=500$ ,  $\mathcal{R}_0$  is the solid black line, the homogeneous finite-population reproductive number,  $\mathcal{R}(N)$ , is the dotted black line, and the dot-dashed lines represent heterogeneous finite-population reproductive numbers with different coefficients of variation. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

compare the results of the two distributions for a fixed coefficient of variation (Fig. 7).

*Heterogeneity in susceptible mixing,  $\mathcal{R}_{sm}(N)$*

$$\mathcal{R}_{sm}(N) = \left\langle \frac{s(y)}{\bar{s}} \left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \right\rangle_y \quad (10)$$

$$\mathcal{R}_{sm}(N) = \left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \quad (11)$$

We solve for  $\mathcal{R}_{sm}(N)$  numerically, once for each of the two distributions of mixing rates, gamma-distributed heterogeneity (Fig. 8) and log-normally distributed heterogeneity (Fig. B2); we compare the results of the two distributions for a fixed coefficient of variation (Fig. 9).

*Heterogeneity in susceptibility probability  $\mathcal{R}_{sp}(N)$*

$$\mathcal{R}_{sp}(N) = \left\langle \frac{s(y)}{\bar{s}} \left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \right\rangle_y \quad (12)$$

$$\mathcal{R}_{sp}(N) = \left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \quad (13)$$

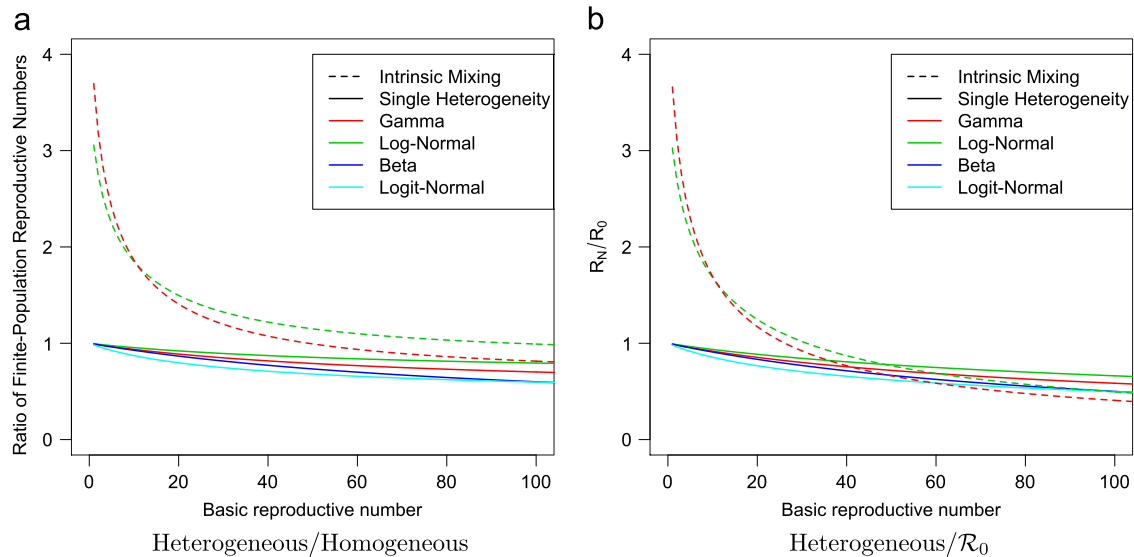
We solve for  $\mathcal{R}_{sp}(N)$  numerically, once for each of the two distributions of mixing rates, beta-distributed heterogeneity (Fig. 10) and logit-normally distributed heterogeneity (Fig. B2); we compare the results of the two distributions for a fixed coefficient of variation (Fig. 11).

Figs. 2 (a), 4(a), 6(a), 8(a), 10(a) and B1(a), B2(a), B2(c), B2(e), B2(g) show how the finite-population reproductive numbers change as  $\mathcal{R}_0$  changes for a fixed population size. Since the finite-population reproductive numbers are limited by the size of the population whereas the infinite reproductive numbers are not, the finite-population reproductive numbers diverge from their corresponding infinite population reproductive number as  $\mathcal{R}_0$  approaches the size of the population. Where they diverge depends on both the distribution and the coefficient of variation of that distribution; although where the finite-population reproductive numbers depend on these factors, the finite-population reproductive number with  $CV=3$  diverges around  $1/5 N$  whereas the homogeneous finite-population reproductive number diverges around  $1/2 N$ . Further, for heterogeneity in intrinsic mixing, the heterogeneous finite-population reproductive numbers are larger than the homogeneous finite-population reproductive number when  $\mathcal{R}_0$  is small relative to the size of the population and the heterogeneous finite-population reproductive numbers are smaller than the homogeneous finite-population reproductive number when  $\mathcal{R}_0$  is large relative to the size of the population. For gamma distributed heterogeneity, this occurs between  $1/3 N$  and  $3/4 N$  (Fig. 2(a)) and at around  $N$  for log-normally distributed heterogeneity (Fig. B3). For the other four types of heterogeneity, the homogeneous finite-population reproductive number is always larger than the heterogeneous finite-population reproductive numbers (Fig. 12).

Figs. 2 (b), 4(b), 6(b), 8(b), 10(b) and B1(b), B2(b), B2(d), B2(f), B2(h) show how the finite-population reproductive numbers vary with the size of the population for fixed  $\mathcal{R}_0$ . These figures show the finite-population reproductive numbers converging on their corresponding infinite population reproductive numbers. The rate at which the finite-population reproductive numbers converge on their corresponding infinite population  $\mathcal{R}_0$  depends on the coefficient of variation.

Figs. 3, 5, 7, 9, and 11 compare the results of the two distributions for each of the finite-population reproductive numbers.





**Fig. 12.** Plot comparing the different distributions. (a) Plot of the ratio of the heterogeneous finite-population reproductive numbers to the homogeneous finite-population reproductive number versus  $\mathcal{R}_0$ ; for each of the four distributions of heterogeneity and for both intrinsic mixing (dashed lines) and simple heterogeneity (solid lines). (b) Plot of the ratio of the heterogeneous finite-population reproductive numbers to  $\mathcal{R}_0$  versus  $\mathcal{R}_0$ ; for each of the four distributions of heterogeneity and for both intrinsic mixing (dashed lines) and simple heterogeneity (solid lines).

Although the results are fairly robust to the distribution, the finite-population reproductive numbers with log-normally and logit-normally distributed heterogeneity are larger than the finite-population reproductive numbers with gamma and beta distributed heterogeneity, respectively, for the biologically relevant parameter range.

#### 4. Discussion

Host heterogeneity has been shown to have a significant effect on disease dynamics (Dye and Hasibeder, 1986; Lloyd-Smith et al., 2005; Alex Perkins et al., 2013; Smith et al., 2007; Diekmann et al., 1990). Of particular interest is the effect of heterogeneity in transmission on  $\mathcal{R}_0$  (Dye and Hasibeder, 1986; Lloyd-Smith et al., 2005; Alex Perkins et al., 2013; Smith et al., 2007). In an infinite population of susceptible hosts, heterogeneity has been shown to increase  $\mathcal{R}_0$ , whereas in a finite population, we show that heterogeneity has a more complicated effect on the reproductive number.

Smith et al. (2007) found that in a finite population, heterogeneity in the attractiveness to mosquitoes decreases the reproductive number; our results for simple heterogeneity (i.e. heterogeneity either mixing or probability for only susceptible or infected) support this (Figs. 4, 6, 8, 10, and B2). However, for heterogeneity in the mixing rate of both infected and susceptible individuals, we find that when the population is large compared to the homogeneous  $\mathcal{R}_0$ , heterogeneity increases  $\mathcal{R}_m(N)$  compared to the homogeneous  $\mathcal{R}_0$ ; and when the population is small relative to  $\mathcal{R}_0$ , heterogeneity decreases  $\mathcal{R}_m(N)$  compared to the homogeneous  $\mathcal{R}_0$  (Figs. 2 and B3).

Compared to the heterogeneous  $\mathcal{R}_0$ , the effect of small population decreases the reproductive number. In general heterogeneity increases the effective mixing rate, because the most susceptible individuals are also the most infectious individuals. This has complex effects in the finite population, when the size of the population is large relative to  $\mathcal{R}_0$ , few people are contacted multiple times, so increasing the mixing rate increases the reproductive number. However, when the size of the population is small relative to  $\mathcal{R}_0$ , many more people are contacted multiple times, absorbing some possible infections, reducing  $\mathcal{R}_m(N)$  (Figs. 2 and B3).

Smith et al. (2007) considered only heterogeneity in attractiveness to mosquitoes. Here, we consider five different types of heterogeneity: the four simple types outlined in the Section 1.1 and heterogeneity in the mixing rate of both susceptible and infected individuals (the last corresponds to Smith's assumptions). We find simple expressions for each of the five finite-population reproductive numbers in terms of the distribution of heterogeneity and the size of the population.

We show that  $\mathcal{R}_0$  is affected by both the choice of the family of distributions of the heterogeneity (e.g., gamma and log-normal) and the specific CV of its distribution. Figs. 3, 5, 7, 9, and 11 show the effect of the distribution on the finite-population reproductive numbers for a fixed coefficient of variation ( $CV=2$ ). While the distribution of heterogeneity for a fixed CV affects the finite-population reproductive numbers, the CV has a larger effect on how much the finite-population reproductive number is changed due to heterogeneity. Figs. 2, 4, 6, 8, 10, B2, and B3 show the effect of the coefficient of variation on the finite-population reproductive numbers. For very small values of the CV, the finite-population reproductive numbers converge on the homogeneous finite-population reproductive number; as CV increases, so does the effect of heterogeneity.

For simple heterogeneity ( $\mathcal{R}_{t_m}(N)$ ,  $\mathcal{R}_{s_m}(N)$ ,  $\mathcal{R}_{t_p}(N)$ , and  $\mathcal{R}_{s_p}(N)$ ), in which heterogeneity always decreases the finite-population reproductive numbers, classical calculations of  $\mathcal{R}_0$  are overestimating the diseases actual reproductive number. And although heterogeneity in the mixing rate has a more complicated effect on the finite-population reproductive number, it has the same implications for control: it suggests that for a disease with a large  $\mathcal{R}_0$  spreading in a small heterogeneous population, the actual reproductive number may be lower than the standard calculation of  $\mathcal{R}_0$ , making control easier than predicted.

#### Contributions

LK and JD conceived and analyzed the model, drafted the paper, and both authors have read and approved the paper.

### Acknowledgments

The authors thank David Earn, and Ben Bolker for useful comments and suggestions. We also thank Juliet Pulliam for helpful discussions about framing the research project.

### Appendix A. Calculation framework

#### A.1. Heterogeneity in transmission mixing $\mathcal{R}_{t_m}(N)$

We calculate the finite-population reproductive number for the case where we allow only transmission mixing rates,  $t_m$ , to vary. We find the expected number of infections from our known infectious individual is

$$\left\langle \frac{\bar{s}t(y)N}{\bar{s}t(y) + \bar{s}_m N} \right\rangle_x \quad (14)$$

$$\frac{\bar{s}t(y)N}{\bar{s}t(y) + \bar{s}_m N} \quad (15)$$

#### A.2. Heterogeneity in transmission probability $\mathcal{R}_{t_p}(N)$

We calculate the finite-population reproductive number for the case where we allow only probability of transmitting an infection per contact,  $t_p$ , to vary.

We find the expected number of infections from the known infectious individual is

$$\left\langle \frac{t(y)\bar{s}N}{t(y)\bar{s} + \bar{s}_m N} \right\rangle_x \quad (16)$$

$$\frac{t(y)\bar{s}N}{t(y)\bar{s} + \bar{s}_m N} \quad (17)$$

#### A.3. Heterogeneity in susceptible mixing, $\mathcal{R}_{s_m}(N)$

We calculate the finite-population reproductive number for the case where we allow only susceptible mixing rates,  $s_m$ , to vary. We find the expected number of new infections as follows:

$$\left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \quad (18)$$

#### A.4. Heterogeneity in susceptibility probability $\mathcal{R}_{s_p}(N)$

We calculate the finite-population reproductive number for the case where we allow only heterogeneity in the probability contracting an infection per contact,  $s_p$ .

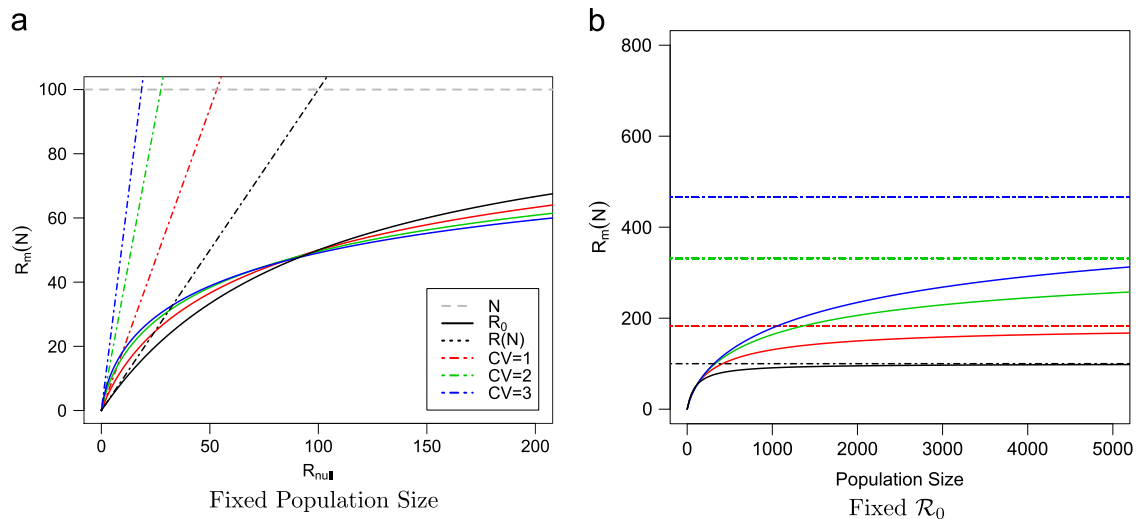
We find the expected number of new infections from a single infected individual of type  $y$  is:

$$\left\langle \frac{s(x)\bar{t}N}{s(x)\bar{t} + \bar{s}_m N} \right\rangle_x \quad (19)$$

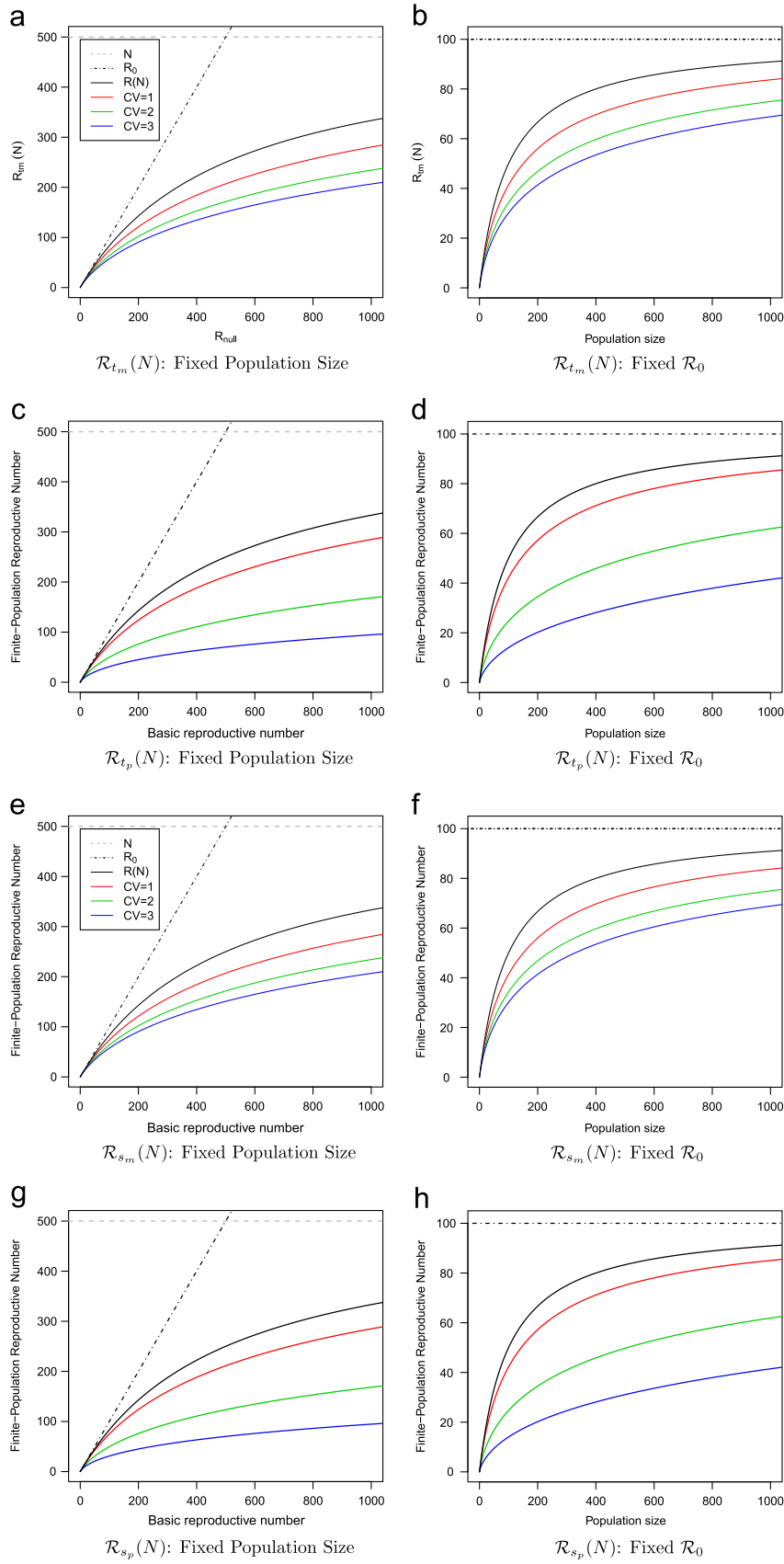
### Appendix B. Results

#### B.1. Additional figures

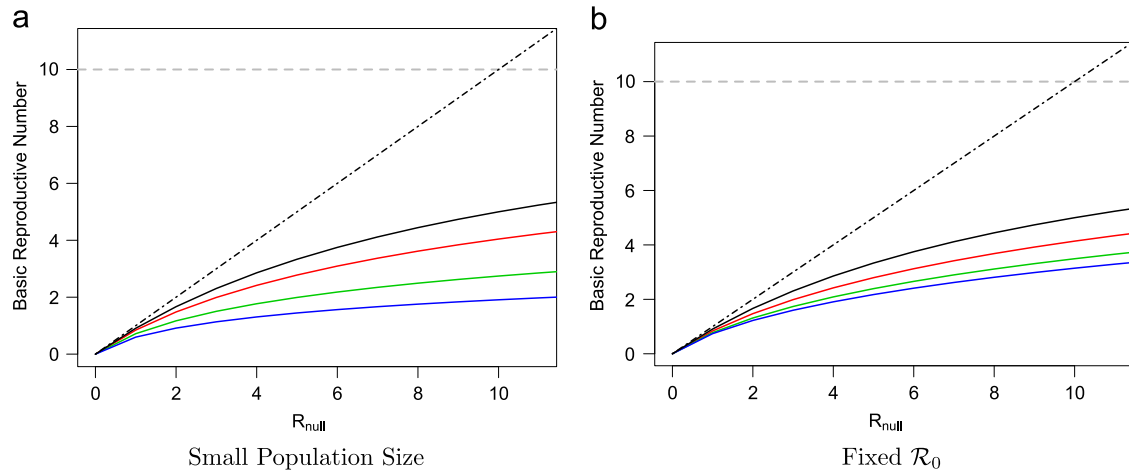
See Figs. B1, B2, and B3.



**Fig. B1.** The finite-population reproductive number,  $\mathcal{R}_m(N)$ , for log-normally distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a) The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=100$  (dashed line) and (b) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 100$  (dot-dashed line).



**Fig. B2.** Plot of the finite-population reproductive numbers, for log-normally distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. (a,c,e,g) the finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=100$  (dashed line) and (b,d,f,h) the finite-population reproductive numbers versus the population size for fixed  $\mathcal{R}_0 = 100$  (dot-dashed line). (a,b)  $\mathcal{R}_{tm}(N)$  (c,d)  $\mathcal{R}_{tp}(N)$  (e,f) and (g,h).



**Fig. B3.** The finite-population reproductive number,  $\mathcal{R}_m(N)$ , for (a) gamma distributed and (b) log-normally distributed heterogeneity. The solid lines are the finite-population reproductive numbers with different coefficients of variation and the dot-dashed lines represent the infinite reproductive numbers with corresponding coefficients of variation. The finite-population reproductive numbers versus the null reproductive numbers,  $\mathcal{R}_{null}$  with a fixed population of size  $N=10$  (dashed line).

## References

- Alex Perkins, T., Scott, Thomas W., Le Menach, Arnaud, Smith, David L., 2013. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. *PLoS Comput. Biol.* 9 (12), e1003327.
- Abu-Raddad, Laith J., Magaret, Amalia S., Celum, Connie, Wald, Anna, Longini Jr, Ira M., Self, Steven G., Corey, Lawrence, 2008. Genital herpes has played a more important role than any other sexually transmitted infection in driving HIV prevalence in Africa. *PLoS One* 3 (5), e2230.
- Ansell, J., Hamilton, K.A., Pinder, M., Walraven, G.E.L., Lindsay, S.W., 2002. Short-range attractiveness of pregnant women to *Anopheles gambiae* mosquitoes. *Trans. R. Soc. Trop. Med. Hyg.* 96 (2), 113–116.
- Bailey, Robert C., Moses, Stephen, Parker, Corette B., Agot, Kawango, Maclean, Ian, Krieger, John N., Williams, Carolyn FM., Campbell, Richard T., Ndiya-Achola, Jeckoniah O., 2007. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *The Lancet* 369 (9562), 643–656.
- Bassett, M.T., Mhloyi, M., 1991. Women and AIDS in Zimbabwe: the making of an epidemic. *Int. J. Health Serv.* 21 (1), 143–156.
- Bean, B., Moore, B.M., Sterner, B., Peterson, L.R., Gerding, D.N., Balfour, H.H., 1982. Survival of influenza viruses on environmental surfaces. *J. Infect. Dis.* 146 (1), 47–51.
- Brooks, Janet, 1996. The sad and tragic life of typhoid Mary. *CMAJ: Can. Med. Assoc. J.* 154 (6), 915.
- Chase-Topping, Margo, Gally, David, Low, Chris, Matthews, Louise, Woolhouse, Mark, 2008. Super-shedding and the link between human infection and livestock carriage of *Escherichia coli* O157. *Nat. Rev. Microbiol.* 6 (12), 904–912.
- Chimbindi, N.Z., McGrath, N., Herbst, K., San Tint, K., Newell, M.L., 2010. Socio-demographic determinants of condom use among sexually active young adults in rural KwaZulu-Natal, South Africa. *Open AIDS J.* 4, 88–95.
- De Benedictis, John, Chow-Shaffer, Esther, Costero, Adriana, Clark, Gary G., Edman, John D., Scott, Thomas W., 2003. Identification of the people from whom engorged *Aedes aegypti* took blood meals in Florida, Puerto Rico, using polymerase chain reaction-based DNA profiling. *Am. J. Trop. Med. Hyg.* 68 (4), 437–446.
- Diekmann, Odo, Heesterbeek, JAP, Metz, Johan A.J., 1990. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 28 (4), 365–382.
- Dushoff, Jonathan, 1999. Host heterogeneity and disease endemicity: a moment-based approach. *Theor. Popul. Biol.* 56 (3), 325–335.
- Dye, C., Hasibeder, G., 1986. Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others. *Trans. R. Soc. Trop. Med. Hyg.* 80 (1), 69–77.
- Earn, David J.D., Andrews, Paul W., Bolker, Benjamin M., 2014. Population-level effects of suppressing fever. *Proc. R. Soc. B: Biol. Sci.* 281 (1778), 20132570.
- Funk, S., Gilad, E., Watkins, C., Jansen, V.A., 2009. The spread of awareness and its impact on epidemic outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* 106 (April (16)), 6872–6877.
- Galvani, Alison P., May, Robert M., 2005. Epidemiology: dimensions of super-spreading. *Nature* 438 (7066), 293–295.
- Galvin, Shannon R., Cohen, Myron S., 2004. The role of sexually transmitted diseases in HIV transmission. *Nat. Rev. Microbiol.* 2 (1), 33–42.
- Glynn, Judith R., Caraël, Michel, Auvert, Betran, Kahindo, Maina, Chege, Jane, Musonda, Rosemary, Kaona, Frederick, Buvé, Anne, et al., 2001. Study Group on the Heterogeneity of HIV Epidemics in African Cities, 2001. Why do young women have a much higher prevalence of HIV than young men? A study in Kisumu, Kenya and Ndola, Zambia. *Aids* 15, S51–S60.
- Caroline Breese Hall, 2007. The spread of influenza and other respiratory viruses: complexities and conjectures. *Clin. Infect. Dis.* 45 (3), 353–359.
- Hethcote, Herbert W., Yorke, James A., Nold, Annett, 1982. Gonorrhea modeling: a comparison of control methods. *Math. Biosci.* 58 (1), 93–109.
- Ijumba, K., Gamielidien, R., Myer, L., Morroni, C., 2004. Sexual risk behaviors are influenced by knowing someone with HIV/AIDS. *S. Afr. Med. J.* 94 (July (7)), 522–523.
- Joseph Hotz Avner Ahituv, V., Philipson, Tomas, 1996. The responsiveness of the demand for condoms to the local prevalence of AIDS. *J. Hum. Resour.* 31 (4), 869–897.
- Kao, P.H., Yang, Ruey-Jen, 2006. Virus diffusion in isolation rooms. *J. Hosp. Infect.* 62 (3), 338–345.
- Karim, Yasmin G., Khalid Ijaz, M., Sattar, Syed A., Margaret Johnson-Lussenburg, C., 1985. Effect of relative humidity on the airborne survival of rhinovirus-14. *Can. J. Microbiol.* 31 (11), 1058–1061.
- Keegan, Lindsay, Dushoff, Jonathan, 2014. Analytic calculation of finite-population reproductive numbers for direct- and vector-transmitted diseases with homogeneous mixing. *Bull. Math. Biol.* (April), 1–12.
- Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond.: Ser. A* 115 (772), 700–721.
- Kilpatrick, A Marm, Daszak, Peter, Jones, Matthew J., Marra, Peter P., Kramer, Laura D., 2006. Host heterogeneity dominates West Nile virus transmission. *Proc. R. Soc. B: Biol. Sci.* 273 (1599), 2327–2333.
- Kline, Susan E., Hedemark, Linda L., Davies, Scott F., 1995. Outbreak of tuberculosis among regular patrons of a neighborhood bar. *N. Engl. J. Med.* 333 (4), 222–227, PMID:7791838.
- Lefèvre, Thierry, Gouagna, Louis-Clément, Roch Dabiré, Kounbo, Elguero, Eric, Fontenille, Didier, Renaud, François, Costantini, Carlo, Thomas, Frédéric, 2010. Beer consumption increases human attractiveness to malaria mosquitoes. *PLoS One* 5 (3), e9546.
- Lengeler, Christian, et al., 2004. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst. Rev.* 2 (2).
- Lindsay, Steve, Ansell, Juliet, Selman, Colin, Cox, Val, Hamilton, Katie, Walraven, Gijs, 2000. Effect of pregnancy on exposure to malaria mosquitoes. *The Lancet* 355 (9219), 1972.
- Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. *Nature* 438 (7066), 355–359.
- MacPhail, C., Terris-Prestholt, F., Kumaranayake, L., Ngoako, P., Watts, C., Rees, H., 2009. Managing men: women's dilemmas about overt and covert use of barrier methods for HIV prevention. *Cult. Health Sex.* 11 (June (5)), 485–497.
- May, Robert M., Anderson, Roy M., 1984. Spatial heterogeneity and the design of immunization programs. *Math. Biosci.* 72 (1), 83–111.
- Miller, William S., Artenstein, Malcolm S., 1967. Aerosol stability of three acute respiratory disease viruses. *Exp. Biol. Med.* 125 (1), 222–227.
- Nold, Annett, 1980. Heterogeneity in disease-transmission modeling. *Math. Biosci.* 52 (3), 227–240.
- Pettifor, A.E., Measham, D.M., Rees, H.V., Padian, N.S., 2004. Sexual power and HIV risk, South Africa. *Emerg. Infect. Dis.* 10 (November (11)), 1996–2004.
- Pitzer, Virginia E., Bowles, Cayley C., Baker, Stephen, Kang, Gagandeep, Balaji, Veeraraghavan, Farrar, Jeremy J., Grenfell, Bryan T., 2014. Predicting the impact of vaccination on the transmission dynamics of typhoid in south Asia: a mathematical modeling study. *PLoS Negl. Trop. Dis.* 8 (1), e2642.
- Port, G.R., Boreham, P.F.L., Bryan, Joan, 1980. The relationship of host size to feeding by mosquitoes of the *Anopheles gambiae* Giles complex (Diptera: Culicidae). *Bull. Entomol. Res.* 70 (01), 133–144.
- Quinn, Thomas C., Overbaugh, Julie, 2005. HIV/AIDS in women: an expanding epidemic. *Science* 308 (5728), 1582–1583.
- Ross, Joshua, 2011. Invasion of infectious diseases in finite homogeneous populations. *J. Theor. Biol.* 289, 83–89.

- Scott, Thomas W., Githeko, Andrew K., Fleisher, Andrew, Harrington, Laura C., Yan, Guiyun, 2006. Dna profiling of human blood in anophelines from lowland and highland sites in western Kenya. *Am. J. Trop. Med. Hyg.* 75 (2), 231–237.
- Shirai, O., Tsuda, Takao, Kitagawa, Shinya, Naitoh, Ken, Seki, Taisuke, Kamimura, Kiyoshi, Morohashi, Masaaki, 2002. Alcohol ingestion stimulates mosquito attraction. *J. Am. Mosq. Control Assoc.* 18 (2), 91–96.
- Shirai, Yoshikazu, Funada, Hisashi, Takizawa, Hisao, Seki, Taisuke, Morohashi, Masaaki, Kamimura, Kiyoshi, 2004. Landing preference of *Aedes albopictus* (Diptera: Culicidae) on human skin among ABO blood groups, secretors or nonsecretors, and ABH antigens. *J. Med. Entomol.* 41 (4), 796–799.
- Smith, David L., Ellis McKenzie, F., Snow, Robert W., Hay, Simon I., 2007. Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLoS Biol.* 5 (February (3)), 531–542.
- Stein, Richard A., 2011. Super-spreaders in infectious diseases. *Int. J. Infect. Dis.* 15 (August (8)), e510–e513.
- Stephens, T.P., McAllister, T.A., Stanford, K., 2009. Perineal swabs reveal effect of super shedders on the transmission of O157: H7 in commercial feedlots. *J. Anim. Sci.* 87 (12), 4151–4160.
- Wawer, Maria J., Gray, Ronald H., Sewankambo, Nelson K., Serwadda, David, Li, Xianbin, Laeyendecker, Oliver, Kiwanuka, Noah, Kigozi, Godfrey, Kiddugavu, Mohammed, Lutalo, Thomas, et al., 2005. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J. Infect. Dis.* 191 (9), 1403–1409.
- Williams, Brian G., Lloyd-Smith, James O., Gouws, Eleanor, Hankins, Catherine, Getz, Wayne M., Hargrove, John, De Zoysa, Isabelle, Dye, Christopher, Auvert, Bertran, 2006. The potential impact of male circumcision on HIV in sub-Saharan Africa. *PLoS Med.* 3 (7), e262.
- Woolhouse, M.E.J., Dye, C., Etard, J.F., Smith, T., Charlwood, J.D., Garnett, G.P., Hagan, P., Hii, J.L.K., Ndhlovu, P.D., Quinell, R.J., 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc. Natl. Acad. Sci.* 94 (1), 338.