Published in final edited form as:
J Math Biol. 2018 December ; 77(6-7): 1795–1831. doi:10.1007/s00285-018-1216-z.

Constrained minimization problems for the reproduction number in meta-population models

Gayane Poghotanyan[1], **Zhilan Feng**[1], **John W. Glasser**[2], and **Andrew N. Hill**[3]

[1]Department of Mathematics, Purdue University, West Lafayette, IN, USA

[2]Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases, Atlanta, GA, USA

[3]Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Atlanta, GA, USA

Abstract

The basic reproduction number ($\mathscr{R}_0$) can be considerably higher in an SIR model with heterogeneous mixing compared to that from a corresponding model with homogeneous mixing. For example, in the case of measles, mumps and rubella in San Diego, CA, Glasser et al. (Lancet Infect Dis 16(5):599–605, 2016. https://doi.org/10.1016/S1473-3099(16)00004-9), reported an increase of 70% in $\mathscr{R}_0$ when heterogeneity was accounted for. Meta-population models with simple heterogeneous mixing functions, e.g., proportionate mixing, have been employed to identify optimal vaccination strategies using an approach based on the gradient of the effective reproduction number ($\mathscr{R}_v$), which consists of partial derivatives of $\mathscr{R}_v$ with respect to the proportions immune p_i in sub-groups i (Feng et al. in J Theor Biol 386:177–187, 2015. https://doi.org/10.1016/j.jtbi.2015.09.006; Math Biosci 287:93–104, 2017. https://doi.org/10.1016/j.mbs.2016.09.013). These papers consider cases in which an optimal vaccination strategy exists. However, in general, the optimal solution identified using the gradient may not be feasible for some parameter values (i.e., vaccination coverages outside the unit interval). In this paper, we derive the analytic conditions under which the optimal solution is feasible. Explicit expressions for the optimal solutions in the case of $n = 2$ sub-populations are obtained, and the bounds for optimal solutions are derived for $n > 2$ sub-populations. This is done for general mixing functions and examples of proportionate and preferential mixing are presented. Of special significance is the result that for general mixing schemes, both $\mathscr{R}_0$ and $\mathscr{R}_v$ are bounded below and above by their corresponding expressions when mixing is proportionate and isolated, respectively.

Keywords

Meta-population model; Convexity of reproduction number; Optimization problem; Vaccination strategy; Epidemiology; 37N25; 49J15; 34H05; 92D30

Zhilan Feng, fengz@purdue.edu.
Gayane Poghotanyan, gpoghota@purdue.edu; John W. Glasser, jglasser@cdc.gov; Andrew N. Hill, ahill2@cdc.gov

1 Introduction

Mechanistic models of pathogen transmission are key public health tools for identifying optimal interventions that can mitigate outbreaks or perhaps even eliminate infectious diseases. However, the utility and credibility of such models hinge on incorporating realistic mixing between sub-populations (i.e., means by which infectious members of one sub-population infect susceptible members of others), which typically is not uniformly random due to preference among age groups, genders, or spatial locations. In fact, models that do not sufficiently account for differences among relevant sub-populations can generate biased or misleading results in situations where evaluations of intervention strategy require incorporation of such heterogeneity and realistic mixing. For example, in the case of measles, mumps and rubella in San Diego, CA, Glasser et al. (2016), reported an increase of 70% in $\mathscr{R}_0$ when heterogeneity was accounted for.

Recently, progress has been made (Glasser et al. 2012; Feng et al. 2017) in extending realistic mixing functions based on earlier work (Nold 1980; Jacquez et al. 1988). The effective reproduction numbers $\mathscr{R}_v$ derived from these meta-population models with non-homogeneous mixing functions are used to identify optimal vaccination strategies by using methods based on the gradients of $\mathscr{R}_v$ (partial derivatives with respect to control parameters) (Feng et al. 2015, 2017). These are constrained optimization problems with the objectives of either minimizing $\mathscr{R}_v$ given limited number of vaccine doses, or minimizing vaccine doses needed to reduce $\mathscr{R}_v$ to a given level. However, the examples considered in these studies focus only on cases where an optimal solution exists and is feasible in the sense that the vaccine coverages lie between 0 and 1. Conditions have not yet been identified to determine parameter regions within which optimal mathematical solutions are indeed feasible. This is the objective of the current paper. Similar optimization problems using $\mathscr{R}_v$ in the context of age-dependent vaccination strategies have been considered in Castillo-Chavez and Feng (1998) and Hadeler and Müller (1996a, b).

When heterogeneous mixing is considered in epidemiological models, preferred mixing is among the commonly used mixing structures (Nold 1980; Jacquez et al. 1988; Glasser et al. 2012), of which proportionate mixing is a special case. In this paper, results and proofs are presented for both preferred and more general mixing. Conditions are determined under which optimal solutions for vaccination strategies exist, some of which are given in explicit expressions depending on model parameters. The proofs for the existence and uniqueness of the optimal vaccination strategy when there are $n = 2$ sub-populations involve some fundamental properties that we establish for the reproduction number $\mathscr{R}_v$ as a function of vaccination coverage $p = (p_1, p_2)$; namely, homogeneity and convexity. For the case of $n > 2$ sub-populations, the proofs of results are based on a convexity result of Friedland (1980/81) for the spectral radius over a class of positive matrices.

For $n = 2$, explicit analytical expressions are derived for the optimal allocation of vaccine $P^* = (p_1^*, p_2^*)$ as well as the minimized reproduction number, $\mathscr{R}_v$. A formula for the critical vaccine doses η_* to achieve $\mathscr{R}_v \leq 1$ is also derived for $n = 2$. For the case of $n > 2$, lower and upper bounds for the minimum of $\mathscr{R}_v$ as well as the minimum vaccine doses are derived.

The organization of this paper is as follows. The main problem is described in Sect. 2, which consists of two constrained optimization problems of Lagrange type. Sect. 3 presents the main results of the optimization problem for $n = 2$ sub-populations. Results for $n > 2$ sub-populations are provided in Sect. 4. In Sect. 5, we discuss the results. Some detailed proofs are included in the "Appendix".

2 Description of the problem

The models considered by Glasser et al. (2016) and Feng et al. (2015, 2017) are of the SIR or SEIR type, i.e., the population is apportioned into disjoint states including susceptible (S), exposed (E), infectious (I), and removed or immune (R), and the models consist of systems of ordinary differential equations (ODEs). These models include one or more types of population heterogeneity (e.g., age, spatial location, activity level, vaccination coverage, preferential mixing, population density, etc.), so they are meta-population models with each sub-population model being an SIR or SEIR type linked by a mixing function. We use the simplest of these models as an example, but similar results apply to other models. The model considered in this paper is described by the ODE system

$$\begin{aligned} \frac{dS_i}{dt} &= (1 - p_i)\theta N_i - (\lambda_i + \theta)S_i \\ \frac{dI_i}{dt} &= \lambda_i S_i - (\gamma + \theta)I_i \\ \frac{dR_i}{dt} &= p_i \theta N_i + \gamma I_i - \theta R_i \\ N_i &= S_i + I_i + R_i \\ \lambda_i &= \beta a_i \sum_{j=1}^{n} c_{ij} I_j \Big/ N_j, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{2.1}$$

where p_i are proportions immunized at entry into sub-population i, γ is the *per capita* recovery rate, θ is the *per capita* rate for entering and leaving sub-population i so that the population size N_i remains constant. The function λ_i is the force of infection, i.e., *per capita* hazard rate of infection of susceptible individuals in sub-population i, in which β is the probability of infection upon contacting an infectious person, a_i is average contact rate (activity) in sub-population i, c_{ij} is the proportion of ith sub-population's contacts that are with members of jth sub-population, and I_j/N_j is the probability that a randomly encountered member of sub-population j is infectious.

One of the most influential factors affecting the reproduction number is the mixing function c_{ij}. Denote the mixing matrix by $C = (c_{ij})$. Typically, the matrix C has to satisfy the following conditions of Busenberg and Castillo-Chavez (1991):

$$c_{ij} \geq 0, \quad i, j = 1, \ldots, n, \quad (2.2)$$

$$\sum_{j=1}^{n} c_{ij} = 1, \quad i = 1, \ldots, n, \quad (2.3)$$

$$a_i N_i c_{ij} = a_i N_j c_{ji}, \quad i, j = 1, \ldots, n. \quad (2.4)$$

A commonly used non-homogeneous mixing function that satisfies conditions (2.2)–(2.4) is the preferred mixing function of Jacquez et al. (1988) given by:

$$c_{ij} = \epsilon_i \delta_{ij} + (1 - \epsilon_i) \frac{(1 - \epsilon_j) a_j N_j}{\Sigma_{k=1}^{n} (1 - \epsilon_k) a_k N_k}, \quad i, j = 1, \ldots, n, \quad (2.5)$$

where δ_{ij} is the Kronecker delta function (taking value 1 when $i = j$, 0 otherwise) and $\epsilon_i \in [0, 1]$ is the fraction of contacts of group i that is reserved for itself (preferential mixing), whereas the complement $(1 - \epsilon_i)$ is distributed among all sub-populations in proportion to the unreserved contacts, including i (proportionate mixing). Special cases arise when: $\epsilon_i = 1$ for all i whence C is the identity matrix (exclusively preferential mixing); $\epsilon_i = 0$ for all i whence $c_{ij} = a_j N_j \big/ \Sigma_k a_k N_k$ (exclusively proportionate mixing). We will refer to this mixing structure throughout the manuscript as Jacquez mixing. More complex examples of mixing matrices $C = (c_{ij})$, such as two-level preferential mixing can be found in Feng et al. (2017).

When Model (2.1) is used, the basic and effective sub-population reproduction numbers, denoted respectively by $\mathscr{R}_{0i}$ and $\mathscr{R}_{vi}$, for sub-population i ($i = 1, 2, \ldots, n$) are given by

$$\mathscr{R}_{0i} = \rho a_i, \quad \mathscr{R}_{vi} = \mathscr{R}_{0i}(1 - p_i), \quad i = 1, 2, \ldots, n, \quad (2.6)$$

where

$$\rho = \frac{\beta}{\gamma + \theta},$$

see e.g. Brauer and Castillo-Chavez (2012, Chapter 10). Following Diekmann et al. (1990) and van den Driessche and Watmough (2002), the next generation matrix (NGM) corresponding to this meta-population model is

$$K_v = \begin{pmatrix} \mathscr{R}_{v1}c_{11} & \mathscr{R}_{v1}c_{12} & \cdots & \mathscr{R}_{v1}c_{1n} \\ \mathscr{R}_{v2}c_{21} & \mathscr{R}_{v2}c_{22} & \cdots & \mathscr{R}_{v2}c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \mathscr{R}_{vn}c_{n1} & \mathscr{R}_{vn}c_{n2} & \cdots & \mathscr{R}_{vn}c_{nn} \end{pmatrix}. \quad (2.7)$$

Then the effective reproduction number for the meta-population is given as

$$\mathscr{R}_v = r(K_v),$$

which is the spectral radius [and the dominant eigenvalue, by the Perron–Frobenius Theorem (Seneta 1973)] of the nonnegative matrix K_v. Let $p = (p_1, p_2, \ldots, p_n)$. Naturally, $\mathscr{R}_v = R_v(p)$ is a function of p. The total number of vaccine doses, denoted by η, is $\eta = \sum_{i=1}^{n} p_i N_i$. For demonstration purposes, we will assume that vaccine efficacy is 100%. In this paper, we focus on identifying the most efficient allocation of vaccine $p = (p_1, p_2, \ldots, p_n) \in [0, 1]^n$ for reducing $\mathscr{R}_v$ with limited vaccine doses η or using fewest doses to achieve $\mathscr{R}_v < 1$ (to prevent outbreaks). More specifically, we consider the following two constrained optimization problems:

(I) Minimize $\mathscr{R}_v = \mathscr{R}_v(p)$, subject to $\ell(p) := \sum_{i=1}^{n} p_i N_i = \eta$, for $p \in [0, 1]^n$.

(II) Minimize $\eta = \sum_{i=1}^{n} p_i N_i$ subject to $\mathscr{R}_v(p) \le 1$, for $p \in [0, 1]^n$.

In this study, we consider the optimization problems only for the case of $\mathscr{R}_0 = \mathscr{R}_v(0) \ge 1$, as there will be no outbreak if $\mathscr{R}_0 < 1$.

Because of the continuity of $\mathscr{R}_v(p)$ and the compactness of $[0, 1]^n$, Problem (**I**) has a solution for any fixed $\eta \in [0, N]$, where $N = N_1 + N_2 + \ldots + N_n$ is the total population. If $P^* = P^*(\eta)$ and $\mathscr{R}_{v\{min\}}(\eta)$ denote the optimal vaccination allocation and the corresponding minimum reproduction number, respectively, then we have

$$\begin{aligned} P^*(\eta) &= (p_1^*(\eta), p_2^*(\eta), \ldots, p_n^*(n)) \in \Omega_p^{(n)}(\eta) \cap [0, 1]^n, \\ \mathscr{R}_{v\{min\}}(\eta) &= \min_{\Omega_p^{(n)}(\eta) \cap [0, 1]^n} \mathscr{R}_v = \mathscr{R}_v \big|_{P^*(\eta)}, \end{aligned} \quad (2.8)$$

where

$$\Omega_p^{(n)}(\eta) := \left\{ (p_1, p_2, \ldots, p_n) : \ell(p) = \eta \right\}.$$

An optimal solution $P^*(\eta)$ to Problem (**I**) that lies in the interior $(0, 1)^n$ of the unit hypercube must also satisfy the following equations:

$$\begin{aligned} \nabla \mathscr{R}_v \big|_{P^*(\eta)} &= \lambda \nabla \ell = \lambda(N_1, \ldots, N_n), \\ \ell \big|_{P^*(\eta)} &= \sum_{i=1}^{n} p_i^*(\eta) N_i = \eta, \end{aligned} \qquad (2.9)$$

where the constant λ is the Lagrange multiplier.

Similarly, noting that the constraint set $\left\{\mathscr{R}_v(p) \le 1\right\} \cap \left[0, 1\right]^n$ is compact and nonempty (as $\mathscr{R}_v(1, 1, \ldots, 1) = 0 \le 1$), Problem (**II**) always has a solution. The minimum value of η, which we denote by η_*, signifies the smallest number of vaccine doses that can prevent outbreaks under an optimal vaccination policy. It is useful practically to have an explicit expression or estimate of the bounds for η_*.

To find η_*, notice that $\mathscr{R}_v(p)$ is a monotonically decreasing function of p_i, and thus, a decreasing function of $\eta = \sum_{i=1}^{n} p_i N_i$. Therefore, recalling also the assumption $\mathscr{R}_v(0) = \mathscr{R}_0 \ge 1$, the inequality constraint $\mathscr{R}_v(p) \le 1$ can be replaced by an equality constraint $\mathscr{R}_v(p) = 1$, and thus,

$$\eta_* = \min_{\left\{\mathscr{R}_v(p) = 1\right\} \cap [0, 1]^n} \ell(p).$$

It follows that η_* is the minimum of $\eta \in [0, N]$ such that $\mathscr{R}_{v\{min\}}(\eta) = 1$ and can be found by solving the equation:

$$\mathscr{R}_{v\{min\}}(\eta_*) = \mathscr{R}_v \big|_{P^*(\eta_*)} = 1. \qquad (2.10)$$

2.1 Notation

Below we introduce some mathematical concepts that will be used throughout the paper. For ease of reference, we also list the quantities that appear, along with their definitions, in Table 1.

- A set E in a vector space is called *convex* if for any $x_0, x_1 \in E$ the convex combination $(1 - t)x_0 + tx_1 \in E$ for any $t \in [0, 1]$.
- We say that a real-valued function f on a convex set E is *convex* if for any $x_0, x_1 \in E$

$$f((1 - t)x_0 + tx_1) \le (1 - t)f(x_0) + tf(x_1), \quad t \in \left[0, 1\right].$$

- We say f is *strictly convex* if the above inequality is strict for all $t \in (0, 1)$.
- Recall that if f is twice continuously differentiable, then f is convex on an open convex set if and only if the Hessian Hess $f = (\partial^2 f/\partial x_i \partial x_j)$ is a nonnegative semi-definite matrix for any $x \in E$. Moreover, f will be strictly convex if Hess f is positive definite. (The converse of this statement is false.)
- All matrices considered in this paper will be over the field of real numbers.
- If $C = (c_{ij})$ is a matrix, then we write $C > 0$ ($C \geq 0$) if $c_{ij} > 0$ ($c_{ij} \geq 0$) for any pair i, j. We also say that such C is *positive* (*nonnegative*). This should not be confused with the notion of positive definite (nonnegative semi-definite) matrices, which are related to the positivity (nonnegativity) of the quadratic form $x \mapsto x^T Cx$ associated with C.
- We say that a square matrix $C = (c_{ij})$ is *essentially nonnegative*[1] if its off-diagonal elements are nonnegative; i.e., $c_{ij} \geq 0$ if $i \neq j$.
- We say that a nonnegative squarematrix C is *irreducible* if for any pair (i, j), $i, j = 1, \ldots, n$, there exists a natural number $m = m(i, j)$ such that the entry in the ith row and jth column of C^m is positive.

3 Results for *n* = 2 sub-populations

In the case of $n = 2$, the NGM (2.7) is a 2 × 2 matrix. From this matrix we obtain the following explicit expression for the reproduction number $\mathcal{R}_v$ as a function of $p = (p_1, p_2)$:

$$\mathcal{R}_v(p) = \frac{1}{2}\left[\mathcal{R}_{v1}c_{11} + \mathcal{R}_{v2}c_{22} + \sqrt{(\mathcal{R}_{v1}c_{11} - \mathcal{R}_{v2}c_{22})^2 + 4\mathcal{R}_{v1}c_{12}\mathcal{R}_{v2}c_{21}}\right].$$

One condition needed for proving the existence of an optimal solution to Problem (**I**) is that the mixing matrix satisfies

$$C > 0, \ |C| = c_{11}c_{22} - c_{12}c_{21} > 0. \qquad (3.1)$$

It is easy to verify that condition (3.1) holds for Jacquez mixing as given in (2.5):

$$|C| = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \epsilon_1\epsilon_2 + \frac{\epsilon_1(1-\epsilon_2)^2 a_2 N_2 + \epsilon_2(1-\epsilon_1)^2 a_1 N_1}{(1-\epsilon_1)a_1 N_1 + (1-\epsilon_2)a_2 N_2} > 0,$$

provided that $\epsilon_i \in (0, 1)$, $a_i > 0$, and $N_i > 0$, $i = 1, 2$.

---

[1] Note that terminology differs among authors. Here, we use the terminology of Nussbaum (1986); the negatives of such matrices are called *M*-matrices by Friedland (1980/81) and are said to have a *Z* sign pattern by van den Driessche and Watmough (2002).

3.1 Statements of the main results

Before we state the main results on the existence and uniqueness of the optimal solutions to Problems (**I**) and (**II**), we give the following critical properties of $\mathscr{R}_v(p_1, p_2)$:

Theorem 3.1 (Key properties of $\mathscr{R}_v(p_1, p_2)$) *Consider* $\mathscr{R}_v = \mathscr{R}_v(p_1, p_2)$ *as a function of* p_1 *and* p_2*, and assume that condition* (3.1) *holds.*

- **(i)** $\mathscr{R}_v(1, 1) = 0$ *and* $\mathscr{R}_v$ *grows linearly on the rays emanating from* (1, 1) *into the square* $[0, 1]^2$ (*See* Fig. 1 *for illustration*).
- **(ii)** $\mathscr{R}_v(p_1, p_2)$ *is convex on* $[0, 1]^2$ *and strictly convex on the constraint set* $\Omega_p^{(2)}(\eta) \cap [0, 1]^2$.

The proof of Theorem 3.1 is provided in "Appendix A.1".

The following theorem describes the solution to Problem (**I**). For ease of presentation, we introduce the following notation:

$$\eta_0 := N - \frac{\kappa_1 N_1 + \kappa_2 N_2}{\max\{\kappa_1, \kappa_2\}}, \tag{3.2}$$

$$\kappa_1 := c_{22}\sqrt{N_1 N_2 \mathscr{R}_{02}} - N_2\sqrt{c_{12}c_{21}\mathscr{R}_{01}\mathscr{R}_{02}},$$

$$\kappa_2 := c_{11}\sqrt{N_1 N_2 \mathscr{R}_{01}} - N_1\sqrt{c_{12}c_{21}\mathscr{R}_{01}\mathscr{R}_{02}},$$

$$\Gamma : (p_1, p_2) = (1, 1) - s(\kappa_1, \kappa_2), \quad s > 0 \quad \left[\text{or } \Gamma : \frac{1 - p_2}{1 - p_1} = \frac{\kappa_2}{\kappa_1}\right].$$

The set Γ describes the ray emanating from (1, 1) in the direction of $-(\kappa_1, \kappa_2)$, to which we will refer as the *critical ray* (see Fig. 1).

Theorem 3.2 (Optimal solution to Problem (**I**) when $n = 2$) *Consider* $\mathscr{R}_v = \mathscr{R}_v(p_1, p_2)$ *as a function of* p_1 *and* p_2*, and let* η_0 *and* κ_i *be given in* (3.2). *Assume that condition* (3.1) *holds.*

- **(i)** *For any given* $\eta \in [0, N]$, *the optimal point* $P^*(\eta)$ *exists and is unique.*
- **(ii)** *The point* $P^*(\eta)$ *lies in the interior of the unit square if and only if*

$$\kappa_1 > 0, \kappa_2 > 0 \tag{3.3}$$

  and $\eta \in (\eta_0, N)$.
- **(iii)** *For* $\eta \in (\eta_0, N)$, *all points* $P^*(\eta)$ *lie on the critical ray* Γ, *defined in* (3.2) (*see* Fig. 1).
- **(iv)** *For each* $\eta \in (\eta_0, N)$, *the explicit formulae for* $P^*(\eta)$ *and* $\mathscr{R}_{v\{min\}}(\eta)$ *are*

$$P^*(\eta) = (1, 1) - \frac{N - \eta}{\kappa_1 N_1 + \kappa_2 N_2}(\kappa_1, \kappa_2) \quad (3.4)$$

$$\mathscr{R}_{v\{min\}}(\eta) = \mid C \mid \mathscr{R}_{01}\mathscr{R}_{02}\sqrt{N_1 N_2}\frac{N - \eta}{\kappa_1 N_1 + \kappa_2 N_2}. \quad (3.5)$$

A proof for Theorem 3.2 is given in “Appendix A.1”.

Remarks A few remarks are in order.

(a) If either of the conditions in (3.3) is violated, or equivalently $\kappa_1 \leq 0$ or $\kappa_2 \leq 0$, then the ray Γ does not intersect the interior of the square $(0, 1)^2$.

(b) The minimum point $P^*(\eta)$ is the intersection of the critical ray Γ and the constraint set $\Omega_p^{(2)}(\eta)$ for each $\eta \in (\eta_0, N)$.

(c) If $\eta \in (0, \eta_0)$, the intersection of Γ and $\Omega_p^{(2)}(\eta)$ lies outside the square $[0, 1]^2$. When $\eta = \eta_0$, the intersection lies on the boundary of the square.

(d) If $\eta \in (0, \eta_0)$, the minimum point $P^*(\eta)$ is one of the boundary points $(\eta/N_1, 0)$ or $(0, \eta/N_2)$ and hence

$$\mathscr{R}_{v\{min\}}(\eta) = \min\left\{\mathscr{R}_v(\eta / N_1, 0), \mathscr{R}_v(0, \eta / N_2)\right\}.$$

(e) Overall, $\mathscr{R}_{v\{min\}}(\eta)$ is a strictly decreasing convex function of η and linear on (η_0, N) (see Fig. 2).

An explicit expression for the optimal solution η_* to Problem (**II**) can be obtained by using Eq. (2.10). We consider two cases depending on the value of

$$\mathscr{R}_{v\{min\}}(\eta_0) = \frac{\mid C \mid \mathscr{R}_{01}\mathscr{R}_{02}\sqrt{N_1 N_2}}{\max\left\{\kappa_1, \kappa_2\right\}}.$$

Theorem 3.3 (Critical number of vaccine doses) *If condition* (3.3) *is satisfied, then the minimum value of* η_* *in Problem* (**II**) *is given below.*

(i) *(Interior minimum) If* $\mathscr{R}_{v\{min\}}(\eta_0) \geq 1$, *then* $\eta_* \geq \eta_0$ *and*

$$\eta_* = N - \frac{\kappa_1 N_1 + \kappa_2 N_2}{\mid C \mid \mathscr{R}_{01}\mathscr{R}_{02}\sqrt{N_1 N_2}}. \quad (3.6)$$

(ii) *(Boundary minimum) If* $\mathscr{R}_{v\{min\}}(\eta_0) \leq 1$, *then* $0 \leq \eta_* \leq \eta_0$ *and*

$$\eta_* = \min\left\{N_1 - \frac{(1 - c_{22}\mathscr{R}_{02})N_1}{c_{11}\mathscr{R}_{01} - \mid C \mid \mathscr{R}_{01}\mathscr{R}_{02}}, N_2 - \frac{(1 - c_{11}\mathscr{R}_{01})N_2}{c_{22}\mathscr{R}_{02} - \mid C \mid \mathscr{R}_{01}\mathscr{R}_{02}}\right\}. \quad (3.7)$$

The proof is given in "Appendix A".

4 Results for $n \geq 2$ sub-populations

In this section, we extend the results for $n = 2$ sub-populations presented in Sect. 3 to the case of $n > 2$. For general mixing matrices $C = (c_{ij})$, due to the complexity of the optimization problems (**I**) and (**II**) when $n > 2$, we are unable to obtain explicit expressions for the optimal solutions. Nevertheless, we can derive lower and upper bounds for the minimum reproduction number $\mathscr{R}_{v\{min\}}(\eta)$ and the minimum vaccine doses η_*. We first present results for general mixing (c_{ij}) satisfying (2.2)–(2.4). We then illustrate that some of the key necessary properties of the mixing matrix can be verified for the Jacquez mixing given in (2.5).

4.1 Preliminaries

Rewrite the NGM given in (2.7) as

$$\begin{aligned} K_v(p) &= \begin{pmatrix} \mathscr{R}_{v1} & & & \\ & \mathscr{R}_{v2} & & \\ & & \ddots & \\ & & & \mathscr{R}_{vn} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \\ &= \operatorname{diag}(\mathscr{R}_{01}(1 - p_1), \ldots, \mathscr{R}_{0n}(1 - p_n))C \end{aligned}$$

The effective reproduction number $\mathscr{R}_v$ for the meta-population is the spectral radius (and the dominant eigenvalue, by Perron–Frobenius Theorem) of the nonnegative matrix $K_v(p)$, i.e., $\mathscr{R}_v(p) = r(K_v(p))$.

Although the focus of this study is on optimal solutions to Problems (**I**) and (**II**), the results presented in the following theorem about the bounds of $\mathscr{R}_v(p)$ and $\mathscr{R}_0 = \mathscr{R}_v(0)$ are significant in more general applications.

Theorem 4.1 (Bounds for $\mathscr{R}_v(p)$) *Let C be a nonnegative, invertible, irreducible matrix such that* $-C^{-1}$ *is essentially nonnegative and the conditions* (2.2)–(2.4) *are satisfied. Then*

(a) *The lower and upper bounds of* $\mathscr{R}_v(p)$ *are:*

$$\sum_{i=1}^{n} \omega_i \mathscr{R}_{vi} \leq \mathscr{R}_v \leq \max\{\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn}\}, \quad \text{where } \omega_i = \frac{a_i N_i}{\Sigma_{k=1}^{n} a_k N_k}. \quad (4.1)$$

(b) *The lower and upper bounds of* $\mathscr{R}_v(p)$ *correspond to the cases of proportionate mixing and isolated mixing, respectively.*

The proof of Theorem 4.1 is given after the proof of Theorem 4.9.

Remarks Theorem 4.1 is stated using the effective reproduction number $\mathscr{R}_v(p)$ for $0 \le p \le 1$. The results holds in particular for the basic reproduction number $\mathscr{R}_0 = \mathscr{R}_v(0)$ that the the lower and upper bounds are $\sum_{i=1}^{n} \omega_i \mathscr{R}_{0i}$ and $\{\mathscr{R}_{01}, \ldots, \mathscr{R}_{0n}\}$.

For the ease of presentation, we introduce the ‘reflected’ variables

$$q_i = 1 - p_i, \quad i = 1, \ldots, n, \quad q = (q_1, \ldots, q_n) \in (0, 1)^n .$$

Note that q_i represents the unvaccinated portion of sub-population $i = 1, \ldots , n$. We also introduce the ‘reflected’ function

$$\begin{aligned} \overline{\mathscr{R}}_v(q) &= \mathscr{R}_v(1 - q_1, \ldots, 1 - q_n) \\ &= r(\operatorname{diag}(\mathscr{R}_{01} q_1, \ldots, \mathscr{R}_{0n} q_n) C) . \end{aligned} \tag{4.2}$$

Note that the formula (4.2) can be used to extend $\overline{\mathscr{R}}_v$ to $[0,\infty)^n$. The constraint hyperplanes $\Omega_p^{(n)}(\eta)$ in Problem (**I**) will transform to

$$\Omega_q^{(n)}(\bar{\eta}) := \{(q_1, q_2, \ldots, q_n) : \ell(q) = \bar{\eta}\},$$

where

$$\bar{\eta} := N - \eta$$

and the optimal point $P^*(\eta)$ will become

$$Q^*(\bar{\eta}) := (1, \ldots, 1) - P^*(\eta) = (1, \ldots, 1) - P^*(N - \bar{\eta}) .$$

For the minimum value of $\overline{\mathscr{R}}_v(q)$ on $\Omega_q^{(n)}(\bar{\eta}) \cap [0, 1]^n$ we will have

$$\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) = \overline{\mathscr{R}}_v \big|_{Q^*(\bar{\eta})} = \mathscr{R}_v \big|_{P^*(\eta)} = \mathscr{R}_{v\{min\}}(\eta) .$$

The following result states key properties of $\overline{\mathscr{R}}_v(q)$, generalizing Theorem 3.1 in the case $n = 2$.

Theorem 4.2 (Convexity and homogeneity of $\overline{\mathscr{R}}_v(q)$) *The function* $\overline{\mathscr{R}}_v(q)$ *is homogeneous of degree 1 on* $q \in [0,1]^n$, *i.e.,*

$$\overline{\mathscr{R}}_v(sq) = s\overline{\mathscr{R}}_v(q), \quad \textit{for all } q \in [0,1]^n \textit{and } s > 0 \textit{ such that } sq \in [0,1]^n .$$

Moreover, $\overline{\mathscr{R}}_v$ *is convex if the matrix* $C = (c_{ij})$ *is invertible and* $-C^{-1}$ *is essentially nonnegative. If additionally C is irreducible, then* $\overline{\mathscr{R}}_v(q)$ *is strictly convex on the constraint set* $\Omega_q^{(n)}(\bar{\eta}) \cap [0,1]^n$ (*see* Fig. 3).

Theorem 4.2 immediately implies the following property.

Theorem 4.3 (Critical ray) *Let C be a nonnegative square invertible irreducible matrix such that* $-\mathrm{C}^{-1}$ *is essentially nonnegative. Let* $\bar{\eta} \in (0, N)$ *be such that* $\overline{\mathscr{R}}_v$ *has an interior relative minimum point* $Q^*(\bar{\eta})$ *on* $\Omega_q^{(n)}(\bar{\eta}) \cap [0,1]^n$. *Then* $Q^*(\bar{\eta})$ *is the unique point that satisfies the Lagrange multiplier condition*

$$\overline{\mathscr{R}}_v \,\Big|_{Q^*(\bar{\eta})} = \lambda(N_1, N_2, \ldots, N_n) .$$

If s > 0 is such that $s\bar{\eta} \in (0, N)$, *then the unique relative minimum point on* $\Omega_q^{(n)}(s\bar{\eta}) \cap [0,1]^n$ *is given by*

$$Q^*(s\bar{\eta}) = sQ^*(\bar{\eta}),$$

provided that this point still lies in the interior of the unit hypercube.[2] *Thus, all interior relative minimum points lie on the critical ray* $\bar{\Gamma}$ *emanating from the origin.*

We start with the following facts about the spectral radius $r(A)$ of a nonnegative matrix A. The first one is rather simple, if not obvious.

Lemma 4.4 *If A is a square matrix and s > 0 then*

$$r(sA) = sr(A) .$$

The second one is also well-known, see e.g. Hill and Longini (2003); Nussbaum (1986).

Lemma 4.5 *If A, B are nonnegative irreducible matrices such that A ≤ B, then*

$$r(A) \le r(B) .$$

---

[2] This will automatically hold for $0 < s < 1$.

The next one is more subtle and is based on a theorem of Friedland (1980/81, Theorem 4.3); see also generalizations of this result in Nussbaum (1986, Sect. 1).

Lemma 4.6 *Let C be a nonnegative invertible squarematrix such that* $-\mathrm{C}^{-1}$ *is essentially nonnegative. Then the mapping*

$$r_C : D \mapsto r(DC)$$

is convex on the set of positive diagonal matrices $D = \operatorname{diag}(d_1, \ldots, d_n)$, $d_i > 0$, *i.e.*,

$$r_C((1-t)D_1 + tD_2) \le (1-t)r_C(D_1) + tr_C(D_2)$$

for any positive matrices D_1 *and* D_2 *and* $t \in [0, 1]$. *Moreover, if additionally C is irreducible,*[3] *then the inequality above is strict for* $t \in (0, 1)$, *unless* $D_2 = sD_1$ *for some* $s > 0$.

Proof of Theorem 4.2 If we denote

$$\bar{K}_v(q) = K_v((1, \ldots, 1) - q) = \operatorname{diag}(\mathscr{R}_{01}q_1, \ldots, \mathscr{R}_{0n}q_n)C,$$

then for homogeneity we just note that

$$\bar{K}_v(sq) = s\bar{K}_v(q)$$

and therefore by Lemma 4.4, we have

$$\overline{\mathscr{R}}_v(sq) = r(\bar{K}_v(sq)) = r(s\bar{K}_v(q)) = sr(\bar{K}_v(q)) = s\overline{\mathscr{R}}_v(q).$$

The convexity of $\overline{\mathscr{R}}_v$ follows from the fact that

$$\overline{\mathscr{R}}_v(q) = r_C(\operatorname{diag}(\mathscr{R}_{01}q_1, \ldots, \mathscr{R}_{0n}q_n)),$$

i.e., $\overline{\mathscr{R}}_v$ is a composition of a linear mapping

$$q \mapsto \operatorname{diag}(\mathscr{R}_{01}q_1, \ldots, \mathscr{R}_{0n}q_n)$$

and a convex function r_C, and is therefore convex, as we impose the condition that $-C^{-1}$ exists and is essentially nonnegative.

---

[3] See the definition in Sect. 2.1.

Finally, the strict convexity of $\overline{\mathscr{R}}_v$ on $\Omega_q^{(n)}(\bar{\eta}) \cap [0, 1]^n$ follows from the strict convexity property of r_C in Lemma 4.6, because no two points on $\Omega_q^{(n)}(\bar{\eta})$ lie on the same ray emanating from the origin.

4.2 Upper and lower bounds of optimal solutions

In this section, we establish bounds on quantities relevant to Problems (**I**) and (**II**) for general mixing matrices C, satisfying (2.2)–(2.4) with an additional property that $-C^{-1}$ is essentially nonnegative. As we saw in Theorem 4.2, the latter condition is needed to guarantee the convexity of $\overline{\mathscr{R}}_v$. On the other hand, conditions (2.2)–(2.4) provide important information on positive eigenvectors of the mixing matrix C, which is instrumental in deriving our bounds.

More specifically, we prove upper and lower bounds for $Q^*(\bar{\eta})$, $\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta})$, and $\bar{\eta}_*$ (or equivalently, $P^*(\eta)$, $\mathscr{R}_{v\{min\}}(\eta)$, and η_*). An interesting feature is that these bounds are independent of the functional form of mixing c_{ij}.

4.2.1 *Equal* per capita *contact rates*—We start with a special case when all *per capita* contact rates a_i are the same. We show that the minimum of $\overline{\mathscr{R}}_v(q)$ on $\Omega_q^{(n)}(\bar{\eta}) \cap [0, \infty)^n$ will occur on the diagonal $q_1 = q_2 = \ldots = q_n$, under the conditions on C that guarantee the convexity of $\overline{\mathscr{R}}_v(q)$. In particular, this will hold for simple mixing matrices given by (2.5) for any choice of $\epsilon_i \in (0, 1)$ and $N_i > 0$.

Theorem 4.7 (Equal per capita contact rates) *Let C be a nonnegative, invertible, irreducible matrix such that* $-\mathrm{C}^{-1}$ *is essentially nonnegative and for which the conditions* (2.2)–(2.4) *are satisfied. Assume additionally that* $a_i = a > 0$, $i = 1, \ldots, n$. *Then the minimum of* $\overline{\mathscr{R}}_v$ *on the intersection of* $\Omega_q^{(n)}(\bar{\eta}) \cap [0, 1]^n$ *for* $\bar{\eta} \in [0, N]$ *is achieved at the point*

$$q_1 = q_2 = \cdots = q_n = \frac{\bar{\eta}}{N}.$$

Thus, the critical ray $\bar{\Gamma}$ *is given by*

$$\bar{\Gamma}: q_1 = q_2 = \cdots = q_n.$$

We will need the following characterization of Friedland (1980/81, Theorem 3.4) for the spectral radius.

Lemma 4.8 *If C is a nonnegative, invertible matrix, such that* $-\mathrm{C}^{-1}$ *is essentially nonnegative, then its spectral radius r (C) is given by*

$$\inf_{\xi \in P_n} \sup_{x > 0} \sum_{i=1}^{n} \xi_i \frac{x_i}{(Cx)_i} = \frac{1}{r(C)},$$

where

$$P_n = \left\{ \xi = (\xi_1, \ldots, \xi_n) : \xi_i \geq 0, \sum_{i=1}^{n} \xi_i = 1 \right\}.$$

Moreover, if C is irreducible, and $u = (u_1, \ldots, u_n)^T > 0$, $v = (v_1, \ldots v_n)^T > 0$ *are right and left eigenvectors of* C, *i.e.,*

$$Cu = r(C)u, \quad C^T v = r(C)v, \quad \sum_{i=1}^{n} u_i v_i = 1,$$

then for $\xi = (u_1 v_1, -, u_n v_n)$ *one has*

$$\sup_{x > 0} \sum_{i=1}^{n} \xi_i \frac{x_i}{(Cx)_i} = \frac{1}{r(C)}.$$

Proof of Theorem 4.7 When $a_i = a$, $i = 1, \ldots, n$, we have the following essential properties of the matrix C

$$\sum_{j=1}^{n} c_{ij} = 1, \quad \sum_{i=1}^{n} N_i c_{ij} = N_j.$$

This means that $u = (1, \ldots, 1)^T$ and $v = (N_1/N, \ldots, N_n/N)^T$ are normalized right and left eigenvectors of C:

$$Cu = u, \quad C^T v = v, \quad \Sigma u_i v_i = 1.$$

By Perron–Frobenius Theorem, we also have $r(C) = 1$. (Note that this holds for any nonnegative matrix satisfying (2.3)). Then by Lemma 4.8 we have

$$1 = \frac{1}{r(C)} = \sup_{x > 0} \sum_{i=1}^{n} \frac{N_i}{N} \frac{x_i}{(Cx)_i}.$$

Suppose now that $q = (q_1, \ldots, q_n) \in (0, 1)^n$ is such that $\sum_{i=1}^{n} q_i N_i = \bar{\eta} \in (0, N)$. Then $\beta = (\beta_1, \ldots, \beta_n) \in P_n$, where

$$\beta_i = q_i N_i \big/ \bar{\eta}, \quad i = 1, \ldots, n.$$

Note, that we also have

$$(\bar{K}_v(q)x)_i = \mathcal{R}_0 q_i (Cx)_i,$$

where $\mathcal{R}_0 = \rho a$ is the common value of $\mathcal{R}_{0i}$, $i = 1, \ldots, n$, which also is the meta-population basic reproduction number. Then, by Lemma 4.8,

$$\begin{aligned} \frac{1}{r(\bar{K}_v)(q)} &\le \sup_{x > 0} \sum_{i = 1} \beta_i \frac{x_i}{(\bar{K}_v(q)x)_i} \\ &= \sup_{x > 0} \sum_{i = 1}^{n} \frac{q_i N_i}{\bar{\eta}} \frac{x_i}{\mathcal{R}_0 q_i (Cx)_i} \\ &= \frac{N}{\mathcal{R}_0 \bar{\eta}} \sup_{x > 0} \sum_{i = 1}^{n} \frac{N_i}{N} \frac{x_i}{(Cx)_i} = \frac{N}{\mathcal{R}_0 \bar{\eta}}. \end{aligned}$$

Note that the use of Lemma 4.8 above is justified because, as with C, $\bar{K}_v(q)$ is positive, invertible, irreducible, and

$$-\bar{K}_v(q)^{-1} = -C^{-1}\mathrm{diag}((\mathcal{R}_0 q_1)^{-1}, \ldots, (\mathcal{R}_0 q_n)^{-1})$$

is essentially nonnegative. Hence,

$$\overline{\mathcal{R}}_v(q) \ge \mathcal{R}_0 \frac{\bar{\eta}}{N}, \quad \text{on } \{\ell(q) = \bar{\eta}\} \cap (0, 1)^n.$$

On the other hand,

$$\bar{K}_v\left(\frac{\bar{\eta}}{N}, \ldots, \frac{\bar{\eta}}{N}\right) = \mathcal{R}_0 \frac{\bar{\eta}}{N} C$$

and therefore

$$\bar{R}_v\left(\frac{\bar{\eta}}{N}, \ldots, \frac{\bar{\eta}}{N}\right) = \mathcal{R}_0 \frac{\bar{\eta}}{N}.$$

By continuity of $\overline{\mathcal{R}}_v$, this completes the proof.

4.2.2 *Arbitrary* per capita *contact rates*—In this section, we first establish the upper and lower bounds for $\overline{\mathcal{R}}_v(q)$ (see Theorem 4.9) and equivalently $\mathcal{R}_v(p)$. The lower bound can be proved using the arguments similar to the proof of Theorem 4.7, and the upper bound follows from the monotonicity of the spectral radius as a function of nonnegative matrices. We then proceed to obtain bounds for the relative minima $\overline{\mathcal{R}}_{v\{min\}}(\bar{\eta})$ and equivalently $\mathcal{R}_{v\{min\}}(\eta)$, as well as a bound for the critical value $\bar{\eta} = \bar{\eta}_*$ that makes $\overline{\mathcal{R}}_{v\{min\}}(\bar{\eta}) \le 1$.

Let $a_{\min}$ denote the minimum of the activities of the sub-populations, i.e.,

$$a_{\min} = \min\left\{a_1, \ldots, a_n\right\}.$$

Note that $a_{\min} > 0$. The results beloware for general mixing matrices (not just Jacquez mixing given in (2.5)).

Theorem 4.9 (Bounds for $\overline{\mathscr{R}}_v(q)$) *Let C be a nonnegative, invertible, irreducible matrix such that* $-\mathrm{C}^{-1}$ *is essentially nonnegative and the conditions* (2.2)–(2.4) *are satisfied. Then the bounds of* $\overline{\mathscr{R}}_v(q)$ *are:*

$$\rho \frac{\Sigma_{i=1}^{n} a_i^2 N_i q_i}{\Sigma_{i=1}^{n} a_i N_i} \leq \overline{\mathscr{R}}_v(q) \leq \rho \max\{a_1 q_1, \ldots, a_n q_n\}, \quad for\ q \in [0,1]^n. \quad (4.3)$$

Moreover, equalities hold if $q = s\left(\frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)$ *for* $s \in [0, a_{\min}]$.

Proof We start with the lower bound in (4.3). The key observation is that

$$\sum_{j=1}^{n} c_{ij} = 1, \quad \sum_{i=1}^{n} a_i N_i c_{ij} = a_j N_j,$$

which gives positive eigenvectors for C and C^T. That is, if $u = (1, \ldots, 1)^T$ and $v = (a_1 N_1/A, \ldots, a_n N_n/A)^T$ with $A = \sum_{i=1}^{n} a_i N_i$, then

$$Cu = u, \quad C^T v = v, \quad u \cdot v = \sum_{i=1}^{n} u_i v_i = 1.$$

Then, by Lemma 4.8,

$$1 = \frac{1}{r(C)} = \sup_{x>0} \sum_{i=1}^{n} \frac{a_i N_i}{A} \frac{x_i}{(Cx)_i} = \frac{\sup_{x>0} \Sigma_{i=1}^{n} a_i N_i \frac{x_i}{(Cx)_i}}{\Sigma_{k=1}^{n} a_k N_k}.$$

Then, for any $q \in (0, 1)^n$, let

$$\beta_i = \frac{a_i^2 N_i q_i}{\Sigma_{k=1}^{n} a_k^2 N_k q_k}$$

and note that $\beta_i > 0$ and $\sum_{i=1}^{n} \beta_i = 1$; i.e., $\beta = (\beta_1, \ldots, \beta_n) \in P_n$. Hence, by Lemma 4.8,

$$\begin{aligned}\frac{1}{\overline{\mathscr{R}}_v(q)} &= \frac{1}{r(\overline{K}_v(q))} \le \sup_{x>0}\sum_{i=1}^{n} \beta_i \frac{x_i}{(\overline{K}_v(q)x)_i}\\ &= \frac{\sup_{x>0}\sum_{i=1}^{n} a_i^2 N_i q_i \frac{x_i}{\rho a_i q_i (Cx)_i}}{\sum_{k=1}^{n} a_k^2 N_k q_k}\\ &= \frac{\sup_{x>0}\sum_{i=1}^{n} a_i N_i \frac{x_i}{(Cx)_i}}{\rho \sum_{k=1}^{n} a_k^2 N_k q_k} = \frac{\sum_{k=1}^{n} a_k N_k}{\rho \sum_{k=1}^{n} a_k^2 N_k q_k},\end{aligned}$$

or equivalently

$$\overline{\mathscr{R}}_v(q) \ge \rho \frac{\sum_{i=1}^{n} a_i^2 N_i q_i}{\sum_{i=1}^{n} a_i N_i}, \quad \text{for any } q \in (0,1)^n,$$

and by continuity also for all $q \in [0, 1]^n$. This proves the lower bound.

The upper bound of $\overline{\mathscr{R}}_v(q)$ can be obtained by noticing that

$$\overline{K}_v(q) = \operatorname{diag}(\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn})C \le \max\{\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn}\}C$$

and applying Lemma 4.5:

$$\begin{aligned}\overline{\mathscr{R}}_v(q) &= r(\overline{K}_v(q))\\ &\le \max\{\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn}\} r(C)\\ &= \max\{\mathscr{R}_{v1}, \ldots, \mathscr{R}_{vn}\} = \rho \max\{a_1 q_1, \ldots, a_n q_n\},\end{aligned} \tag{4.4}$$

where we have used that $r(C) = 1$.

To establish the equalities in (4.3), it is easy to see that the upper and lower bounds are the same when $q = s(1/a_1, \ldots, 1/a_n)$ for constant $s \in (0, a_{\min})$. The restriction on s guarantees that the point q is in $(0, 1)^n$. Then, by continuity, equality holds also for the endpoint values of s.

The proof is completed.

We now use the results above to prove Theorem 4.1.

Proof of Theorem 4.1 To prove Theorem 4.1(a), i.e., the bounds for $\mathscr{R}_v(p)$ in (4.1), recall that $q_i = 1 - p_i$, $\mathscr{R}_v(p) = \overline{\mathscr{R}}_v(q)$, and $\mathscr{R}_{vi} = \rho a_i q_i$ for $i = 1, 2, \ldots, n$. Substitution of these relationships into (4.3) yields (4.1).

For (b), note that when themixing $C = (c_{ij})$ is proportionate, $c_{ij} = \omega_j$. In this case, the NGM has rank 1 and its dominant eigenvalue is the sum of the diagonal elements. That is, $\mathscr{R}_v(p) = \sum_{i=1}^{n} \omega_i \mathscr{R}_{vi}$. Note also that in this case of isolated mixing, i.e., $\epsilon_i = 1$ for all i in (2.5), $C = I$ and the NGM is $\mathrm{diag}(\mathscr{R}_{v1}, \mathscr{R}_{v2}, \ldots, \mathscr{R}_{vn})$, for which $\mathscr{R}_v = \max\{\mathscr{R}_{v1}, \mathscr{R}_{v1}, \ldots, \mathscr{R}_{v1}\}$.

It is clear that for the proportionate mixing and isolated mixing, the corresponding reproduction numbers coincide with the lower and upper bounds of $\mathscr{R}_v$ given in (4.1). Thus, while Theorem 4.1, as stated, is not formally applicable to the proportionate and isolated mixing functions, the bounds of $\mathscr{R}_v$ in (4.1) correspond to these two extreme cases.

This completes the proof.

It is easy to verify that, when $a_1 = a_2 = \cdots = a_n$, Theorem 4.9 contains the conclusion of Theorem 4.7.

Theorems 4.1 and 4.9 can be used to derive the lower and upper bounds for the minimum reproduction number $\mathscr{R}_{v\{min\}}(\eta)$. Introduce the following notation:

$$\begin{aligned}
\pi_i &:= N_i / N, \quad 1 \le i \le n \text{Population fraction of sub-population} i; \\
\mathscr{S} &:= \sum_{i=1}^{n} (1 - p_i)\pi_i \quad \text{Population fraction unvaccinated;} \\
\widehat{\mathscr{R}}_0 &:= \sum_{i=1}^{n} \mathscr{R}_{0i}\pi_i \quad \text{Population weighted reproduction number;} \\
\mathscr{R}_0^{\diamond} &:= \left(\sum_{i=1}^{n} \frac{1}{\mathscr{R}_{0i}}\pi_i\right)^{-1} \quad \text{Harmonic mean of} \mathscr{R}_{0i} \text{weighted by sub-population fractions } \pi_i; \\
\widetilde{\mathscr{R}}_0 &:= \min_i \mathscr{R}_{0i}^2 \Big/ \widehat{\mathscr{R}}_0 \quad \text{Analogous to a scaled reproduction number}
\end{aligned} \tag{4.5}$$

The following results provide the lower and upper bounds for the minimum $\mathscr{R}_{v\{min\}}(\eta)$ in Problem (**I**):

Theorem 4.10 *Assume that the conditions of Theorem* 4.9 *hold. Let* $\eta < N$ (or $\bar{\eta} > 0$), *and let* S, $\widehat{\mathscr{R}}_0$, $\mathscr{R}_0^{\diamond}$, *and* $\widetilde{\mathscr{R}}_0$ *be defined in* (4.5).

(a) *The bounds of* $\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta})$ *for* $q \in \Omega_q^{(n)}(\bar{\eta}) \cap [0,1]^n$ *are:*

$$\frac{\rho a_{\min}^2 \bar{\eta}}{\Sigma_{i=1}^{n} a_i N_i} \le \overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) \le \frac{\rho \bar{\eta}}{\Sigma_{i=1}^{n} N_i \big/ a_i} \text{for } \bar{\eta} \le \frac{\min_i\{\mathscr{R}_{0i}\}}{\mathscr{R}_0^{\diamond}} N. \tag{4.6}$$

(b) *The bounds of* $\mathscr{R}_{v\{min\}}(\eta)$ *for* $p \in \Omega_p^{(n)}(\eta) \cap [0,1]^n$ *are:*

$$\widetilde{\mathscr{R}}_0\mathscr{S} \le \mathscr{R}_{v\{min\}}(\eta) \le \mathscr{R}_0^{\diamond}\mathscr{S}\,. \quad (4.7)$$

Proof (a) To prove the lower bound, note first that $\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) \le \overline{\mathscr{R}}_v(q)$ for all $q \in$ $q \in \Omega_q^{(n)}(\bar{\eta}) \cap [0,1]^n$. Note also that from the inequality in Theorem 4.9 we have

$$\overline{\mathscr{R}}_v(q) \ge \frac{\rho \Sigma_{i=1}^n a_i^2 N_i q_i}{\Sigma_{i=1}^n a_i N_i} \ge \frac{\rho a_{\min}^2 \Sigma_{i=1}^n N_i q_i}{\Sigma_{i=1}^n a_i N_i} = \frac{\rho a_{\min}^2 \bar{\eta}}{\Sigma_{i=1}^n a_i N_i}\,.$$

This proves the lower bound in (4.6).

For the upper bound in (a), choose $s > 0$ so that $Q = s(1/a_1, \dots, 1/a_n) \in \Omega_q^{(n)}(\bar{\eta})$; i.e., $s = \bar{\eta} \Big/ \left(\sum_{i=1}^n N_i \big/ a_i\right)$. Note that $Q \in [0,1]^n$ if $0 \le s \le a_{\min}$ or $0 \le \bar{\eta} \le a_{\min} \sum_{i=1}^n N_i \big/ a_i$. Then, by the case of equality in (a) of Theorem 4.9,

$$\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) \le \bar{R}_v(Q) = \rho s = \frac{\rho \bar{\eta}}{\Sigma_{i=1}^n N_i \big/ a_i}\,.$$

This completes the proof of (a).

To prove (b), note that $\mathscr{R}_{v\{min\}}(\eta) = \overline{\mathscr{R}}_{v\{min\}}(\bar{\eta})$. Note also that the left-hand side (LHS) and right-hand side (RHS) of inequality (4.6) can be re-expressed in terms of sub-population reproduction numbers $\mathscr{R}_{0i} = \rho a_i$. For the LHS of (4.6),

$$\text{LHS} = \frac{\min_{1 \le i \le n} \mathscr{R}_{0i}^2}{\widehat{\mathscr{R}}_0} \times \sum_{i=1}^n (1 - p_i)\pi_i = \widetilde{\mathscr{R}}_0\mathscr{S}\,.$$

It follows that

$$\mathscr{R}_{v\{\min\}}(\eta) \ge \widetilde{\mathscr{R}}_0\mathscr{S}\,. \quad (4.8)$$

For the RHS of (4.6), note that

$$\mathscr{R}_{v\{\min\}}(\eta) \le \frac{1}{\Sigma_{i=1}^n \pi_i \big/ \mathscr{R}_{0i}} \times \sum_{i=1}^n (1 - p_i)\pi_i = \mathscr{R}_0^{\diamond}\mathscr{S}\,. \quad (4.9)$$

From (4.8) to (4.9) we obtain (4.7). This completes the proof of (b).

We can now deduce the results for the upper bound of the critical number of vaccine doses η_* in Problem (**II**):

Theorem 4.11 (Critical number of vaccine doses) *Let C satisfy the same conditions as in Theorem* 4.9, *and let* $\mathscr{R}_0^{\diamond}$ *be defined in* (4.5).

(a) *the upper bound for the minimum vaccine dose is given by:*

$$\eta_* \leq N - \min\{1/\rho, a_1, \ldots, a_n\} \sum_{i=1}^{n} N_i / a_i. \quad (4.10)$$

(b) *If* $\mathscr{R}_{0i} > 1$ *for all i , then the inequality* (4.10) *can be re-written as*

$$\frac{\eta_*}{N} \leq 1 - \frac{1}{\mathscr{R}_0^{\diamond}}. \quad (4.11)$$

Proof (a) We prove the inequality (4.10) by considering two cases.

Case 1: Assume first that $1/\rho \leq a_{\min}$. In that case, the upper bound in Theorem 4.10 is applicable for $\bar{\eta} = (1/\rho)\sum_{i=1}^{n} N_i / a_i$, which gives

$$\overline{\mathscr{R}}_{v\{\min\}}(\bar{\eta}) \leq 1.$$

Thus,

$$\bar{\eta}_* \geq \frac{1}{\rho} \sum_{i=1}^{n} N_i \Big/ a_i.$$

Case 2: Assume now that $1/\rho \leq a_{\min}$. Then the upper bound in Theorem 4.10 for $\bar{\eta} = a_{\min}\sum_{i=1}^{n} N_i / a_i$ gives

$$\overline{\mathscr{R}}_{v\{\min\}}(\bar{\eta}) \leq \rho a_{\min} \leq 1.$$

Thus,

$$\bar{\eta}_* \geq a_{\min} \sum_{i=1}^{n} N_i / a_i.$$

Combining cases 1 and 2, we obtain

$$\bar{\eta}_* \geq \min\{1/\rho, a_1, \ldots, a_n\} \sum_{i=1}^{n} N_i / a_i,$$

and recalling that $\bar{\eta}_* = N - \eta_*$, we complete the proof of (a).

For (b), note that the RHS of (4.10) can be re-written as

$$N - \frac{1}{\rho} \sum_{i=1}^{n} \frac{N_i}{a_i} = N - \sum_{i=1}^{n} \frac{N_i}{\mathcal{R}_{0i}} = N\left(1 - \sum_{i=1}^{n} \frac{1}{\mathcal{R}_{0i}} \pi_i\right).$$

Thus, the inequality (4.10) can be re-written as

$$\frac{\eta_*}{N} \leq 1 - \sum_{i=1}^{n} \frac{1}{\mathcal{R}_{0i}} \pi_i = 1 - \frac{1}{\mathcal{R}_0^{\diamond}}.$$

This completes the proof.

Remarks

- **(i)** Note that $\mathcal{R}_0^{\diamond}$ and $\widetilde{\mathcal{R}}_0$ (see (4.5)) are weighted basic reproduction numbers, and the factor S is the fraction of the overall population that remains susceptible. In light of this, we see that the lower and upper bounds for $\mathcal{R}_{v\{min\}}(\eta)$ in (4.7) take the familiar form of an effective reproduction number.
- **(ii)** The lower and upper bounds in (4.7) are equal if the activities a_i for subpopulations i are all the same. Note that $\mathcal{R}_{0i} = \rho a_i = \beta a_i / (\theta + \gamma)$. Thus, $\mathcal{R}_{0i}$ are the same when a_i are the same for all i. Then, from (4.5) we see that $\widetilde{\mathcal{R}}_0 = \mathcal{R}_0^{\diamond}$, which implies that the inequalities in (4.7) become equalities.
- **(iii)** For the upper bound of η_*, if $a_i = a$ are all the same, we have $\mathcal{R}_0^{\diamond} = \mathcal{R}_0$, in which case the upper bound in (4.11) becomes $1 - 1/\mathcal{R}_0$. This is similar to the usual formula for the critical vaccination fraction $p_c = 1 - 1/\mathcal{R}_0$, for which the number of vaccinated is $\eta_c = p_c N = N(1 - 1/\mathcal{R}_0)$.

4.3 Example: the case of Jacquez mixing

As an example ofmixing functions that satisfy the conditions described in Lemma 4.6, we consider the c_{ij} for the meta-population model in Feng et al. (2015), which is the Jacquez mixing as given in (2.5).

Proposition 4.12 *Let the matrix* $C = (c_{ij})$ *be given by* (2.5) *with* $\epsilon_i \in (0, 1)$, $i = 1, \ldots, n$. *Then* C *is invertible and* $B = C^{-1} = (b_{ij})$ *is given by*

$$b_{ij} = \delta_{ij}\epsilon_i^{-1} + \frac{(1-\epsilon_i^{-1})(1-\epsilon_j^{-1})a_jN_j}{\Sigma_{k=1}^{n}(1-\epsilon_k^{-1})a_kN_k}, \quad i, j = 1, \ldots, n.$$

In particular, $-C^{-1}$ *is essentially nonnegative.*

Remark 4.13 It is also clear that under conditions of Proposition 4.12, C is positive and thus irreducible.

Proof It will be sufficient to show that $CB = I$ (the identity matrix), or

$$\sum_{j=1}^{n} c_{ij}b_{jk} = \delta_{ik}, \quad \text{for all } i, k = 1, \ldots, n.$$

To simplify computations, let

$$\mu_c = \sum_{l=1}^{n} (1-\epsilon_l)a_lN_l, \quad \mu_b = \sum_{l=1}^{n} (1-\epsilon_l^{-1})a_lN_l.$$

Then

$$\begin{aligned}
\sum_{j=1}^{n} c_{ij}b_{jk} &= \sum_{j=1}^{n}\left[\delta_{ij}\epsilon_i + \frac{(1-\epsilon_i)(1-\epsilon_j)a_jN_j}{\mu_c}\right] \times \left[\delta_{jk}\epsilon_j^{-1} + \frac{(1-\epsilon_j^{-1})(1-\epsilon_k^{-1})a_kN_k}{\mu_b}\right] \\
&= \sum_{j=1}^{n} \delta_{ij}\delta_{jk}\epsilon_i\epsilon_j^{-1} + \frac{1}{\mu_c}\sum_{j=1}^{n} \delta_{jk}(1-\epsilon_i)(1-\epsilon_j)\epsilon_j^{-1}a_jN_j + \frac{1}{\mu_b}\sum_{j=1}^{n} \delta_{ij}\epsilon_i(1-\epsilon_j^{-1})(1-\epsilon_k^{-1})a_kN_k + \frac{1}{\mu_c\mu_b}\sum_{j=1}^{n} (1-\epsilon_i)[(1-\epsilon_i)(1-\epsilon_j^{-1})](1-\epsilon_k^{-1})a_jN_ja_kN \\
&= \delta_{ik} - \frac{(1-\epsilon_i)(1-\epsilon_k^{-1})a_kN_k}{\mu_c} - \frac{(1-\epsilon_i)(1-\epsilon_k^{-1})a_kN_k}{\mu_b} + \frac{\Sigma_{j=1}^{n}[(1-\epsilon_j) + (1-\epsilon_j^{-1})]a_jN_j}{\mu_c\mu_b}(1-\epsilon_i)(1-\epsilon_k^{-1})a_kN_k \\
&= \delta_{ik} + \left(-\frac{1}{\mu_c} - \frac{1}{\mu_b} + \frac{\mu_c + \mu_b}{\mu_c\mu_b}\right)(1-\epsilon_i)(1-\epsilon_k^{-1})a_kN_k = \delta_{ik},
\end{aligned}$$

where we have used above that

$$(1-\epsilon_j)(1-\epsilon_j^{-1}) = (1-\epsilon_j) + (1-\epsilon_j^{-1}).$$

This completes the proof.

5 Discussion

The main goal of this study was to solve Problems (**I**) and (**II**), which identify the most efficient allocation of limited vaccines using a meta-population model for vaccine-preventable infectious diseases. Although we demonstrated the results using Model (2.1), the approach can be applied to other meta-population models for vaccine-preventable diseases.

Model (2.1) incorporates various heterogeneities such as in activity, contacts between sub-populations (mixing), vaccination coverage, and size of each sub-population. We considered general mixing functions that satisfy conditions (2.2)–(2.4), including the special case of Jacquez mixing in (2.5) and special cases of this namely, proportionate mixing ($\epsilon_i = 0$ for all i) and preferential mixing ($\epsilon_i = 1$ for all i). However, it would be more challenging to consider a similar approach to meta-population models that are less-tractable analytically than Model (2.1), particularly when the NGM has a more complicated structure. For example, when a model includes additional factors such as aging from one age-group to the next, multi-level mixing (e.g., age and spatial), and heterogeneity in infectivity and susceptibility, the effective reproduction number $\mathscr{R}_v$ as a function of vaccination coverage will be more difficult to study from an analytic point of view.

The optimization problem is based on reducing the effective reproduction number $\mathscr{R}_v$ (if $\mathscr{R}_0 > 1$) by determining the optimal combination of vaccine coverages $p = (p_1, p_2, \ldots , p_n)$. Because the parameters p_i must be between 0 and 1, the optimal solution $P^*(\eta)$ needs to be in the unit hypercube. Even in the case of $n = 2$ sub-populations, the solution of Problems (**I**) and (**II**) is not trivial. For $n > 2$, the most challenging task is to show the convexity of $\mathscr{R}_v(p)$; Theorem 4.2. This proves a conjecture of Hill and Longini (2003), although those authors did not consider the structure of the mixing matrix C specified by conditions (2.2)–(2.4). Our proofs are facilitated by using the 'reflected' quantities $q_i = 1 - p_i$ and $\overline{\mathscr{R}}_v(q_1, q_2, \ldots , q_n) = \mathscr{R}_v(p_1, p_2, \ldots , p_n)$. For ease of presentation, we first illustrated results for the simpler case of $n = 2$ sub-populations, and then extended them to $n > 2$ sub-populations.

In the case of $n = 2$, explicit formulae are obtained for the optimal solutions when mixing is proportionate or preferential. For Problem (**I**), the optimal solution $P^*(\eta)$ for a given number of vaccine doses η and the minimized reproduction number $\mathscr{R}_{v\{min\}}(\eta)$ are described as functions of model parameters (Theorem 3.2). For Problem (**II**), an analytical formula for the minimum vaccine doses η_* that will reduce $\mathscr{R}_v$ to below 1 is provided (Theorem 3.3).

Another interesting finding is that, for any number of vaccine doses η in the constraint, the optimal vaccine coverage $P^*(\eta)$ lies along the 'critical' ray Γ. In addition, when $n = 2$, for the optimal solution $P^*(\eta)$ to be in the unit square $[0, 1]^2$, available vaccine doses must satisfy $\eta_0 < \eta < N$, where the lower bound η_0 is determined by model parameters. For $\eta \leq \eta_0$, the optimal strategy will be to vaccinate only one sub-population.

Extension of these results to the case $n > 2$ is complicated by the fact that no explicit formulae are available. Nevertheless, using results for the spectral radius of nonnegative matrices, in Sects. 4.1 and 4.2, we obtain bounds of $\mathscr{R}_v$ for an arbitrary mixing C that satisfies the conditions in Theorem 4.1. An interesting finding is that for a large class of mixing matrix C (not just Jacquez), the proportionate mixing gives the smallest $\mathscr{R}_v$ while the isolated mixing (no mixing between sub-populations) gives the largest $\mathscr{R}_{vi}$ given $\mathscr{R}_{vi}$ (see Theorem 4.1). It is clear that these conclusions hold particularly for $\mathscr{R}_0 = \mathscr{R}_v(0)$. The facts that population heterogeneities tend to increase $\mathscr{R}_0$ and that models assuming proportionate mixing generate lower values of $\mathscr{R}_0$ have been suggested by other researchers (Adler 1992; Andersson and Britton 1998; Diekmann et al. 2012).

We also establish bounds on the relative minima $\mathscr{R}_{v\{min\}}(\eta)$ and the critical vaccine dose η_* (see Theorems 4.10 and 4.11). Interpretations of those bounds are provided in terms of biological quantities such as weighted reproduction numbers. In particular, we see that the lower and upper bounds are products of the weighted reproduction number $\widetilde{\mathscr{R}}_0$ and the harmonic mean $\mathscr{R}_0^\diamond$, respectively, with the fraction unvaccinated S (see (4.7)). Thus, both bounds are in the familiar form of effective reproduction numbers. Moreover, the bounds are equal when all sub-populations have the same activity (a_i). A similar interpretation holds for the upper bound of η_*, in which case the usual basic reproduction number $\mathscr{R}_0$ is replaced by the harmonic mean $\mathscr{R}_0^\diamond$ of sub-population reproduction numbers the $\mathscr{R}_{0i}$ weighted by sub-population fractions π_i (see also (4.11)).

Acknowledgements

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention or other institutions with which they are affiliated. We thank the anonymous reviewers for comments and suggestions, which helped improve the presentation of the manuscript.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

A Appendix

In this appendix, we provide detailed proofs for Theorems 3.1–3.3 in Sect. 3 and illustrate an example of these results in the case of Jacquez mixing.

A.1 Proofs of Theorems 3.1–3.3

To prove these theorems, we first prove several propositions. Instead of working with the function $\mathscr{R}_v(p_1, p_2)$ along the rays emanating from the point (p_1, p_2) (1, 1), it is much easier to consider the ‘reflected’ variables:

$$q_i = 1 - p_i, \quad i = 1, 2, \quad q = (q_1, q_2),$$

and the corresponding rays emanating from the point $(q_1, q_2) = (0, 0)$ into the unit square $[0, 1]^2$. The ‘reflected’ function $\overline{\mathscr{R}}_v$ is then given by

$$\overline{\mathscr{R}}_v(q_1, q_2) = \mathscr{R}_v(1 - q_1, 1 - q_2).$$

The quantities corresponding to those mentioned directly after Problem (**I**) in terms of q and the reflected function are

$$\begin{aligned}\bar{\eta} &:= q_1 N_1 + q_2 N_2 = N - \eta,\\ \ell(q) &:= q_1 N_1 + q_2 N_2 = \bar{\eta},\\ Q^*(\bar{\eta}) &:= (1,1) - P^*(\eta) = (1,1) - P^*(N - \bar{\eta}).\end{aligned}$$

In addition, at the optimal points P^* or Q^*, we have

$$\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) = \overline{\mathscr{R}}_v \big|_{Q^*(\bar{\eta})} = \mathscr{R}_v \big|_{P^*(\eta)} = \mathscr{R}_{v\{min\}}(\eta),$$

and we know that $\overline{\mathscr{R}}_v$ satisfies the equation

$$\nabla \overline{\mathscr{R}}_v \big|_{Q^*(\bar{\eta})} = \lambda(N_1, N_2),$$

or, equivalently,

$$\nabla \overline{\mathscr{R}}_v \big|_{Q^*(\bar{\eta})} \cdot (N_2, -N_1) = 0,$$

provided $Q^*(\bar{\eta})$ is in the interior of the unit square.

A.1.1 Reflected function $\overline{\mathscr{R}}_v$ and its properties

The reflected function $\overline{\mathscr{R}}_v$ is given explicitly by

$$\overline{\mathscr{R}}_v(q) = \frac{1}{2}\Big[\mathscr{R}_{01}c_{11}q_1 + \mathscr{R}_{02}c_{22}q_2 + \sqrt{(\mathscr{R}_{01}c_{11}q_1 - \mathscr{R}_{02}c_{22}q_2)^2 + 4\mathscr{R}_{01}\mathscr{R}_{02}c_{12}c_{21}q_1q_2}\Big] \quad \text{(A.1)}$$

for $q \in [0, 1]^2$. Note that the formula (A.1) can be used to naturally extend $\overline{\mathscr{R}}_v$ to the first quadrant $q \geq 0$ (i.e., q_1, $q_2 \geq 0$). It can be checked that the function $\overline{\mathscr{R}}_v$ is homogeneous of degree 1, i.e.,

$$\overline{\mathscr{R}}_v(sq) = s\overline{\mathscr{R}}_v(q), \quad s > 0, q \geq 0. \quad \text{(A.2)}$$

Geometrically, this means that $\overline{\mathscr{R}}_v$ grows linearly on the rays emanating from the origin, as illustrated in Fig. 4.

We next compute the first derivatives of $\overline{\mathscr{R}}_v$:

$$\frac{\partial \overline{\mathscr{R}}_v}{\partial q_i} = \frac{1}{2}\left[\mathscr{R}_{0i}c_{ii} + \frac{\mathscr{R}_{0i}^2 c_{ii}^2 q_i + (2c_{12}c_{21} - c_{11}c_{22})\mathscr{R}_{01}\mathscr{R}_{02}q_j}{\sqrt{(\mathscr{R}_{01}c_{11}q_1 - \mathscr{R}_{02}c_{22}q_2)^2 + 4\mathscr{R}_{01}\mathscr{R}_{02}c_{12}c_{21}q_1q_2}}\right] \quad \text{(A.3)}$$

for $i, j = 1, 2$ and $i \neq j$. Note that the function $\nabla\overline{\mathscr{R}}_v = (\partial\overline{\mathscr{R}}_v / \partial q_1, \partial\overline{\mathscr{R}}_v / \partial q_2)$ is homogeneous of degree 0:

$$\nabla\overline{\mathscr{R}}_v(sq) = \nabla\overline{\mathscr{R}}_v(q), \quad s > 0, q \geq 0, \quad \text{(A.4)}$$

i.e., $\nabla\overline{\mathscr{R}}_v$ is constant on rays emanating from the origin.

Further, we compute the second derivatives of $\overline{\mathscr{R}}_v$. By direct computation, we have the following formula for the Hessian

$$\text{Hess } \overline{\mathscr{R}}_v = k\begin{pmatrix} q_2^2 & -q_1q_2 \\ -q_1q_2 & q_1^2 \end{pmatrix},$$

where

$$k = \frac{c_{12}c_{21} \mid C \mid \mathscr{R}_{01}^2\mathscr{R}_{02}^2}{\left[(\mathscr{R}_{01}c_{11}q_1 - \mathscr{R}_{02}c_{22}q_2)^2 + 4\mathscr{R}_{01}\mathscr{R}_{02}c_{12}c_{21}q_1q_2\right]^{3/2}}.$$

Note that $k > 0$ by (3.1) and therefore Hess $\overline{\mathscr{R}}_v$ is a nonnegative semi-definite matrix, by Sylvester's criterion. Consequently, $\overline{\mathscr{R}}_v$ is a convex function of (q_1, q_2). We explicitly note here that $\overline{\mathscr{R}}_v$ is not a strictly convex function, as it grows linearly on the rays emanating from the origin. However, as we show below, it is strictly convex in certain directions. To be more precise, let $u = (u_1, u_2)$ be a unit vector, and consider the second derivative of $\overline{\mathscr{R}}_v$ in the direction u. We have

$$\frac{\partial^2\overline{\mathscr{R}}_v}{\partial u^2} = k(u_1^2q_2^2 + u_2^2q_1^2 - 2u_1u_2q_1q_2) = k(u_1q_2 - u_2q_1)^2 \geq 0,$$

whereupon

$$\frac{\partial^2\overline{\mathscr{R}}_v}{\partial u^2} > 0 \quad \text{unless} u \parallel (q_1, q_2).$$

(Here, $u \parallel v$ indicates that vectors u and v are parallel.) In particular, $\overline{\mathscr{R}}_v$ is strictly convex in the direction $(N_2, -N_1)$.

To proceed, recall that

$$\Omega_q^{(2)}\bar{\eta} := \left\{(q_1, q_2) : \ell(q) = \bar{\eta}\right\},$$

and consider again the constraint set

$$\Omega_q^{(2)}(\bar{\eta}) \cap [0, 1]^2, \quad \text{for} \bar{\eta} \in \left[0, N\right].$$

This is a line segment parallel to the vector $(N_2, -N_1)$ with endpoints on the boundary of $[0, 1]^2$. We will denote the left and right endpoints (with respect to the direction $(N_2, -N_1)$) by $Q_1(\bar{\eta})$ and $Q_2(\bar{\eta})$, respectively. It is easy to see that

$$\begin{aligned} Q_1(\bar{\eta}) &\in (\{0\} \times [0, 1]) \cup ([0, 1] \times \{1\}), \\ Q_2(\bar{\eta}) &\in ([0, 1] \times \{0\}) \cup (\{1\} \times [0, 1]). \end{aligned}$$

Proposition A.1 *The function* $\overline{\mathscr{R}}_v(q)$ *is strictly convex on the intersections* $\Omega_q^{(2)}(\bar{\eta}) \cap [0, 1]^2$, *for* $\bar{\eta} \in (0, N)$ *if condition* (3.1) *holds. Consequently its minimum will occur at an interior point* $Q^*(\bar{\eta})$ *if and only if*

$$\nabla \overline{\mathscr{R}}_v \,|_{Q_1(\bar{\eta})} \cdot (N_2, -N_1) < 0, \quad \text{(A.5)}$$

$$\nabla \overline{\mathscr{R}}_v \,|_{Q_2(\bar{\eta})} \cdot (N_2, -N_1) > 0, \quad \text{(A.6)}$$

where $Q_1(\bar{\eta})$ *and* $Q_2(\bar{\eta})$ *are the left and right endpoints of* $\Omega_q^{(2)}(\bar{\eta}) \cap [0, 1]^2$ *(see Fig. 5). Moreover,* $Q^*(\bar{\eta})$ *is the unique point on* $\Omega_q^{(2)}(\bar{\eta}) \cap [0, 1]^2$ *such that*

$$\nabla \overline{\mathscr{R}}_v \,|_{Q^*(\bar{\eta})} \cdot (N_2, -N_1) = 0, \quad \text{(A.7)}$$

From the homogeneity properties (A.2) and (A.4), we also have the following proposition.

Proposition A.2 *Under the assumptions of Proposition* A.1, *let* $\bar{\eta} \in (0, N)$ be such that $Q^*(\bar{\eta})$ *is an interior point. Then, for* $s > 0$ *such that* $s\bar{\eta} \in (0, N)$, *the minimum point of* $\overline{\mathscr{R}}_v$ *on* $\Omega_q^{(2)}(s\,\bar{\eta}) \cap [0, 1]^2$ *is given by*

$$Q^*(s\bar{\eta}) = sQ^*(\bar{\eta}),$$

provided that this point still lies in the interior of the unit square.[4] *In other words, all interior minimum points lie on a ray emanating from the origin. We will denote this ray by* $\bar{\Gamma}$ *and call it the* (reflected) critical ray. *Moreover, by Proposition* A.1

$$\bar{\Gamma} \cap [0,1]^2 = \{q : \nabla \overline{\mathscr{R}}_v(q) \cdot (N_2, -N_1) = 0\} \cap [0,1]^2 .$$

Proof For $Q^*(\bar{\eta})$ we have (see Fig. 5)

$$\ell(Q^*(\bar{\eta})) = \bar{\eta} \quad \text{and} \quad \nabla \overline{\mathscr{R}}_v \big|_{Q^*(\bar{\eta})} \cdot (N_2, -N_1) = 0 .$$

But then

$$\ell(sQ^*(\bar{\eta})) = s\bar{\eta}$$

and by (A.4)

$$\nabla \overline{\mathscr{R}}_v \big|_{sQ^*(\bar{\eta})} \cdot (N_2, -N_1) = \nabla \overline{\mathscr{R}}_v \big|_{Q^*(\eta)} \cdot (N_2, -N_1) = 0,$$

implying that $Q^*(\bar{\eta})$ is the critical point on the constraint with constant $s\,\bar{\eta}$ and thus that

$$Q^*(s\bar{\eta}) = sQ^*(\bar{\eta}) .$$

The rest then follows from Proposition A.1.

A.1.2 Endpoint conditions

We now write the endpoint conditions (A.5)–(A.6) using the explicit formulae for the derivatives of $\overline{\mathscr{R}}_v$. Because of Proposition A.2, to verify (A.5)–(A.6), it will be sufficient to verify them for the constraint value $s\,\bar{\eta}$ with a small $s > 0$. Thus, without loss of generality, we may assume that $\bar{\eta}$ itself is small. In that case, the intersection points $Q_1(\bar{\eta})$ and $Q_2(\bar{\eta})$ of $\Omega_q^{(2)}(\bar{\eta})$ with $\partial([0, 1]^2)$ will lie on the left and bottom sides of the square, i.e.,

$$Q_1(\bar{\eta}) \in \left\{0\right\} \times (0,1), \quad Q_2(\bar{\eta}) \in (0,1) \times \left\{0\right\} .$$

From (A.3) we have

---

[4] This will automatically hold for $0 < s < 1$.

$$\nabla \overline{\mathscr{R}}_v\,|_{q_1 = 0} = \left(\frac{c_{12}c_{21}}{c_{22}}\mathscr{R}_{01}, \mathscr{R}_{02}c_{22}\right),$$

$$\nabla \overline{\mathscr{R}}_v\,|_{q_2 = 0} = \left(\mathscr{R}_{01}c_{11}, \frac{c_{12}c_{21}}{c_{11}}\mathscr{R}_{02}\right)$$

and, therefore, the conditions (A.5)–(A.6) will take the form

$$\frac{c_{12}c_{21}}{c_{22}}\mathscr{R}_{01}N_2 - \mathscr{R}_{02}c_{22}N_1 < 0, \quad \mathscr{R}_{01}c_{11}N_2 - \frac{c_{12}c_{21}}{c_{11}}\mathscr{R}_{02}N_1 > 0$$

which is equivalent to (3.3).

A.1.3 The critical ray $\bar{\Gamma}$

We next characterize the critical ray $\bar{\Gamma}$. Namely, we identify the intersection point of $\bar{\Gamma}$ with $\partial([0, 1]^2)$, which we denote by Q_0.

For this, consider the function

$$\phi(q) = \nabla \overline{\mathscr{R}}_v(q) \cdot (N_2, -N_1).$$

Because $\phi = 0$ on $\bar{\Gamma}$ (see Proposition A.1), we must have

$$\phi(Q_0) = 0.$$

Note that conditions (A.5) and (A.6) are equivalent to

$$\phi(0, 1) < 0, \quad \phi(1, 0) > 0.$$

Thus, the location of Q_0 will depend on the sign of

$$\phi(1, 1) = \nabla \overline{\mathscr{R}}_v(1, 1) \cdot (N_2, -N_1).$$

That is, $Q_0 = (1, q_2^0) \in \{1\} \times (0, 1]$ if $\phi(1, 1) \leq 0$, and $Q_0 = (q_1^0, 1) \in (0, 1] \times \{1\}$ if $\phi(1, 1) \geq 0$. We will use this fact to find Q_0 and characterize $\bar{\Gamma}$.

Proposition A.3 *Assume that conditions* (3.1) *and* (3.3) *hold, where* κ_1, κ_2 *are as in* (3.2). *Then the intersection* $Q_0 = (q_1^0, q_2^0)$ *of the critical ray* $\bar{\Gamma}$ *with* $\partial([0, 1]^2)$ *has the property*

$$\frac{q_2^0}{q_1^0} = \frac{\kappa_2}{\kappa_1}.$$

Consequently, the critical ray $\bar{\Gamma}$ is given by

$$\bar{\Gamma}: \frac{q_2}{q_1} = \frac{\kappa_2}{\kappa_1}.$$

Proof Recall that $\phi(q) = \nabla\overline{\mathscr{R}}_v(q) \cdot (N_2, -N_1)$ and consider the case $\phi(1, 1) \le 0$ first. In that case, $Q_0 = (1, q_2^0)$, where q_2^0 is found from the equation $\phi(1, q_2) = 0$. This equation can be reduced to a quadratic equation for q_2, which has two roots, given by the formulae

$$q_2^{(-)} = \lambda\rho\frac{c_{11}\rho - \sqrt{c_{12}c_{21}}\lambda}{c_{22}\lambda - \sqrt{c_{12}c_{21}}\rho}, \quad q_2^{(+)} = \lambda\rho\frac{c_{11}\rho + \sqrt{c_{12}c_{21}}\lambda}{c_{22}\lambda + \sqrt{c_{12}c_{21}}\rho},$$

where

$$\lambda = \sqrt{\frac{N_1}{N_2}}, \quad \rho = \sqrt{\frac{\mathscr{R}_{01}}{\mathscr{R}_{02}}}.$$

Then q_2^0 equals either $q_2^{(-)}$ or $q_2^{(+)}$ (the reduction to a quadratic equation may have introduced a false root). Plugging the formulae above into ϕ, it can be verified that $\phi\left(1, q_2^{(-)}\right) = 0$ if and only if

$$c_{22}\lambda^2 + c_{11}\rho^2 \ge 2\lambda\rho\sqrt{c_{12}c_{21}}.$$

This inequality is indeed satisfied, because of the condition (3.1):

$$c_{22}\lambda^2 + c_{11}\rho^2 \ge 2\lambda\rho\sqrt{c_{22}c_{11}} \ge 2\lambda\rho\sqrt{c_{12}c_{21}}.$$

On the other hand, the verification of the equation $\phi\left(1, q_2^{(+)}\right) = 0$ results in the condition $c_{22}\lambda^2 = c_{11}\rho^2$, which implies that $q_2^{(+)} = q_2^{(-)}$. Thus, in either case, we can conclude that $q_2^0 = q_2^{(-)}$. Hence,

$$\frac{q_2^0}{q_1^0} = q_2^0 = \lambda\rho\frac{c_{11}\rho - \sqrt{c_{12}c_{21}}\lambda}{c_{22}\lambda - \sqrt{c_{12}c_{21}}\rho},$$

which is the same quantity as κ_2/κ_1 as in the statement of the proposition. This finishes the proof in this case.

The case $\phi(1, 1) \ge 0$ is considered similarly and we obtain exactly the same value for the ratio q_2^0 / q_1^0.

We next identify the value of the constraint $\bar{\eta} = \bar{\eta}_0$, which corresponds to the intersection point $Q_0 = (q_1^0, q_2^0)$ of $\bar{\Gamma}$ and $\partial([0, 1]^2)$.

Proposition A.4 *Let* Q_0 *be as in Proposition* A.3. *Then*

$$\bar{\eta}_0 = \ell(Q_0) = \frac{\kappa_1 N_1 + \kappa_2 N_2}{\max\left\{\kappa_1, \kappa_2\right\}}.$$

Proof If $\kappa_1 \geq \kappa_2$ then $q_1^0 = 1$ and $q_2^0 = \kappa_2 \Big/ \kappa_1$ and thus

$$\bar{\eta}_0 = \ell(Q_0) = N_1 + \frac{\kappa_2}{\kappa_1} N_2 = \frac{\kappa_1 N_1 + \kappa_2 N_2}{\kappa_1}.$$

On the other hand, if $\kappa_1 \leq \kappa_2$ then $q_2^0 = 1$ and $q_1^0 = \kappa_1 \Big/ \kappa_2$ and thus

$$\bar{\eta}_0 = \ell(Q_0) = \frac{\kappa_1}{\kappa_2} N_1 + N_2 = \frac{\kappa_1 N_1 + \kappa_2 N_2}{\kappa_2}.$$

Combining the two cases, we obtain the stated formula.

A.1.4 Explicit formulae for minima

Proposition A.5 *For* $0 < \bar{\eta} < \bar{\eta}_0$, *we have the following explicit formulae for* $Q^*(\bar{\eta})$ *and the minimum of* $\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta})$:

$$\begin{aligned} Q^*(\bar{\eta}) &= \frac{\bar{\eta}}{\kappa_1 N_1 + \kappa_2 N_2}(\kappa_1, \kappa_2), \\ \overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) &= |\, C \,| \mathscr{R}_{01} \mathscr{R}_{02} \sqrt{N_1 N_2} \frac{\bar{\eta}}{\kappa_1 N_1 + \kappa_2 N_2}. \end{aligned}$$

Proof We know that $Q^*(\bar{\eta}) = (\kappa_1, \kappa_2)s$, where $s > 0$ can be found from the constraint $\ell(\kappa_1 s, \kappa_2 s) = \bar{\eta}$, which gives

$$s = \frac{\bar{\eta}}{\kappa_1 N_1 + \kappa_2 N_2}.$$

This proves the first formula. To establish the second, we first use the homogeneity of $\overline{\mathscr{R}}_v$:

$$\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) = \overline{\mathscr{R}}_v \,\big|_{Q^*(\bar{\eta})} = \overline{\mathscr{R}}_v(\kappa_1 s, \kappa_2 s) = \overline{\mathscr{R}}_v(\kappa_1, \kappa_2) s.$$

Furthermore, by direct calculations, one can show that

$$\overline{\mathscr{R}}_v(\kappa_1, \kappa_2) = |C| \mathscr{R}_{01}\mathscr{R}_{02}\sqrt{N_1 N_2},$$

which completes the proof.

A.1.5 Proofs of the Theorems

The proofs for Theorems 3.1 and 3.2 can be completed by combining Propositions A.1–A.5, and by writing their statements in terms of the original variables p_i and the constraint value η.

To prove Theorem 3.3, we consider several cases. When $\mathscr{R}_{v\{min\}}(\eta_0) \geq 1$, we simply use the formula for $\mathscr{R}_{v\{min\}}(\eta)$ in Theorem 3.2 for $\eta_0 < \eta < N$ to find η_*. When $\mathscr{R}_{v\{min\}}(\eta_0) \leq 1$, the minimum points $P^*(\eta)$ will be on the boundary of the square for $0 < \eta < \eta_0$ and we solve the equations

$$\mathscr{R}_v(p_1, 0) = 1, \quad \mathscr{R}_v(0, p_2) = 1,$$

or more precisely in variables (q_1, q_2):

$$\mathscr{R}_v(q_1, 1) = 1, \quad \mathscr{R}_v(1, q_2) = 1,$$

from which we get

$$\begin{aligned} q_1 &= \frac{(1 - c_{22}\mathscr{R}_{02})}{c_{11}\mathscr{R}_{01} - |C| \mathscr{R}_{01}\mathscr{R}_{02}}, \\ q_2 &= \frac{(1 - c_{11}R_{01})}{c_{22}\mathscr{R}_{02} - |C| \mathscr{R}_{01}\mathscr{R}_{02}}. \end{aligned}$$

The corresponding values of η_* are

$$\eta_* = \min\left\{N_1(1 - q_1), N_2(1 - q_2)\right\}.$$

A.2 Example: the case of Jacquez preferred mixing

The conditions in (3.3), which guarantee that the critical ray Γ passes through the interior of the unit square, hold for general mixing functions that satisfy (2.2)–(2.4). These conditions may simplify when specific functions are considered. In this section, we consider the Jacquez preferred mixing given in (2.5).

A.2.1 Interpretation of conditions for the interior critical ray

To verify the conditions (3.3), we consider two cases, one for homogeneous activity ($a_1 = a_2$) and other for heterogeneous activity ($a_1 \neq a_2$).

Case 1: $a_1 = a_2$. In this case, the inequality $\kappa_1 > 0$ can be rewritten as

$$(1-\epsilon_1)(1-\epsilon_2)N_2 < \epsilon_2(1-\epsilon_1)N_1 + [\epsilon_2(1-\epsilon_2) + (1-\epsilon_2)^2]N_2$$

or

$$-\epsilon_1(1-\epsilon_2)N_2 < \epsilon_2(1-\epsilon_1)N_1,$$

which is always satisfied. Similarly, one can verify that $\kappa_2 > 0$ as well and therefore conditions (3.3) will always hold if $a_1 = a_2$.

Case 2: $a_1 \neq a_2$. In this case, the inequality $\kappa_1 > 0$ can be rewritten as

$$(1-\epsilon_1)(1-\epsilon_2)a_1N_2 < \epsilon_2[(1-\epsilon_1)a_1N_1 + (1-\epsilon_2)a_2N_2] + (1-\epsilon_2)^2a_2N_2$$

and simplifying further to

$$(1-\epsilon_2)\left[(1-\epsilon_1)a_1 - a_2\right]N_2 < \epsilon_2(1-\epsilon_1)a_1N_1\,.$$

Note that this inequality will readily hold if $(1 - \epsilon_1)a_1 - a_2 \leq 0$ (which will happen, e.g., if $a_1 \leq a_2$), as the left-hand side will be nonpositive and the right-hand side positive. If, however, $(1 - \epsilon_1)a_1 - a_2 > 0$, then the above inequality will transform to

$$\frac{N_2}{N_1} < \frac{\epsilon_2}{1-\epsilon_2}\frac{(1-\epsilon_1)a_1}{(1-\epsilon_1)a_1 - a_2}\,.$$

By repeating this analysis with interchanged indices, we summarize the results above in the following proposition.

Proposition A.6 *Suppose the mixing matrix C is given by* (2.5) *with* $n = 2$.

- **(i)** *If* $a_1 = a_2$, *then condition* (3.3) *holds for any* N_i, $\epsilon_i \in (0, 1)$, $i = 1, 2$.
- **(ii)** *If* $1 - \epsilon_1 \leq a_2/a_1 \leq 1/(1 - \epsilon_2)$, *then* (3.3) *holds for any* N_1, N_2.
- **(iii)** *If* $a_2/a_1 < 1 - \epsilon_1$, *then* (3.3) *becomes*

$$\frac{N_2}{N_1} < \frac{\epsilon_2}{1-\epsilon_2}\frac{(1-\epsilon_1)a_1}{(1-\epsilon_1)a_1 - a_2}\,.$$

(iv) *If* $1/(1 - \epsilon_2) < a_2/a_1$, *then* (3.3) *becomes*

$$\frac{1-\epsilon_1}{\epsilon_1}\frac{(1-\epsilon_2)a_2 - a_1}{(1-\epsilon_2)a_2} < \frac{N_2}{N_1}\,.$$

A.2.2 Simplified expressions at the optimal point

Some of the explicit expressions for the optimal solution provided in previous subsections hold for more general mixing functions $C = (c_{ij})$. These expressions may be simplified when the preferred mixing given in (2.5) is used. These simplified expressions are described in the following propositions. Let

$$\alpha_i := a_j\left[(1-\epsilon_j)(a_j-(1-\epsilon_j)a_i)N_j + a_i\epsilon_j(1-\epsilon_i)N_i\right], \quad \text{(A.8)}$$

for $i, j = 1, 2$ and $i \neq j$. Note that the α_i differ from the κ_i in (3.2) by a constant positive factor. In particular, $\alpha_2/\alpha_1 = \kappa_2/\kappa_1$.

Proposition A.7 *Consider the mixing function given in* (2.5). *Let condition* (3.3) *be satisfied as described in Proposition* A.6. *Then the relative minima* $Q^*(\bar{\eta})$ *will be interior points if and only if*

$$0 < \bar{\eta} < \bar{\eta}_0 = \frac{\alpha_1 N_1 + \alpha_2 N_2}{\max\left\{\alpha_1, \alpha_2\right\}},$$

and will lie on the critical ray

$$\bar{\Gamma}: \frac{q_2}{q_1} = \frac{\alpha_2}{\alpha_1}.$$

Moreover, the following explicit formulae hold for $0 < \bar{\eta} < \bar{\eta}_0$:

$$\begin{aligned} Q^*(\bar{\eta}) &= \frac{(\alpha_1, \alpha_2)}{\alpha_1 N_1 + \alpha_2 N_2}\bar{\eta}, \\ \mathscr{R}_{v\{min\}}(\bar{\eta}) &= \frac{\rho a_1 a_2 (a_2 N_2 \epsilon_1 (1-\epsilon_2) + a_1 N_1 (1-\epsilon_1)\epsilon_2)}{\alpha_1 N_1 + \alpha_2 N_2}\bar{\eta}. \end{aligned} \quad \text{(A.9)}$$

We next note that, for α_i defined in (A.8), we have

$$\alpha_2 - \alpha_1 = (a_1 - a_2)\left[a_1 N_1 (1-\epsilon_1) + a_2 N_2 (1-\epsilon_2)\right],$$

and therefore,

$$\max\{\alpha_1, \alpha_2\} = \begin{cases} \alpha_1, & a_1 \leq a_2 \\ \alpha_2, & a_1 \geq a_2. \end{cases}$$

In particular, $\alpha_1 = \alpha_2$ if $a_1 = a_2$. This implies the following particular case of Proposition A. 7, which is especially interesting as the minimum points and values do not depend on ϵ_i ($i =$

1, 2) although the function $\overline{\mathscr{R}}_v$ and its level sets do (see Fig. 6). We also remark that this is essentially a version of Theorem 4.7 for $n = 2$, but its proof is more elementary.

Proposition A.8 *Let C be the Jacquez mixing given in* (2.5), *and let*

$$a_1 = a_2 =: a\,.$$

Then for all possible values of constants $\epsilon_i \in (0, 1)$ *and* $N_i > 0$ ($i = 1, 2$), *the critical ray coincides with the diagonal*

$$\bar{\Gamma}: q_1 = q_2,$$

and the expressions in (A.9) *simplify to*

$$Q^*(\bar{\eta}) = \frac{\bar{\eta}}{N}, \quad \overline{\mathscr{R}}_{v\{min\}}(\bar{\eta}) = \rho a \frac{\bar{\eta}}{N}, \quad 0 < \bar{\eta} < N\,. \qquad \text{(A.10)}$$

References

Adler FR (1992) The effects of averaging on the basic reproduction ratio. Math Biosci 111(1):89–98 [PubMed: 1515741]

Andersson H, Britton T (1998) Heterogeneity in epidemic models and its effect on the spread of infection. J Appl Probab 35(3):651–661

Brauer F, Castillo-Chavez C (2012) Mathematical models in population biology and epidemiology, texts in applied mathematics, 2nd edn. Springer, New York https://doi.org/10.1007/978-1-4614-1686-9

Busenberg S, Castillo-Chavez C (1991) A general solution of the problem of mixing of subpopulations and its application to risk- and age-structured epidemic models for the spread of AIDS. IMA J Math Appl Med Biol 8(1):1–29 [PubMed: 1875096]

Castillo-Chavez C, Feng Z (1998) Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. Math Biosci 151(2):135–154 [PubMed: 9711046]

Diekmann O, Heesterbeek JAP, Metz JAJ (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. https://doi.org/10.1007/BF00178324 [PubMed: 2117040]

Diekmann O, Heesterbeek H, Britton T (2012) Mathematical tools for understanding infectious disease dynamics. Princeton University Press, Princeton

Feng Z, Hill AN, Smith PJ, Glasser JW (2015) An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing. J Theor Biol 386:177–187. https://doi.org/10.1016/j.jtbi.2015.09.006 [PubMed: 26375548]

Feng Z, Hill AN, Curns AT, Glasser JW (2017) Evaluating targeted interventions via meta-population models with multi-level mixing. Math Biosci 287:93–104. https://doi.org/10.1016/j.mbs.2016.09.013 [PubMed: 27671169]

Friedland S (1980/81) Convex spectral functions. Linear Multilinear Algebra 9(4):299–316. https://doi.org/10.1080/03081088108817381

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. https://doi.org/10.1016/j.mbs.2011.10.001 [PubMed: 22037144]

Glasser JW, Feng Z, Omer SB, Smith PJ, Rodewald LE (2016) The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study. Lancet Infect Dis 16(5):599–605. https://doi.org/10.1016/S1473-3099(16)00004-9 [PubMed: 26852723]

Hadeler K, Müller J (1996a) Vaccination in age structured populations I: the reproduction number. Models Infect Human Dis Struct Relat Data 6:90

Hadeler K, Müller J (1996b) Vaccination in age structured populations II: Optimal strategies. Models for infectious human diseases: their structure and relation to data pp 102–114

Hill AN, Longini IM, Jr (2003) The critical vaccination fraction for heterogeneous epidemic models. Math Biosci 181(1):85–106. https://doi.org/10.1016/S0025-5564(02)00129-3 [PubMed: 12421553]

Jacquez JA, Simon CP, Koopman J, Sattenspiel L, Perry T (1988) Modeling and analyzing HIV transmission: the effect of contact patterns. Math Biosci 92(2):119–199. https://doi.org/10.1016/0025-5564(88)90031-4

Nold A (1980) Heterogeneity in disease-transmission modeling. Math Biosci 52(3–4):227–240. https://doi.org/10.1016/0025-5564(80)90069-3

Nussbaum RD (1986) Convexity and log convexity for the spectral radius. Linear Algebra Appl 73:59–122. https://doi.org/10.1016/0024-3795(86)90233-8

Seneta E (1973) Non-negative matrices. An introduction to theory and applications. Halsted Press, New York

van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Dedicated to the memory of John Jacquez. Math Biosci 180(1–2):29–48. https://doi.org/10.1016/S0025-5564(02)00108-6 [PubMed: 12387915]

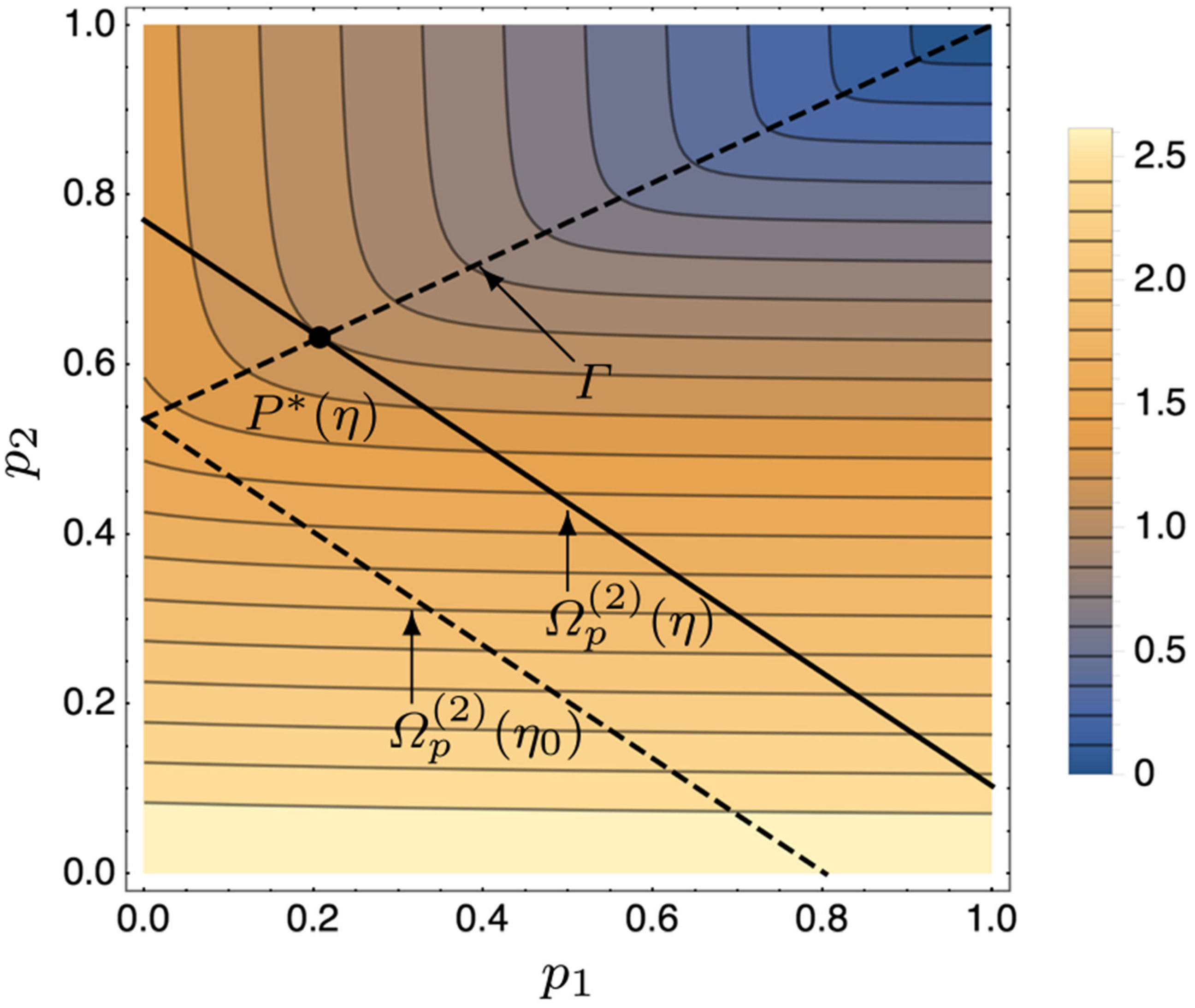

Fig. 1.

Plot of contour curves of $\mathscr{R}_v(p_1, p_2)$ in the case of $n = 2$ sub-populations and depiction of the optimal point $P^*(\eta)$. Γ is the critical ray. $\Omega_p^{(2)}(\eta)$ is the constraint line. η_0 is the greatest lower bound of $0 < \eta < N$ such that $P^*(\eta) \in (0, 1)^2$ only if $\eta_0 < \eta < N$. This figure illustrates that $P^*(\eta)$ always lies on Γ for such η

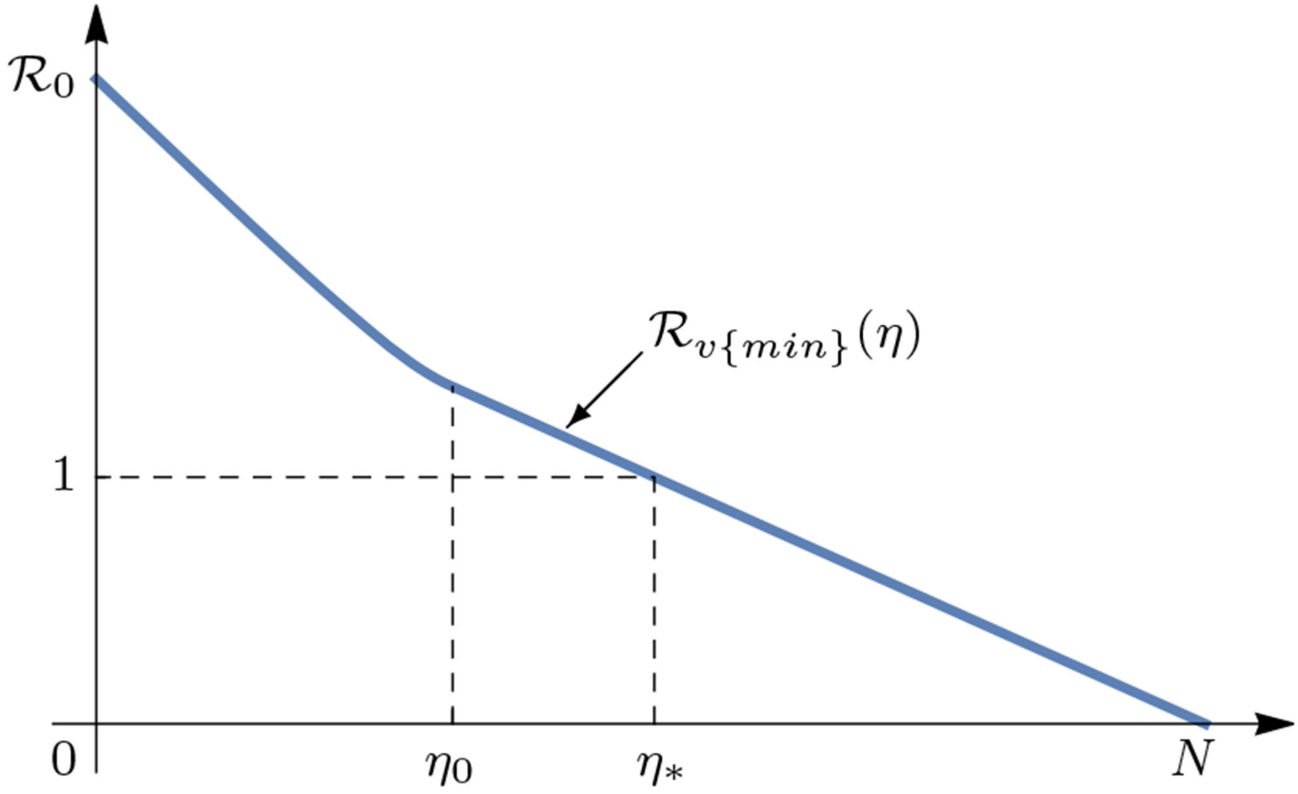

Fig. 2.

Depiction of the minimized reproduction number $\mathscr{R}_{v\{min\}}(\eta)$ as a function of η. The illustration corresponds to the case when $\mathscr{R}_{v\{min\}}(\eta_0) \geq 1$

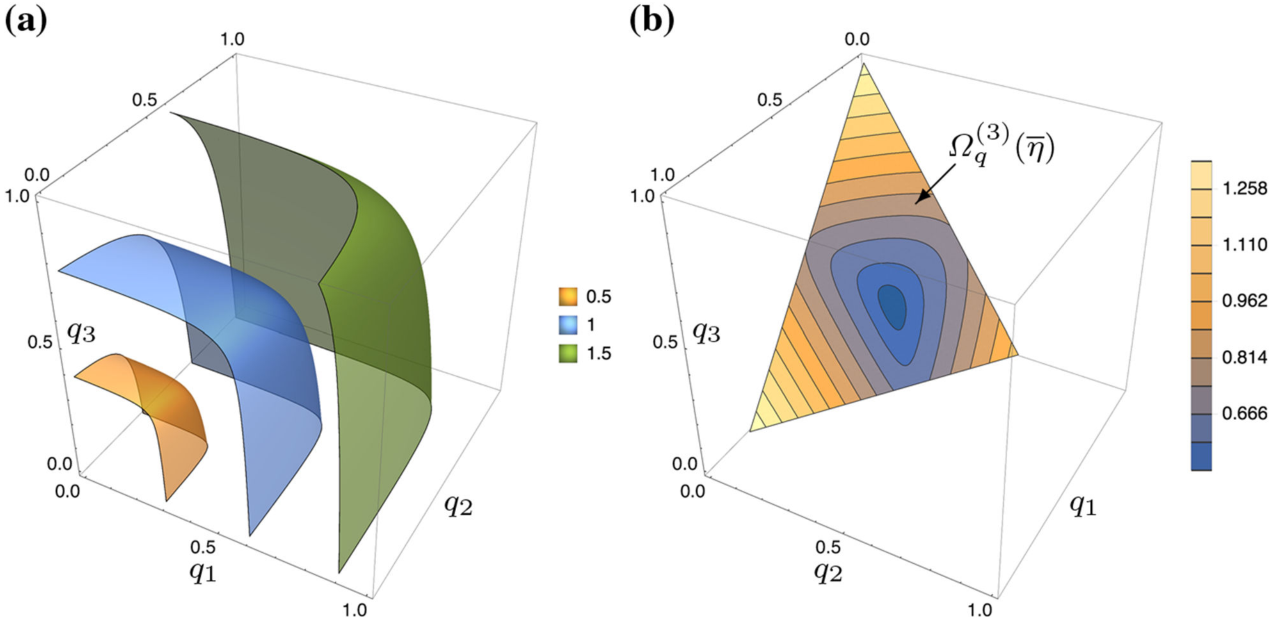

Fig. 3.

a is a contour plot of $\overline{\mathscr{R}}_v(q) = \mathscr{R}_v((1, 1, 1) - q)$ in $[0, 1]^3$. **b** illustrates the restriction of level sets to $\Omega_q^{(3)}(\bar{\eta}) \cap [0, 1]^3$

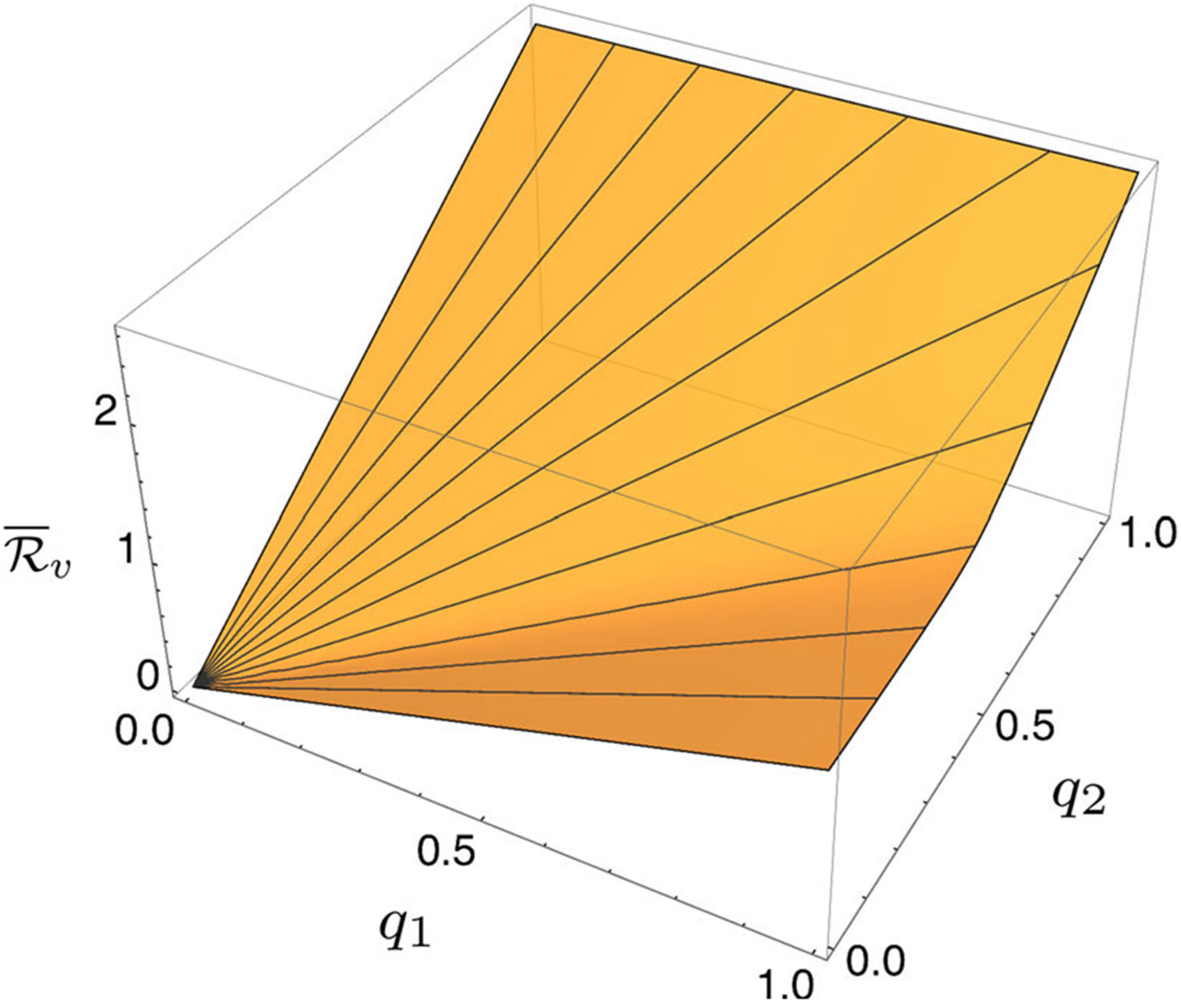

Fig. 4.

The graph of $\overline{\mathscr{R}}_v(q_1, q_2)$. $\overline{\mathscr{R}}_v$ grows linearly on the rays emanating from the origin

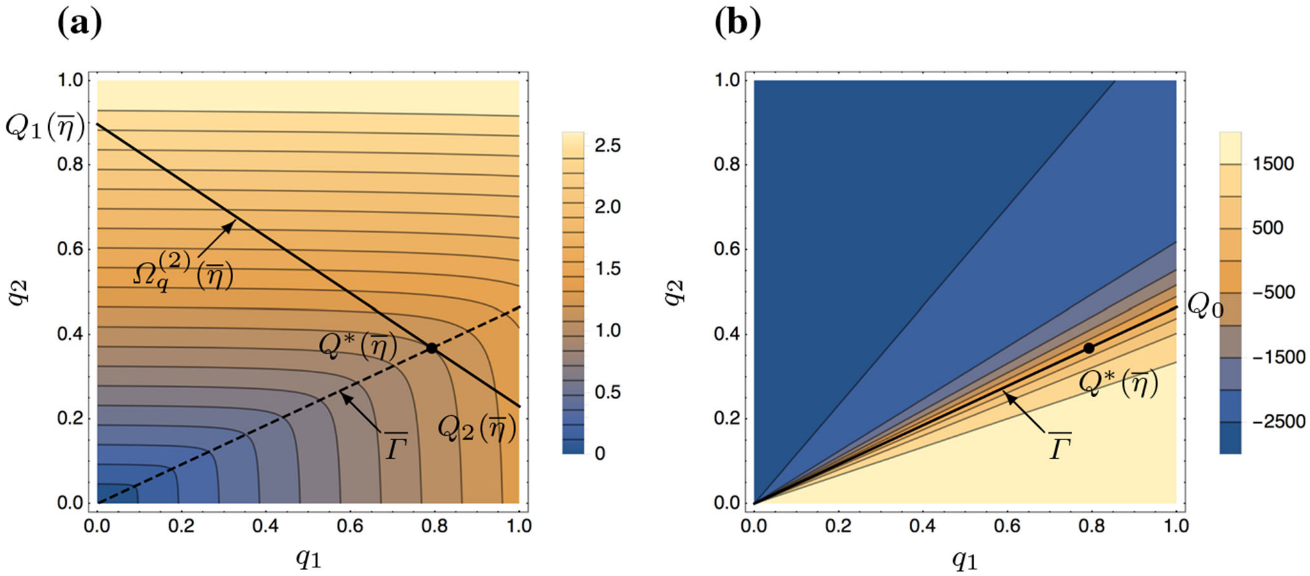

Fig. 5.

Level sets of **a** $\overline{\mathscr{R}}_v$ and **b** $\phi = \nabla\overline{\mathscr{R}}_v \cdot (N_2, -N_1)$. The point $Q^*(\bar{\eta})$ in (**b**) is the relative minimum point of $\overline{\mathscr{R}}_v$ on $\Omega_q^{(2)}(\bar{\eta}) \cap [0,1]^2$. $\bar{\Gamma}$ is the set of all such minima

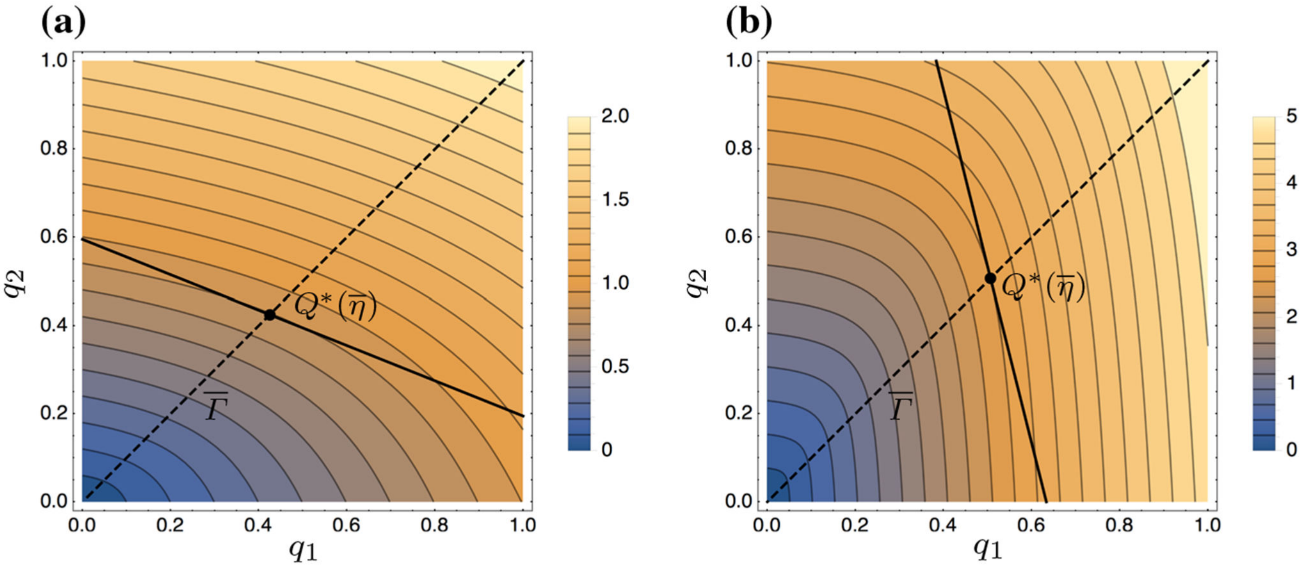

Fig. 6.

The critical ray $\bar{\Gamma}$ coincides with the main diagonal when $a_1 = a_2$ as in Proposition A.8, for any values of $\epsilon_i \in (0, 1)$, $N_i > 0$, $i = 1, 2$. This figure illustrates two cases with values a $\epsilon_1 = 0.1$, $\epsilon_2 = 0.5$, $N_1 = 400$, $N_2 = 1000$ and **b** $\epsilon_1 = 0.8$, $\epsilon_2 = 0.2$, $N_1 = 1200$, $N_2 = 300$

Table 1

Parameters and symbols with their definitions

Symbol	Description
a_i	*Per capita* contact rate of members of sub-population i
γ	*Per capita* rate of recovery
θ	*Per capita* rate of entering and exiting a sub-population
β	Probability of infection on contact
ρ	$= \beta/(\gamma + \theta)$
N_i	Size of sub-population i
N	$= N_1 + N_2 + \ldots + N_n$. Total population
c_{ij}	Proportion of contacts of individuals in group i that are with group j
C	$= (c_{ij})$. Mixing matrix
$r(A)$	Spectral radius of the matrix A
ϵ_i	Fraction of contacts of group i reserved for itself
n	Number of sub-populations in the meta-population
p_i	Proportion of sub-population i that is vaccinated
p	$= (p_1, p_2, \ldots, p_n)$
$\mathscr{R}_{0i}$	$= \rho a_i$. Basic reproduction number of sub-population i
$\mathscr{R}_{vi}$	$= \mathscr{R}_{0i}(1 - p_i)$. Effective reproduction number of sub-population i
$\mathscr{R}_0$	Meta-population basic reproduction number
$K_v(p)$	$\mathrm{diag}(\mathscr{R}_{v1}, \ldots, \mathscr{R}vn)C$. Next generation matrix (NGM)
$\mathscr{R}_v(p)$	$= r(K_v(p))$. Meta-population effective reproduction number
η	$= \sum_{i=1}^{n} p_i N_i$. Total number of vaccine doses
$\Omega_p^{(n)}(\eta)$	$= \{(p_1, p_2, \ldots, p_n) : \ell(p) = \eta\}$
$P^*(\eta)$	$= (p_1^*(\eta), p_2^*(\eta), \ldots, p_n^*(\eta))$. Optimal allocation of vaccine
$\mathscr{R}_{v\{min\}}(\eta)$	Minimum of $\mathscr{R}_v$ for a given number of vaccine doses η
η_*	Minimum doses for achieving $\mathscr{R}_v \leq 1$
η_o	Infimum of all $\eta \in (0, N)$ such that $P^*(\eta) \in (0, 1)^n$
q_i	$= 1 - p_i$
q	$= (q_1, q_2, \ldots, q_n)$
$\bar{\eta}$	$= \sum_{i=1}^{n} q_i N_i = N - \eta$
$\Omega_q^{(n)}(\bar{\eta})$	$= \{(q_1, q_2, \ldots, q_n): \ell(q) = \bar{\eta}\}$
$Q^*(\bar{\eta})$	$= (1, 1, \ldots, 1) - p^*(\eta)$
$\overline{\mathscr{R}}_v(q)$	$= \mathscr{R}_v(1 - q_1, 1 - q_2, \ldots, 1 - q_n)$

Symbol	Description
$\overline{\mathscr{R}}_{v\{min\}}(\bar{\eta})$	$= \mathscr{R}_{v\{min\}}(\eta)$
$\Gamma, \bar{\Gamma}$	Critical and ‘reflected’ critical rays (see Theorems 3.2 and 4.3)
$\boldsymbol{\pi}_i$	$= N_i/N$. Fraction of population belong to sub-population i
S	$= \sum_{i=1}^{n}(1-p_i)\pi_i$. Fraction of the population that is unvaccinated
$\mathscr{R}_0^{\diamond}$	$= (\sum_{i=1}^{n}\pi_i \Big/ \mathscr{R}_{0i})^{-1}$. Harmonic mean of $\mathscr{R}_{0i}$ weighted by $\boldsymbol{\pi}_i$
$\hat{R}_0$	$= \sum_{i=1}^{n}\mathscr{R}_{0i}\pi_i \geq \mathscr{R}_0^{\diamond}$. Population weighted reproduction number
$\widetilde{\mathscr{R}}_0$	$= \min_i \mathscr{R}_{0i}^2 \Big/ \widehat{\mathscr{R}}_0$. Analogous to a scaled reproduction number

$i, j = 1, 2, \dots, n$.