HERD IMMUNITY UNDER INDIVIDUAL VARIATION AND REINFECTION

ANTONIO MONTALBÁN, RODRIGO M. CORDER, AND M. GABRIELA M. GOMES

Abstract. We study a SEIR model considered by Gomes et al. [7] and Aguas et al. [2] where different individuals are assumed to have different levels of susceptibility or exposure to infection. Under this heterogeneity assumption, epidemic growth is effectively suppressed when the percentage of population having acquired immunity surpasses a critical level - the herd immunity threshold - that is lower than in homogeneous populations. We find explicit formulas to calculate herd immunity thresholds and stable configuration, and explore extensions of the model.

1. Introduction

We analyse a SEIR (Susceptible-Exposed-Infectious-Recovered) model considered in [7] where each of the compartments S, E, I, and R are split into continuum many compartments $S(x)$, $E(x)$, $I(x)$, and $R(x)$ for $x \in \mathbb{R}^+$. We are modeling a situation where each individual has a *level of susceptibility* x. This individual will start in compartment $S(x)$ and stay within the compartments $S(x)$, $E(x)$, $I(x)$, and $R(x)$ the whole time. They may infect or be infected by individuals in other compartments. We will consider two types of models:

In the **variable susceptibility case** the susceptibility of an individual at level x is proportional to x, or, in other words, if you compare an individual at level x and an individual at level y, the one at level x is $\frac{x}{y}$ times more likely to get infected than the one with susceptibility y. One may interpret this as variation in biological susceptibility which may be due to genetics, epigenetics or life history.

In the **variable connectivity case** the propensities for an individual at level x to acquire infection and transmit to others are both proportional to x, or, in other words, if you compare an individual at level x and an individual at level y, the one at level x is $\frac{x}{y}$ times more likely to get infected than the one in level y and also $\frac{x}{y}$ times more likely to infect someone else once infected. One may interpret this as variation due to the connectivity, i.e., individuals that have many contacts are both more likely to get infected and to infect others.

For each x, we have a system of the form:

$$\boxed{\text{S(x)}} \xrightarrow{x\lambda} \boxed{\text{E(x)}} \xrightarrow{\delta} \boxed{\text{I(x)}} \xrightarrow{\gamma} \boxed{\text{R(x)}}$$

where λ is the force of infection which is formulated differently in the *variable susceptibility* or the *variable connectivity* cases:

$$\text{Variable susceptibility:} \qquad \lambda = \beta \int I(x)\, dx,$$

$$\text{Variable connectivity:} \qquad \lambda = \beta \int x\, I(x)\, dx.$$

[0]Saved: June 28, 2020
Compiled: October 27, 2020

The system is given by the equations:

$$\dot{S}(x) = -x\lambda S(x), \tag{1}$$

$$\dot{E}(x) = x\lambda S(x) - \delta E(x), \tag{2}$$

$$\dot{I}(x) = \delta E(x) - \gamma I(x), \tag{3}$$

$$\dot{R}(x) = \gamma I(x). \tag{4}$$

We assume that the system has been scaled such that the total population is 1. The initial condition satisfies $S_0(x) = (1-\epsilon)q(x)$, $E_0(x) = \epsilon q(x)$ and $I_0(x) = R_0(x) = 0$, where $q(x)$ is a distribution with mean 1 and coefficient of variation CV, and $0 < \epsilon \ll 1$ is a small scalar to seed the epidemic.

We use subindex t to denote the time at which we are considering these compartments: $S_t(x)$, $E_t(x)$, $I_t(x)$ and $R_t(x)$. We use S_t to denote the conjunction of all the compartments $S_t(x)$ for $x \in \mathbb{R}^+$. We thus have $S_t = \int_0^{+\infty} S_t(x)dx$. Same with E_t, I_t and R_t.

We will use the first three momenta of $S_t(x)$, that we denote S_t, $\bar{S}_t$ and $\overline{\overline{S}}_t$:

$$S_t = \int S_t(x)dx, \qquad \bar{S}_t = \int xS_t(x)dx, \quad \text{and} \quad \overline{\overline{S}}_t = \int x^2 S_t(x)dx.$$

When infection is absent ($\epsilon = 0$), we have $S_0 = 1$, $\bar{S}_0 = 1$ and $\overline{\overline{S}}_0 = 1 + CV^2$. But note that $S_t(x)$ is not a probability density function for $\epsilon > 0$ as S_t becomes less than 1. The quotient $S_t(x)/S_t$ will be a probability distribution for $\epsilon > 0$ and all t with first and second momenta $\bar{S}_t/S_t$ and $\overline{\overline{S}}_t/S_t$ which decrease as time passes. We will see that in the case where the initial configuration $q(x)$ is a gamma distribution, all the distributions $S_t(x)/S_t$ will be also be gamma distributions with the same coefficient of variation CV but with lower mean.

Similarly, we define the moments R_t, $\bar{R}_t$ and $\overline{\overline{R}}_t$ for the recovered compartment, and the same with E and I. Notice for instance that λ_t is equal to βI_t and to $\beta \bar{I}_t$ in the variable susceptibility and variable connectivity cases respectively.

Here we describe key epidemiological quantities when system of equations (1)-(4) is adopted. The basic reproductive number $\mathcal{R}_0$ is the number of new infections expected to be caused by an infected individual in a totally susceptible population. It depends on characteristics of both the pathogen and the population. When this number is above 1, the introduction of infection in a virgin population is expected to generate an epidemic. This is followed by almost exponential growth in cumulative infections which decelerates gradually as susceptibles are depleted. The effective reproduction number $\mathcal{R}_{\text{eff}}$ is a time-dependent quantity that defines the average number of new infections caused by an infected individual at time t. $\mathcal{R}_{\text{eff}}$ coincides with $\mathcal{R}_0$ at the beginning of an epidemic but declines as individuals are removed from the susceptible subpopulation by infection and immunity. When $\mathcal{R}_{\text{eff}}$ crosses 1 towards lower values, we say that herd immunity $\mathcal{H}$ has been attained.

We obtain precise formulas for the effective reproductive number $\mathcal{R}_{\text{eff}}$ and the herd immunity threshold $\mathcal{H}$ in terms of the momenta of the susceptible population. We will see that in the case when $q(x)$ is a gamma distribution, we can calculate this momenta in terms of S_t and get an exact formula for $\mathcal{H}$ in terms only of the basic reproductive number $\mathcal{R}_0$ and CV.

In the **variable susceptibility** case we will get that

$$\mathcal{R}_0 = \frac{\beta}{\gamma} \quad \text{and} \quad \mathcal{R}_{\text{eff}} = \bar{S}_t \, \frac{\beta}{\gamma}.$$

This implies that herd immunity is achieved when $\bar{S}_t < 1/\mathcal{R}_0$. If we assume that $q(x)$ is a gamma distribution the proportion of individuals that have been infected by the time herd

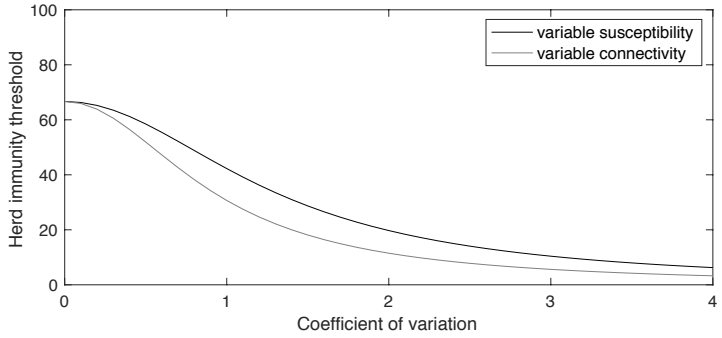


FIGURE 1. Herd immunity threshold. Curves generated using formulas (5) for variable susceptibility and (6) for variable connectivity, with $\mathcal{R}_0 = 3$.

immunity is reached is

$$\mathcal{H} = 1 - \mathcal{R}_0^{\frac{-1}{1+CV^2}}. \tag{5}$$

In the **variable connectivity** case we will get that

$$\mathcal{R}_0 = (1 + CV^2)\,\frac{\beta}{\gamma} \quad \text{and} \quad \mathcal{R}_{\text{eff}} = \overline{\overline{S}}_t\,\frac{\beta}{\gamma}.$$

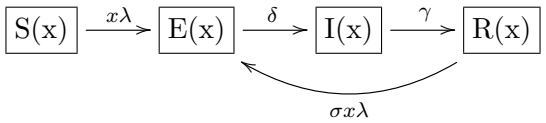
This implies that herd immunity is achieved when $\overline{\overline{S}}_t < (1 + CV^2)/\mathcal{R}_0$. If we assume that $q(x)$ is a gamma distribution the proportion of individuals that have been infected by the time herd immunity is reached is

$$\mathcal{H} = 1 - \mathcal{R}_0^{\frac{-1}{1+2CV^2}}. \tag{6}$$

We call this model, *the basic model*, and provide graphical representations for the corresponding $\mathcal{H}$ formulas in Figure 1 (variable susceptibility in black and variable connectivity in grey), showing a monotonic decrease as CV increases ([7], [4], [12]). We will see now a few variations.

1.1. **The model with reinfections.** In Aguas et al. [2], the authors consider an extension of the model where immunity after recovery is not fully protective, but only partially. A factor σ is added to represent the quotient of the probability of getting reinfected after recovery over the probability of getting infected while fully susceptible.

The model now looks like this:



with λ as above. The extended system is given by the equations:

$$\dot{S}(x) = -x\lambda S(x), \tag{7}$$

$$\dot{E}(x) = x\lambda(S(x) + \sigma R(x)) - \delta E(x), \tag{8}$$

$$\dot{I}(x) = \delta E(x) - \gamma I(x), \tag{9}$$

$$\dot{R}(x) = \gamma I(x) - \sigma x\lambda R(x). \tag{10}$$

The basic reproductive number is calculated exactly as in the no-reinfections case, but the effective reproductive now depends not only on the distribution of $S_t(x)$, but also on the distribution of $\mathcal{R}_{\text{eff}}(x)$. We will see that

- $\mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma\bar{R}_t) \cdot \beta/\gamma$ in the variable susceptibility case, and
- $\mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma\overline{\overline{R}}_t) \cdot \beta/\gamma$ in the variable connectivity case.

The system exhibits newer dynamics in comparison with the basic case. Depending on whether σ is below or above $1/\mathcal{R}_0$ (known as the reinfection threshold [9, 8]) we get that either the disease dies out after a while and a certain proportion of the population never gets infected, or continues endemically and every individual is eventually infected and then reinfected over and over again. Pursuing these considerations on longer time scales, however, would require a further extension to the model to account for rates of birth and mortality due to general causes. This is beyond the scope of the present article.

1.2. **A variation with a carrier stage.** In the original Gomes et al. model [7], the exposed compartments are not simply a latent stage but a carrier stage where individuals are infectious but to a lesser degree than individual in the infectious compartments. What changes is the force of infection λ:

$$\text{Variable susceptibility:} \qquad \lambda_t = \beta \int \rho E_t(x) + I_t(x)dx = \beta(\rho E_t + I_t),$$

$$\text{Variable connectivity:} \qquad \lambda_t = \beta \int x(\rho E_t(x) + I_t(x))dx = \beta(\rho \bar{E}_t + \bar{I}_t).$$

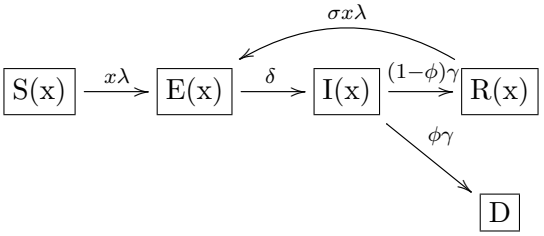
We will then get that the basic and effective reproductive numbers are as follows:

$$\text{Variable susceptibility:} \quad \mathcal{R}_0 = \beta\left(\frac{\rho}{\delta} + \frac{1}{\gamma}\right) \quad \text{and} \quad \mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma\bar{R}_t)\beta\left(\frac{\rho}{\delta} + \frac{1}{\gamma}\right),$$

$$\text{Variable connectivity:} \quad \mathcal{R}_0 = (1 + CV^2)\beta\left(\frac{\rho}{\delta} + \frac{1}{\gamma}\right) \quad \text{and} \quad \mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma\overline{\overline{R}}_t)\beta\left(\frac{\rho}{\delta} + \frac{1}{\gamma}\right).$$

In analogy with the previous models, we still get $\mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma\bar{R}_t) \cdot \mathcal{R}_0$ and $\mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma\overline{\overline{R}}_t) \cdot \mathcal{R}_0/(1 + CV^2)$ respectively. The formulas for the herd immunity threshold in terms of $\mathcal{R}_0$ remain the same.

1.3. **A variation with a death rate.** Aguas et al. [2] have one more feature: a death rate ϕ meaning that a proportion ϕ of the individuals that come out of $I(x)$ go to a new compartment D. The model now looks like this:



We will not analyse this model in detail. The inclusion of death causes the total population to decline over time. This should typically be balanced by births, requiring a more elaborate model. Let us just say that if the reinfection parameter σ is below the reinfection threshold $1/\mathcal{R}_0$, the infection goes extinct slightly faster than before, and if the reinfection parameter is above the reinfection threshold, everybody eventually dies due to the infection according to this model.

2. Effective reproductive number in the basic model

Let us start by studying the basic model with no reinfections, no carrier stage, and no death rate.

Definition 2.1. The effective reproductive number $\mathcal{R}_{\text{eff}}$ at time t is defined to be the average number of secondary infections caused by an infected individual over their entire infectious period.

We make an approximation by assuming that while the individual is contagious, the susceptible population does not change. That is, we disregard the fact that since the susceptible population declines, this individual infects more people at the beginning than at the end of their infection. The decline in the susceptible population is usually slow enough compared to the length of the infectious period so that this does not make a big difference. Formally, we will use $\mathcal{R}_{\text{eff}}$ to analyze stable configurations, where this assumption holds, so our results will be precise.

Here we provide the derivations of the formulas presented in the introduction.

Let's look first at the **variable susceptibility** case: Consider an individual who gets infected at time t (i.e., moving from S to E at time t). They will eventually move to I where they will then spend, on average $1/\gamma$ days. While in I, they will infect $\beta \int y S_t(y) dy$ other individuals each day. We thus get

$$\mathcal{R}_{\text{eff}} = \frac{1}{\gamma}\beta\left(\int y(S_t(y))dy\right) = \bar{S}_t \cdot \frac{\beta}{\gamma}.$$

In particular, we get that $\mathcal{R}_0 = \frac{\beta}{\gamma}$ and that $\mathcal{R}_{\text{eff}} = \bar{S}_t \cdot \mathcal{R}_0$.

Let's look now at the **variable connectivity** case: Consider an individual who gets infected at time t (i.e., that enters E at time t). It matters now what kind of individual they are, i.e., what connectivity level they have, because individuals at different levels will infect different numbers of people.

Let $p(x)$ be the probability distribution function measuring the probability that this individual has connectivity level x . Their probability of becoming infected (i.e., of entering the $E(x)$ compartment) is $x\lambda$. Thus, the value of $p(x)$ is proportional to $xS_t(x)$. We get

$$p(x) = \frac{xS_t(x)}{\int yS_t(y)dy} = x\frac{S_t(x)}{\bar{S}_t}.$$

As above, an individual who enters E will eventually move to I where they are then going to spend, on average, $1/\gamma$ days, and where they will infect $\beta \int yS_t(y)\ dy$ other individuals each day. We thus get

$$\begin{aligned}\mathcal{R}_{\text{eff}} &= \int \Big(x\ p(x)\Big)\Big(\int yS_t(y)dy\Big)\ \beta/\gamma\ dx \\ &= \int \Big(\frac{x^2\ S_t(x)}{\bar{S}_t}\Big) \qquad \bar{S}_t \qquad \beta/\gamma\ dx.\end{aligned}$$

Moving things around we get

$$\mathcal{R}_{\text{eff}} = \overline{\overline{S}}_t \cdot \beta/\gamma.$$

In particular we get that $\mathcal{R}_0 = \overline{\overline{q}} \cdot \beta/\gamma = (1 + CV^2) \cdot \beta/\gamma$ and that $\mathcal{R}_{\text{eff}} = \overline{\overline{S}}_t \cdot \mathcal{R}_0/(1 + CV^2)$.

3. Herd immunity

Suppose we have a population with no infected individuals, so that all individuals are either susceptible or recovered. We say that this population *has herd immunity* if an introduction of the disease (i.e., a tiny increase in E) does not trigger an epidemic. Mathematically, this means that any small enough deviation from the configuration with no infected individuals will quickly converge back to a configuration with no infected individuals. More formally, a configuration with $E(x) = I(x) = 0$ *has herd immunity* if, for every $\epsilon > 0$, there is a $\delta > 0$ such that if we modify the configuration to one that is at a distance less than δ from the original configuration, the system then converges to a configuration that is at a distance less than ϵ.[1]

If we visualize the dynamical system as modeling individuals who move between the compartments $S(x)$, $E(x)$, $I(x)$ and $R(x)$, it is not hard to see that a configuration with $E(x) = I(x) = 0$ has herd immunity if and only if

$$\mathcal{R}_{\text{eff}} \leq 1.$$

We say that a configuration with no infected individuals is *at the herd immunity threshold* if $\mathcal{R}_{\text{eff}} = 1$. Usually, in SEIR models with no variability, configurations are determined by the value of $1 - S = R$, which is the number of people in the recovered compartment, and the herd immunity threshold is defined as the value of $1 - S$, the unique configuration that is at the herd immunity threshold, a value that is well-know to be equal to $1 - 1/\mathcal{R}_0$ [3]. But, for the heterogenous models, there are many configurations which are at the herd immunity threshold. One such configuration is given by $S(x) = q(x)/\mathcal{R}_0$ for all x. This would be obtained, for instance, if one vaccinates a proportion $1 - 1/\mathcal{R}_0$ of the total population randomly without taking into account susceptibility levels. When immunity is acquired naturally, however, individuals with higher susceptibility get infected earlier and a configuration which has herd immunity is reached before a proportion $1 - 1/\mathcal{R}_0$ of the total population is infected.

For now, let us observe that in the **variable susceptibility** case, herd immunity is achieved when $\bar{S}_t = 1/\mathcal{R}_0$, and that in the **variable connectivity** case, herd immunity is reached when $\overline{\overline{S}}_t = (1 + CV^2)/\mathcal{R}_0$

Next, we will see how we can calculate $\bar{S}_t$ and $\overline{\overline{S}}_t$ under the assumption that the original distribution is a gamma distribution.

4. Starting with the gamma distribution

Let us start this section studying how the distribution of susceptible compartments evolves, and then see how this fits nicely in the case where individual variation in susceptibility or connectivity is gamma distributed.

4.1. The evolution of the susceptible compartments.

We claim that at every t, there is a number $k_t \in \mathbb{R}$ that only depends on t, such that

$$S_t(x) = q(x) \cdot e^{-x \cdot k_t}. \tag{11}$$

This holds in all models: in both the variable susceptibility case and the variable connectivity case (with different values for k_t), in the case with reinfection, with carrier stage, and with a death rate.

[1] For a distance in the space of configurations, one may use the L_1 distance, i.e., the sum of the integrals of the differences of the compartments $\int |S(x) - S'(x)| + |E(x) - E'(x)| + |I(x) - I'(x)| + |R(x) - R'(x)|dx$.

Let's start by proving equation (11). From the SEIR equation for $\frac{\partial S_t(x)}{\partial t}$ we get

$$\frac{1}{S_t(x)}\frac{\partial S_t(x)}{\partial t} = x \cdot \lambda_t.$$

Integrate with respect to t:

$$\int_0^t \frac{1}{S_t(x)}\frac{\partial S_t(x)}{\partial t} dt = \int_0^t x \cdot \lambda_t \; dt.$$

Then by substitution

$$\int_0^t \frac{1}{S_t(x)} dS_t(x) = x \cdot \int_0^t \lambda_t \; dt.$$

Evaluating the integrals and letting $k_t = -\int_0^t \lambda_t dt$:

$$\ln(S_t(x)) - \ln(q(x)) = -xk_t$$

from which (11) follows.

4.2. **The gamma distribution.** We use the following notation for the *gamma probability density function*:

$$\text{Gamma}_{a,b}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}.$$

The gamma distribution has mean a/b and coefficient of variation $CV = 1/\sqrt{a}$.

Since our initial distribution $q(x)$ has mean 1, we are using a gamma distribution with $a = b$, i.e., $q(x) = \text{Gamma}_{a,a}(x)$.

Substituting $q(x)$ for $\text{Gamma}_{a,a}(x)$ in (11) we get

$$\begin{aligned} S_t(x) &= \frac{a^a}{\Gamma(a)} x^{a-1} e^{-x(a+k_t)} \\ &= \left(\frac{a}{a+k_t}\right)^a \frac{(a+k_t)^a}{\Gamma(a)} x^{a-1} e^{-x(a+k_t)} \end{aligned}$$

$$\boxed{S_t(x) = \left(\frac{a}{a+k_t}\right)^a \cdot \text{Gamma}_{a,a+k_t}(x)}$$

Using that

$$\begin{aligned} \int \text{Gamma}_{a,a+k_t}(x) dx &= 1 \\ \int x \, \text{Gamma}_{a,a+k_t}(x) dx &= \frac{a}{a+k_t} \\ \int x^2 \, \text{Gamma}_{a,a+k_t}(x) dx &= \frac{a(a+1)}{(a+k_t)^2}, \end{aligned}$$

we calculate S_t, $\bar{S}_t$ and $\overline{\overline{S}}_t$:

$$S_t = \left(\frac{a}{a+k_t}\right)^a \cdot \int \text{Gamma}_{a,a+k_t}(x) dx = \left(\frac{a}{a+k_t}\right)^a$$

$$\bar{S}_t = \left(\frac{a}{a+k_t}\right)^a \cdot \int x \, \text{Gamma}_{a,a+k_t}(x) dx = \left(\frac{a}{a+k_t}\right)^a \cdot \frac{a}{a+k_t} = \left(\frac{a}{a+k_t}\right)^{a+1}$$

$$\overline{\overline{S}}_t = \left(\frac{a}{a+k_t}\right)^a \cdot \int x^2 \, \text{Gamma}_{a,a+k_t}(x) dx = \left(\frac{a}{a+k_t}\right)^a \cdot \frac{a(a+1)}{(a+k_t)^2} = \left(\frac{a}{a+k_t}\right)^{a+2} \frac{a+1}{a}$$

From the above we get

$$\bar{S}_t = (S_t)^{\frac{a+1}{a}} \quad \text{and} \quad \overline{\overline{S}}_t = (S_t)^{\frac{a+2}{a}} \cdot \frac{a+1}{a}.$$

Using that $a = 1/CV^2$, we get $(a+1)/a = 1 + CV^2$, and we can rewrite these as:

$$\bar{S}_t = (S_t)^{1+CV^2} \quad \text{and} \quad \overline{\overline{S}}_t = (S_t)^{1+2CV^2} \cdot (1 + CV^2).$$

4.3. **Putting all together.** Recall that the herd immunity threshold $\mathcal{H}$ is $1 - S_t$ at the time t when $\mathcal{R}_{\text{eff}} = 1$ in the case where there are almost no people infected.

In the **variable susceptibility** case we had that herd immunity is achieved when $\mathcal{R}_{\text{eff}} = \bar{S}_t \cdot \mathcal{R}_0$. Thus, when $\mathcal{R}_{\text{eff}} = 1$, we have $\mathcal{R}_0^{-1} = \bar{S}_t = (S_t)^{1+CV^2}$. It follows that

$$\mathcal{H} = 1 - \mathcal{R}_0^{\frac{-1}{1+CV^2}}.$$

In the **variable connectivity** case we had that herd immunity is achieved when $\mathcal{R}_{\text{eff}} = \overline{\overline{S}}_t \mathcal{R}_0/(1 + CV^2)$. Thus, when $\mathcal{R}_{\text{eff}} = 1$, we get $\mathcal{R}_0^{-1} = \overline{\overline{S}}_t/(1 + CV^2) = (S_t)^{1+2CV^2}$. It follows that

$$\mathcal{H} = 1 - \mathcal{R}_0^{\frac{-1}{1+2CV^2}}.$$

5. The model with reinfection

5.1. **Effective reproductive number.** Let's consider now the model with a reinfection factor σ as described in the introduction. We will see that $\mathcal{R}_{\text{eff}}$ depends not only on $S_t(x)$ but also on $R_t(x)$. When we consider configurations with no infected individuals, we will have that $R_t(x) = q(x) - S_t(x)$ and we will be able to express $\mathcal{R}_{\text{eff}}$ in terms of $S_t(x)$ only.

The formulas for the effective reproductive number $\mathcal{R}_{\text{eff}}$ at time t are

- $\mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma \bar{R}_t) \cdot \beta/\gamma$ in the variable susceptibility case, and
- $\mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma \overline{\overline{R}}_t) \cdot \beta/\gamma$ in the variable connectivity case.

In particular, we get $\mathcal{R}_0 = \beta/\gamma$ and $\mathcal{R}_0 = (1 + CV^2) \cdot \beta/\gamma$ as in the case with no reinfection.

The derivation of these formulas is essentially the same as the derivations in Section 2. There are two differences: One, that each level-x individual infects $\beta \cdot (\bar{S}_t + \sigma \bar{R}_t)$ or $x\beta \cdot (\bar{S}_t + \sigma \bar{R}_t)$ other people in each day spent in I, in the respective cases, instead of $\beta \cdot (\bar{S}_t)$ or $x\beta \cdot (\bar{S}_t)$ Two, that when we consider an individual that gets infected in the variable connectivity case, the probability that they are a level-x individual is proportional to $xS(x) + \sigma x R(x)$ instead of $xS(x)$.

5.2. **Herd immunity.** Recall that a configuration with no infected individuals has herd immunity if and only if $\mathcal{R}_{\text{eff}} \leq 1$. Assuming that no one is infected, that is $R(x) = q(x) - S(x)$, we get $\bar{R} = 1 - \bar{S}$ and $\overline{\overline{R}} = \overline{\overline{q}} - \overline{\overline{S}}$. We can then understand the configurations at herd immunity in terms of $\bar{S}$ and $\overline{\overline{S}}$.

In the **variable susceptibility** case a configuration with no infected individuals is at the herd immunity threshold if and only if $1/\mathcal{R}_0 = \bar{S} + \sigma(1 - \bar{S})$, and hence, if and only if

$$\bar{S} = \frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma}.$$

In the **variable connectivity** case a configuration with no infected individuals is at the herd immunity threshold if and only if $1/\mathcal{R}_0 = (\overline{\overline{S}} + \sigma(\overline{\overline{q}} - \overline{\overline{S}}))/(1 + CV^2)$, and hence, if and only if

$$\overline{\overline{S}} = (1 + CV^2)\frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma}.$$

A point worth mentioning is that the behavior of herd immunity differ from the basic model. When one runs the basic model, herd immunity is achieved when the number of infections start to decline. In the case with reinfections, the number of infections may decline earlier. The reason is that when E_t and I_t are non-zero, the individuals in those compartments are not susceptible to get infected at that moment, but will become partially susceptible once they recover. In other words, we may have $\mathcal{R}_{\text{eff}} < 1$ but if we move all those infected people instantly to R_t, we could create a situation where $\mathcal{R}_{\text{eff}} > 1$. In practical terms, imagine strong isolation measures are imposed at that moment and all those infected people recover without infecting anyone else. Then, when the measures are lifted, since R_t becomes larger, $\mathcal{R}_{\text{eff}}$ is larger and any reappearance of the disease my trigger a small epidemic. Thus, we need the condition $\mathcal{R}_{\text{eff}} \leq 1$ evaluated using $R_t(x) = q(x) - S_t(x)$ to ensure that new introduction will not trigger epidemics.

5.3. **The reinfection threshold.** For the formulas above to make sense, it is necessary that we have

$$\sigma < \mathcal{R}_0^{-1}.$$

That is, the *reinfection factor* σ has to be **below** $\mathcal{R}_0^{-1}$, a critical value known as the *reinfection threshold* [9, 8]. If this is verified, then all configuration with no infected individuals and satisfying the conditions above (either $\bar{S} = (\mathcal{R}_0^{-1} - \sigma)/(1 - \sigma)$ or $\overline{\overline{S}} = (1 + CV^2)(\mathcal{R}_0^{-1} - \sigma)/(1 - \sigma)$ depending on the case) are configurations which have herd immunity in the sense that any increase of the infected compartments quickly extinguishes as $\mathcal{R}_{\text{eff}}$ won't go above 1 again.

If the reinfection factor is **above** the reinfection threshold, $\mathcal{R}_{\text{eff}}$ will be greater than 1 in any such configurations, so there won't be any configuration with no infected individuals and with herd immunity. This implies that there will always be a portion of the population infected, and hence that the population of susceptible individuals will eventually be completely depleted. The equilibrium configuration will now have $S = 0$.

Suppose we have an equilibrium configuration when $\sigma > \mathcal{R}_0^{-1}$. We will then have $S = 0$, and hence $E(x) + I(x) + R(x) = q(x)$ for all x. Setting $\dot{I}(x)$ to zero, we get that $E(x) = I(x) \cdot \gamma/\delta$ and from this we derive that $I(x) = (q(x) - R(x)) \cdot \delta/(\delta + \gamma)$ and $E(x) = (q(x) - R(x)) \cdot \gamma/(\delta + \gamma)$. Using that $\mathcal{R}_{\text{eff}} = 1$ at any equilibrium configuration, we get that either $\bar{R} = \frac{1}{\sigma \mathcal{R}_0}$ or $\overline{\overline{R}} = \frac{1}{\sigma \mathcal{R}_0}$.

5.4. **Starting with a gamma distribution.** Recall from Section 3 that if we start with a gamma distribution for $q(x)$, we get that $S_t(x)/S_t$ remains a gamma distribution for all t, and that $\bar{S}_t = S_t^{1+CV^2}$ and $\overline{\overline{S}}_t = S_t^{1+2CV^2}(1 + CV^2)$. We can then obtain the values of S_t at the moment when herd immunity is achieved, and then calculate $\mathcal{H} = 1 - S_t$.

In the **variable susceptibility** case we get:

$$\mathcal{H} = 1 - \left(\frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma}\right)^{\frac{1}{1+CV^2}}.$$

In the **variable connectivity** case we get:

$$\mathcal{H} = 1 - \left(\frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma}\right)^{\frac{1}{1+2CV^2}}.$$

Curves assuming a selection of values for σ are represented graphically in Figure 2 alongside typical epidemic curves. Note the critical behaviour at the reinfection threshold ($\sigma = 1/Ro$) in red, which separates the regime where individual immunity is sufficiently potent for herd immunity to be achievable from the regime where endemicity will establish without natural immunity.

6. The model with a carrier stage

Recall that, in this model, individuals in the E compartments are infectious, but by a factor ρ with respect to the individuals in I. This model is sometimes used when one wants to differentiate pre-symptomatic form symptomatic stages, which is usually closer to what the real word data represents, and, at the same time, allow for pre-symptomatic transmission.

The calculation of the effective reproductive number $\mathcal{R}_{\text{eff}}$ is slightly different. The difference is that now one has to add the time an individual is in E to the infectious period, multiplied by the factor ρ. The average time an individual spends in E is $1/\delta$. We then get

- $\mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma \bar{R}_t) \cdot \beta(\rho/\delta + 1/\gamma)$ in the variable susceptibility case, and
- $\mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma \overline{\overline{R}}_t) \cdot \beta(\rho/\delta + 1/\gamma)$ in the variable connectivity case,

where $\bar{S}_t$ and $\overline{\overline{S}}_t$ are the momenta of $S_t(x)$ defined above, and the same with R_t.

In particular, we get $\mathcal{R}_0 = \beta(\rho/\delta + 1/\gamma)$ and $\mathcal{R}_0 = (1 + CV^2) \cdot \beta(\rho/\delta + 1/\gamma)$ respectively as in the previous case. Also as in the previous cases we get $\mathcal{R}_{\text{eff}} = (\bar{S}_t + \sigma \bar{R}_t) \cdot \mathcal{R}_0$ and $\mathcal{R}_{\text{eff}} = (\overline{\overline{S}}_t + \sigma \overline{\overline{R}}_t) \cdot \mathcal{R}_0/(1 + CV^2)$. We then get the same formulas for the herd immunity threshold in terms of $\mathcal{R}_0$ and CV as before.

7. Discussion

The concept of herd immunity was developed in the context of vaccination programs ([10],[6]). Defining the percentage of the population that must be immune to cause infection incidences to decline, the herd immunity threshold ($\mathcal{H}$) provides a useful target for vaccination coverage. In idealized scenarios of vaccines delivered at random and individuals mixing at random, $\mathcal{H}$ is given by a simple formula ($\mathcal{H} = 1 - 1/\mathcal{R}_0$) which, in the case of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), for example, suggests that 60-80% of subjects chosen randomly from the population would need be immunized to halt spread considering estimates of $\mathcal{R}_0$ between 2.5 and 5. This formula would not apply if vaccination programmes were designed to prioritize more susceptible or exposed individuals and, similarly, it does not apply to infection-induced immunity given that natural infection does not occur at random. Individuals who are more susceptible or more exposed are more prone to be infected and become immune, providing greater community protection than random vaccination [5]. In our model, the herd immunity threshold becomes $\mathcal{H} = 1 - (1/\mathcal{R}_0)^{1/(1+CV^2)}$ in the case of variable susceptibility, and $\mathcal{H} = 1 - (1/\mathcal{R}_0)^{1/(1+2CV^2)}$ in the case of variable exposure, which decline sharply when coefficients of variation increase from 0 to 2, remaining below 20% for more variable populations. The magnitude of the decline depends on what property is heterogeneous and how it is distributed among individuals, but the downward trend is robust provided susceptibility or exposure to infection are variable (Figure 1) and acquired immunity is efficacious enough to keep transmission below the reinfection threshold (i.e. $\sigma < 1/\mathcal{R}_0$) (Figure 2).

If immunity was not potent enough to keep the system below the reinfection threshold then herd immunity would not be attainable and the disease would persist in stable endemicity, irrespective of individual variation. Respiratory viruses, however, are typically associated with epidemic dynamics below the reinfection threshold, characterized by seasonal epidemics

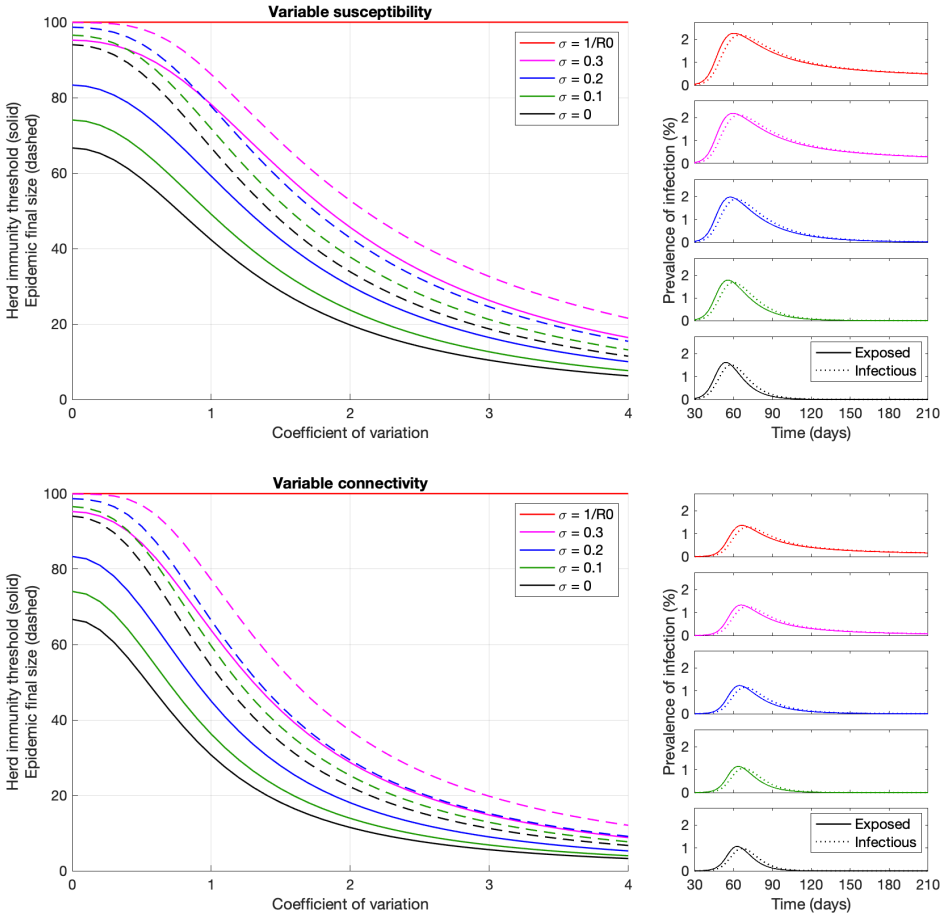


FIGURE 2. Herd immunity threshold and epidemic final size with reinfection. Curves in the main panels were generated with the SEIR model with reinfection (7)-(10) assuming $\mathcal{R}_0 = 3$ and gamma-distributed susceptibility (top) or connectivity (bottom). Efficacy of naturally acquired immunity is captured by a reinfection parameter σ, potentially ranging between $\sigma = 0$ (100% efficacy) and $\sigma = 1$ (0 efficacy). This illustration depicts final epidemic sizes and associated herd immunity thresholds $\mathcal{H}$ for 5 values of σ: $\sigma = 0$ (black); $\sigma = 0.1$ (green); $\sigma = 0.2$ (blue); $\sigma = 0.3$ (magenta); and $\sigma = 1/3$ (red). Above $\sigma = 1/\mathcal{R}_0$ (reinfection threshold ([9],[8]) the infection becomes stably endemic and there is no herd immunity threshold. Representative epidemics of the regime $\sigma \leq 1/\mathcal{R}_0$ are shown in small panels on the right. All depicted dynamics are based on the rightmost CVs represented on the main panel.

intertwined with quiescent periods. As several candidate vaccines against SARS-CoV-2 are undergoing clinical trials, we note that the reinfection threshold informs not only the requirements on naturally acquired immunity but, similarly, it sets a target for how efficacious a vaccine needs to be in order to effectively interrupt transmission ([9],[8]). Specifically, given an estimated value of $\mathcal{R}_0$ we should aim for a vaccine efficacy higher than $1 - 1/\mathcal{R}_0$ (60% or

80% if $\mathcal{R}_0$ is 2.5 or 5, respectively) irrespective of whether vaccination programmes prioritize or not those individuals who are more susceptible or exposed to infection.

Overdispersed infectiousness has been vastly discussed in the context of SARS viruses (e.g. [11] featuring SARS-CoV-1 and [1] addressing SARS-CoV-2). The otherwise unexplained occurrence of explosive outbreaks is often attributed to superspreading events which are largely due to chance. We emphasize, however, that infectiousness alone does not respond to selection as susceptibility or exposure do, and therefore does not accelerate the acquisition of herd immunity on a large scale. Models with individual variation in infectiousness perform equivalently to homogeneous versions when implemented deterministically. They diverge when stochasticity is added in the sense that disease extinction becomes more likely and outbreaks become rarer and more explosive, but this an entirely different phenomenon to that presented in this paper.

References

[1] D. Adam, et al., Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. *Nat. Med.* (2020). https://doi.org/10.1038/s41591-020-1092-0

[2] R. Aguas, R. M. Corder, J. G. King, G. Gonçalves, M. U. Ferreira, M. G. M. Gomes, Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. *medRxiv* 10.1101/2020.07.23.20160762.

[3] R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press Inc., New York (1991).

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

[5] M. J. Ferrari, S. Bansal, L. A. Meyers, O. N. Bjornstad, Network frailty and the geometry of herd immunity. *Proc. R. Soc.* B 273, 2743-2748 (2006).

[6] P. Fine, K. Eames, D. L. Heymann, "Herd immunity": a rough guide, *Clin. Infect. Dis.* 52, 911-916 (2011).

[7] M. G. M. Gomes, R. M. Corder, J. G. King, K. E. Langwig, C. Souto-Maior, J. Carneiro, G. Gonçalves, C. Penha-Gonçalves, M. U. Ferreira, R. Aguas, Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *medRxiv* 10.1101/2020.04.27.20081893.

[8] M. G. M. Gomes, E. Gjini, J. S. Lopes, C. Souto-Maior, C. Rebelo, A theoretical framework to identify invariant thresholds in infectious disease epidemiology. *J. Theor. Biol.* 395, 97-102 (2016).

[9] M. G. M. Gomes, L. J. White, G. F. Medley, Infection, reinfection, and vaccination under suboptimal immune protection: Epidemiological perspectives. *J. Theor. Biol.* 228, 539-549 (2004).

[10] G. Gonçalves, Herd immunity: recent uses in vaccine assessment. *Expert Rev. Vaccines* 7, 1493-1506 (2008).

[11] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence. *Nature* 438, 355-359 (2005).

[12] A. V. Tkachenko, S. Maslov, A. Elbanna, G. N. Wong, Z. J. Weiner, N. Goldenfeld, Persistent heterogeneity not short-term overdispersion determines herd immunity to COVID-19. *medRxiv* 10.1101/2020.07.26.20162420.

Department of Mathematics, University of California, Berkeley, USA
Email address: antonio@math.berkeley.edu
URL: www.math.berkeley.edu/∼antonio

Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, Brazil
Email address: rodrigo.corder@usp.br

Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK
Email address: Gabriela.Gomes@strath.ac.uk
URL: www.strath.ac.uk/staff/gomesgabrieladr