

Epidemics with partial immunity to reinfection

Guy Katriel

Biomathematics Unit, Faculty of Life Sciences, Tel Aviv University, Israel

ARTICLE INFO

Article history:

Received 12 June 2010

Received in revised form 19 September 2010

Accepted 21 September 2010

Available online 27 September 2010

Keywords:

Epidemics

Partial immunity

Reinfection threshold

Final size formulas

ABSTRACT

We obtain analytical results about epidemics generated by the partial immunity model of Gomes et al. [3], in which infection confers partial immunity to reinfection. When the demographic process is excluded, the behavior switches from epidemic to endemic as the basic reproduction number \mathcal{R}_0 crosses the reinfection threshold $\mathcal{R}_0 = \frac{1}{\sigma}$. We derive formulas for two quantities characterizing the size of the epidemic below the reinfection threshold: the attack rate A , which is the fraction of the population infected at least once, and the final size Z , which is the average number of infections per individual. We also derive a system of differential equations which can be used to obtain more detailed information, such as the fraction of the population infected n times throughout the epidemic, for every n .

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

In this work we study epidemics generated in a model of an infection which induces partial immunity, introduced by Gomes et al. [3]. In this model, after recovering from infection, individuals gain partial immune protection, that is they can be infected again, but with a lower probability. Denoting the proportions of susceptible, infected, and recovered individuals in the population by S , I , R ($S + I + R = 1$) we thus have

$$S' = -\beta SI, \quad (1)$$

$$I' = \beta(S + \sigma R)I - \gamma I, \quad (2)$$

$$R' = -\sigma\beta RI + \gamma I, \quad (3)$$

where β measures the rate of transmission of the infection to susceptible individuals, $\sigma \in (0, 1)$ is the reduction in susceptibility of individuals who have been infected at least once before, and γ is the rate of recovery. We refer to [3,4] for a discussion of the relevance of this model to actual diseases.

We consider the Eqs. (1)–(3) together with initial conditions $S(0)$, $I(0)$, $R(0)$. $I(0)$ is the initial proportion of infected individuals in the population, which is usually very small, and indeed one is often most interested in the limit $I(0) \rightarrow 0$. $R(0)$ is the fraction of the population which is partially immune at the beginning of the epidemic, perhaps due to infection with the pathogen during a previous epidemic. We develop our results for general initial conditions, though in many cases it is useful to consider the case $R(0) = 0$ (no one partially immune before the epidemic), and sometimes also $I(0) = 0$, in which some of the results assume a simpler form.

Gomes et al. [3], also included births and deaths at rate μ in the equations:

$$S' = \mu(1 - S) - \beta SI, \quad (4)$$

$$I' = \beta(S + \sigma R)I - (\gamma + \mu)I, \quad (5)$$

$$R' = -\sigma\beta RI + \gamma I - \mu R. \quad (6)$$

Our aim here is to study the dynamics *without* the birth and death terms (that is, the case $\mu = 0$). We thus decouple the epidemic process from longer time-scale demographic dynamics. One advantage of doing so is that the concept of ‘reinfection threshold’, introduced in [3], becomes particularly clear. Indeed, setting

$$\mathcal{R}_0 = \frac{\beta}{\gamma},$$

we have that, when $\mathcal{R}_0 > \frac{1}{\sigma}$ one has an endemic equilibrium

$$S^* = 0, \quad I^* = 1 - \frac{1}{\sigma\mathcal{R}_0}, \quad R^* = \frac{1}{\sigma\mathcal{R}_0}, \quad (7)$$

which will be proved to be globally stable. On the other hand when $\mathcal{R}_0 < \frac{1}{\sigma}$ an endemic equilibrium does not exist. Thus $\mathcal{R}_0 = \frac{1}{\sigma}$ is the reinfection threshold, above which the disease transmission is strong enough to sustain the disease endemically, even *without* the replenishment of susceptibles provided by births.

Our main focus will be on what occurs below the reinfection threshold, that is in the case $0 < \mathcal{R}_0 < \frac{1}{\sigma}$. Two situations can arise: if $\mathcal{R}_0 > 1$, then introduction of an arbitrarily small fraction of infected individuals into a totally susceptible population will lead to an epidemic, which will eventually die out. If $\mathcal{R}_0 \leq 1$, no epidemic will occur. These basic facts are proved in Section 2. In Fig. 1 we plot the epidemic curves for $\mathcal{R}_0 = 1.5$ and three values

E-mail address: haggaiika@yahoo.com

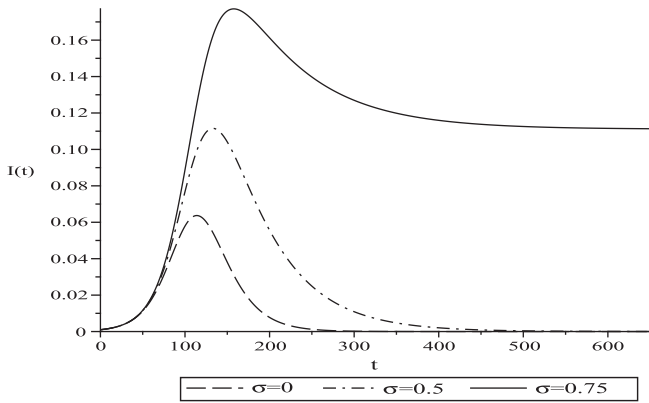


Fig. 1. Graphs $I(t)$ for $\gamma = 0.1, \beta = 0.15(\mathcal{R}_0 = 1.5)$, for $\sigma = 0, 0.5, 0.75, S(0) = 0.999, I(0) = 0.001, R(0) = 0$.

of σ . When $\sigma = 0, 0.5$ we are below the reinfection threshold, so that we have an epidemic which dies down, while in the case $\sigma = 0.75 > \frac{2}{3} = \frac{1}{\mathcal{R}_0}$ we are above the reinfection threshold, and we see $I(t)$ approaching the endemic state $I = \frac{1}{9}$.

In Section 3 we will study the size of the epidemic and its dependence on the parameters \mathcal{R}_0, σ . For the present model, there are in fact two distinct quantities which measure the size of the epidemic: the attack rate A , which is the fraction of the population that is infected during the epidemic, and the final size Z , which is the average number of infections per individual in the population. For the standard SIR model ($\sigma = 0$), in which each individual can be infected only once, these two quantities coincide, but here, due to reinfections, we have $Z > A$. We obtain formulas for A and Z in terms of \mathcal{R}_0, σ , and the initial conditions of the epidemic. These formulas are different from the ‘standard’ final-size relation which is valid for a large variety of epidemiological models [6]. Of course, as we shall show explicitly, when $\sigma \rightarrow 0$, so that the model approaches the standard SIR model, the values of A, Z given by the formulas developed here approach the value given by the final-size formula for the SIR model.

In Section 4 we show that it is possible to obtain more detailed information about the epidemics generated in the model considered here, for example to study the size of the population which is infected precisely n times, for $n \geq 1$, by formulating a system of differential equations for variables which count the reinfected populations.

When demography is included in the model, as in (4)–(6), then an endemic equilibrium exists for any $\mathcal{R}_0 > 1$, that is both below and above the reinfection threshold. However, as Gomes et al. [3] have shown, the reinfection threshold is still manifested in that the fraction of the population infected in the endemic state sharply increases near the reinfection threshold. The fact that in the model including demography the reinfection threshold does not correspond to an actual bifurcation has led Breban and Blower [2] to state that ‘the reinfection threshold does not exist’. As pointed out in [5], the disagreement here is one of terminology. In any case, in the model without demography considered here, the reinfection threshold does appear as bifurcation in the strict sense, from epidemic dynamics to endemic dynamics, as noted above.

The study of the epidemic in the case $1 < \mathcal{R}_0 < \frac{1}{\sigma}$ for the model without demography (1)–(3) is highly relevant also for understanding the behavior of the model with demography (4)–(6) in the case of a newly-emerging disease. Since the demographic process typically occurs on a much longer time scale than the epidemic process, that is $\mu \ll \gamma, \beta$, upon introduction of a pathogen into the population the infection curve for the model with demography will be virtually identical to that of the model without

demography, and only much later will the difference between the two models be manifested, in that in the model with demography the infection curve will approach the endemic equilibrium while in the model without demography it will decay to 0.

2. The two thresholds

We begin with a result that is key to understanding the behavior of (1)–(3), and to deriving the final-size formulas in Section 3.

Lemma 1. Define:

$$V(S, R) = \left[\frac{1}{\mathcal{R}_0 \sigma} - R \right] \frac{1}{S^\sigma}.$$

Then V is a constant of motion for (1)–(3), that is, $V(S(t), R(t))$ is constant in time for any solution $(S(t), I(t), R(t))$ of (1)–(3).

Proof. Using (1) and (3) we have

$$\begin{aligned} [V(S, R)]' &= -\sigma \left[\frac{1}{\mathcal{R}_0 \sigma} - R \right] S^{-\sigma-1} S' - S^{-\sigma} R' \\ &= \beta \sigma \left[\frac{1}{\mathcal{R}_0 \sigma} - R \right] S^{-\sigma-1} SI - S^{-\sigma} [-\sigma \beta RI + \gamma I] = 0. \quad \square \end{aligned}$$

The expression $V(S, R)$ was ‘pulled out of the hat’, and shown to have the required property. Let us show how this expression can be found. Dividing (1) by (3) we obtain

$$\begin{aligned} \frac{1}{S} \frac{dS}{dR} &= \frac{1}{\sigma R - \mathcal{R}_0^{-1}} \Rightarrow \frac{d}{dR} \log(S) = \frac{1}{\sigma R - \mathcal{R}_0^{-1}} \\ \Rightarrow \log(S) &= \frac{1}{\sigma} \log(\sigma R - \mathcal{R}_0^{-1}) + c \Rightarrow \left[R - \frac{1}{\mathcal{R}_0 \sigma} \right] \frac{1}{S^\sigma} = C, \end{aligned}$$

leading to $V(S, R)$.

Theorem 1. Assume $0 < \sigma < 1$. Let $S(t), I(t), R(t)$ be the solution of (1)–(3) with initial conditions $I(0) > 0, S(0) + I(0) + R(0) = 1$.

Then

(i) (Endemic case). If $\mathcal{R}_0 > \frac{1}{\sigma}$ then,

$$\lim_{t \rightarrow \infty} (S(t), I(t), R(t)) = \left(0, 1 - \frac{1}{\sigma \mathcal{R}_0}, \frac{1}{\sigma \mathcal{R}_0} \right). \quad (8)$$

(ii) If $\mathcal{R}_0 \leq \frac{1}{\sigma}$ then $\lim_{t \rightarrow \infty} I(t) = 0$, and,

(a) (Epidemic case). If $[S(0) + \sigma R(0)] \mathcal{R}_0 > 1$, $I(t)$ increases monotonically for $t \in [0, t^*]$, where $t^* > 0$, up to the maximal value,

$$I(t^*) = 1 - \frac{1}{\mathcal{R}_0 \sigma} + \left(\frac{1}{\sigma} - 1 \right) \left[\frac{\mathcal{R}_0^{-1} - \sigma R(0)}{S(0)^\sigma} \right]^{\frac{1}{1-\sigma}},$$

and then decreases monotonically to 0 for $t \in [t^*, \infty)$.

(b) (No epidemic). In case $[S(0) + \sigma R(0)] \mathcal{R}_0 \leq 1, I(t)$ decreases monotonically to 0.

Note that if we take $R(0) = 0$, that is no individuals are (partially) immune at the beginning, then the threshold condition in part (ii) of the above theorem is $S_0 \mathcal{R}_0 > 1$, and if moreover we take $I_0 \rightarrow 0$, that is $S_0 = 1$, we get the familiar threshold condition $\mathcal{R}_0 > 1$.

Part (ii) of the theorem shows, in particular, that the epidemic curve is unimodal, that is it has a unique local maximum which is the global maximum. Thus this model never generates a ‘second wave’ of the epidemic, despite the reinfections.

Proof. We first note that (1) implies that $S(t)$ is strictly decreasing, and bounded from below by 0. Therefore the limit

$$S(\infty) = \lim_{t \rightarrow \infty} S(t), \tag{9}$$

exists.

From Lemma 1 we have that $V(S(t), R(t)) = V(S(0), R(0))$ for all t , that is

$$\left[\frac{1}{\mathcal{R}_0 \sigma} - R(t) \right] \frac{1}{S^\sigma(t)} = \left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \frac{1}{S(0)^\sigma},$$

or

$$R(t) = \frac{1}{\mathcal{R}_0 \sigma} - \left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \left(\frac{S(t)}{S(0)} \right)^\sigma, \tag{10}$$

which, together with (9), implies that the limit

$$R(\infty) = \lim_{t \rightarrow \infty} R(t),$$

exists. Since $S + I + R = 1$, (10) implies

$$I(t) = 1 - \frac{1}{\mathcal{R}_0 \sigma} + \left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \left(\frac{S(t)}{S(0)} \right)^\sigma - S(t), \tag{11}$$

so that the limit

$$I(\infty) = \lim_{t \rightarrow \infty} I(t), \tag{12}$$

also exists. Clearly $(S(\infty), I(\infty), R(\infty))$ must be an equilibrium point of (1)–(3).

(i) Assume now that $\mathcal{R}_0 > \frac{1}{\sigma}$. We first prove that,

$$I(\infty) > 0. \tag{13}$$

Assume by way of contradiction that $I(\infty) = 0$. Using (2) we then have

$$\begin{aligned} \lim_{t \rightarrow \infty} [\log(I(t))] &= \lim_{t \rightarrow \infty} \beta[S(t) + \sigma R(t)] - \gamma \\ &= \beta[S(\infty) + \sigma R(\infty) - \mathcal{R}_0^{-1}] \\ &> \beta[S(\infty) + \sigma R(\infty) - \sigma] = \beta(1 - \sigma)S(\infty) \geq 0, \end{aligned}$$

which implies that $\log(I(t)) \rightarrow \infty$, hence $I(t) \rightarrow \infty$, as $t \rightarrow \infty$, a contradiction. We thus have (13). Since there is a unique equilibrium (S^*, I^*, R^*) of (4) and (5) with $I^* > 0$, given by (7), we have (8).

(ii) Assume now that $1 < \mathcal{R}_0 \leq \frac{1}{\sigma}$. Then, since we have already shown that $(S(t), I(t), R(t))$ converges to an equilibrium, and since the only equilibria are those with $I = 0$, we must have $I(\infty) = 0$.

Now:

(a) Assume $[S(0) + \sigma R(0)]\mathcal{R}_0 > 1$. From (2) we have,

$$\begin{aligned} I'(0) &= \beta(S(0) + \sigma R(0))I(0) - \gamma I(0) \\ &= \beta[S(0) + \sigma R(0) - \mathcal{R}_0^{-1}]I(0) > 0, \end{aligned}$$

so that $I(t)$ is increasing for sufficiently small t . Since $I(\infty) = 0$, this implies that $I(t)$ has a maximum point. To show that this maximum point is unique, we note that (11) can be written as $I(t) = f(S(t))$ where

$$f(x) = 1 - \frac{1}{\mathcal{R}_0 \sigma} + \left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \left(\frac{x}{S(0)} \right)^\sigma - x.$$

Since $f(x)$ is a concave function it has a unique maximum point, and since $S(t)$ is monotone decreasing this implies that the same is true for $I(t) = f(S(t))$. The maximum point x^* of f can be found explicitly by solving $f'(x^*) = 0$, obtaining

$$x^* = \left[\frac{\mathcal{R}_0^{-1} - \sigma R(0)}{S(0)^\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{14}$$

Therefore the maximal value of $I(t)$ is given by

$$\begin{aligned} I(t^*) &= f(S(t^*)) = f(x^*) \\ &= 1 - \frac{1}{\mathcal{R}_0 \sigma} + \left(\frac{1}{\sigma} - 1 \right) \left[\frac{\mathcal{R}_0^{-1} - \sigma R(0)}{S(0)^\sigma} \right]^{\frac{1}{1-\sigma}}. \end{aligned}$$

(b) Assume $[S(0) + \sigma R(0)]\mathcal{R}_0 \leq 1$. This implies that $S(0) \leq x^*$, where x^* is given by (14). Therefore, since $S(t)$ is decreasing and f is a decreasing function (f is concave) we have,

$$f'(S(t)) \geq f'(S(0)) \geq f'(x^*) = 0, \tag{15}$$

hence

$$I'(t) = f'(S(t))S'(t) \leq 0,$$

for all $t > 0$. \square

3. Size of the epidemic below the reinfection threshold

It now becomes interesting to study the size of the epidemic in the case $1 < \mathcal{R}_0 < \frac{1}{\sigma}$. In fact there are two quantities which characterize the epidemic size:

- Attack rate A – the total fraction of the population ever infected during the epidemic.
- Final size Z – total number of infections divided by the population (in other words, the average number of infections per individual).

In the SIR model ($\sigma = 0$) these quantities coincide, but in a model with reinfections we have $A < Z$ since Z counts individuals according to their number of infections. In particular, while $A < 1$ always, Z can be greater than 1.

The quantities A, Z characterizing the size of the epidemic below the reinfection threshold should be distinguished from the equilibrium value of the fraction of infectives in the model (4)–(6) including births, as discussed in [3], obtained by setting the left-hand sides of (4)–(6) to 0 and solving these algebraic equations. In the case $\mu = 0$, which is considered here, the equilibrium value of I would be 0, since there is no endemic equilibrium. In the case $\mu > 0$ the attack rate A is not well-defined since the infection is endemic even below the reinfection threshold, unlike in the case $\mu = 0$ for which a decaying epidemic occurs. However, the quantities A, Z computed for the case $\mu = 0$ are still of relevance in the case that there are births ($\mu > 0$): while the equilibrium value of I will describe the fraction of the population which is infected at each point in time after the endemic equilibrium has been established, the quantity A will describe, to a very good approximation, the size of the epidemic that will occur when the pathogen is first introduced into the population. Consider for example the case $\gamma = 0.1, \beta = 0.15, \sigma = 0.5$, which, assuming time is measured in days, describes an epidemic with an infectious period of 10 days and $\mathcal{R}_0 = 1.5$, with the susceptibility of those recovered reduced by 50%. We are below the reinfection threshold, so that in the case $\mu = 0$, and starting from a totally susceptible population, we get an epidemic which is plotted as the middle curve in Fig. 1. The attack rate for this epidemic is 88.8% which can be computed from the Eq. (17) below, or by direct simulation. If we now consider the case of a birth rate of 5% a year (that is $\mu = \frac{0.05}{365}$), the curve $I(t)$ will look virtually identical to the curve for the case $\mu = 0$ over the time span of two years plotted in Fig. 1, with 85.6% of the population infected after 500 days, close to the value A obtained in the case $\mu = 0$. The very low endemic equilibrium, in which $I = 0.0018$, is only established after many years, after a series of oscillations (smaller epidemics). Thus for the model (4)–(6) with $\mu > 0$, below the reinfection threshold, the size of the large initial epidemic is

predicted by the value A , and the fraction of infectives at each point in time after a long time is predicted by the equilibrium value of I .

We now derive the precise formulas for A and Z for the model (1)–(3).

Theorem 2. Assume $0 < \sigma < 1$, $I(0) > 0$ and

$$1 < \mathcal{R}_0 < \frac{1}{\sigma}. \tag{16}$$

(i) The attack rate A is the unique solution $A \in (0, 1)$ of the equation

$$1 - \mathcal{R}_0 \sigma A = \left[\frac{1 - R(0) - (1 - \mathcal{R}_0 \sigma R(0))A}{S(0)} \right]^\sigma. \tag{17}$$

(ii) The final size Z is given by

$$Z = \frac{1}{\mathcal{R}_0 \sigma} \log \left(\frac{1}{1 - \mathcal{R}_0 \sigma A} \right). \tag{18}$$

Proof. (i) We decompose the fraction A of the population infected during the epidemic into three components

$$A = I(0) + A_S + A_R, \tag{19}$$

where $I(0)$ is the fraction of the population that was infected at the beginning of the epidemic, A_S is the fraction of the population consisting of individuals who were initially susceptible and were infected during the epidemic, and A_R is the fraction of the population consisting of individuals who were initially partially immune and were infected during the epidemic. $I(0)$ is given, so we need to compute A_S and A_R . We first note that

$$A_S = S(0) - S(\infty), \tag{20}$$

that is the fraction of susceptibles infected is obtained by subtracting the fraction of susceptibles after the epidemic from that before the epidemic. From (11) we have

$$I(\infty) = \left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \frac{S(\infty)^\sigma}{S(0)^\sigma} + 1 - \frac{1}{\mathcal{R}_0 \sigma} - S(\infty), \tag{21}$$

and since, by (16), we have, from Theorem 1, $I(\infty) = 0$, we get

$$\left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \frac{S(\infty)^\sigma}{S(0)^\sigma} - S(\infty) = \frac{1}{\mathcal{R}_0 \sigma} - 1. \tag{22}$$

From (20) and (22) we have

$$\left[\frac{1}{\mathcal{R}_0 \sigma} - R(0) \right] \left[1 - \frac{A_S}{S(0)} \right]^\sigma + A_S = S(0) + \frac{1}{\mathcal{R}_0 \sigma} - 1, \tag{23}$$

so we have an equation satisfied by A_S .

We now compute A_R , the fraction of the population consisting of individuals who are partially immune at time $t=0$ and who become infected during the epidemic. Note that, in contrast with the case of A_S , we cannot compute A_R as the difference of the fraction of individuals in the class R between times $t=0$ and $t=\infty$, since, in contrast to the class S , the class R is replenished by individuals recovering from infection. To calculate A_R , we denote by $R_0(t)$ the fraction of the population consisting of those who are partially immune at time $t=0$ and who remain partially immune at time t . Since these individuals are infected at rate $\sigma\beta I(t)$, the equation describing the evolution of $R_0(t)$ is

$$R'_0 = -\beta\sigma I(t)R_0, R_0(0) = R(0). \tag{24}$$

We then have

$$A_R = R(0) - R_0(\infty). \tag{25}$$

Solving (24) we have

$$R_0(t) = R(0)e^{-\beta\sigma \int_0^t I(s)ds},$$

and taking $t \rightarrow \infty$ we have

$$R_0(\infty) = R(0)e^{-\beta\sigma \int_0^\infty I(t)dt}. \tag{26}$$

From (1) we get

$$\log(S(\infty)) - \log(S(0)) = \int_0^\infty \frac{S'(t)}{S(t)} dt = -\beta \int_0^\infty I(t)dt,$$

so that, together with (20) we have

$$\int_0^\infty I(t)dt = \frac{1}{\beta} \log \left(\frac{S(0)}{S(0) - A_S} \right). \tag{27}$$

Substituting (27) into (26) we get

$$R_0(\infty) = R(0)e^{-\sigma \log \left(\frac{S(0)}{S(0) - A_S} \right)} = R(0) \left(1 - \frac{A_S}{S(0)} \right)^\sigma,$$

and together with (25) we have

$$A_R = R(0) \left[1 - \left(1 - \frac{A_S}{S(0)} \right)^\sigma \right]. \tag{28}$$

We now note that (23) can be rewritten as

$$1 - \left(1 - \frac{A_S}{S(0)} \right)^\sigma = \mathcal{R}_0 \sigma \frac{A_S + I(0)}{1 - \mathcal{R}_0 \sigma R(0)},$$

from which it follows that (28) can be rewritten as

$$A_R = \mathcal{R}_0 \sigma R(0) \frac{A_S + I(0)}{1 - \mathcal{R}_0 \sigma R(0)}.$$

Thus, going back to (19) we have

$$A = I(0) + A_S + \mathcal{R}_0 \sigma R(0) \frac{A_S + I(0)}{1 - \mathcal{R}_0 \sigma R(0)} = \frac{A_S + I(0)}{1 - \mathcal{R}_0 \sigma R(0)}. \tag{29}$$

Solving (29) for A_S in terms of A , we have

$$A_S = A(1 - \mathcal{R}_0 \sigma R(0)) - I(0), \tag{30}$$

and substituting (30) into (23) we finally get the Eq. (17) for A .

Let us show that (17) has a unique solution $A \in (0, 1)$. Setting

$$g(A) = \left[\frac{1 - R(0) - (1 - \mathcal{R}_0 \sigma R(0))A}{S(0)} \right]^\sigma + \mathcal{R}_0 \sigma A - 1,$$

so that the g is well-defined for

$$A \leq \frac{1 - R(0)}{1 - \mathcal{R}_0 \sigma R(0)}, \tag{31}$$

we need to show that $g(A)$ has a unique positive zero. We have,

$$g(0) = \left[\frac{1 - R(0)}{S(0)} \right]^\sigma - 1 = \left[1 + \frac{I(0)}{S(0)} \right]^\sigma - 1 > 0, \tag{32}$$

and, using (16),

$$g \left(\frac{1 - R(0)}{1 - \mathcal{R}_0 \sigma R(0)} \right) = \frac{\mathcal{R}_0 \sigma - 1}{1 - \mathcal{R}_0 \sigma R(0)} < 0.$$

The above and the intermediate value theorem imply that g has positive zero (of course this fact does not really need proving, since we have already shown that the attack rate A is a solution). To show that the positive zero of g is unique, we note first that, since $\sigma < 1$, we have $\lim_{A \rightarrow -\infty} g(A) = -\infty$, which together with (32) implies that g has a negative zero. However, since it is easily seen that g is concave, it has at most two zeros, and since we have proven the existence of a positive and a negative zero, we see that the positive zero is unique.

(ii) The fraction of the population infected during an infinitesimal time interval $[t, t + dt]$ is $\beta(S(t) + \sigma R(t))I(t)dt$, so that the number of infections divided by the total population is

$$Z = \beta \int_0^\infty (S(t) + \sigma R(t))I(t)dt,$$

and using (2) we get

$$Z = \gamma \int_0^\infty I(t)dt, \tag{33}$$

hence from (27) and (33) we get

$$Z = \frac{1}{\mathcal{R}_0} \log \left(\frac{S(0)}{S(0) - A_S} \right). \tag{34}$$

Substituting (30) into (34), we get

$$Z = \frac{1}{\mathcal{R}_0} \log \left(\frac{S(0)}{1 - R(0) - A(1 - \mathcal{R}_0\sigma R(0))} \right), \tag{35}$$

and since, by (17), we have

$$\frac{S(0)}{1 - R(0) - A(1 - \mathcal{R}_0\sigma R(0))} = (1 - \mathcal{R}_0\sigma A)^{-\frac{1}{\sigma}},$$

(35) can be rewritten as (18). \square

A particularly relevant case is when we start with an infinitesimal fraction of infective individuals. Of course if we take $I(0) = 0$ as the initial condition in (1)–(3), the solution will be $I(t) = 0$. However, in (17) the limit $I(0) \rightarrow 0$ is well-defined and obtained by setting $I(0) = 0$, that is $R(0) = 1 - S(0)$, giving

$$1 - \mathcal{R}_0\sigma A = \left[1 - \frac{1 - \mathcal{R}_0\sigma R(0)}{1 - R(0)} A \right]^\sigma. \tag{36}$$

Note that this equation always has the trivial solution $A = 0$, but we are interested in positive solutions. We show below that there exists a non-zero solution of (36) if and only if $[S(0) + \sigma R(0)]\mathcal{R}_0 > 1$. In particular, in the case of an initially totally susceptible population $S(0) = 1$, we have the condition $\mathcal{R}_0 > 1$ for an epidemic when the initial number of infectives is infinitesimal, the well-known threshold, and since in this case $R(0) = 0$, (36) then reduces to

$$A = \frac{1}{\mathcal{R}_0\sigma} (1 - [1 - A]^\sigma).$$

Proposition 1. Assume (16), and $I(0) = 0$. Then (36) has a solution $A \in (0, 1)$ if and only if $[S(0) + \sigma R(0)]\mathcal{R}_0 > 1$, and this solution is unique.

Proof. Set

$$h(A) = \left[1 - \frac{1 - \mathcal{R}_0\sigma R(0)}{1 - R(0)} A \right]^\sigma + \mathcal{R}_0\sigma A - 1,$$

defined for A satisfying (31). Then $h(0) = 0$, and using (16),

$$h \left(\frac{1 - R(0)}{1 - \mathcal{R}_0\sigma R(0)} \right) = \frac{\mathcal{R}_0\sigma - \mathcal{R}_0\sigma R(0)}{1 - \mathcal{R}_0\sigma R(0)} - 1 < 0. \tag{37}$$

Now, using the assumption $I(0) = 0$,

$$h'(0) = \mathcal{R}_0\sigma - \sigma \frac{1 - \mathcal{R}_0\sigma R(0)}{1 - R(0)} = \sigma \frac{[S(0) + \sigma R(0)]\mathcal{R}_0 - 1}{S(0)}.$$

Thus if $[S(0) + \sigma R(0)]\mathcal{R}_0 > 1$ then $h'(0) > 0$, so that $h(A)$ is positive for $A > 0$ sufficiently small, so, together with (37), we conclude that it has a positive zero. Since h is concave it can have at most two zeros, and since $A = 0$ is a zero, the non-trivial zero we have found is unique.

Assuming now that $[S(0) + \sigma R(0)]\mathcal{R}_0 \leq 1$, we have $h'(0) \leq 0$, and since h is concave this implies that $h'(A) < 0$ for all $A > 0$ satisfying (31), so that h is decreasing, and since $h(0) = 0$ there is no positive zero. \square

Plots of A and Z as functions of \mathcal{R}_0 , for $S(0) = 1, I(0) \rightarrow 0, \sigma = \frac{1}{3}$, made using (36) and (18), are shown in Fig. 2. At the reinfection threshold, that is when $\sigma\mathcal{R}_0 = 1$, the solution of (17) is $A = 1$, and (18) then gives $Z = +\infty$. Thus as \mathcal{R}_0 approaches the reinfection threshold $\frac{1}{\sigma}$, A approaches 1, and Z approaches $+\infty$. This is in accordance with the fact that above the reinfection threshold we are in the endemic situation in which infection persists forever, so that everyone eventually gets reinfected infinitely many times, and we have $A = 1, Z = +\infty$, although at each point in time only a certain fraction of the population, approaching the endemic value $I^* = 1 - \frac{1}{\sigma\mathcal{R}_0}$ (see (7)) is infected. Thus, in the version of the model considered here, in which replenishment of the susceptible pool by new births is ignored, the reinfection threshold is manifested in a different way than in the original model of Gomes et al. [3] (4)–(6), in which an endemic equilibrium exists both below and above the reinfection threshold. While in the model including births the reinfection threshold appears as a sharp (though continuous) increase in the equilibrium value of the fraction of the population infected as one crosses the reinfection threshold, in the model considered here the reinfection threshold appears as a saturation of the attack rate at the value $A = 1$, and the ‘explosion’ in the average number of infections per individual $Z \rightarrow \infty$, concurrently with the ‘birth’ of an endemic equilibrium (7).

It is interesting to connect the final size formula (17), which is valid for $0 < \sigma < 1$, with the familiar final-size formula for the SIR model ($\sigma = 0$) [6]. Taking logs on both sides of (17) we have

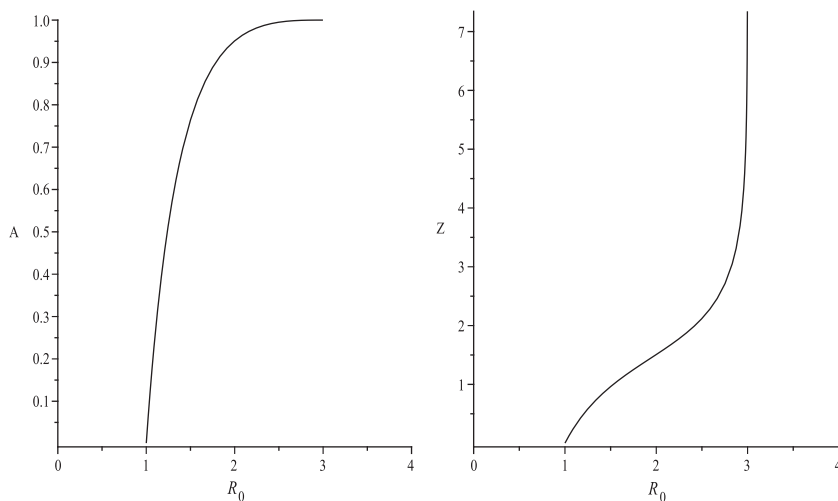


Fig. 2. A and Z as functions of \mathcal{R}_0 , for $S(0) = 1, I(0) = 0, \sigma = \frac{1}{3}$.

$$\frac{1}{\sigma} \log(1 - \mathcal{R}_0 \sigma A) = \log \left[\frac{1 - R(0) - (1 - \mathcal{R}_0 \sigma R(0))A}{S(0)} \right], \quad (38)$$

and taking the limit $\sigma \rightarrow 0$, noting that $\lim_{\sigma \rightarrow 0} \frac{1}{\sigma} \log(1 - \mathcal{R}_0 \sigma A) = -\mathcal{R}_0 A$, we get

$$-\mathcal{R}_0 A = \log \left[\frac{1 - R(0) - A}{S(0)} \right],$$

or equivalently

$$A = I(0) + S(0)(1 - e^{-\mathcal{R}_0 A}),$$

and in the particular case $I(0) = 0, S(0) = 1$ we get the familiar formula

$$A = 1 - e^{-\mathcal{R}_0 A}. \quad (39)$$

Note also that taking the limit $\sigma \rightarrow 0$ in (18) leads to $Z = A$, as expected.

4. Counting reinfections

We have considered above the total number of infected individuals, as well as the total number of infections. There are, however, some more detailed questions that can be asked. For example:

- At a particular point in the course of the epidemic, what is the number of individuals who are infected for the first, second, third, (etc.) times?
- At the end of the epidemic, how is the population divided among those have never been infected, those who have been infected once, twice, etc.?

To attempt to answer such questions, we formulate differential equations for the following time-dependent quantities:

- $I_n(t)$ ($n = 1, 2, \dots$) is the fraction of the population which is infected for the n th time.
- $R_n(t)$ ($n = 1, 2, \dots$) is the fraction of the population which is recovered (and thus partially immune) and which has been infected n times in the past.
- $R_0(t)$ is the fraction of the population which was partially immune at time $t = 0$ and has not been infected so far.

These quantities satisfy the following differential equations:

$$I'_1 = \beta(S + \sigma R_0)I - \gamma I_1, \quad (40)$$

$$I'_n = \sigma \beta R_{n-1} I - \gamma I_n, n \geq 2, \quad (41)$$

$$R'_0 = -\sigma \beta R_0 I, \quad (42)$$

$$R'_n = -\sigma \beta R_n I + \gamma I_n, n \geq 1, \quad (43)$$

where S, I are the solutions of (1) and (2).

We explain each of these equations in turn:

- (i) Eq. (40) describes the compartment I_1 consisting of individuals infected for the first time. The inflow into this compartment consists of susceptibles who get infected for the first time, given by βSI , and of partially immune individuals who get infected for the first time, given by $\sigma \beta R_0 I$. The outflow from this compartment consists of those recovering from their first infection, given by γI_1 .
- (ii) Eq. (41) describes, for each $n \geq 2$, the compartment I_n of individuals infected for the n th time. The inflow into this compartment consists of partially immune individuals who have been infected $n - 1$ times before, given by $\sigma \beta R_{n-1} I$. The outflow consists of recoveries.
- (iii) Eq. (42) describes the compartment consisting of individuals who were partially immune at the beginning of the epi-

demical, and there is only outflow from this compartment, consisting of partially immune individuals being infected for the first time, given by $\sigma \beta R_0 I$.

- (iv) Eq. (43) describes the compartment consisting of individuals who have recovered from their n th infection ($n \geq 1$). The inflow into this compartment comes from the recoveries coming from compartment I_n , given by γI_n , and the outflow consists of individuals being reinfected for the $n + 1$ th time, given by $\sigma \beta R_n I$.

Let us note that the system (40)–(43) is similar to a system of equations studied by Abu-Raddad & Ferguson [1] (Eq. 3.1), which models a population in which many different strains of a pathogen are circulating, and in which the susceptibility of an individual to a new strain depends only on the number of strains that the individual has already been infected with. In that context I_n denotes the fraction of the population that has been infected with n strains in the past. In [1] the focus is on studying the long-time equilibrium of the system, with susceptibles replenished by births. Here we use (40)–(43) to examine the epidemic dynamics, in the case that there is no endemic equilibrium.

The initial conditions for I_n, R_n are

$$I_1(0) = I(0), \quad I_n(0) = 0, \quad n \geq 2, \quad (44)$$

$$R_0(0) = R(0), \quad R_n(0) = 0, \quad n \geq 1. \quad (45)$$

As a consistency check, let us set

$$\hat{I}(t) = \sum_{n=1}^{\infty} I_n(t) \quad \hat{R}(t) = \sum_{n=0}^{\infty} R_n(t),$$

and use (40)–(43) to compute

$$\begin{aligned} \hat{I}' &= \left(\sum_{n=1}^{\infty} I_n(t) \right)' = \beta(S + \sigma R_0)I - \gamma I_1 + \sum_{n=2}^{\infty} [\sigma \beta R_{n-1} I - \gamma I_n] \\ &= \beta I(S + \sigma \hat{R}) - \gamma \hat{I}, \end{aligned}$$

$$\hat{R}' = \left(\sum_{n=0}^{\infty} R_n(t) \right)' = -\sigma \beta R_0 I + \sum_{n=1}^{\infty} [-\sigma \beta R_n I + \gamma I_n] = -\sigma \beta I \hat{R} + \gamma \hat{I},$$

and from (44) and (45) we have $\hat{I}(0) = I(0), \hat{R}(0) = R(0)$, and we see that \hat{I}, \hat{R} satisfy (1) and (3) with the same initial conditions as I and R , so by uniqueness of solutions to initial value problems for ODE's we have $\hat{I}(t) = I(t), \hat{R}(t) = R(t)$. Thus we have shown that

$$I(t) = \sum_{n=1}^{\infty} I_n(t), \quad R(t) = \sum_{n=0}^{\infty} R_n(t).$$

Note that (40)–(43) is an infinite system of linear equations with time-varying coefficients S, I , where S, I are given as the solutions of (1)–(3) with some initial conditions $S(0), I(0)$. Since the equations at level n depend only on the variables $I_k, 1 \leq k \leq n$ and $R_k, 0 \leq k \leq n$, this is a triangular system, and thus one can easily solve it numerically up to any order n . In Fig. 3 we plotted the curves $I_n(t)$, computed numerically by solving the system (1)–(3) simultaneously with (40)–(43), for n up to 20 (however only the first few are large enough to be visible in the plot). The epidemic is thus built up as a sum of a series of epidemics $I_n(t)$, each of which peaks later than the previous one. It is interesting to note that although in general the sum of several curves with peaks at different locations can be multi-peaked, in the case of the current model, as we have proved in Section 2, the curve $I(t)$ has a unique peak, so that the reinfection mechanism described by this model does not give rise to 'second waves'.

It is of interest to compute the fraction of the population A_n which is infected *exactly* n times during the epidemic. For $n \geq 1$ this is the fraction of the population which remains in the R_n compartment at $t = \infty$

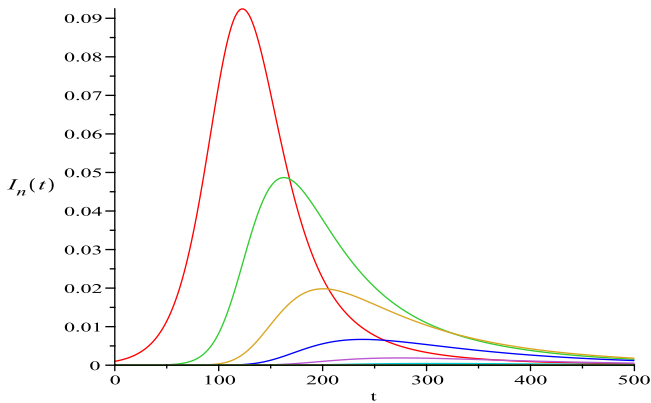


Fig. 3. Fraction of the population $I_n(t)$ infected for the n th time ($n = 1, 2, \dots$), $\beta = 0.15, \gamma = 0.1 (\mathcal{R}_0 = 1.5), \sigma = 0.6$, for $S(0) = 0.999, I(0) = 0.001$.

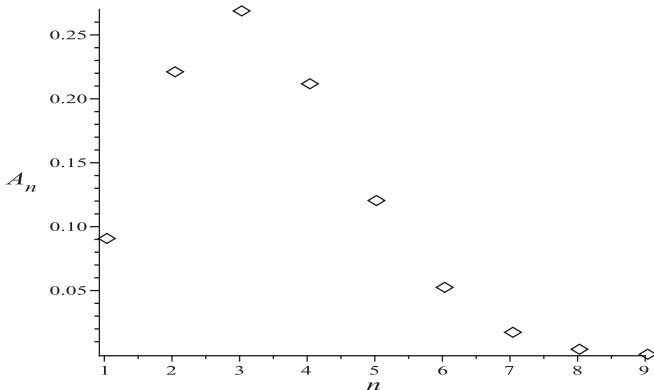


Fig. 4. Fraction of the population infected n times throughout the epidemic, for $\beta = 2, \gamma = 1 (\mathcal{R}_0 = 2), \sigma = 0.48. S(0) = 0.999, I(0) = 0.001, R(0) = 0$.

$$A_n = \lim_{t \rightarrow \infty} R_n(t).$$

By numerically integrating the coupled system (1)–(3), (40)–(43) up to a high value of t , we can calculate A_n . In Fig. 4 we show results of such a calculation. Given statistical data on infections and reinfections, one could use results like these to examine whether the model (1)–(3) can account for the epidemic, and to fit the parameters.

In Fig. 5 we examine the fraction A_1 of the population infected precisely once during an epidemic as a function of the parameter σ , plotting several such curves for different values of \mathcal{R}_0 , always starting with a nearly totally susceptible population, and computing A_1 numerically by integrating the coupled system (1)–(3), (40)–(43) up to a high value of t . Note that while the total attack rate A always increases with σ (as is intuitively obvious and can also be shown analytically using (17)), the effect of changing σ on A_1 is not obvious, since increasing σ also increases the number of individuals infected more than once, which has a negative effect on A_1 . What we discover is that there are two possible behaviors of the curve $A_1 = A_1(\sigma)$, depending on the value of \mathcal{R}_0 . For \mathcal{R}_0 sufficiently large, like the case $\mathcal{R}_0 = 2$ in the plot, the number of people

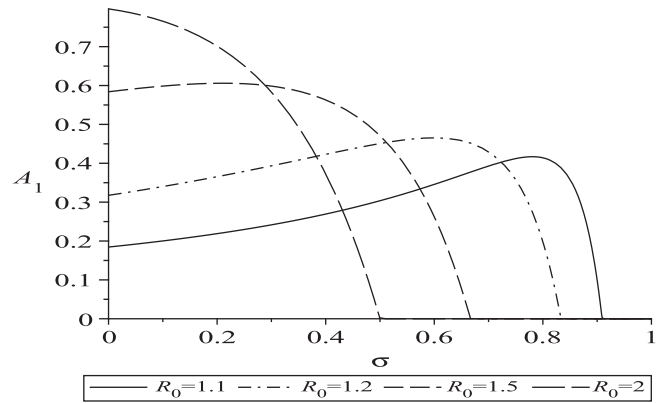


Fig. 5. Fraction A_1 of the population infected precisely once as a function of σ , for different values of $\mathcal{R}_0. S(0) = 0.999, I(0) = 0.001, R(0) = 0$.

infected once decreases monotonically with σ . For smaller values of \mathcal{R}_0 the number of people infected precisely once first increases with σ and reaches a maximum, and then decreases. In all cases A_1 reaches the value 0 at the reinfection threshold $\sigma = \frac{1}{\mathcal{R}_0}$, as can be expected, since beyond this threshold everyone is infected infinitely many times. The transition between the two types of behavior occurs around $\mathcal{R}_0 = 1.75$.

It would be very nice to have explicit equations for the quantities A_n , in terms of the parameters, in the spirit of the formulas for A and Z obtained in Section 3. Such formulas would enable to analytically study phenomena such as the transition in the shape of the curves $A_1(\sigma)$ as described above. However, we have not been able to obtain such equations for A_n , although we have the Eq. (17) for their sum A . We thus leave it as an open problem to either obtain such formulas, or to prove that such formulas do not exist, that is, that in some precise sense the quantities A_n cannot be expressed in terms elementary functions of the parameters.

Funding

The author acknowledges support of EU-FP7 grant Epiwork.

Acknowledgment

I thank Lewi Stone for useful discussions.

References

- [1] L.J. Abu-Raddad, N.M. Ferguson, Characterizing the symmetric equilibrium of multi-strain host-pathogen systems in the presence of cross immunity, *J. Math. Biol.* 50 (2005) 531.
- [2] R. Breban, S. Blower, Letter to the editor, *J. Theor. Biol.* 235 (2005) 151.
- [3] M.G.M. Gomes, L.J. White, G.F. Medley, Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives, *J. Theor. Biol.* 228 (2004) 539.
- [4] M.G.M. M Gomes, A.O. Franco, M.C. Gomes, G.F. Medley, The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy, *Proc. R. Soc. Lond. B* 271 (2004) 617.
- [5] M.G.M. Gomes, L.J. White, G.F. Medley, The reinfection threshold, *J. Theor. Biol.* 236 (2005) 111.
- [6] J. Ma, D.J.D. Earn, Generality of the final size formula for an epidemic of a newly invading infectious disease, *Bull. Math. Biol.* 68 (2006) 679.