



Herd immunity under individual variation and reinfection

Antonio Montalbán¹ · Rodrigo M. Corder² · M. Gabriela M. Gomes^{3,4} 

Received: 10 August 2021 / Revised: 9 May 2022 / Accepted: 10 June 2022 /
Published online: 30 June 2022
© The Author(s) 2022

Abstract

We study a susceptible-exposed-infected-recovered (SEIR) model considered by Aguas et al. (In: Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics, 2021), Gomes et al. (In: J Theor Biol. 540:111063, 2022) where individuals are assumed to differ in their susceptibility or exposure to infection. Under this heterogeneity assumption, epidemic growth is effectively suppressed when the percentage of the population having acquired immunity surpasses a critical level - the herd immunity threshold - that is lower than in homogeneous populations. We derive explicit formulas to calculate herd immunity thresholds and stable configurations, especially when susceptibility or exposure are gamma distributed, and explore extensions of the model.

Keywords SEIR model · Individual variation · Selection · Effective reproduction number · Herd immunity threshold

Mathematics Subject Classification 92D30

✉ M. Gabriela M. Gomes
Gabriela.Gomes@strath.ac.uk
<https://www.strath.ac.uk/staff/gomesgabrieladr>

Antonio Montalbán
antonio@math.berkeley.edu
<https://www.math.berkeley.edu/~antonio>

Rodrigo M. Corder
rodrigo.corder@berkeley.edu

- ¹ Department of Mathematics, University of California, Berkeley, USA
- ² Department of Epidemiology and Biostatistics, University of California, Berkeley School of Public Health, Berkeley, USA
- ³ Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK
- ⁴ Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

1 Introduction

Understanding and predicting the dynamics and control of infectious diseases relies on representative models, whether conceptual or mathematical. Mathematical modelling was established in infectious diseases over a century ago, with the seminal works of Ross (1916), Ross and Hudson (1917), Kermack and McKendrick (1927) and others. Propelled by the discovery of aetiological agents for infectious diseases, and the germ theory, models have focused on the complexities of pathogen transmission and evolution (Heesterbeek 2015). It has recurrently been noted for over a century, however, that these models tend to overpredict transmission potential and overestimate the impact of control measures which may be explained by limitations in capturing the effects of heterogeneity (Kermack and McKendrick 1927; McKendrick 1939; Gart 1968, 1971; Ball 1985; Anderson et al. 1986; Pastor-Satorras and Vespignani 2001; Miller et al. 2012; Gomes et al. 2022).

Here we analyze a set of susceptible-exposed-infected-recovered (SEIR) models presented in Gomes et al. (2022), Aguas et al. (2021) where each of the compartments S, E, I and R is expanded into continuum many compartments $S(x)$, $E(x)$, $I(x)$ and $R(x)$, where $x \in \mathbb{R}^+$ is a trait that varies among individuals. Specifically we model a situation where each individual has a level of susceptibility or exposure (connectivity) x , starting in compartment $S(x)$ and staying within the compartments $S(x)$, $E(x)$, $I(x)$ and $R(x)$ the whole time. This individual may infect or be infected by others irrespective of their trait value x assuming random mixing (Anderson and May 1991; Diekmann et al. 2013). We will consider two types of models:

A **variable susceptibility** case where the susceptibility of an individual at level x is proportional to x , or, in other words, if we compare an individual at level x and an individual at level y , the one at level x is x/y times more likely to get infected than the one with susceptibility y . We may interpret this as variation in biological susceptibility which may be due to genetics, epigenetics or life history.

A **variable connectivity** case where 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 we compare an individual at level x and an individual at level y , the one at level x is x/y times more likely to get infected than the one in level y and also x/y times more likely to infect someone else once infected. This is interpreted as individuals with many contacts being both more likely to get infected and to infect others.

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

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

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

$$\begin{aligned} \text{Variable susceptibility: } \lambda &= \beta \int I(x) dx, \\ \text{Variable connectivity: } \lambda &= \beta \int x I(x) dx. \end{aligned}$$

Note that λ varies with time, as it depends on the time-dependent infected population.

The dynamics of compartments $S(x)$, $E(x)$, $I(x)$ and $R(x)$ are governed by the infinite system of ordinary differential equations:

$$\frac{dS(x)}{dt} = -\lambda x S(x), \tag{1}$$

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

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

$$\frac{dR(x)}{dt} = \gamma I(x). \tag{4}$$

We assume that the system has been scaled such that the total population is 1. The initial conditions for variables $S(x, t)$, $E(x, t)$, $I(x, t)$ and $R(x, t)$, satisfy $S(x, 0) = (1 - \epsilon) q(x)$, $E(x, 0) = \epsilon q(x)$ and $I(x, 0) = R(x, 0) = 0$, where $0 < \epsilon \ll 1$ is a small scalar to seed the epidemic, and $q(x)$ is a probability density function with mean 1 and coefficient of variation ν :

$$\int xq(x) dx = 1 \quad \text{and} \quad \sqrt{\int (x - 1)^2q(x) dx} = \nu. \tag{5}$$

We use $S(t)$ to denote the integral over all susceptibility levels, $S(x, t)$, for $x \in \mathbb{R}^+$. We thus have $S(t) = \int_0^{+\infty} S(x, t)dx$. Same with $E(t)$, $I(t)$ and $R(t)$.

We will use the first three moments of $S(x, t)$, that we denote $S(t)$, $\bar{S}(t)$ and $\bar{\bar{S}}(t)$:

$$S(t) = \int S(x, t) dx, \quad \bar{S}(t) = \int xS(x, t) dx \quad \text{and} \quad \bar{\bar{S}}(t) = \int x^2S(x, t) dx. \tag{6}$$

When infection is absent ($\epsilon = 0$), we have $S(0) = 1$, $\bar{S}(0) = 1$ and $\bar{\bar{S}}(0) = 1 + \nu^2$. But note that $S(x, t)$ is not a probability density function for $\epsilon > 0$ as $S(t)$ becomes less than 1. The quotient $S(x, t)/S(t)$ as a function of x for fixed t will be a probability density function for $\epsilon > 0$ and all t with first and second moments $\bar{S}(t)/S(t)$ and $\bar{\bar{S}}(t)/S(t)$ which decrease over time. When the initial configuration $q(x)$ is a gamma distribution, all the distributions $S(x, t)/S(t)$ are also gamma with the same coefficient of variation ν but with lower mean (see Appendix A), an argument which enables mathematical derivations to advance further when traits are assumed to be gamma distributed (Novozhilov 2008).

Similarly, we define the moments $R(t)$, $\bar{R}(t)$ and $\bar{\bar{R}}(t)$ for the recovered compartment, and the same with E and I . Notice for instance that $\lambda(t)$ is written as $\beta I(t)$ and $\beta \bar{I}(t)$ in the variable susceptibility and variable connectivity cases, respectively.

Here we describe key epidemiological quantities when system of equations (Eqs. 1–4) is adopted. The basic reproduction number \mathcal{R}_0 is the average number of secondary infections generated by an infected individual in a totally susceptible population. It depends on characteristics of both the pathogen and the host population. When this

number is below 1 no epidemics are expected. When \mathcal{R}_0 is above 1, however, 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}_{eff} is a time-dependent quantity loosely defined as the number of secondary infections generated by a typical infected individual when the susceptibility of the population is as at time t . \mathcal{R}_{eff} coincides with \mathcal{R}_0 at the beginning of an epidemic (when the population is totally susceptible) but declines as individuals are removed from the susceptible pool by infection and immunity. As \mathcal{R}_{eff} crosses 1 towards lower values, the epidemic subsides and future reintroductions of infection are not expected to generate new outbreaks as long as population immunity is maintained.

We derive formulas for the effective reproduction number \mathcal{R}_{eff} and the herd immunity threshold \mathcal{H} in terms of moments S , \bar{S} and $\bar{\bar{S}}$, of the susceptible population. When $q(x)$ is a gamma distribution, \bar{S} and $\bar{\bar{S}}$ can be formulated in terms of S and we get an exact formula for \mathcal{H} in terms only of the basic reproduction number \mathcal{R}_0 and the coefficient of variation ν . In this case we can also reduce the infinite system (Eqs. 1–4) to a finite system of ordinary differential equations in S , E , I and R with nonlinear transmission (exactly when the variable trait is susceptibility and approximately in the case of variable connectivity). In the case of variable connectivity, we provide an exact derivation of a finite system in the variables \bar{S} , \bar{E} , \bar{I} and \bar{R} .

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

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

and consequently, $\mathcal{R}_{\text{eff}} = \mathcal{R}_0 \bar{S}$. This implies that the population is above the herd immunity threshold when $\bar{S} < 1/\mathcal{R}_0$. When we assume that $q(x)$ is a gamma distribution, the proportion of individuals that have been infected by the time the herd immunity threshold is reached is deduced as:

$$\mathcal{H} = 1 - \mathcal{R}_0^{-\frac{1}{1+\nu^2}}, \tag{7}$$

and system (Eqs. 1–4) can be reduced to:

$$\frac{dS}{dt} = -\beta I S^{1+\nu^2}, \tag{8}$$

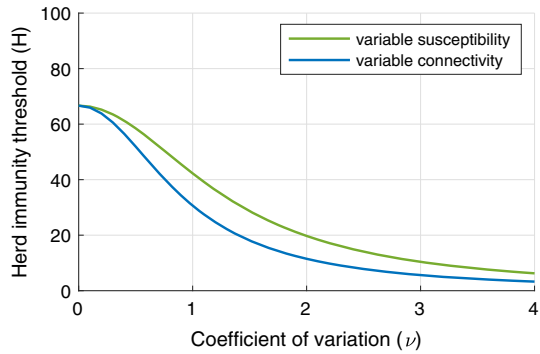
$$\frac{dE}{dt} = \beta I S^{1+\nu^2} - \delta E, \tag{9}$$

$$\frac{dI}{dt} = \delta E - \gamma I, \tag{10}$$

$$\frac{dR}{dt} = \gamma I. \tag{11}$$

These equations are exact.

Fig. 1 Herd immunity threshold. Curves generated using formulas (Eq. 7) for gamma distributed susceptibility and (Eq. 12) for gamma distributed connectivity, with $\mathcal{R}_0 = 3$



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

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

and consequently, $\mathcal{R}_{\text{eff}} = \mathcal{R}_0 / (1 + \nu^2) \bar{\bar{S}}$. This implies that the population is above the herd immunity threshold when $\bar{\bar{S}} < (1 + \nu^2) / \mathcal{R}_0$. When we assume that $q(x)$ is a gamma distribution, the proportion of individuals that have been infected by the time the herd immunity threshold is reached is deduced as:

$$\mathcal{H} = 1 - \mathcal{R}_0^{-\frac{1}{1+\nu^2}}. \tag{12}$$

In (Eqs. 28–31), we derive a closed system for the variable connectivity model on the variables \bar{S} , \bar{E} , \bar{I} and \bar{R} . But this does not directly allow the model to be fitted to incidence data provided by routine surveillance. For finding equations that determine the system only on the variables S , E , I and R , in the variable connectivity case we need to make an approximation as detailed in Appendix B. The resulting system (Eqs. 22–25) is shown to approximate the original (Eqs. 1–4) when the infectious period is small as in acute infectious diseases.

In Fig. 1 we provide graphical representations for the \mathcal{H} formulas corresponding to the basic model introduced in this section (variable susceptibility in green and variable connectivity in blue), showing a monotonic decrease as the coefficient of variation ν increases (Gomes et al. 2022).

In Sects. 2 and 3 we provide general derivations for effective reproduction numbers and herd immunity thresholds, while Sect. 4 is focussed on special cases when traits x are gamma distributed. Towards the end of the paper, we analyze two extensions of the basic model. In Sect. 5, we consider a model with reinfection where immunity after recovery is not fully protective but only partially. In Sect. 6, we consider a model with a carrier state, which (Gomes et al. 2022; Aguas et al. 2021) apply to the coronavirus disease (COVID-19) pandemic. There, the exposed compartments are not simply latent but a carrier state where individuals are infectious but to a lesser degree

than individuals in the fully infectious compartment. In each case we derive formulas for herd immunity thresholds especially when the initial trait distribution is gamma.

2 Effective reproduction number

The *effective reproduction number* at time t is defined as the number of secondary infections caused by a typical infected individual over their entire infectious period in an idealized situation, as described below.

Before proceeding with its derivation we need to make two considerations. The first concerns the evolving susceptible pool. We assume that during the period in which an individual is contagious, the density of susceptibles is frozen in time. This means that we disregard the fact that since the susceptible population declines, this individual infects less at the end than at the beginning of their infection. In acute infections, the decline in the susceptible population is usually slow compared to the rate of recovery from infection so, in practice, the impact of the assumption is negligible. Moreover, when \mathcal{R}_{eff} is used to analyze stable configurations, such as in the derivation of herd immunity thresholds, the assumption holds and hence has no effect on the results. This consideration pertains to both variable susceptibility and variable connectivity models.

The second concerns the infectivity profile of the infected population at time t . We define:

\mathcal{R}_{eff} at time t as the average number of secondary infections generated by an individual who *becomes infected* at time t . This average is taken over the pool of individuals that go from S to E at time t .

When $\mathcal{R}_{\text{eff}} < 1$, infection is not expected to invade an infection-free population. Further details on this concept are discussed in Appendix B. In the remaining of this section we derive explicit formulas for \mathcal{R}_{eff} .

First, the **variable susceptibility** case: Consider an individual who gets infected (more precisely, exposed and consequently infected) at time t (i.e., moves from S to E at time t). This individual will eventually move to I and spend on average $1/\gamma$ days there. While in I , the individual will infect an average of $\beta \int y S(y, t) dy$ others per day. We thus get:

$$\mathcal{R}_{\text{eff}}(t) = \frac{\beta}{\gamma} \int y S(y, t) dy = \frac{\beta}{\gamma} \bar{S}(t). \quad (13)$$

In particular, we get:

$$\mathcal{R}_0 = \frac{\beta}{\gamma}, \quad (14)$$

and consequently:

$$\mathcal{R}_{\text{eff}}(t) = \mathcal{R}_0 \bar{S}(t). \quad (15)$$

Second, the **variable connectivity** case: Consider again an individual who gets infected at time t . It now matters what trait value x this individual has because it determines how many others they will infect.

Let $p(x, t)$ be the density function measuring the probability at time t that this individual has connectivity level x . The probability of becoming infected (i.e., of entering the E compartment) is $x\lambda(t)$. Thus, the value of $p(x, t)$ is proportional to $xS(x, t)$:

$$p(x, t) = x \frac{S(x, t)}{\bar{S}(t)}.$$

As above, an individual who enters E will eventually move to I, spend on average $1/\gamma$ days there, and infect an average of $\beta \int y S(y, t) dy$ others per day. We thus get:

$$\begin{aligned} \mathcal{R}_{\text{eff}}(t) &= \int \frac{\beta}{\gamma} \left(\int y S(y, t) dy \right) x p(x, t) dx \\ &= \frac{\beta}{\gamma} \int \bar{S}(t) \frac{x^2 S(x, t)}{\bar{S}(t)} dx = \frac{\beta}{\gamma} \bar{\bar{S}}(t). \end{aligned} \tag{16}$$

In particular, we get:

$$\mathcal{R}_0 = \bar{\bar{q}} \frac{\beta}{\gamma} = (1 + \nu^2) \frac{\beta}{\gamma}, \tag{17}$$

and consequently:

$$\mathcal{R}_{\text{eff}}(t) = \frac{\mathcal{R}_0}{1 + \nu^2} \bar{\bar{S}}(t). \tag{18}$$

Expressions for \mathcal{R}_0 , such as (Eq. 14) and (Eq. 17), have been known for decades (Anderson and May 1991; Diekmann et al. 2013; Woolhouse et al. 1997). Worth highlighting, however, is that disproportionately less attention has been given to variable susceptibility than to variable connectivity due to the coefficient of variation ν not affecting the formula explicitly in the former case but only in the latter. It will be important to realise, however, that variation in susceptibility is just as impactful when we consider quantities such as herd immunity thresholds and inferences of \mathcal{R}_0 from observational data (Gomes et al. 2022; Aguas et al. 2021).

3 Herd immunity threshold

Suppose we have a population with no infected individuals, so that all individuals are either susceptible or recovered. The population is said to be at or above the herd immunity threshold for a pathogen if its susceptibility profile to that pathogen is such that a new introduction of infection (i.e., a small increase in E or I) does not trigger

an outbreak. By inspection on the differential equations (Eqs. 1–4), we see that a configuration with $E(x) = I(x) = 0$ satisfies this condition if and only if

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

In the variable susceptibility case it is equivalent to formulate the herd immunity threshold in terms of suppression of future outbreaks (as adopted here) or in terms of an unmitigated epidemic passing its peak infection prevalence. With variable connectivity, however, this equivalence does not hold as explained in Sect. 2.

A configuration with no infected individuals is then said to be at the herd immunity threshold if and only if $\mathcal{R}_{\text{eff}} = 1$. In SEIR models with no individual variation, configurations with no infected individuals are determined by the values of S and $R = 1 - S$, and the herd immunity threshold is defined as the value of $1 - S$ at the unique configuration with $\mathcal{R}_{\text{eff}} = 1$. This value is well-known to be equal to $1 - 1/\mathcal{R}_0$ (Anderson and May 1991; Diekmann et al. 2013). With individual variation in susceptibility or exposure to infection, however, there are many configurations which satisfy $\mathcal{R}_{\text{eff}} = 1$. One such configuration is given by $S(x) = q(x)/\mathcal{R}_0$ for all x . This could be obtained, for instance, by vaccinating a proportion $1 - 1/\mathcal{R}_0$ of the total population randomly without taking into account susceptibility or exposure levels (Fine et al. 2011). When immunity is acquired naturally, however, individuals with higher susceptibility or exposure tend to be infected earlier and the herd immunity threshold is reached before the susceptible population is as low as $1/\mathcal{R}_0$ of the total. For now, let us retain that for the basic heterogeneous models considered here the herd immunity threshold is reached when $\bar{S} = 1/\mathcal{R}_0$ in the **variable susceptibility** case (Eq. 15) and $\bar{\bar{S}}/(1 + \nu^2) = 1/\mathcal{R}_0$ with **variable connectivity** (Eq. 18).

Next, we will see how we can derive $\bar{S}(t)$ and $\bar{\bar{S}}(t)$ under the assumption that the initial distribution is gamma.

4 Case of the gamma distribution

Here we study how the distribution of the trait x within the susceptible compartment evolves generally, and then specify to the case where individual variation in susceptibility or connectivity is gamma distributed. We also refer to related work in Neipel et al. (2020), Novozhilov (2008).

4.1 Evolution of the susceptible compartment

From the SEIR equation for $dS(x, t)/dt$ we get:

$$\frac{1}{S(x, t)} \frac{dS(x, t)}{dt} = -x \lambda(t).$$

Integrating with respect to t we get:

$$S(x, t) = q(x) e^{-x k_t} \quad \text{where} \quad k_t = \int_0^t \lambda(u) du. \tag{19}$$

This holds in both variable susceptibility and variable connectivity models (with different values for k_t). It also holds in the cases with reinfection and with carrier state considered in Sects. 5 and 6.

4.2 Gamma distributed traits

The key observation here is that since $S(x, t) = q(x) e^{-x k_t}$, we have that $S(x, t)/S(t)$ remains a gamma distribution at all values of t . This enables the derivations in Appendix A of explicit formulas for the moments \bar{S} (Eq. A.1) and $\bar{\bar{S}}$ (Eq. A.2) in terms of S and the shape parameter α , when susceptibility or connectivity are gamma distributed.

Using that the coefficient of variation is $\nu = 1/\sqrt{\alpha}$ we rewrite the respective formulas as:

$$\bar{S}(t) = S(t)^{1+\nu^2}, \tag{20}$$

$$\bar{\bar{S}}(t) = (1 + \nu^2) S(t)^{1+2\nu^2}. \tag{21}$$

To derive the reduced system in the **variable susceptibility** case (Eqs. 8–11) we integrate the equations in (Eqs. 1–4) over the susceptibility domain and apply (Eqs. 13 and 20) to get:

$$\frac{dS}{dt} = -\beta I \bar{S} = -\beta I S^{1+\nu^2}, \tag{22}$$

$$\frac{dE}{dt} = \beta I \bar{S} - \delta E = \beta I S^{1+\nu^2} - \delta E, \tag{23}$$

$$\frac{dI}{dt} = \delta E - \gamma I, \tag{24}$$

$$\frac{dR}{dt} = \gamma I. \tag{25}$$

This closed system in S, E, I and R has been used to fit epidemic curves of COVID-19 (Gomes et al. 2022; Aguas et al. 2021). Recalling that the herd immunity threshold \mathcal{H} is $1 - S(t)$ at the time t when $\mathcal{R}_{\text{eff}} = 1$ and that $\mathcal{R}_{\text{eff}}(t) = \mathcal{R}_0 \bar{S}(t)$ (Eq. 15), we get that $\mathcal{R}_0^{-1} = \bar{S}(t) = S(t)^{1+\nu^2}$ at the herd immunity threshold when susceptibility is gamma distributed. Hence:

$$\mathcal{H} = 1 - \mathcal{R}_0^{-\frac{1}{1+\nu^2}}. \tag{26}$$

In the **variable connectivity** case we have that $\mathcal{R}_{\text{eff}}(t) = \mathcal{R}_0/(1 + \nu^2) \bar{\bar{S}}(t)$ (Eq. 18). Thus, when $\mathcal{R}_{\text{eff}} = 1$, we get $\mathcal{R}_0^{-1} = \bar{\bar{S}}(t)/(1 + \nu^2) = S(t)^{1+2\nu^2}$ when connectivity is gamma distributed. Hence:

$$\mathcal{H} = 1 - \mathcal{R}_0^{-\frac{1}{1+2\nu^2}}. \tag{27}$$

Multiplying the original (Eqs. 1–4) by x , integrating over the connectivity domain and applying (Eqs. 16 and 21) we get:

$$\frac{d\bar{S}}{dt} = -\beta \bar{I} \bar{S} = -(1 + \nu^2) \beta \bar{I} \bar{S}^{\frac{1+2\nu^2}{1+\nu^2}}, \tag{28}$$

$$\frac{d\bar{E}}{dt} = \beta \bar{I} \bar{S} - \delta \bar{E} = (1 + \nu^2) \beta \bar{I} \bar{S}^{\frac{1+2\nu^2}{1+\nu^2}} - \delta \bar{E}, \tag{29}$$

$$\frac{d\bar{I}}{dt} = \delta \bar{E} - \gamma \bar{I}, \tag{30}$$

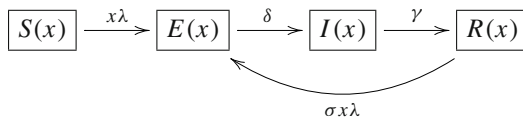
$$\frac{d\bar{R}}{dt} = \gamma \bar{I}. \tag{31}$$

Mathematically this is a tractable closed system in \bar{S} , \bar{E} , \bar{I} and \bar{R} . However, these variables are not convenient for practical data fitting and parameter estimation. In Appendix B we propose an approximation in the variables S , E , I and R .

5 Model with reinfection

Here we consider an extension of the model considering that immunity after recovery is not fully protective, but only partially. A factor σ , with $0 \leq \sigma \leq 1$, 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 is represented diagrammatically as:



with λ as in the basic models (without reinfection) studied above. The extended model is given by the equations:

$$\frac{dS(x)}{dt} = -\lambda x S(x), \tag{32}$$

$$\frac{dE(x)}{dt} = \lambda x (S(x) + \sigma R(x)) - \delta E(x), \tag{33}$$

$$\frac{dI(x)}{dt} = \delta E(x) - \gamma I(x), \tag{34}$$

$$\frac{dR(x)}{dt} = \gamma I(x) - \sigma \lambda x R(x). \tag{35}$$

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 (Gomes et al. 2004, 2016)) 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 reinfected repeatedly.

5.1 Effective reproduction number

The basic reproduction number is calculated exactly, as in the absence of reinfection, but the effective reproduction now depends not only on the distribution of $S(x, t)$, but also on the distribution of $R(x, t)$. When we consider configurations with no infected individuals, we will have that $R(x, t) = q(x) - S(x, t)$ and will be able to express $\mathcal{R}_{\text{eff}}(t)$ in terms of $S(x, t)$ only.

The formulas for the effective reproduction number \mathcal{R}_{eff} at time t are:

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

The derivation of these formulas is essentially as the derivations in Sect. 2, with two differences. First, each individual with trait value x infects $\beta(\bar{S}(t) + \sigma \bar{R}(t))$ or $\beta x(\bar{\bar{S}}(t) + \sigma \bar{\bar{R}}(t))$ others per day spent in l , in the respective cases, instead of $\beta \bar{S}(t)$ or $\beta x \bar{S}(t)$. Second, when we consider an individual that gets infected in the variable connectivity case, the probability that this individual has trait value x is proportional to $x(S(x, t) + \sigma R(x, t))$ instead of $xS(x, t)$.

5.2 Herd immunity threshold

Recall that a configuration with no infected individuals is at or above the herd immunity threshold 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 $\bar{\bar{R}} = \bar{\bar{q}} - \bar{\bar{S}}$, where $\bar{\bar{q}} = \int x^2 q(x) dx = 1 + v^2$. We can then understand the configurations at the herd immunity threshold in terms of \bar{S} and $\bar{\bar{S}}$.

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

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

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

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

5.3 Reinfection threshold

The formulas above require:

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

That is, the reinfection factor σ has to be below \mathcal{R}_0^{-1} , a critical value known as the reinfection threshold (Gomes et al. 2004, 2016). If this is verified, then all configurations with no infected individuals and satisfying the conditions above (either $\bar{S} = (\mathcal{R}_0^{-1} - \sigma)/(1 - \sigma)$ or $\bar{S} = (1 + \nu^2)(\mathcal{R}_0^{-1} - \sigma)/(1 - \sigma)$ depending on the case) are herd immunity threshold configurations in the sense that when susceptibility is at that level or lower, infection reintroductions will not trigger new outbreaks as \mathcal{R}_{eff} will not increase above 1.

If the reinfection factor is so high that the population is above the reinfection threshold, \mathcal{R}_{eff} will be greater than 1 in any such configurations, so there will not be any configuration with no infected individuals which is at the herd immunity threshold. 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 infection becomes endemic. The equilibrium configuration will now have $S(x) = 0$ for all x . In these situations the level of endemicity will depend on how much resistance the population is able to mount and maintain.

5.4 Case of the gamma distribution

Recall from Sect. 3 that, when the initial distribution $q(x)$ is gamma, we get that $S(x, t)/S(t)$ remains a gamma distribution for all t , and that $\bar{S}(t) = S(t)^{1+\nu^2}$ and $\bar{\bar{S}}(t) = (1 + \nu^2) S(t)^{1+2\nu^2}$. We can then obtain the values of $S(t)$ at the time when the herd immunity threshold is reached, and then calculate \mathcal{H} as $1 - S(t)$ for that particular t .

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

$$\mathcal{H} = 1 - \left(\frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma} \right)^{\frac{1}{1+\nu^2}}. \tag{36}$$

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

$$\mathcal{H} = 1 - \left(\frac{\mathcal{R}_0^{-1} - \sigma}{1 - \sigma} \right)^{\frac{1}{1+2\nu^2}}. \tag{37}$$

Curves assuming a selection of values for σ are represented graphically in Fig. 2. Note the critical behaviour at the reinfection threshold ($\sigma = 1/\mathcal{R}_0$) in red, which separates the regime where individual immunity is sufficiently potent for the herd immunity threshold to be achievable from the regime where endemicity will establish.

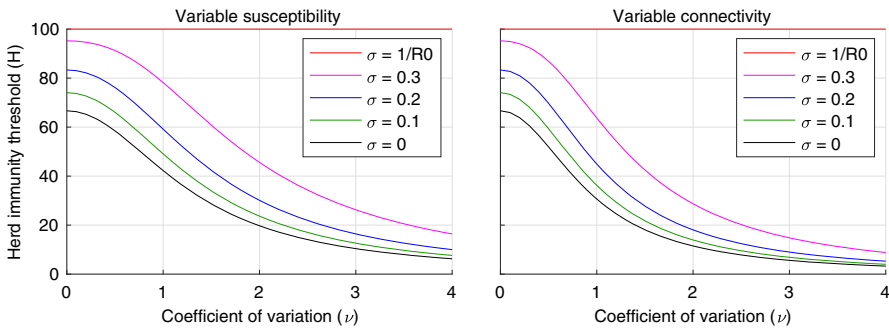


Fig. 2 Herd immunity threshold with reinfection. Curves correspond to the SEIR model with reinfection (Eqs. 32–35) assuming $\mathcal{R}_0 = 3$ and gamma distributed susceptibility (left, Eq. 36) or connectivity (right, Eq. 37). Efficacy of naturally acquired immunity is captured by a reinfection parameter σ , potentially ranging between $\sigma = 0$ (100% efficacy) and $\sigma = 1$ (0 efficacy). Five values of the reinfection parameter are depicted: $\sigma = 0$ (black); $\sigma = 0.1$ (green); $\sigma = 0.2$ (blue); $\sigma = 0.3$ (magenta); and $\sigma = 1/\mathcal{R}_0$ (red). Above $\sigma = 1/\mathcal{R}_0$ (reinfection threshold Gomes et al. 2004, 2016) the infection becomes stably endemic and there is no herd immunity threshold (colour figure online)

6 Model with a carrier state

In applications to the COVID-19 pandemic (Gomes et al. 2022; Aguas et al. 2021) the exposed compartments are not simply a latent state but a carrier state where individuals are infectious but to a lesser degree than those in the fully infectious compartment. Building on the reinfection model (Eqs. 32–32) we now introduce parameter $\rho \leq 1$ to denote the ratio of infectiousness between exposed and fully infectious individuals. What changes is the force of infection λ :

Variable susceptibility: $\lambda(t) = \beta \int \rho E(x, t) + I(x, t) dx = \beta (\rho E(t) + I(t)),$

Variable connectivity: $\lambda(t) = \beta \int x(\rho E(x, t) + I(x, t)) dx = \beta (\rho \bar{E}(t) + \bar{I}(t)).$

The basic and effective reproduction numbers become:

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

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

The calculation of the effective reproduction number \mathcal{R}_{eff} is slightly different. The difference is that now we have to add the time an individual is incubating the infection

in E to the infectious period, multiplied by the factor ρ . Since the average time an individual spends in E is $1/\delta$ we get:

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

where $\bar{S}(t)$ and $\bar{\bar{S}}(t)$ are the moments of $S(x, t)$ defined above, and the same with $R(x, t)$.

In particular, we get $\mathcal{R}_0 = \beta(\rho/\delta + 1/\gamma)$ and $\mathcal{R}_0 = (1 + \nu^2)\beta(\rho/\delta + 1/\gamma)$, respectively, and $\mathcal{R}_{\text{eff}} = \mathcal{R}_0 (\bar{S} + \sigma \bar{R})$ and $\mathcal{R}_{\text{eff}} = \mathcal{R}_0/(1 + \nu^2) (\bar{\bar{S}} + \sigma \bar{\bar{R}})$. Finally, we obtain the same formulas for the herd immunity threshold in terms of \mathcal{R}_0 and ν (Sect. 3, without reinfection) or more generally \mathcal{R}_0 , ν and σ (Sect. 5, with reinfection).

7 Discussion

After completion of this work, similar ideas for capturing individual variation with mean-field epidemic models were elaborated (Britton et al. 2020; Di Lauro et al. 2021; Kawagoe et al. 2021; Neipel et al. 2020; Rose et al. 2021; Tkachenko et al. 2021). These recent developments were largely prompted by the COVID-19 pandemic. In agreement with our results, Rose et al. (2021), Neipel et al. (2020) find that when susceptibility is initially gamma distributed it remains so through the course of the epidemic, leading naturally to power-law behaviour in the force of infection (Novozhilov 2008). In addition the authors show that other initial distributions converge towards gamma through the process of contagion. Other authors (Kawagoe et al. 2021) derive epidemic final sizes assuming alternative distributions of susceptibility, considering in addition that infectivity may exhibit some correlation with susceptibility (such as in the variable connectivity models analyzed here). They compare numerical results for gamma and lognormal distributions with those obtained using an empirical distribution of individual contacts derived from cell phone geolocation data. Focussing on variable connectivity (Britton et al. 2020; Di Lauro et al. 2021) address herd immunity thresholds using age-structured compartmental models. Additionally (Di Lauro et al. 2021) consider a variety of non-pharmaceutical intervention scenarios, emphasizing subtle results when interventions change the contact network. Another group of authors (Tkachenko et al. 2021) distinguishes between persistent and transient individual variation to highlight that only the former is subject to the kind of selection that lowers epidemic final sizes and herd immunity thresholds.

Despite being prompted by COVID-19 none of the above references attempted to quantify the individual variation which was under selection by the force of infection and hence contributed to lower epidemic final sizes and herd immunity thresholds. This is done in our associated work (Gomes et al. 2022; Aguas et al. 2021), where the problem is inverted and selectable variation is inferred from its effects on epidemic patterns. Although we and others had previously adopted similar approaches to other infectious diseases (Finkenstadt and Grenfell 2000; Smith et al. 2005; Bellan et al. 2015; Corder et al. 2020) the work has been processed more cautiously during the pandemic due to the greater implications that estimating lower herd immunity thresholds and epidemic final sizes might have for policies and behaviors in this context.

The concept of herd immunity was originally developed in the context of vaccination programs (Gonçalves 2008; Fine et al. 2011). Defining the percentage of the population that must be immunised to cause infection prevalence to decline, the concept has provided a useful target for vaccination coverage. In hypothetical scenarios of vaccines delivered at random and individuals mixing at random, the herd immunity threshold is given by the simple expression $(1 - \mathcal{R}_0^{-1})$ when immunity is fully protective. Concretely, for \mathcal{R}_0 between 2.5 and 5, this would indicate that 60-80% random subjects would need be immunized to prevent spread of infection. This formula would not apply, however, if vaccination programmes were designed to prioritize more connected individuals, for instance (Elbasha and Gumel 2021). Similarly, it does not apply when immunity is acquired in response to natural infection, which does not occur at random. Individuals who are more susceptible or more exposed to infection are prone to be infected and become immune earlier. As a result, earlier episodes contribute disproportionately to herd immunity as they remove highly susceptible and exposed subjects from the susceptible pool (Ferrari et al. 2006; Britton et al. 2020; Kawagoe et al. 2021; Neipel et al. 2020; Rose et al. 2021; Gomes et al. 2022; Aguas et al. 2021). In our basic models, the herd immunity threshold becomes $\mathcal{H} = 1 - \mathcal{R}_0^{-1/(1+\nu^2)}$ in the case of gamma distributed susceptibility, and $\mathcal{H} = 1 - \mathcal{R}_0^{-1/(1+2\nu^2)}$ with gamma distributed connectivity (exposure), which decline sharply when coefficients of variation (ν) increase from 0 to 2, remaining below 20% for more variable populations in a particular illustration where $\mathcal{R}_0 = 3$ (Fig. 1). 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 that acquired immunity is efficacious enough to keep transmission below the reinfection threshold (Fig. 2) ($\sigma < \mathcal{R}_0^{-1}$, where σ is the susceptibility of individuals who have recovered relative to their respective susceptibility prior to infection) (Gomes et al. 2004, 2016). In our reinfection models, herd immunity thresholds are derived as $\mathcal{H} = 1 - \left((\mathcal{R}_0^{-1} - \sigma) / (1 - \sigma) \right)^{1/(1+\nu^2)}$ in the case of gamma distributed susceptibility, and $\mathcal{H} = 1 - \left((\mathcal{R}_0^{-1} - \sigma) / (1 - \sigma) \right)^{1/(1+2\nu^2)}$ with gamma distributed connectivity, when $\sigma < \mathcal{R}_0^{-1}$. If immunity is not potent enough to keep the system below the reinfection threshold then a herd immunity threshold is not attainable and the disease persists in stable endemicity, irrespective of individual variation.

Finally, we stress that the herd immunity threshold (\mathcal{H}) is a theoretical framework to assess epidemic potential to the same extent that \mathcal{R}_0 is a theoretical framework. Their interdependence shows that if \mathcal{R}_0 increases due to evolution of the infectious agent, for example, so does \mathcal{H} . Also, if new susceptibles enter the population through birth or other processes, or if immunity wanes or is evaded by pathogen lineages, a previously acquired herd immunity status may be lost leaving the population prone to subsequent outbreaks. Furthermore, if transmission has a marked seasonal pattern the same level of immunity may place the population above threshold in low season and below threshold in high season, in a cyclical manner. Although \mathcal{H} is not as immediately applicable as often implied, it is a more informative measure of epidemic potential than \mathcal{R}_0 given that it accounts for variation in susceptibility or exposure (ν) in addition

to average transmissibility \mathcal{R}_0 . The more accurately we know \mathcal{H} the better we can assess trade-offs and inform public health policy.

Acknowledgements We thank Paul McKeigue and Joel Miller for valuable discussions and two anonymous reviewers for constructive criticisms and suggestions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A. Derivations of moments S , \bar{S} and $\bar{\bar{S}}$ for gamma distributed traits

In this appendix we derive explicit formulas for \bar{S} and $\bar{\bar{S}}$, in the cases of variable susceptibility and connectively respectively, in terms of S when the traits are gamma distributed.

We use conventional notation for the gamma probability distribution:

$$\text{Gamma}_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},$$

noting that α and β are the shape and rate parameters, respectively. The notation β clashes with the transmission coefficient which is also conventionally denoted by the same Greek letter, but its appearance as rate parameter in this appendix ceases once we condition our gamma distributions to having mean $\alpha/\beta = 1$, i.e., $\beta = \alpha$, so confusion can be precluded.

Given an initial gamma distribution with mean 1, explicitly $q(x) = \text{Gamma}_{\alpha,\alpha}(x)$, which we substitute in (Eq. 19), we get

$$\begin{aligned} S(x, t) &= \frac{\alpha^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x(\alpha+k_t)} \\ &= \left(\frac{\alpha}{\alpha+k_t} \right)^\alpha \frac{(\alpha+k_t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x(\alpha+k_t)} \\ &= \left(\frac{\alpha}{\alpha+k_t} \right)^\alpha \cdot \text{Gamma}_{\alpha,\alpha+k_t}(x). \end{aligned}$$

Using the equalities $\int \text{Gamma}_{\alpha,\alpha+k_t}(x) dx = 1$, $\int x \text{Gamma}_{\alpha,\alpha+k_t}(x) dx = \alpha/(\alpha+k_t)$ and $\int x^2 \text{Gamma}_{\alpha,\alpha+k_t}(x) dx = [\alpha(\alpha+1)]/[(\alpha+k_t)^2]$, we calculate $S(t)$, $\bar{S}(t)$

and $\bar{S}(t)$:

$$\begin{aligned}
 S(t) &= \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha \cdot \int \text{Gamma}_{\alpha, \alpha+k_I}(x) dx = \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha, \\
 \bar{S}(t) &= \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha \cdot \int x \text{Gamma}_{\alpha, \alpha+k_I}(x) dx = \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha \\
 &\quad \cdot \frac{\alpha}{\alpha + k_I} = \left(\frac{\alpha}{\alpha + k_I}\right)^{\alpha+1}, \\
 \bar{\bar{S}}(t) &= \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha \cdot \int x^2 \text{Gamma}_{\alpha, \alpha+k_I}(x) dx \\
 &= \left(\frac{\alpha}{\alpha + k_I}\right)^\alpha \cdot \frac{\alpha(\alpha + 1)}{(\alpha + k_I)^2} = \left(\frac{\alpha}{\alpha + k_I}\right)^{\alpha+2} \cdot \frac{\alpha + 1}{\alpha}.
 \end{aligned}$$

From the above we get

$$\bar{S}(t) = S(t)^{\frac{\alpha+1}{\alpha}}, \tag{A.1}$$

$$\bar{\bar{S}}(t) = S(t)^{\frac{\alpha+2}{\alpha}} \cdot \frac{\alpha + 1}{\alpha}. \tag{A.2}$$

Appendix B. Approximate variable connectivity model in S, E, I, R variables

In (Eqs. 28–31), we derived a closed system for the variable connectivity model on the variables \bar{S} , \bar{E} , \bar{I} , \bar{R} . However, we would like to have a closed system on the variables S, E, I, R, as we did in the variable susceptibility case to enable direct fitting to incidence data (Eqs. 22–25) for parameter estimation and scenario projections. We propose a system of approximate equations which works well when the infectious period is small as in acute infectious diseases (Eqs. B.8–B.11 below).

The justification for the approximation is based on the assumption that the following two quantities are very close to each other when the infectious period is small compared to the length of the epidemic. At each time t we consider the following two factors, that we denote $f(t)$ and $t(t)$

- $f(t)$ is the average infectivity of a typical individual who *becomes infected* at time t . In other words, this is the average value of trait x taken over the pool of individuals that go from S to E at time t .
- $t(t)$ is the average infectivity of a typical individual who *is infectious* at time t . In other words, this is the average value of trait x taken over the pool of individuals that are in compartment I at time t .

Both f and t are formulated below (Eq. B.2). The distinction between them is only meaningful in models where infectivity varies among individuals in such a way that those infected in different days tend to have different infectivities — this occurs in our variable connectivity model (as infectivity is positively correlated with susceptibility,

the selection which reduces susceptibility in the susceptible pool over time also reduces infectivity) but not in the case of variable susceptibility (where all individuals have the same infectivity).

Note that we use the factor f in our calculation of \mathcal{R}_{eff} in (Eq. 16), where we multiply β by the average infectivity f (same as average connectivity), the average susceptibility \bar{S} , and the length of the infectious period ($1/\gamma$):

$$\mathcal{R}_{\text{eff}} = \frac{\beta}{\gamma} \cdot f \cdot \bar{S}. \tag{B.1}$$

Let us also note that the “ R -number” which was popularized in the COVID-19 pandemic, loosely described as the number of new infections caused by a typical individual who is infectious at time t , is defined in terms of t rather than f . This more empirical notion, formulated as $\mathcal{R}_t = (\beta/\gamma) \cdot t \cdot \bar{S}$, is not always coincident with the basic reproduction number \mathcal{R}_{eff} , introduced in textbooks as the spectral radius of the next generation operator (Diekmann et al. 2013) and adopted in this paper. The two notions differ when the average infectivity of individuals infected in a particular day is different to the average infectivity of those who have been infected in another day. However, this difference is of little consequence to variable connectivity models when the infectious period is small and of no consequence when variation is in susceptibility (see Fig. 3).

We claim that the values of f and t are given by the following formulas:

$$f = \bar{S}/\bar{S} \quad \text{and} \quad t = \bar{l}. \tag{B.2}$$

The second equality follows from the formula for average infectivity, $\int (xI(x, t)/l)dx$, for every time t . The first equality was essentially derived in Sect. 2 when we calculated \mathcal{R}_{eff} for variable connectivity (Eq. 16). Within that calculation we had to consider the infectivity of an individual who becomes infected at time t , which is equal to $\int xp(x, t)dx$, where $p(x, t)$ is the density function measuring the probability that such an individual has connectivity level x . We saw that since $p(x, t)$ is proportional to $xS(x, t)$, we have that $p(x, t) = xS(x, t)/\bar{S}(t)$. It follows then that $f(t) = \int x^2S(x, t)/\bar{S}(t)dx$, for every t , which is the expression above.

We now can write our assumption as:

$$\bar{S}/\bar{S} \approx \bar{l}. \tag{B.3}$$

With this we can derive the approximate closed system of equations on the variables S, E, I, R . Observe that if we integrate the Eqs. (1)–(4) over x and use that $\lambda = \beta \bar{l}$, we obtain

$$\frac{dS}{dt} = -\beta \bar{l} \bar{S} \tag{B.4}$$

$$\frac{dE}{dt} = \beta \bar{l} \bar{S} - \delta E \tag{B.5}$$

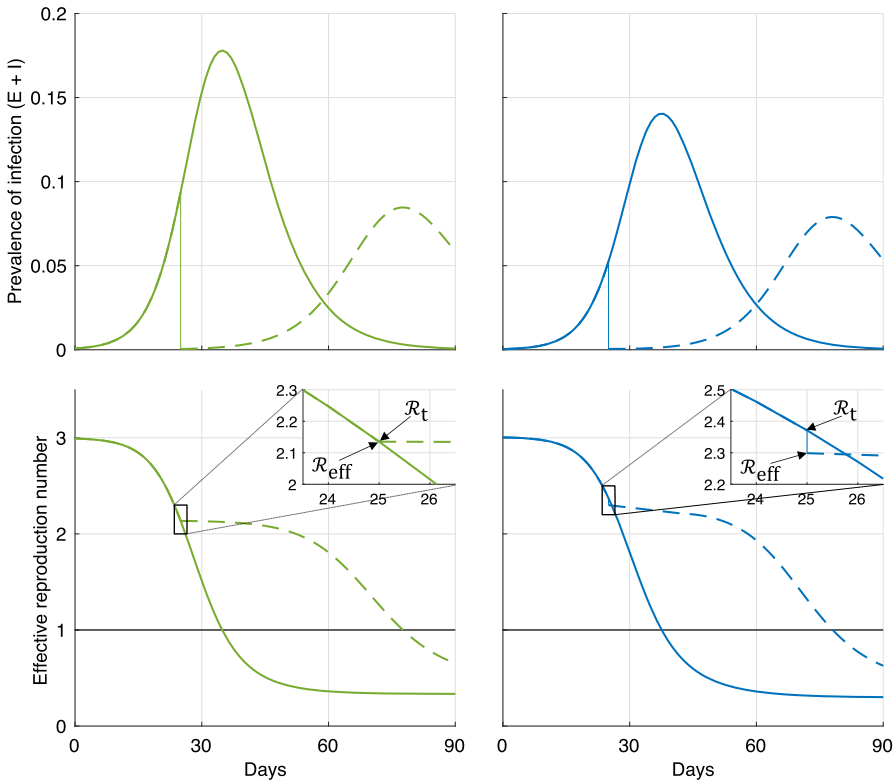


Fig. 3 Effective reproduction number. Top panels show epidemic curves generated by running system (Eqs. 1–4) with explicit gamma distributed susceptibility (left) or connectivity (right). Solid curves are unmodified epidemics while dashed show the outcome of hypothetically clearing all infections instantaneously on day 25 and progressing with a small infectious seed. Bottom panels show the respective \mathcal{R}_t values, calculated as $(\beta/\gamma) \cdot \bar{S}$ in the case of variable susceptibility and $(\beta/\gamma) \cdot (\bar{I}/I) \cdot \bar{S}$ in the case of variable connectivity, with or without the intervention on day 25. The empirical index \mathcal{R}_t coincides with the effective reproduction number \mathcal{R}_{eff} in the variable susceptibility case but is slightly higher when variation is in connectivity. This is highlighted by zooming into the areas marked by black rectangles (see insets). Parameters: $\mathcal{R}_0 = 3$; $\delta = 1/4$ per day; $\gamma = 1/3$ per day; and $\nu = 1$ (colour figure online)

$$\frac{dI}{dt} = \delta E - \gamma I \tag{B.6}$$

$$\frac{dR}{dt} = \gamma I. \tag{B.7}$$

Then, using that $\bar{I} \bar{S} \approx I \bar{S}$ (Eq. B.3), and that with gamma distributed traits we have $\bar{S} = (1 + \nu^2) S(t)^{1+2\nu^2}$ (Eq. 21), we get the approximate closed system we seek:

$$\frac{dS}{dt} \approx -(1 + \nu^2) \beta I S^{1+2\nu^2}, \tag{B.8}$$

$$\frac{dE}{dt} \approx (1 + \nu^2) \beta I S^{1+2\nu^2} - \delta E, \tag{B.9}$$

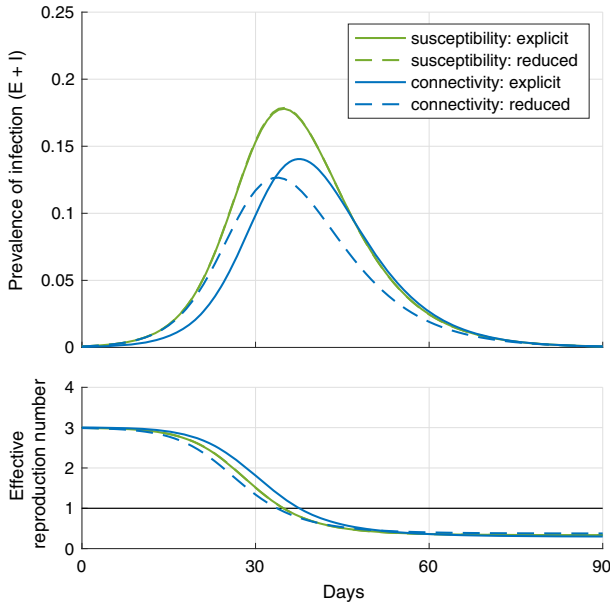


Fig. 4 Epidemics with gamma distributed susceptibility or connectivity. Solid curves were generated by running a discretized version of the infinite system (Eqs. 1–4) with explicit gamma distributed susceptibility (green) or connectivity (blue). Dashes curves are the corresponding outputs of the reduced systems (Eqs. 22–25) and (Eqs. B.8–B.11). Prevalences of infection ($E + I$) over time are depicted on the top panel and effective reproduction numbers on the bottom. Parameters: $\mathcal{R}_0 = 3$; $\delta = 1/4$ per day; $\gamma = 1/3$ per day; and $\nu = 1$ (colour figure online)

$$\frac{dI}{dt} = \delta E - \gamma I, \tag{B.10}$$

$$\frac{dR}{dt} = \gamma I. \tag{B.11}$$

Figure 4 illustrates epidemic curves generated with system (Eqs. 1–4) with an explicit gamma distribution (mean 1 and coefficient of variation $\nu = 1$) for susceptibility (solid green) and connectivity (solid blue). Technically the distribution was discretized in n bins and a system of $4n$ ODEs was run to generate the curves. Dashed curves were added to the plots to enable comparisons with the corresponding results generated by the reduced systems (Eqs. 22–25 and B.8–B.11, respectively). The bottom panel shown the associated effective reproduction numbers \mathcal{R}_{eff} in the case of the reduced systems, and the empirical proxy \mathcal{R}_t when explicit distributions are implemented (recall that the two coincide when variation is in susceptibility). All runs were conducted with basic reproduction number $\mathcal{R}_0 = 3$, incubation period ($1/\delta = 4$ days), infectious period ($1/\gamma = 3$ days) and coefficient of individual variation ($\nu = 1$). Notice that the reduction is exact in the variable susceptibility case but not with variable connectivity. In a related study we fitted both explicit and reduced models to real epidemics and obtained similar parameter estimates and conclusions irrespective of which version was adopted.

References

- Aguas R, Gonçalves G, Ferreira MU, Gomes MGM (2021) Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. *medRxiv*<https://doi.org/10.1101/2020.07.23.20160762v4>
- Anderson RM, May RM (1991) *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press Inc., New York
- Anderson RM, Medley GF, May RM, Johnson AM (1986) A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA J Math Appl Med Biol* 3:229–263
- Ball F (1985) Deterministic and stochastic epidemic models with several kinds of susceptibles. *Adv Appl Probab* 17:1–22
- Bellan SE, Dushoff J, Galvani AP, Meyers LA (2015) Reassessment of HIV-1 Acute Phase Infectivity: Accounting for Heterogeneity and Study Design with Simulated Cohorts. *PLoS Med* 12:e1001801
- Britton T, Ball F, Trapman PA (2020) A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science* 369:846–849
- Corder RM, Ferreira MU, Gomes MGM (2020) Modelling the epidemiology of residual *Plasmodium vivax* malaria in a heterogeneous host population: a case study in the Amazon Basin. *PLoS Comput Biol* 16:e1007377
- Di Lauro F, Berthouze L, Dorey MD, Miller JC, Kiss IZ (2021) The Impact of Contact Structure and Mixing on Control Measures and Disease-Induced Herd Immunity in Epidemic Models: A Mean-Field Model Perspective. *Bull Math Biol* 83:1–25
- Diekmann O, Heesterbeek JAP, Britton T (2013) *Mathematical tools for Understanding Infectious Disease Dynamics*. Princeton University Press, Princeton, New Jersey
- Elbasha EH, Gumel AB (2021) Vaccination and herd immunity thresholds in heterogeneous populations. *J Math Biol* 83:239
- Ferrari MJ, Bansal S, Meyers LA, Bjornstad ON (2006) Network frailty and the geometry of herd immunity. *Proc R Soc B* 273:2743–2748
- Finkenstadt BF, Grenfell BT (2000) Time series modelling of childhood diseases: a dynamical systems approach. *Appl Statist* 49:187–205
- Fine P, Eames K, Heymann DL (2011) “Herd immunity”: a rough guide. *Clin Infect Dis* 52:911–916
- Gart JJ (1968) The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics* 24:557–566
- Gart JJ (1971) The statistical analysis of chain-binomial epidemic models with several kinds of susceptibles. *Biometrics* 28:921–930
- Gomes MGM, Ferreira MU, Corder RM, King JG, Souto-Maior C, Penha-Gonçalves C, Gonçalves G, Chikina M, Pegden W, Aguas R (2022) Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *J Theor Biol* 540:111063
- Gomes MGM, Gjini E, Lopes JS, Souto-Maior C, Rebelo C (2016) A theoretical framework to identify invariant thresholds in infectious disease epidemiology. *J Theor Biol* 395:97–102
- Gomes MGM, White LJ, Medley GF (2004) Infection, reinfection, and vaccination under suboptimal immune protection: Epidemiological perspectives. *J Theor Biol* 228:539–549
- Gonçalves G (2008) Herd immunity: recent uses in vaccine assessment. *Expert Rev Vaccines* 7:1493–1506
- Heesterbeek H et al (2015) Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347:aaa4339
- Kawagoe K, Rychnovsky M, Chang S, Huber G, Li LM, Miller J, Pnini R, Veytsman B, Yllanes D (2021) Epidemic dynamics in inhomogeneous populations and the role of superspreaders. *Phys Rev Res* 3:033283
- Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A* 115:700–721
- McKendrick AG (1939) The dynamics of crowd infection. *Edinb Med J* 47:117–136
- Miller JC, Slim AC, Volz EM (2012) Edge-based compartmental modelling for infectious disease spread. *J R Soc Interface* 9:890–906
- Neipel J, Bauermann J, Bo S, Harmon T, Jülicher F (2020) Power-Law population heterogeneity governs epidemic waves. *PLoS ONE* 15:e0239678
- Novozhilov AS (2008) On the spread of epidemics in a closed heterogeneous population. *Math Biosci* 215:177–185

- Pastor-Satorras R, Vespignani A (2001) Epidemic dynamics and endemic states in complex networks. *Phys Rev E* 63:066117
- Rose C, Medford AJ, Goldsmith CF, Vegge T, Weitz JS, Peterson AA (2021) Heterogeneity in susceptibility dictates the order of epidemic models. *J Theor Biol* 528:110839
- Ross R (1916) An application of the theory of probabilities to the study of a priori pathometry, Part I. *Philos Trans R Soc Lond A* 92:204–230
- Ross R, Hudson HH (1917) An application of the theory of probabilities to the study of a priori pathometry, Part II. *Philos Trans R Soc Lond A* 93:212–225
- Smith DL, Dushoff J, Snow RW, Hay SI (2005) The entomological inoculation rate and its relation to the prevalence of *Plasmodium falciparum* infection in African children. *Nature* 438:492–495
- Tkachenko AV, Maslov S, Elbanna A, Wong GN, Weiner ZJ, Goldenfeld N (2021) Time-dependent heterogeneity leads to transient suppression of the COVID-19 epidemic, not herd immunity. *Proc Natl Acad Sci USA* 118:e2015972118
- Woolhouse MEJ, Dye C, Etard J-F, Smith T, Charlwood JD, Garnett GP, Hagan P, Hii JLK, Ndhlovu PD, Quinell RJ, Watts CH, Chandiwana SK, Anderson RM (1997) Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proc Natl Acad Sci USA* 94:338–342

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.