Journal of Mathematical Biology (2021) 83:73
https://doi.org/10.1007/s00285-021-01686-z

Mathematical Biology

Vaccination and herd immunity thresholds in heterogeneous populations

Elamin H. Elbasha[1] · Abba B. Gumel[2,3]

Received: 26 February 2021 / Revised: 23 August 2021 / Accepted: 21 October 2021 /
Published online: 8 December 2021

Abstract

It has been suggested, without rigorous mathematical analysis, that the classical vaccine-induced herd immunity threshold (HIT) assuming a homogeneous population can be substantially higher than the minimum HIT obtained when considering population heterogeneities. We investigated this claim by developing, and rigorously analyzing, a vaccination model that incorporates various forms of heterogeneity and compared it with a model that considers a homogeneous population. By employing a two-group vaccination model in heterogeneous populations, we theoretically established conditions under which heterogeneity leads to different HIT values, depending on the relative values of the contact rates for each group, the type of mixing between the groups, the relative vaccine efficacy, and the relative population size of each group. For example, under biased random mixing assumption and when vaccinating a given group results in disproportionate prevention of higher transmission *per capita*, we show that it is optimal to vaccinate that group before vaccinating the other groups. We also found situations, under biased assortative mixing assumption, where it is optimal to vaccinate more than one group. We show that regardless of the form of mixing between the groups, the HIT values assuming a heterogeneous population are always lower than the HIT values obtained from a corresponding model with a homogeneous population. Using realistic numerical examples and parametrization (e.g., assuming assortative mixing together with vaccine efficacy of 95% and the value of the basic reproduction number, $\mathcal{R}_0$, of the model set at $\mathcal{R}_0 = 2.5$), we demonstrate that the HIT value generated from a model that considers population heterogeneity (e.g., biased assortative mixing) is significantly lower (40%) compared with a HIT value of 63% obtained if the model uses homogeneous population.

✉ Abba B. Gumel
agumel@asu.edu

[1] Merck & Co., Inc., Kenilworth, NJ, USA

[2] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA

[3] Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria 0002, South Africa

Keywords Basic reproduction number · Herd immunity threshold · Heterogeneity · Homogeneous population · Mixing pattern · SVIR model

Mathematics Subject Classification 92D30 · 60J85 · 60J28

1 Introduction

Since the identification of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the subsequent devastating impact of the 2019 novel coronavirus pandemic (COVID-19), a pertinent question being asked by the global public health and scientific research communities has been what is the minimum fraction of the unvaccinated susceptible population that needs to be immunized for the pandemic to end? It has been widely reported, in the literature for the mathematical modeling of infectious diseases, that the vaccine-derived herd immunity threshold (HIT) coverage needed to effectively eliminate an infectious disease, denoted by v^*, is given by the formula $v^* = 1 - \frac{1}{\mathcal{R}_0}$, where $\mathcal{R}_0$ is *the basic reproduction number* of the model, defined as the average number of secondary infections caused by a typical infective individual over his or her infectious period when introduced into a completely susceptible population (Randolph and Barreiro 2020). For example, if $\mathcal{R}_0 = 2.5$, then HIT can be achieved if at least 60% of the unvaccinated susceptible population is immunized (and, consequently, the disease can be eliminated from the population). However, this derivation for HIT is based on making several modeling assumptions regarding the properties of the vaccine and characteristics of the population. Specifically, it assumes that the vaccine efficacy to protect against the acquisition of infection is perfect and lasts throughout the lifetime of the vaccinated individual (i.e., the vaccine does not wane).

McLean and Blower (McLean and Blower 1993), and other researchers (Scherer and McLean 2002; Magpantay 2017), derived modified HIT formulae under various assumptions for vaccine properties. Their derivation is based on considering the following scenario for a cohort (childhood) vaccination program. Suppose a hypothetical vaccine against a childhood disease is efficacious in a fraction, ε, of recipients and confers full protection that wanes at a rate ω *per* unit time in a population with average life expectation of $\frac{1}{\mu}$, so that the average duration of vaccine-induced protection is $\frac{1}{\mu+\omega}$. In this case, the critical vaccination coverage (i.e., the proportion of newborns that need to be vaccinated to achieve herd immunity, consequently leading to disease elimination), denoted by v^{**}, is given by (Scherer and McLean 2002)

$$v^{**} = \left(\frac{\mu+\omega}{\mu}\right) \times \left(\frac{1}{\varepsilon}\right) \times \left(1 - \frac{1}{\mathcal{R}_0}\right) = \left(\frac{\mu+\omega}{\mu}\right) \times \left(\frac{1}{\varepsilon}\right) v^*,$$

where v^* is as defined above. Thus, the herd immunity threshold (v^*) needs to be adjusted upward to reflect the imperfect efficacy of the vaccine ($0 < \varepsilon < 1$) and the fraction of the average lifetime a vaccinated individual remains protected ($1/(\mu+\omega) \div 1/\mu$). For example, with $\mathcal{R}_0 = 2.5$ and a vaccine with protective efficacy of only 70% (i.e., $\varepsilon = 0.7$) that lasts 90% of a lifetime (i.e., $\frac{1}{\mu+\omega} \div \frac{1}{\mu} = 0.9$) is used in the population, at least 95.2% of the population of the unvaccinated susceptible newborns needs to be

vaccinated to achieve disease elimination (i.e., to achieve HIT), since $v^{**} = 0.952$ in this case.

The above derivations assume a homogeneous, well-mixed population. However, the transmission of many infectious diseases (such as SARS-CoV-2) occurs in a diverse heterogeneous population. Hence, a more realistic approach to carry out the above derivations and computations will be to account for the relevant heterogeneities (Britton et al. 2020). In other words, the derivations and computations need to be carried out for the case where the total population is sub-divided into multiple groups with similar characteristics, such as age, contact patterns, infectious period, or social, cultural, demographic, or geographic factors. For example, several mathematical models for disease transmission employ different mixing patterns, such as those between different age groups (Jacquez et al. 1996; Hethcote 2000). These contact patterns are typically parametrized using empirical or synthetic social contact matrices estimated from population-based surveys (Mossong et al. 2008).

2 Formulation of disease transmission model in heterogeneous populations

The multigroup vaccination model for the transmission dynamics of an infectious disease (such as SARS-CoV-2) in a heterogeneous population typically follows a standard susceptible-vaccinated-exposed-infected-recovered (SVEIR) Kermack-McKendrick-type compartmental modeling formulation (Hethcote 2000; Kermack and McKendrick 1927). In this formulation, the total population at time t (denoted by $N(t)$) is sub-divided into m distinct homogeneous groups. Each group is further sub-divided into five disjoint (mutually exclusive) classes or compartments of unvaccinated susceptible (S(t)), vaccinated susceptible ($V(t)$), exposed ($E(t)$), infectious ($I(t)$), and recovered/removed ($R(t)$), so that $N_i = S_i + V_i + E_i + I_i + R_i$, with $i = 1,2,\ldots, m$. Mathematically speaking, 'exposed individuals' are those who are newly infected with the pathogen but are not yet able to transmit the pathogen to other individuals (i.e., they are not infectious yet) (Hethcote 2000).

The resulting heterogenous multigroup SVEIR model, for the transmission dynamics of a disease in m heterogeneous groups or populations, is given by the following deterministic system of nonlinear differential equations (where a dot represents differentiation with respect to time t):

$$\begin{aligned}
\dot{S}_i(t) &= \Lambda_i + \omega_i V_i(t) - \lambda_i(t) S_i(t) - (\mu_i + \xi_i) S_i(t),\\
\dot{V}_i(t) &= \xi_i S_i(t) - (1-\varepsilon_i)\lambda_i(t) V_i(t) - (\mu_i + \omega_i) V_i(t),\\
\dot{E}_i(t) &= \lambda_i(t) S_i(t) + (1-\varepsilon_i)\lambda_i(t) V_i(t) - (\mu_i + \sigma_i) E_i(t),\\
\dot{I}_i(t) &= \sigma_i E_i(t) - (\mu_i + \gamma_i) I_i(t),\\
\dot{R}_i(t) &= \gamma_i I_i(t) - \mu_i R_i(t),
\end{aligned} \tag{1}$$

where the group-specific *force of infection*, $\lambda_i(t)$, is given by

$$\lambda_i(t) = \sum_{j=1}^{m} \beta_i a_i c_{ij} \frac{I_j(t)}{N_j(t)}, \tag{2}$$

with β_i as the transmission probability *per* contact for group i, a_i is the average number of contacts that an individual of group i has during a certain period of time (called group-specific activity level), and c_{ij} is the proportions of contacts that members of group i have with other individuals of group j. Mixing should meet the following closure relation (Glasser et al. 2012):

$$a_i c_{ij} N_i(t) = a_j c_{ji} N_j(t).$$

That is, the total number of contacts that individuals of group i have with other individuals of group j during a certain period of time should equal the total number of contacts that individuals of group j have with other individuals of group i. In the heterogeneous multigroup model (1), heterogeneity between groups is captured through differences in demographic rates (i.e., birth and death rates), transmission probability *per* contact, contact rates, progression and recovery rates, and vaccine efficacy and waning rates.

Adding all the equations of the heterogeneous multigroup model (1) gives the following equation for the rate of change of the total population:

$$\dot{N}_i(t) = \Lambda_i - \mu_i N_i(t).$$

In the heterogenous multigroup model (1), Λ_i is the *per capita* recruitment (birth) into the population, ω_i is the vaccine waning rate, $\lambda_i(t)$ is the force of infection, μ_i is the natural death rate (i.e., $1/\mu_i$ is the average lifespan of a person in group i) and ξ_i is the vaccination rate. Furthermore, $0 < \varepsilon_i < 1$ is the protective efficacy of the vaccine (against the acquisition of infection), σ_i is the rate at which exposed individuals develop clinical symptoms of the disease (i.e., $1/\sigma_i$ is the latency period of the disease), and γ_i is the recovery rate. Some of the main assumptions made in the formulation of the heterogeneous multigroup model (1) are:

(a) Within-group homogeneous mixing (i.e., although the model considers m heterogeneously mixed groups, contact patterns within each group is homogeneous).
(b) Exponentially distributed waiting time in each epidemiological compartment.
(c) The vaccine is not perfect (i.e., $0 < \varepsilon_i < 1$), and the protection offered by the vaccine wanes over time (i.e., $\omega > 0$). In addition, the vaccine has no therapeutic benefits.
(d) No disease-induced mortality (so that the total population in each group remains constant).
(e) Recovery induces permanent immunity against acquisition of future infection.
(f) All heterogenous groups can be identified and their relevant parameters (e.g., average number of contacts, contact pattern, transmission rate, and protective efficacy of the vaccine) are known.

2.1 Disease-free equilibrium of general model in heterogeneous populations

The special case of the heterogeneous multigroup model (1) where only a single group is considered (i.e., the model (1) with $m = 1$), has been subjected to rigorous mathematical analysis in the literature. Specifically, results for its well-posedness, invariance of its solutions, and existence and asymptotic stability of its equilibria (disease-free and endemic) have been established (Hethcote 2000; Gumel et al. 2020). These results are not repeated in this study.

The heterogeneous multigroup model (1) has a unique disease-free equilibrium, given by

$$\left(S_i^*, V_i^*, E_i^*, I_i^*, R_i^*\right) = \left(\frac{\Lambda_i\left(\omega_i + \mu_i\right)}{\mu_i\left(\omega_i + \mu_i + \xi_i\right)}, \frac{\Lambda_i \xi_i}{\mu_i\left(\omega_i + \mu_i + \xi_i\right)}, 0, 0, 0\right),$$
$$i = 1, 2, \ldots, m.$$

It is convenient to assume that the population in each group i has reached a stationary (equilibrium) state, such that

$$N_i(0) = N_i^* = \frac{\Lambda_i}{\mu_i}, \ with \ \ i = 0, 1, 2, \ \ldots, \ m$$

It is also convenient to work with the fraction of the population in each group. For example, the proportion of individuals in group i that are vaccinated (at the disease-free equilibrium) is given by:

$$v_i^* = \frac{V_i^*}{N_i^*} = \frac{\xi_i}{\left(\omega_i + \mu_i + \xi_i\right)}.$$

Thus, adding the fractions of all compartments of group i gives n_i, where

$$n_i^* = N_i^* \Big/ \sum_{j=1}^{m} N_i^*$$

is the fraction of total number of individuals in group i relative to the total population.

3 Analysis of model in heterogeneous populations with two risk groups

Here, we focus on deriving analytic expressions for HITs for a special case of the heterogeneous population model (1) that considers only two risk groups (i.e., the heterogenous multigroup model (1) with $m = 2$). For the resulting two-group model, we first derive an expression for *the vaccination reproduction number*, denoted by $\mathcal{R}_v$, and then find the minimum proportion of the unvaccinated susceptible individuals in the community that need to be vaccinated to reduce $\mathcal{R}_v$ to a value less than 1, so that HIT can be achieved.

3.1 Computation of vaccination and basic reproduction numbers for the two-group model

The next-generation operator method (van den Driessche and Watmough 2002; Diekmann et al. 2010) can be used to compute the vaccination reproduction number (and, subsequently, the basic reproduction number) of the special case of the heterogeneous multigroup model (1) with m = 2. It follows, using the notation in Driessche and Watmough (2002) on the resulting two-group model, that the associated non-negative matrix of new infection terms (F) and the M-Matrix (V) of linear transition terms in the infected compartments are given, respectively, by (where N_1^* and N_2^* are the total population of group 1 and group 2, respectively; similarly, v_1^* and v_2^* represent the HIT for groups 1 and 2, respectively):

$$F = \begin{pmatrix} 0 & 0 & \frac{\beta_1 a_1 c_{11}(S_1^*+(1-\varepsilon_1)V_1^*)}{N_1^*} & \frac{\beta_1 a_1 c_{12}(S_1^*+(1-\varepsilon_1)V_1^*)}{N_2^*} \\ 0 & 0 & \frac{\beta_2 a_2 c_{21}(S_2^*+(1-\varepsilon_2)V_2^*)}{N_1^*} & \frac{\beta_2 a_2 c_{22}(S_2^*+(1-\varepsilon_2)V_2^*)}{N_2^*} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and,

$$V = \begin{pmatrix} \mu_1+\sigma_1 & 0 & 0 & 0 \\ 0 & \mu_2+\sigma_2 & 0 & 0 \\ -\sigma_1 & 0 & \gamma_1+\mu_1 & 0 \\ 0 & -\sigma_2 & 0 & \gamma_2+\mu_2 \end{pmatrix},$$

from which it follows that the *vaccination reproduction number* of the two-group model is given by (where ρ is the spectral radius)

$$\mathcal{R}_v = \rho\left(FV^{-1}\right) = \frac{1}{2}\left(\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_2}\right), \tag{3}$$

where,

$$\Delta_1 = \left(1 - v_1^*\varepsilon_1\right)\mathcal{R}_{11} + \left(1 - v_2^*\varepsilon_2\right)\mathcal{R}_{22},$$

$$\Delta_2 = \left(1 - v_1^*\varepsilon_1\right)\left(1 - v_2^*\varepsilon_2\right)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22}).$$

In deriving Eq. (3), we utilized the following definition of the constituent basic reproduction numbers associated with disease transmission between individuals in group i with individuals in group j:

$$\mathcal{R}_{ij} = \beta_i a_i c_{ij} \frac{\sigma_j}{(\gamma_j + \mu_j)(\mu_j + \sigma_j)}, \tag{4}$$

where the index i represents group i and the index j represents group j. To obtain the *basic reproduction number* ($\mathcal{R}_0$) associated with the two-group model, we set the vaccination coverage rates in the expression for $\mathcal{R}_v$, given by (3), to zero (i.e., we set $v_1^* = v_2^* = 0$ in (3)). This gives,

$$\mathcal{R}_0 = \frac{1}{2}\left(\mathcal{R}_{11} + \mathcal{R}_{22} + \sqrt{\mathcal{R}_{1,1}^2 + 4\mathcal{R}_{1,2}\mathcal{R}_{2,1} - 2\mathcal{R}_{1,1}\mathcal{R}_{2,2} + \mathcal{R}_{2,2}^2}\right). \tag{5}$$

3.2 Herd immunity thresholds for a two-group model with heterogeneous populations

For the computation of herd immunity threshold of a two-group model (such as (1) with $m = 2$), the objective is to find the values of the respective HITs, v_1^* and v_2^*, such that the total vaccine coverage (i.e., the proportion of individuals in the community that is vaccinated), given by

$$\frac{V_1^* + V_2^*}{N_1^* + N_2^*} = \left(\frac{N_1^*}{N_1^* + N_2^*}\right)\frac{V_1^*}{N_1^*} + \left(\frac{N_2^*}{N_1^* + N_2^*}\right)\frac{V_2^*}{N_2^*} = n_1^* v_1^* + n_2^* v_2^*,$$

is at its minimum and vaccination reproduction number,$\mathcal{R}_v$, given by (3), is less than or equal to one. Formally, the associated optimization problem can be written as choosing v_1^* and v_2^* to

$$\text{minimize}(n_1^* v_1^* + n_2^* v_2^*)$$

subject to,

$$0 \leq v_1^* \leq 1, 0 \leq v_2^* \leq 1, \mathcal{R}_v \leq 1,$$

where $\mathcal{R}_v$ is given by Eq. (3). The solution of this nonlinear optimization problem will be characterized using a geometrical approach. Specifically, we compare the shape of the curve depicting values of vaccination coverage where the vaccination reproduction number is equal to one ($i.e.$, $\mathcal{R}_v = 1$) with the contour lines (or level sets) $n_1^* v_1^* + n_2^* v_2^*$ (as illustrated in Fig. 1). Each contour line represents the locus of vaccination coverage combinations (v_1^*, v_2^*) that yield the same level of total vaccination coverage at the population level. The blue contour lines correspond to lower levels of total vaccination coverage when moving in the southwestern direction toward the origin.

The solution of the equation of the orange curve $\mathcal{R}_v = 1$ yields a value of v_2^* as a function of v_1^*:

$$v_2^* = \frac{\mathcal{R}_{22} + (1 - v_1^*\varepsilon_1)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22} + \mathcal{R}_{11}) - 1}{\varepsilon_2\left[\mathcal{R}_{22} + (1 - v_1^*\varepsilon_1)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})\right]}. \tag{6}$$

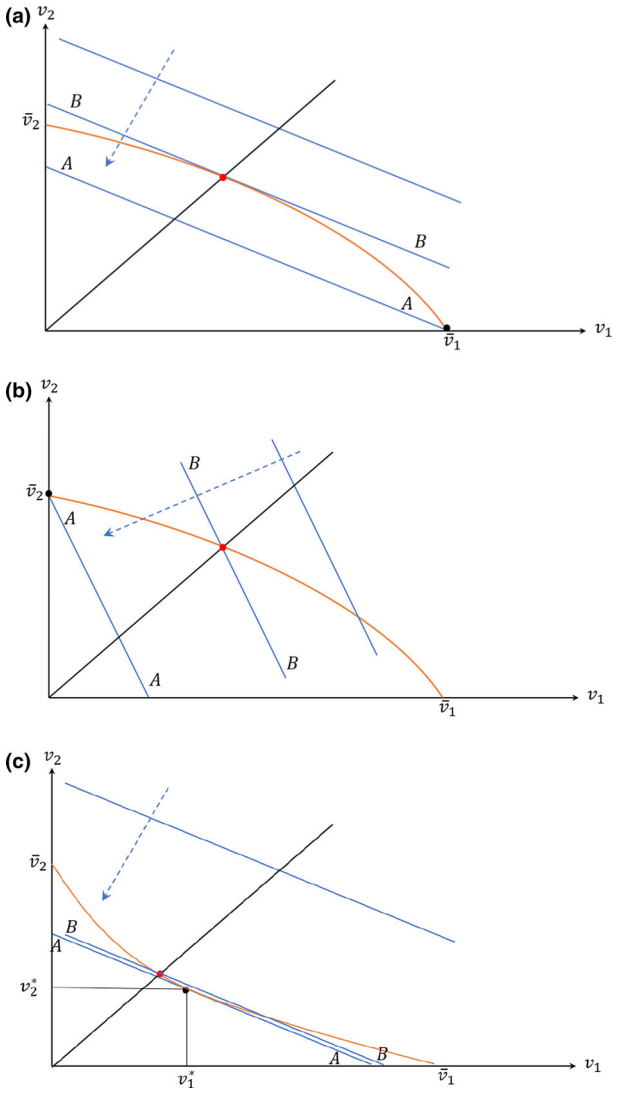

Fig. 1 Vaccination threshold values for group 1 and group 2. The orange curve shows values of vaccination coverage where the vaccinated reproductive number is equal to one: $\mathcal{R}_v = 1$. The slope of this curve when intersects with the y-axis is given by $-\frac{\varepsilon_1 \mathcal{R}_{12}\mathcal{R}_{21}}{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}$. The blue level curves show different values of total vaccination coverage going down in the direction of the origin: $n_1^* v_1^* + n_2^* v_2^*$. The slope of these level curves is $-\frac{n_1^*}{n_2^*}$. Under biased random mixing (**a**) when the blue line is flatter than the orange curve when it intersects the y-axis, the closest blue line to the origin that intersect the orange curve occurs when $v_1^* = \bar{v}_1 = \frac{\mathcal{R}_{11}+\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22}-1}{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}$, $v_2^* = 0$; (**b**) when the blue line is steeper than the orange curve when it intersects the x-axis, the closest blue line to the origin that intersect the orange curve occurs when $v_1^* = 0$, $v_2^* = \bar{v}_2 = \frac{\mathcal{R}_{11}+\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22}-1}{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}$. Under biased assortative mixing (**c**) the optimum occurs when the blue line is tangent to the orange curve. In the homogeneous population model ($v_1^* = v_2^*$), the optimum occurs at the intersection of the black 45-degree line with the orange curve

It follows from Eq. (6) that this function intersects the x-axis ($v_2^* = 0$) and y-axis ($v_1^* = 0$) at

$$\begin{aligned} v_1^* = \bar{v}_1 &= \frac{\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22} - 1}{\varepsilon_1(\mathcal{R}_{11} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})}, \\ v_2^* = \bar{v}_2 &= \frac{\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22} - 1}{\varepsilon_2(\mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})}, \end{aligned} \tag{7}$$

respectively. The slope of the orange curve (Fig. 1) is given by

$$\frac{dv_2^*}{dv_1^*} = -\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{\varepsilon_2[\mathcal{R}_{22} + (1 - v_1\varepsilon_1)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})]^2}. \tag{8}$$

The slope of the orange curve, at the point of intersection with the y-axis (i.e., $v_1^* = 0$), is given by

$$\frac{dv_2^*}{dv_1^*}|_{v_1^*=0} = -\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{\varepsilon_2(\mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})^2}, \tag{9}$$

and the slope of the orange curve when it intersects with the x-axis (i.e., $v_2^* = 0$) is given by

$$\frac{dv_2^*}{dv_1^*}|_{v_1^*=\bar{v}_1} = -\frac{\varepsilon_1(\mathcal{R}_{11} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22})^2}{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}}. \tag{10}$$

The equations of the blue level curves are: $n_1^*v_1^* + n_2^*v_2^* = A$, where A is an arbitrary constant (Fig. 1). The slope of these level curves is

$$\frac{dv_2^*}{dv_1^*} = -\frac{n_1^*}{n_2^*}. \tag{11}$$

Depending on the situation, the geometrical approach involves comparing the slopes of the curves given by Eqs. (8)–(11). We consider the following two scenarios, based on to which group(s) the vaccine is prioritized to.

Scenario 1: Vaccinating only one group

We start with the case where the value of $\mathcal{R}_v$ can be brought down to or below one by vaccinating either group 1 alone or group 2 alone. That is, the values of $\bar{v}_1$ and $\bar{v}_2$ given in Eq. (7) are less than 1.

We note that the type of mixing between groups (i.e., c_{ij}) determines the sign of the expression $\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22}$. By using Eq. (4), we can establish that

$$sign(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{11}\mathcal{R}_{22}) = sign(c_{12}c_{12} - c_{11}c_{22}).$$

We distinguish two types of mixing.

3.2.1 Biased random mixing: $\mathcal{R}_{12}\mathcal{R}_{21} \geq \mathcal{R}_{11}\mathcal{R}_{22}$

When transmission occurs more because of mixing between groups rather than because of mixing within the groups, we call this type of mixing *biased* toward *random mixing*. The *separable proportionate mixing* is a special type of biased random mixing (Hethcote 2000; Glasser et al. 2012). We consider the two boundary scenarios below

i. Vaccinating group 1 disproportionately contributes more to prevention of *per-capita* transmission. Here, the following inequality holds:

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} > \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*}.$$

Figure 1a depicts the situation where the absolute value of the slope of the orange curve when it intersects the y-axis (Eq. (9)) is greater than the absolute value of the slope of the blue line (Eq. (11)) (i.e., blue line is flatter). That is (hence the inequality above),

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2} > \frac{n_1^*}{n_2^*}.$$

In this case, the closest blue line to the origin that intersect the orange curve happens when

$$v_1^* = \bar{v}_1 = \frac{\mathcal{R}_{11}+\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22}-1}{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})},\ v_2^* = 0,$$

and the herd immunity threshold is given by

$$n_1^*v_1^* + n_2^*v_2^* = n_1^*\frac{\mathcal{R}_{11}+\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22}-1}{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}. \tag{12}$$

ii. Vaccinating group 1 disproportionately contributes less to prevention of per-capita transmission:

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} \leq \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*}.$$

In this case, the absolute value of the slope of the orange curve when it intersects the y-axis (Eq. (9)) is less than or equal the absolute value of the slope of the blue line (Eq. (11)) (i.e., blue line is steeper) (illustrated in Fig. 1b), and the herd immunity threshold is

$$v_1^* = 0,\ v_2^* = \bar{v}_2,$$

$$n_1^*v_1^* + n_2^*v_2^* = n_2^*\frac{\mathcal{R}_{11}+\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22}-1}{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}. \tag{13}$$

In summary, the above results show that, under the biased random mixing assumption, achieving herd immunity entails exclusively optimizing vaccination coverage among the group that results in relatively more prevention of *per-capita* disease transmission.

3.2.2 Biased assortative mixing: $\mathcal{R}_{12}\mathcal{R}_{21} < \mathcal{R}_{11}\mathcal{R}_{22}$

When transmission occurs more due to mixing within groups rather than due to mixing between groups, we refer to this type of mixing as *biased* toward *assortative mixing*. In this case, the optimization problem has both boundary (corner) and interior solutions. We start with the two boundary solutions followed by the interior solution.

i. Vaccinating group 1 disproportionately contributes more to prevention of *per-capita* transmission:

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} > \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*},$$
$$\frac{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}}{n_2^*} < \frac{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_1^*}.$$

In this case, the absolute values of the slope of the orange curve when it intersects the y-axis (Eq. (9)) and the x-axis (Eq. (10)) are greater than the absolute value of the slope of the blue line (Eq. (11)) (i.e., blue line is flatter). In this case,

$$v_1^* = \bar{v}_1,\ v_2^* = 0,$$

and the herd immunity threshold is given by Eq. (12). In other words, we achieve optimal results by allocating all necessary vaccine resources to group 1.

ii. Vaccinating group 1 disproportionately contributes less to prevention of *per-capita* transmission. Here, we have:

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} < \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*},$$
$$\frac{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}}{n_2^*} > \frac{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_1^*}.$$

In this case the absolute values of the slope of the orange curve when it intersects the y-axis (Eq. (9)) and the x-axis (Eq. (10)) are less than the absolute value of the slope of the blue line (Eq. (11)) (i.e., blue line is steeper). In this case,

$$v_1^* = 0, v_2^* = \bar{v}_2,$$

and the herd immunity threshold is given by Eq. (13). Here, optimal results are achieved by allocating all vaccine resources to group 2.

Scenario 2: Vaccinating both groups

This scenario includes one interior solution and 2 boundary solutions. For this scenario, we do not need to assume that $\mathcal{R}_v$ can be brought down to one by vaccinating either group 1 alone or group 2 alone.

i. Interior solution. For this solution to exist, the following inequalities must hold:

$$\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} > \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*},$$

and,

$$\frac{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}}{n_2^*} > \frac{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_1^*}.$$

When the absolute value of the slope of the blue line (Eq. (11)) is between the absolute values of the slope of the orange curve when it intersects the y-axis (Eq. (9)) and the x-axis (Eq. (10)), it is optimal to vaccinate both groups (illustrated in Fig. 1c). The interior solution is obtained when the slope of the orange curve (Eq. (8)) is equal to the slope of the blue curve (Eq. (11)). Thus, by equating the right-hand side of Eqs. (8) and (11), and solving for v_1^* and using the resulting solution in Eq. (6), we have:

$$v_1^* = \frac{n_1^*\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{22}\mathcal{R}_{11})-\sqrt{n_1^*\varepsilon_2 n_2^*\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}}{n_1^*\varepsilon_2\varepsilon_1(\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}$$

and,

$$v_2^* = \frac{\varepsilon_1 n_2^*(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{22}\mathcal{R}_{11})-\sqrt{n_1^*\varepsilon_2 n_2^*\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}}{\varepsilon_1 n_2^*\varepsilon_2(\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}.$$

It follows from the assumption of biased assortative mixing, and the inequalities above, that a positive interior solution exists if the following inequalities hold:

$$\mathcal{R}_{12}\mathcal{R}_{21} < \mathcal{R}_{11}\mathcal{R}_{22},$$

$$\frac{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}} < \frac{n_1^*}{n_2^*} < \frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}, \tag{14}$$

$$0 \le v_1^* \le 1, 0 \le v_2^* \le 1.$$

The herd immunity threshold is

$$n_1^*v_1^* + n_2^*v_2^* = \frac{\varepsilon_1 n_2^*(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{22}\mathcal{R}_{11})+\varepsilon_2 n_1^*(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{22}\mathcal{R}_{11})-2\sqrt{\varepsilon_1 n_2^*\varepsilon_2 n_1^*\mathcal{R}_{12}\mathcal{R}_{21}}}{\varepsilon_1\varepsilon_2(\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})}. \tag{15}$$

ii. Boundary solutions.
The case where $\mathcal{R}_v$ cannot be brought down to one by vaccinating either group 1 alone or group 2 alone (i.e., the values of either $\bar{v}_1$ or $\bar{v}_2$ are greater than 1) always leads to boundary solutions. For example, when $\frac{\varepsilon_1\mathcal{R}_{12}\mathcal{R}_{21}}{n_1^*} < \frac{\varepsilon_2(\mathcal{R}_{22}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_2^*}$ and $\frac{\varepsilon_2\mathcal{R}_{12}\mathcal{R}_{21}}{n_2^*} > \frac{\varepsilon_1(\mathcal{R}_{11}+\mathcal{R}_{12}\mathcal{R}_{21}-\mathcal{R}_{11}\mathcal{R}_{22})^2}{n_1^*}$ and $\bar{v}_2 > 1$, we first maximize vaccination coverage among group 2 and then find the value of coverage among group 1 that brings $\mathcal{R}_v$ to one. In this case,

$$v_1^* = \frac{\mathcal{R}_{11} + (1-\varepsilon_2)(\mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}) - 1}{\varepsilon_1[\mathcal{R}_{11} + (1-\varepsilon_2)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}]}, v_2^* = 1,$$

and the herd immunity threshold is

$$n_1^* v_1^* + n_2^* v_2^* = n_1^* \frac{\mathcal{R}_{11} + (1-\varepsilon_2)(\mathcal{R}_{22} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}) - 1}{\varepsilon_1[\mathcal{R}_{11} + (1-\varepsilon_2)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}]} + n_2^*. \quad (16)$$

Similarly, when $\bar{v}_1 > 1$,we may have,

$$v_1^* = 1, v_2^* = \frac{\mathcal{R}_{22} + (1-\varepsilon_1)(\mathcal{R}_{11} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}) - 1}{\varepsilon_2[\mathcal{R}_{22} + (1-\varepsilon_1)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}]},$$

and the herd immunity threshold is

$$n_1^* v_1^* + n_2^* v_2^* = n_1^* + n_2^* \frac{\mathcal{R}_{22} + (1-\varepsilon_1)(\mathcal{R}_{11} + \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}) - 1}{\varepsilon_2[\mathcal{R}_{22} + (1-\varepsilon_1)(\mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_{22}\mathcal{R}_{11}]}. \quad (17)$$

Thus, the optimal solutions and HIT values for the heterogeneous two-group vaccination model are summarized as follows:

1. The heterogeneous two-group vaccination model has an interior solution if the inequalities given in (14) are satisfied. The corresponding HIT value is given by Eq. (15);
2. The heterogeneous two-group vaccination model has four boundary solutions if any of the inequalities given in (14) does not hold. The associated HIT value is given by Eqs. (12), (13), (16), and (17);
3. The heterogeneous two-group vaccination model has no solution otherwise.

These analyses show that, for a two-group vaccination model in heterogeneous populations (such as (1) with m = 2), the optimum vaccination program depends on the relative values of the constituent reproduction numbers of the model (i.e., $\mathcal{R}_{12}$, $\mathcal{R}_{21,}\mathcal{R}_{11}, \mathcal{R}_{22}$), the relative vaccine efficacy, and the relative population sizes of the two groups. The values of the constituent reproduction numbers are determined by the type of mixing allowed or assumed between the two groups. When the mixing between the groups is biased towards random mixing, achieving herd immunity entails restricting vaccination coverage to the group that results in relatively more prevention of *per-capita* transmission. If herd immunity cannot be achieved by vaccinating one of the two groups alone, vaccination coverage among the group that results in relatively

more prevention of *per-capita* transmission should be maximized before vaccinating the other group. These scenarios occur under biased assortative mixing too, but scenarios involving vaccinating both groups are more common, including interior optimum where coverage of any of the two groups is less than 100%.

4 Analysis of model with a homogeneous population

The heterogeneous multigroup vaccination model (1) can be reduced to that with homogeneous population by assuming that each individual in any of the two groups is identical with every other individual in the community. That is, we achieve homogeneity by setting $c_{ij} = c, a_i = a, \xi_i = \xi, \varepsilon_i = \varepsilon, \mu_i = \mu, \Lambda_i = \Lambda,\ \sigma_i = \sigma,\ \gamma_i = \gamma$ for all i and j into the model (1). Specifically, the vaccination model with a homogeneous population is obtained from system (1) by dropping the group subscript i and re-defining the *force of infection* as

$$\lambda(t) = nca\beta \frac{I(t)}{N(t)}.$$

Using the next-generation operator method (van den Driessche and Watmough 2002; Diekmann et al. 2010), the *vaccination reproduction number* associated with the resulting homogeneous model, also denoted by $\mathcal{R}_v$, is given by:

$$\mathcal{R}_v = \left(1 - \frac{\varepsilon\xi}{\mu + \xi + \omega}\right)\mathcal{R}_0, \tag{18}$$

where the basic reproduction number, $\mathcal{R}_0$, is now given by

$$\mathcal{R}_0 = \frac{nca\beta\sigma}{(\gamma + \mu)(\mu + \sigma)}. \tag{19}$$

Since the vaccination coverage at the disease-free equilibrium (v^*) for the multigroup model with homogeneous population is

$$v^* = \frac{\sum_j^m V_j^*}{\sum_j^m N_j^*} = \frac{\xi}{\mu + \xi + \omega}, \tag{20}$$

it follows, by using (18) in (20), that

$$\mathcal{R}_v = \left(1 - \varepsilon v^*\right)\mathcal{R}_0. \tag{21}$$

It should be noted that this relationship between $\mathcal{R}_v$ and $\mathcal{R}_0$ (Eq. (21)) in the homogeneous population model can be obtained directly from the formula of the vaccination reproduction number of the heterogeneous population model (1) by substituting $\varepsilon_i = \varepsilon$ and $v_i^* = v^*,\ i = 1, 2,$ into Eq. (3) and using the definition of $\mathcal{R}_0$ given by Eq. (5).

Setting $\mathcal{R}_v$ (from (21)) to one and solving for v^* gives the following threshold value needed to achieve herd immunity for the model with homogeneous population (Scherer and McLean 2002; Hethcote 2000):

$$v^* = \frac{1}{\varepsilon} \times \left(1 - \frac{1}{\mathcal{R}_0}\right). \tag{22}$$

Thus, in constructing the homogeneous population model corresponding to the heterogeneous model (1) for comparison, we followed previous studies and matched the two modeling types (i.e., homogeneous vs. heterogeneous population) according to the expressions or values of their respective basic reproduction number, $\mathcal{R}_0$ (Andreasen 2011; Clancy and Pearce Christopher 2013).

4.1 Comparison of HIT values using homogeneous and heterogeneous population models under proportionate mixing

One of the simplest types of mixing in disease epidemiology is the separable proportionate mixing, in which the contacts of a person of group i are distributed over those of other groups in proportion to the activity levels and sizes of the other groups (Hethcote 2000). Thus, with separable proportionate mixing, proportions of contacts that members of group i have with group j, c_{ij}, is given by

$$c_{ij} = \frac{a_j N_j^*}{\sum_{k=1}^{m} a_k N_k^*},$$

where a_j and N_j^* are as defined before.

By substituting this definition of c_{ij} into the formulae for the constituent reproduction numbers in Eq. (4), it can be shown that the assumption of proportionate mixing implies that $\mathcal{R}_{12}\mathcal{R}_{21} = \mathcal{R}_{11}\mathcal{R}_{22}$, which, in turn, implies $\Delta_2 = 0$. Thus, for this scenario of proportionate mixing, the vaccination reproduction number reduces to,

$$\mathcal{R}_v = \left(1 - v_1^*\varepsilon_1\right)\mathcal{R}_{11} + \left(1 - v_2^*\varepsilon_2\right)\mathcal{R}_{22}.$$

We now seek to answer the question: what are the values of v_1^* and v_2^* such that the total vaccination coverage, $n_1^* v_1^* + n_2^* v_2^*$, is at its minimum and $\mathcal{R}_v \leq 1$? The answer depends on the relationship between the ratio of constituent reproduction numbers ($\frac{\mathcal{R}_{11}}{\mathcal{R}_{22}}$) adjusted by efficacy and the ratio of population ($\frac{n_1^*}{n_2^*}$) in the two groups. The solution of this simple linear programming problem is a special case of the biased random mixing where $\mathcal{R}_{12}\mathcal{R}_{21} = \mathcal{R}_{11}\mathcal{R}_{22}$. As before, there are three scenarios:

Scenario (i): Vaccinating group 1 disproportionately contributes more to prevention of *per-capita* transmission: $\frac{\varepsilon_1\mathcal{R}_{11}}{n_1^*} > \frac{\varepsilon_2\mathcal{R}_{22}}{n_2^*}$.

In this case,

$$v_1^* = \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_1\mathcal{R}_{11}}, \; v_2^* = 0,$$

and the herd immunity threshold is

$$n_1^* v_1^* + n_2^* v_2^* = n_1^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_1 \mathcal{R}_{11}}. \tag{23}$$

Scenario (ii): Vaccinating group 1 disproportionately contributes less to prevention of *per-capita* transmission: $\frac{\varepsilon_1 \mathcal{R}_{11}}{n_1^*} < \frac{\varepsilon_2 \mathcal{R}_{22}}{n_2^*}$.

In this case,

$$v_1^* = 0, \ v_2^* = \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_2 \mathcal{R}_{22}},$$

and the herd immunity threshold is

$$n_1^* v_1^* + n_2^* v_2^* = n_2^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_2 \mathcal{R}_{22}}. \tag{24}$$

Scenario (iii): Vaccinating both groups contribute equally to prevention of transmission: $\frac{\varepsilon_1 \mathcal{R}_{11}}{\varepsilon_2 \mathcal{R}_{22}} = \frac{n_1^*}{n_2^*}$

In this case, values of v_1^* and v_2^* such that

$$n_1^* v_1^* + n_2^* v_2^* = n_1^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_1 \mathcal{R}_{11}} = n_2^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon_2 \mathcal{R}_{22}} \tag{25}$$

will yield the minimum fraction that need to be vaccinated to achieve herd immunity.

To facilitate the comparison of the herd immunity thresholds between homogeneous and heterogeneous population models, we need to make sure efficacy in the two models is the same. One approach is to assume that vaccine efficacy does not vary across the two groups. That is, we set $\varepsilon_1 = \varepsilon_2 = \varepsilon$.

If $\frac{\varepsilon_1 \mathcal{R}_{11}}{n_1^*} > \frac{\varepsilon_2 \mathcal{R}_{22}}{n_2^*}$, the threshold vaccine coverage under heterogeneous population model is (given by the right-hand side of Eq. (23)):

$$n_1^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon \mathcal{R}_{11}},$$

and that for the homogeneous population model is given by

$$v^* = \frac{\mathcal{R}_0 - 1}{\varepsilon \mathcal{R}_0} = \frac{\mathcal{R}_{11} + \mathcal{R}_{22} - 1}{\varepsilon(\mathcal{R}_{11} + \mathcal{R}_{22})},$$

since $\mathcal{R}_0 = \mathcal{R}_{11} + \mathcal{R}_{22} > 1$ under proportionate mixing.

It can be shown that

$$n_1^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon \mathcal{R}_{11}} < v^* = \frac{\mathcal{R}_{11} + \mathcal{R}_{22} - 1}{\varepsilon(\mathcal{R}_{11} + \mathcal{R}_{22})}.$$

Upon simplifications, the above inequality holds if

$$n_1^* \frac{1}{\mathcal{R}_{11}} < \frac{1}{\mathcal{R}_{11} + \mathcal{R}_{22}},$$

Noting (and using) our starting assumption $\frac{\mathcal{R}_{11}}{\mathcal{R}_{22}} > \frac{n_1^*}{n_2^*}$, it follows that $\frac{\mathcal{R}_{22}}{\mathcal{R}_{11}}+1 < \frac{n_2^*}{n_1^*}+1$ or (upon further algebraic manipulation)

$$n_1^* \frac{1}{\mathcal{R}_{11}} < \frac{1}{\mathcal{R}_{11} + \mathcal{R}_{22}}.$$

Therefore, it follows, from the above inequality, that the threshold vaccine coverage in the heterogeneous population model, under scenario (i), is always less than that in the corresponding homogeneous population model.

Similarly, if $\frac{\varepsilon_1 \mathcal{R}_{11}}{n_1^*} < \frac{\varepsilon_2 \mathcal{R}_{22}}{n_2^*}$, we can follow the same approach above and show that the threshold vaccine coverage under heterogeneous population model, under scenario (ii), given by Eq. (24) is always lower than that under the homogeneous population:

$$n_1^* v_1^* + n_2^* v_2^* = n_2^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon \mathcal{R}_{22}} < v^* = \frac{\mathcal{R}_{11} + \mathcal{R}_{22} - 1}{\varepsilon(\mathcal{R}_{11} + \mathcal{R}_{22})}.$$

When $\frac{\varepsilon_1 \mathcal{R}_{11}}{n_1^*} = \frac{\varepsilon_2 \mathcal{R}_{22}}{n_2^*}$, it follows, under equal vaccine efficacy, that $\frac{\mathcal{R}_{11}}{\mathcal{R}_{22}}+1 = \frac{n_1^*}{n_2^*}+1$ or

$$n_2^* \frac{1}{\mathcal{R}_{22}} = \frac{1}{\mathcal{R}_{11} + \mathcal{R}_{22}}.$$

Thus, the threshold vaccine coverage for heterogeneous population model, under scenario (iii), given by Eq. (25) is always equal to that under the homogeneous population:

$$n_1^* v_1^* + n_2^* v_2^* = n_2^* \frac{(\mathcal{R}_{11} + \mathcal{R}_{22} - 1)}{\varepsilon \mathcal{R}_{22}} = v^* = \frac{\mathcal{R}_{11} + \mathcal{R}_{22} - 1}{\varepsilon(\mathcal{R}_{11} + \mathcal{R}_{22})}.$$

Therefore, under the form of proportionate mixing between groups, we show analytically that the HIT value in the heterogeneous population model is always less than or equal to the HIT value in a homogeneous population model. This result is consistent with that reported by Britton et al. (Britton et al. 2020), which showed, via numerical simulation of an age-structured model with mixing rates fitted to social activity, that if the basic reproduction number of the model is $\mathcal{R}_0 = 2.5$, the HIT is 43%, which is significantly lower than the HIT value of 60% obtained for the corresponding model that uses homogenous immunization of the population.

4.2 Comparison of HIT values using homogeneous and heterogeneous population models under general mixing

To show that the HIT value in the heterogeneous population model is always less than or equal to the HIT value in a homogeneous population model under general mixing between groups, we utilize the geometrical approach depicted in Fig. 1. Recall that the blue level curve represents total vaccination coverage. Thus, line *AA* corresponds to HIT value in heterogeneous population model, whereas *BB* corresponds to HIT value in the homogeneous population model. We consider scenarios with two boundary solutions and one interior solution:

i. Vaccinating group 1 only (Fig. 1a). In this case, line *AA* is closer to the origin than line *BB*. Hence, the HIT value in heterogeneous population model is lower than the HIT value in the homogeneous population model.
ii. Vaccinating group 2 only (Fig. 1b). In this case, line *AA* is closer to the origin than line *BB*. Hence, the HIT value in heterogeneous population model is lower than the HIT value in the homogeneous population model.
iii. Vaccinating both group 1 and group 2 (Fig. 1c). In this case, line *AA* is closer to the origin than line *BB*. Hence, the HIT value in heterogeneous population model is lower than the HIT value in the homogeneous population model.

5 Numerical analysis of HIT values in heterogeneous and homogeneous population models

Figure 2 numerically illustrates the different scenarios leading to different HIT values in the heterogeneous population model and compare them with the HIT values in a corresponding homogeneous population model. The orange curve shows values of vaccination coverage where $\mathcal{R}_v = 1$ and the blue level curves show different values of total vaccination coverage going down in the direction of the origin ($n_1^* v_1^* + n_2^* v_2^*$). In a homogeneous population model, vaccination coverage (determined by the intersection with the orange curve) is uniform across the two groups as shown by the dotted black 45-degree line. In all scenarios, we set vaccine efficacy to 95% across the two groups.

(a) The chosen parameter values represent a situation of biased random mixing between groups, and $\bar{v}_1 < 1$ and $\bar{v}_2 < 1$. As a result, the blue line is steeper than the orange curve when the latter intersects the y-axis, and the closest blue line to the origin that intersect the orange curve happens when $v_2^* = \bar{v}_2 = 0.69$, $v_1^* = 0$ (Fig. 2a). Given that group 2 represents only 25% of the population, the overall HIT value is just 17.2% compared with the HIT value in a homogeneous population model of 49.3%.
(b) The chosen parameter values represent a scenario of biased assortative mixing between groups,$v_1 < 1$ and $\bar{v}_2 < 1$, and the inequalities in (14) are satisfied. As a result, it is optimal to vaccinate both groups ($v_2^* = 0.789$, $v_1^* = 0.594$) for an overall HIT value of 64.3% (Fig. 2b). The HIT value in a homogeneous population model is 64.9%.

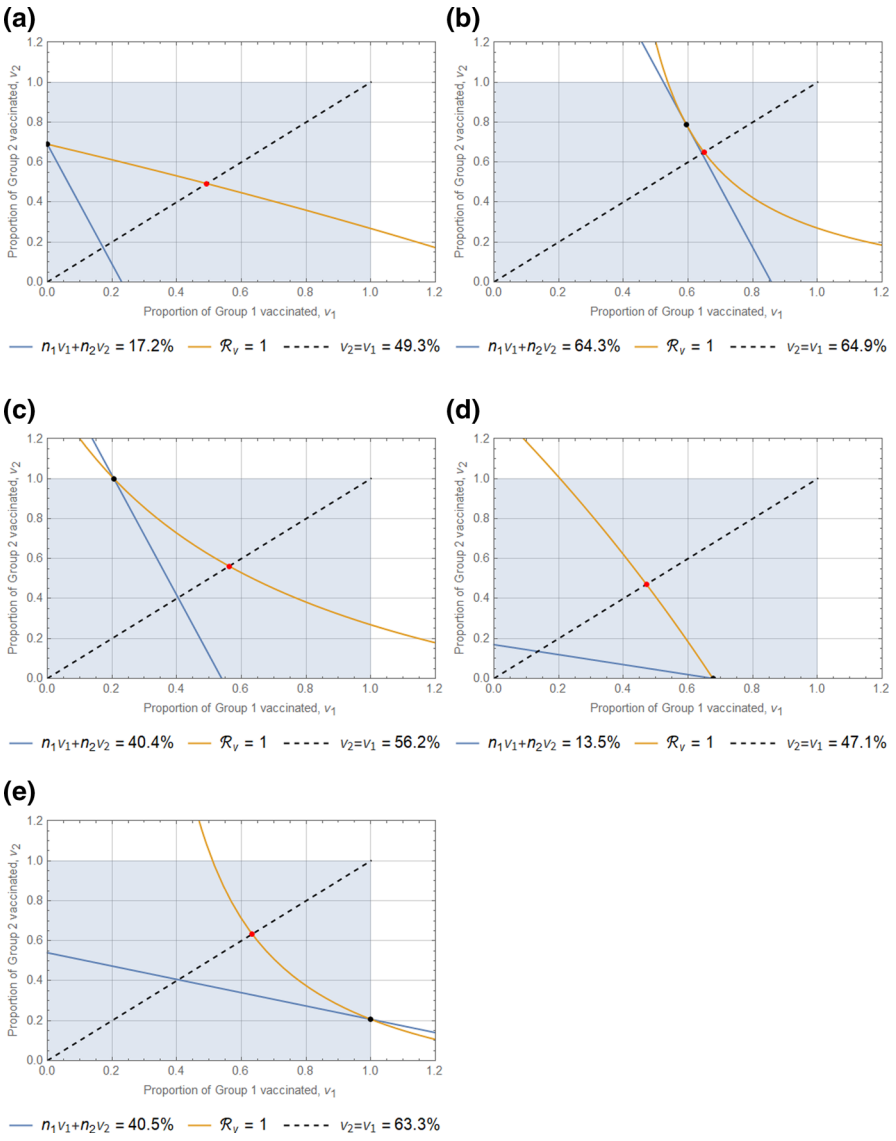

Fig. 2 Illustration of optimal least vaccine coverage determination by group that satisfies the constraint $\mathcal{R}_v = 1$ (orange curve). Parameter values: $\varepsilon = 0.95, \varepsilon_1 = 0.95, \varepsilon_2 = 0.95$. (**a**–**c**) $\mathcal{R}_{12} = 1.0$, $\mathcal{R}_{21} = 0.8, \mathcal{R}_{22} = 1.3, N_1^* = 0.75, N_2^* = 0.25$. (**a**) $\mathcal{R}_{11} = 0.5$, $\mathcal{R}_0 = 1.88$; (**b**) $\mathcal{R}_{11} = 2$, $\mathcal{R}_0 = 2.6$; (**c**) $\mathcal{R}_{11} = 1.2$, $\mathcal{R}_0 = 2.16$; (**d**) $\mathcal{R}_{12} = 0.8$, $\mathcal{R}_{21} = 1.0, \mathcal{R}_{22} = 0.5, N_1^* = 0.75, N_2^* = 0.25, \mathcal{R}_{11} = 0.5$, $\mathcal{R}_0 = 1.88$; (**e**) $\mathcal{R}_{12} = 0.8$, $\mathcal{R}_{21} = 1.0, \mathcal{R}_{22} = 0.5, N_1^* = 0.2, N_2^* = 0.8, \mathcal{R}_{11} = 1.9$, $\mathcal{R}_0 = 2.51$

(c) The chosen parameter values represent a situation of biased assortative mixing between groups, and $\bar{v}_1 > 1$ and $\bar{v}_2 > 1$. As a result, the blue line is steeper than the orange curve, and it is optimal to vaccinate all of group 2 and 20.6% of group 1 ($v_2^* = 1,\ v_1^* = 0.206$) for an overall HIT value of 40.4% (Fig. 2c). The HIT value in a homogeneous population model is 56.2%.

(d) The chosen parameter values represent a scenario of biased random mixing between groups and $\bar{v}_1 < 1$ and $\bar{v}_2 > 1$. As a result, the blue line is flatter than the orange curve, and it is optimal to vaccinate group 1 only ($v_2^* = 0,\ v_1^* = 0.713$) for an overall HIT value of 13.5% (Fig. 2d). The HIT value in a homogeneous population model is 47.1%.

(e) The chosen parameter values represent a situation of biased assortative mixing between groups and $\bar{v}_1 > 1$ and $\bar{v}_2 > 1$. As a result, the blue line is flatter than the orange curve, and it is optimal to vaccinate all of group 1 and 20.6% of group 2 ($v_1^* = 1,\ v_2^* = 0.207$) for an overall HIT value of 40.5% (Fig. 2e). The HIT value in a homogeneous population model is 63.3%.

Figure 3 illustrates different optimal solutions and HIT values for various values of the basic reproduction numbers. The figure shows the situation where, relative to its small size, group 2 is contributing more to transmission for low values of $\mathcal{R}_{11}$ and $\mathcal{R}_0$, and there is a need to vaccinate more people from group 2. As $\mathcal{R}_{11}$(and $\mathcal{R}_0$) increases, necessary vaccination coverage among group 2 increases until all of group 2 is vaccinated. As $\mathcal{R}_{11}$(and $\mathcal{R}_0$) increases further, both groups are vaccinated, but vaccination coverage among group 1 increases; whereas that among group 2 decreases. Of note, the herd immunity threshold for the homogeneous population (red curve) is consistently higher, but the difference between the two shrinks as the basic reproduction number increases.

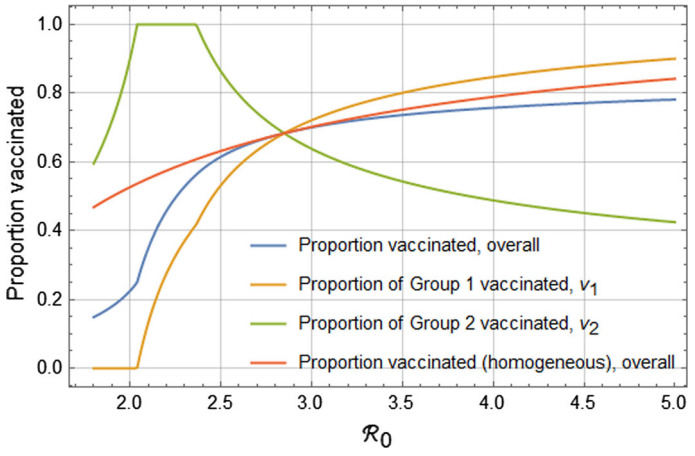

Fig. 3 Vaccine coverage above which herd immunity is achieved by group and basic reproduction number. Parameters values: $\mathcal{R}_{12} = 1.0,\ \mathcal{R}_{21} = 0.8, \mathcal{R}_{22} = 1.3, \varepsilon = 0.95, \varepsilon_1 = 0.95, \varepsilon_2 = 0.95, \mathrm{N}_1^* = 0.75, \mathrm{N}_2^* = 0.25$. Using the definition of $\mathcal{R}_0$, we chose $\mathcal{R}_{11} = \left(\mathcal{R}_0^2 - \mathcal{R}_{12}\mathcal{R}_{21} - \mathcal{R}_0\mathcal{R}_{22}\right)/(\mathcal{R}_0 - \mathcal{R}_{22})$.

6 Discussion

Our theoretical results, for a heterogeneous two-group vaccination model, suggest that deriving the vaccine-induced herd immunity threshold (HIT) in a heterogeneous population model is much more complex than deriving HIT for the corresponding model with homogeneous population. Our study shows that the HIT for each vaccinated group depends on the relative values of the constituent reproduction numbers, the relative vaccine efficacy, and the relative population sizes of the two groups. The values of the reproduction numbers are determined by the level and duration of infectiousness of a contact for each group, contact rates for each group, as well as the type of mixing between the two groups. We show that, under biased random mixing assumption and when vaccinating a given group results in disproportionate prevention of higher transmission *per capita*, it is optimal to prioritize vaccination of that group before vaccinating the other groups. We also found situations, under biased assortative mixing assumption, where it is optimal to vaccinate more than one group.

We show that population heterogeneities tend to result in lower HIT values, compared with the corresponding case with homogeneous population. This is true under both proportionate and other types of mixing among heterogeneous populations. Using realistic numerical examples and parametrizations (e.g., assuming biased assortative mixing with vaccine efficacy of 95% and basic reproduction number, $\mathcal{R}_0$, set at $\mathcal{R}_0 = 2.5$), we illustrate this finding, where the HIT value considering heterogeneity is shown to be significantly lower (40%) compared with a HIT value assuming a homogeneous population of 63%. It should be noted that the above parametrizations are consistent with the transmission dynamics of SARS-CoV-2 in the US where the Pfizer-BioNTech COVID-19 (BNT162b2) vaccine and the Moderna COVID-19 (mRNA-1273) vaccine (with estimated protective efficacy of 95%) are used (United States Food and Drug Administration 2020a, b).

Although our rigorous theoretical analysis is based on using a heterogeneous two-group vaccination model, our findings can be extended to models of multiple (more than two) heterogeneous groups. Although, admittedly more complex than that of a two-group heterogenous model, the rigorous analysis of more realistic models with many (more than two) heterogeneous groups can be conducted using the methods introduced and/or used in this study.

In our heterogeneous multi-group model, we assume that the multiple heterogenous groups can be identified, and that the relevant parameters of the model associated with these groups are known or can be estimated from available data. For a number of disease types and transmission settings, heterogeneities can be accounted for by easily stratifying the total population in terms of groups with similar characteristics, such as demographic, social, cultural, or geographic factors. Data to estimate the sizes of these groups are readily available. For example, population census data can be used to stratify populations according to age or sex. Data from social surveys, epidemiological studies, and clinical trials can be used to estimate model parameters such as mixing preferences (Wallinga et al. 2006), risk of acquiring infection, and vaccine efficacy and duration of protection. However, some populations may not be readily categorizable in terms of such heterogeneities and evidence on some parameter values may be lacking. For example, it may not be practically feasible to identify groups who disproportionately

contribute more to transmission than others (e.g., super spreaders). In addition, as is the case with many epidemiological outcomes predicted using modeling, estimated HIT values obtained from modeling are contingent on the availability, validity, and generalizability of parameter estimations (Metcalf et al. 2015a; Holmdahl and Buckee 2020).

It has been recognized, since the 1970s (Smith 1970; Dietz et al. 1975), that HIT, under homogeneous mixing population and sterilizing vaccine-derived immunity, follows a simple formula: the critical proportion of the population that must be vaccinated to achieve elimination should exceed $1-\frac{1}{\mathcal{R}_0}$ (Fine et al. 2011; Fine 1993; Metcalf et al. 2015b). Research by McLean and colleagues (McLean and Blower 1993; Scherer and McLean 2002) has addressed the complexities of imperfect immunity and vaccine-derived duration of protection, and derived HIT values when vaccination does not confer perfect, long-lasting immunity against infection to all recipients.

Although the importance of population heterogeneity and its effect on HIT values has been emphasized (Fox et al. 1971; Anderson and May 1985), rigorous theoretical work on herd immunity using mathematical models that consider heterogeneous populations is very rare. Most of the work that considers the complications induced by heterogeneity relied heavily on numerical simulations (largely owing to the fact that models that incorporate heterogeneity tend to be not readily amenable, or tractable, to rigorous mathematical analysis). For example, Britton et al. (Britton et al. 2020) show, via numerical simulations of a disease transmission model with basic reproduction number $\mathcal{R}_0$ set at 2.5, that the classical HIT value assuming a homogeneous population is substantially higher (60%) than the minimum HIT obtained when considering population heterogeneities in an age-structured population with mixing rates fitted to social activity (43%).

Our rigorous mathematical analysis supports the conclusion that the HIT values assuming a heterogeneous population are always lower than the HIT values obtained from a corresponding model with a homogeneous population. In addition, we show that there may not be a unique HIT for populations. For example, under biased random mixing assumption and when vaccinating a given group results in disproportionate prevention of higher transmission *per capita*, we show that it is optimal to vaccinate this group in its entirety before vaccinating the other groups. We also found situations where it is optimal to vaccinate more than one group at different rates.

Acknowledgements ABG acknowledges the support, in part, of the Simons Foundation (Award #585022) and the National Science Foundation (DMS-2052363). The authors are grateful to the anonymous reviewers and the handling editor for their very constructive comments.

References

Anderson RM, May RM (1985) Vaccination and herd immunity to infectious diseases. Nature 318(6044):323–329

Andreasen V (2011) The final size of an epidemic and its relation to the basic reproduction number. Bull Math Biol 73:2305–2321

Britton T, Ball F, Trapman P (2020) A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 369(6505):846–849

Clancy D, Pearce Christopher J (2013) The effect of population heterogeneities upon spread of infection. J Math Biol 67:963–987
Diekmann O, Heesterbeek JA, Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models. J R Soc Interface 7(47):873–885
Dietz K (1975) Transmission and control of arbovirus diseases. In: Ludwig D, Cooke KL (eds) Epidemiology. Society for Industrial and Applied Mathematics, Philadelphia PA, pp 104–121
Fine PEM (1993) Herd immunity: history, theory, practice. Epidemiol Rev 15:265–302
Fine P, Eames K, Heymann DL (2011) Herd immunity: a rough guide. Clin Infect Dis 52(7):911–916
Fox JP, Elveback L, Scott W, Gatewood L, Ackerman E (1971) Herd immunity: basic concept and relevance to public health immunization practices. Am J Epidemiol 141(3):187–197
Glasser J, Feng Z, Moylan A, Del Valle S, Castillo-Chavez C (2012) Mixing in age-structured population models of infectious diseases. Math Biosci 235(1):1–7
Gumel AB, Iboi EA, Ngonghala CN, Elbasha EH (2020) A primer on using mathematics to understand COVID-19 dynamics: modeling, analysis and simulations. Infect Dis Model 6:148–168
Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42:599–653
Holmdahl I, Buckee C (2020) Wrong but useful: what Covid-19 epidemiologic models can and cannot tell us. N Engl J Med 383(4):303–305
Jacquez JA, Simon CP, Koopman JS (1996) Core groups and the R0's for subgroups in heterogeneous SIS models. In: Mollison D (ed) Epidemic models: their structure and relation to data. Cambridge University Press, Cambridge, UK, pp 279–301
Kermack WO, McKendrick AG (1927) Contributions to the mathematical theory of epidemics, part 1. Proc Roy Soc Lond Ser A 115:700–721
Magpantay FMG (2017) Vaccine impact in homogeneous and age-structured models. J Math Biol 75(6–7):1591–1617
McLean AR, Blower SM (1993) Imperfect vaccines and herd immunity to HIV. Proc R Soc Lond B 253:9–13
Metcalf CJ, Andreasen V, Bjørnstad ON, Eames K, Edmunds WJ, Funk S, Hollingsworth TD, Lessler J, Viboud C, Grenfell BT (2015a) Seven challenges in modeling vaccine preventable diseases. Epidemics 10:11–15
Metcalf CJE, Ferrari M, Graham AL, Grenfell BT (2015b) Understanding Herd Immunity. Trends Immunol 36(12):753–755
Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R et al (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5:e74
Randolph H, Barreiro L (2020) Herd immunity: understanding COVID-19. Immunity 52(5):737–741
Scherer A, McLean A (2002) Mathematical models of vaccination. Br Med Bull 62:187–199
Smith CEG (1970) Prospects of the control of disease. Proc Roy Soc Med 63:1181–1190
United States Food and Drug Administration (2020a) FDA Takes Key Action in Fight Against COVID-19 By Issuing Emergency Use Authorization for First COVID-19 Vaccine. https://www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19. Accessed on 17 June 2021
United States Food and Drug Administration (2020b) FDA briefing document pfizer-BioNTech COVID-19 Vaccine. https://www.fda.gov/media/144245/download. Accessed on 17 June 2021
van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180:29–48
Wallinga J, Teunis P, Kretzschmar M (2006) Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. Am J Epidemiol 164(10):936–944