Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

Mathematical Biosciences and Engineering, 16(2): 813–830.
DOI: 10.3934/mbe.2019038
Received: 23 January 2018
Accepted: 23 November 2018
Published: 17 January 2019

Research article

The change of susceptibility following infection can induce failure to predict outbreak potential by $\mathcal{R}_0$

Yukihiko Nakata[1†*] **and Ryosuke Omori**[2†]

[1] Department of Mathematical Sciences, Shimane University, 1060 Nishikawatsu-cho, Matsue, Japan
[2] Division of Bioinformatics, Research Center for Zoonosis Control, Hokkaido University, Sapporo, Hokkaido, Japan, JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332–0012, Japan

* **Correspondence:** Email: ynakata@riko.shimane-u.ac.jp

Abstract: Time-varying individual host susceptibility to a disease due to waning and boosting of immunity is known to induce rich long-term behavior of the disease transmission dynamics. Simultaneously, the impact of the time-varying heterogeneity of host susceptibility on the short-term behavior of epidemics is not well-studied despite the availability of a large amount of epidemiological data on short-term epidemics. This paper proposes a parsimonious mathematical model describing the short-term transmission dynamics by taking into account waning and enhancing susceptibility following the infection. In addition to the common classification in the standard SIR model, i.e., "no epidemic" as $\mathcal{R}_0 \leq 1$ or normal epidemic as $\mathcal{R}_0 > 1$, the proposed model also shows the "delayed epidemic" class when an epidemic takes off after the negative slope of the epidemic curve at the initial phase of the epidemic. The condition for each of the three classes is derived based on the obtained explicit solution for the proposed model.

Keywords: epidemic model; short-term disease transmission dynamics; boosting of immunity; final epidemic size

1. Introduction

Modelling is often used to understand the transmission dynamics of infectious diseases. The susceptible-infective-removed (SIR) model is known to be the simplest model to describe these transmission dynamics [1, 10]. The SIR model describes the transmission of pathogens from infected to susceptible individuals and the removal of infected individuals from the targeted host population due to the establishment of immunity or the host's death or immigration. Due to the wide variation in the natural history of pathogens, many extensions of the basic SIR model have been proposed so far.

[†]Both authors have equally contributed to this paper.

An important extension is the inclusion of the evolution of susceptibility to an infection along with a pathogen. The basic SIR model assumes that the host immunity perfectly protects the host from reinfection over time and reinfection can never occur. However, reinfection events are frequently observed for many infectious diseases such as tuberculosis [31] and those caused by coronavirus [20], respiratory syncytial virus [15] and hepatitis C virus [30]. One of the important mechanisms of reinfection is waning immunity. Decreased herd immunity by waning immunity of individuals induces re-emergence of an epidemic, and boosting immunity by re-vaccination is required to control the epidemic [4]. Another mechanism is imperfection of immunity by an infection event. A booster dose of a vaccine is required to establish a high enough immunity level to protect hosts from reinfection [29], which implies that multiple exposures to the pathogen is required to establish the high enough immunity level. Moreover, the enhancement of susceptibility to reinfection is also observed for several infectious diseases [32, 24].

An epidemic model incorporating variable susceptibility of recovered individuals was formulated by Kermack and McKendrick [22, 23]. However, the authors did not obtain a clear biological conclusion [8, 18]. In [17, 18] the author performed stability analysis for the Kermack and McKendrick's reinfection model formulated as a system of partial differential equations. The existence and bifurcation of the endemic equilibrium was analyzed in detail. Destabilization of the endemic equilibrium was shown to be possible for epidemic models with waning immunity [9, 16, 28]. Previous modeling studies showed that waning and natural-boosting immunity by exposing to pathogens can trigger a counter-intuitive effect of vaccination [27]. It is suggested that waning immunity in vaccinated hosts can trigger backward bifurcation of the endemic equilibrium [2, 6, 26]. Estimating vaccine effectiveness is essential to control epidemics. However, vaccine effectiveness reflects the complicated epidemiological dynamics which is scaled by waning and natural-boosting immunity, e.g., waning and boosting of immunity can induce periodic epidemic outbreaks as part of the long-term behavior of the epidemic [3].

Unlike the long-term behavior of an epidemic, its short-term behavior with waning and boosting of immunity is not well understood despite many field data of short-term epidemics have been analyzed using models that do not consider waning and boosting of immunity. For the short-term behavior of epidemics, the dynamics with the constant immune protection rate against reinfection has been studied so far. However, waning and boosting of immunity change the immune protection rate. In [21] the author analyzed transient dynamics of a reinfection epidemic model, while ignoring the demographic process in the model. In the model, reinfection of recovered individuals occurs assuming that these individuals have suitable susceptibility to the disease. It was shown that the disease transmission dynamics changes qualitatively when the basic reproduction number crosses the reinfection threshold. See also [13] for the model with demographic process.

This paper presents a mathematical model taking into account waning and enhancing susceptibility following an infection. We suppose that reinfection can reduce the immune protection level, e.g., as it is the case in antibody-dependent enhancement [5].

The timescale of natural-waning immunity, which is not due to reinfection, is relatively longer than that of the transmission dynamics e.g., the minimum annual waning rate of immunity is -2.9% for rubella and -1.6% for measles [25], compared to the infectious periods of 11 days for rubella [11] and 14 days for measles [1]. Hence the proposed model focuses only on waning and enhancing susceptibility triggered following an infection.

Since the waning and boosting of immunity can induce periodic infection outbreak on the long-term

disease transmission dynamics ([3]), a complicated epidemic curve can be observed in the model for the short-term disease transmission dynamics as well. In this paper, an explicit solution for a number of infective individuals in terms of a number of susceptible individuals is obtained. Consequently, the impact of waning and enhancing susceptibility following an infection on the short-term behavior of the transmission dynamics is explored. The shape of the short-term epidemic curve is analyzed in detail.

The paper is organized as follows. In Section 2 an epidemic model taking into account waning and enhancing susceptibility following an infection is formulated as a nonlinear system of differential equations. The model includes the standard SIR epidemic model and the reinfection epidemic model studied in [21] as special cases. Section 3 studies the disease transmission dynamics when the basic reproduction number $\mathcal{R}_0$ exceeds one. The epidemic peak is shown to be one, as is the case for the standard SIR epidemic model. Section 4 considers the disease transmission dynamics when $\mathcal{R}_0 \leq 1$. It demonstrates that the epidemic occurs even if $\mathcal{R}_0 \leq 1$ due to the enhancement of susceptibility of recovered individuals. The shape of the epidemic curve is analyzed in detail. In Section 5, the relatiton of final epidemic size with the model parameters is derived from the obtained explicit solutions. Plans for future work are discussed in Section 6.

2. An epidemic model with waning and enhancing susceptibility following infection

First of all let us introduce the epidemic model studied in [21]. In the model it is assumed that the infectious disease induces partial immunity. Denote by $S(t)$, $I(t)$ and $R(t)$ the proportions of susceptible population, infective population and recovered population at time t, respectively. The partial immunity model is formulated as

$$S'(t) = -\beta S(t)I(t), \tag{2.1a}$$
$$I'(t) = \beta S(t)I(t) + \beta\sigma R(t)I(t) - \gamma I(t), \tag{2.1b}$$
$$R'(t) = \gamma I(t) - \beta\sigma R(t)I(t). \tag{2.1c}$$

The positive parameters β and γ are the transmission coefficient and the recovery rate, respectively. The parameter σ is the relative susceptibility of recovered individuals, who have been infected at least once and have recovered from the infection. We obtain the standard SIR epidemic model, if $\sigma = 0$, i.e., recovered individuals are completely protected from the infection.

In this paper the partial immunity model (2.1) is modified as follows. When a recovered individual is exposed to the force of infection, immunity is boosted with probability $1 - \alpha$ so that one obtains permanent immunity to the disease, while one contracts the disease again with probability α. The partial immunity model (2.1) is modified as

$$S'(t) = -\beta S(t)I(t), \tag{2.2a}$$
$$I'(t) = \beta S(t)I(t) - \gamma I(t) + \beta\sigma\alpha I(t)R(t), \tag{2.2b}$$
$$R'(t) = \gamma I(t) - \beta\sigma I(t)R(t), \tag{2.2c}$$
$$B'(t) = \beta\sigma(1-\alpha)I(t)R(t) \tag{2.2d}$$

with the following initial conditions

$$S(0) = S_0 > 0,\ I(0) = I_0 > 0,\ R(0) = R_0 \geq 0,\ B(0) = B_0 \geq 0,$$

$$S_0 + I_0 + R_0 + B_0 = 1.$$

Here $B(t)$ denotes the proportion of population with permanent immunity at time t. We obtain the model (2.1) by $\alpha = 1$ and the SIR model by $\alpha = 0$. Throughout the paper, we assume the following two conditions

$$0 < \sigma, \tag{2.3}$$

$$0 < \alpha < 1. \tag{2.4}$$

We first prove a basic result for asymptotic behavior of the solution.

Lemma 1. *There exist* $\lim_{t\to\infty} S(t)$, $\lim_{t\to\infty} I(t)$, $\lim_{t\to\infty} R(t)$ *and* $\lim_{t\to\infty} B(t)$. *It holds that*

$$\lim_{t\to\infty} I(t) = 0. \tag{2.5}$$

Proof. One easily sees that S, R and B are respectively monotone bounded functions. Specifically S is a monotone decreasing function, while B is a monotone increasing function. Therefore, S, R and B tend to some constants. Since $S(t) + I(t) + R(t) + B(t) = 1$ holds for $t \geq 0$, $\lim_{t\to\infty} I(t)$ also exists. We now claim that (2.5) holds. From (2.2a), (2.2b) and (2.2c) one has

$$S'(t) + I'(t) + \alpha R'(t) = -\gamma(1-\alpha) I(t).$$

Suppose that $\lim_{t\to\infty} I(t) > 0$. Integrating the above equation, we derive a contradiction. Hence (2.5) holds. □

3. One epidemic peak for $\mathcal{R}_0 > 1$

We define the basic reproduction number by

$$\mathcal{R}_0 := b(S_0 + \alpha\sigma R_0),$$

where $b = \frac{\beta}{\gamma}$. The basic reproduction number is the expected number of secondary cases produced by one infective individual in the expected *one* infectious period, $\frac{1}{\gamma}$ in the initial phase of epidemic. Noting that both susceptible and recovered populations, which compose the initial host population, have susceptibility to the disease, we may call $\mathcal{R}_0$ the basic reproduction number, although $\mathcal{R}_0$ is conventionally called the *effective* reproduction number [19].

From (2.2a) and (2.2b) one obtains the following

$$\frac{dI}{dS} = -\frac{1}{bS}(\mathcal{R}(S,R) - 1) \tag{3.1}$$

and

$$\frac{dI(t)}{dt} = \gamma I(t)(\mathcal{R}(S(t), R(t)) - 1), \tag{3.2}$$

where

$$\mathcal{R}(S,R) := b(S + \alpha\sigma R),\ S \geq 0,\ R \geq 0.$$

Noting that $\mathcal{R}(S_0, R_0) = \mathcal{R}_0$, it is easy to see that

$$\begin{aligned}\mathcal{R}_0 > 1 &\Leftrightarrow I'(0) > 0,\\ \mathcal{R}_0 = 1 &\Leftrightarrow I'(0) = 0,\\ \mathcal{R}_0 < 1 &\Leftrightarrow I'(0) < 0,\end{aligned}$$

i.e., if $\mathcal{R}_0 > 1$ then the epidemic curve initially grows, while if $\mathcal{R}_0 < 1$ then the epidemic curve initially decays.

First we show that $R(t)$ can be expressed in terms of $S(t)$.

Proposition 1. *It holds that*

$$R(t) = \frac{1}{\sigma b}\left(1 - (1 - \sigma b R_0)\left(\frac{S(t)}{S_0}\right)^{\sigma}\right),\ t \geq 0. \tag{3.3}$$

Proof. Assume that $1 - \sigma b R_0 \neq 0$ holds. From the equations (2.2a) and (2.2c) we have

$$\frac{dR}{dS} = -\frac{1 - \sigma b R}{bS}. \tag{3.4}$$

Using the separation of variables, we obtain

$$\left(\frac{S(t)}{S_0}\right)^{\sigma} = \frac{1 - \sigma b R(t)}{1 - \sigma b R_0}, \tag{3.5}$$

thus (3.3) follows. It is easy to see that the equality in (3.3) also holds, if $1 - \sigma b R_0 = 0$. □

From Lemma 1 we have

$$R(t) = r(S(t)), \tag{3.6}$$

where

$$r(S) := \frac{1}{\sigma b}\left(1 - (1 - \sigma b R_0)\left(\frac{S}{S_0}\right)^{\sigma}\right),\ 0 \leq S \leq S_0. \tag{3.7}$$

Then we can show that $I(t)$ and $B(t)$ are explicitly expressed in terms of $S(t)$ and $R(t)$.

Proposition 2. *It holds that*

$$I(t) = I_0 + (S_0 - S(t)) + (1 - \alpha)\frac{1}{b}\ln\left(\frac{S(t)}{S_0}\right) + \alpha(R_0 - R(t)), \tag{3.8}$$

$$B(t) = B_0 + (1 - \alpha)\left\{-\frac{1}{b}\ln\left(\frac{S(t)}{S_0}\right) + (R_0 - R(t))\right\} \tag{3.9}$$

for $t \geq 0$.

Proof. From (2.2a) and (2.2b) we have

$$\frac{dI}{dS} = -1 + \frac{1}{bS} - \alpha\sigma\frac{R}{S}. \tag{3.10}$$

One can compute, using (3.6) and (3.7),

$$\begin{aligned}\int \frac{\sigma R}{S} dS =& \sigma \int \frac{1}{S}\left[\frac{1}{\sigma b}\left(1-(1-\sigma b R_0)\left(\frac{S}{S_0}\right)^{\sigma}\right)\right] dS \\ =& \frac{1}{b}\left[\int \frac{1}{S} dS - (1-\sigma b R_0)\int \left(\frac{S}{S_0}\right)^{\sigma}\frac{1}{S} dS\right].\end{aligned}$$

From (3.5) and (3.4) in the proof of Lemma 1 we get

$$\begin{aligned}\int_{S_0}^{S(t)} \left(\frac{S}{S_0}\right)^{\sigma}\frac{1}{S} dS &= -b\int_{R_0}^{R(t)} \left(\frac{1-\sigma b R}{1-\sigma b R_0}\right)\frac{1}{1-\sigma b R} dR = -b\frac{1}{1-\sigma b R_0}\int_{R_0}^{R(t)} dR. \\ &= -b\frac{1}{1-\sigma b R_0}\left(R(t)-R_0\right).\end{aligned}$$

Therefore, we get

$$\int_{S_0}^{S(t)} \frac{\sigma R}{S} dS = \frac{1}{b}\ln\left(\frac{S(t)}{S_0}\right) + \left(R(t)-R_0\right).$$

Thus we obtain (3.8).

From (2.2a) and (2.2b) we have

$$\frac{dB}{dS} = -(1-\alpha)\sigma\frac{R}{S}. \tag{3.11}$$

Similarly, one obtains (3.9). □

To analyze the epidemic curve, we study the function $\mathcal{R}(S,R)$ with $R = r(S)$. Let

$$\hat{\mathcal{R}}(S) := \mathcal{R}(S, r(S)).$$

We then compute the first and second derivatives of $\hat{\mathcal{R}}$:

$$\hat{\mathcal{R}}'(S) = b\left(1+\alpha\sigma r'(S)\right), \tag{3.12}$$

$$\hat{\mathcal{R}}''(S) = b\alpha\sigma r''(S). \tag{3.13}$$

From the equation (3.4) in Lemma 1, it is easy to obtain the following result.

Lemma 2. *One has*

$$r'(S) = -\frac{1}{b}\left(1-\sigma b R(0)\right)\frac{S^{\sigma-1}}{S_0^{\sigma}}, \tag{3.14}$$

$$r''(S) = (\sigma-1)\frac{1}{S}r'(S). \tag{3.15}$$

Note that, for $\sigma = 1$ or $1-\sigma b R_0 = 0$, r is a constant function. For $\sigma \neq 1$ and $1-\sigma b R_0 \neq 0$, , $r$is either monotone increasing or decreasing function, thus $\hat{\mathcal{R}}$ has at most one extremum.

We now show the standard epidemic case if $\mathcal{R}_0 > 1$ holds.

Proposition 3. *Let us assume that* $\mathcal{R}_0 > 1$ *holds. Then*

$$\hat{\mathcal{R}}(0) = \alpha < 1 < \hat{\mathcal{R}}(S_0) = \mathcal{R}_0 \tag{3.16}$$

holds and there exists a unique root of

$$\hat{\mathcal{R}}(S) = 1,\ 0 < S < S_0.$$

Proof. It is easy to see that (3.16) holds. First assume that $r'(S) \geq 0$ for $0 \leq S \leq S_0$. Then, from (3.12), one can see that $\hat{\mathcal{R}}$ is an increasing function. Thus we obtain the conclusion. Next assume that $r'(S) < 0$ for $0 \leq S \leq S_0$. By (3.13) and Lemma 2, one sees that $\hat{\mathcal{R}}$ has at most one minimum for $0 \leq S \leq S_0$. Therefore, from (3.16), we obtain the conclusion. □

Then, from Proposition 3, Lemma 1 and (3.2), we obtain the following result.

Theorem 1. *Let us assume that* $\mathcal{R}_0 > 1$ *holds. Then there is a* $t_p > 0$ *such that I is monotonically increasing for* $t \in \left(0, t_p\right)$ *and monotonically decreasing for* $t > t_p$. *It holds* $\lim_{t\to\infty} I(t) = 0$.

4. Delayed epidemic for $\mathcal{R}_0 \leq 1$

In the standard SIR model, when $\mathcal{R}_0 \leq 1$ holds, then the epidemic curve monotonically decreases and infective population tends to 0 eventually as time goes to infinity. The situation changes in the model with waning and enhancing susceptibility following infection(2.2), due to the susceptibility of the recovered individuals. In particular, if $\sigma > 1$ then there is a possible delayed outbreak as the recovered population increases which will induce the epidemic later even if $\mathcal{R}_0 \leq 1$. The basic reproduction number, which characterizes the initial dynamics, is not a sufficient criterion to determine the outbreak due to the recovered population.

First let us consider a simple case that $\sigma \leq 1$ holds. We have the standard scenario: if $\mathcal{R}_0 \leq 1$ then the epidemic does not occur. Subsequently we study the disease transmission dynamics when $\sigma > 1$. We show that enhancement of susceptibility after the infection can induce an epidemic later.

4.1. $\sigma \leq 1$

We show that the infective population is monotonically decreasing for $t \geq 0$, similar to the SIR model, when $\mathcal{R}_0 \leq 1$.

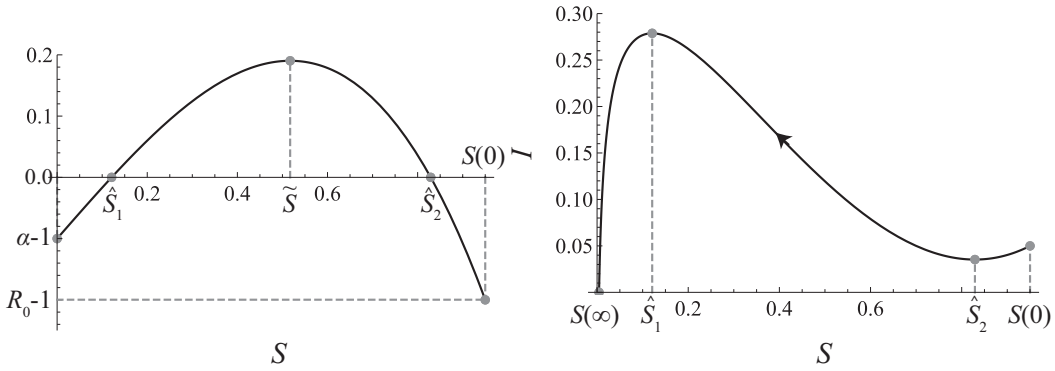
Proposition 4. *Let us assume that* $\mathcal{R}_0 \leq 1$ *and* $\sigma \leq 1$ *holds. Then*

$$\hat{\mathcal{R}}(S) \leq 1,\ 0 \leq S \leq S_0.$$

Proof. Note that

$$\hat{\mathcal{R}}(0) = \alpha < 1,\ \hat{\mathcal{R}}(S_0) = \mathcal{R}_0 \leq 1 \tag{4.1}$$

holds. First let us consider the case that $\sigma = 1$ or $1 - \sigma b R_0 = 0$. In these cases $\mathcal{R}(S)$ is a linear function of S, thus we obtain the conclusion from (4.1). Let us assume that $\sigma < 1$ and $1 - \sigma b R_0 \neq 0$ hold. If



(a) Graph of $\hat{\mathcal{R}}(S) - 1$

(b) Phase portrait of the solution $(I((t), S(t))$ in the (I, S) plane

Figure 1. The graph of $\hat{\mathcal{R}}(S) - 1$ for $0 \leq S \leq S_0$ is shown in (A). Here $\mathcal{R}_0 \leq 1$ and $\sigma > 1$. The parameters are chosen such that (4.4) and $\hat{\mathcal{R}}(\tilde{S}) > 1$ hold. The parametric curve $t \to (I(t), S(t))$ in the (I, S) plane is plotted in (B). It is shown that I has a local minimum at $S = \hat{S}_2$ and local maximum at $S = \hat{S}_1$.

$1 - \sigma b R_0 < 0$, then one has $r'(S) \geq 0$ for $0 < S < S_0$, thus $\hat{\mathcal{R}}$ is an increasing function. We then obtain the conclusion from (4.1). Next assume that $1 - \sigma b R_0 > 0$. One sees that

$$\lim_{S \downarrow 0} r'(S) = -\infty \implies \lim_{S \downarrow 0} \hat{\mathcal{R}}'(S) = -\infty.$$

Then $r'(S) < 0$ holds for $0 < S < S_0$. By Lemma 2, one sees that $\hat{\mathcal{R}}$ has at most one minimum for $0 \leq S \leq S_0$. Therefore we obtain the conclusion. □

Then, from Propositions 4, Lemma 1 and (3.2), we obtain the following result.

Theorem 2. *Let us assume that $\mathcal{R}_0 \leq 1$ and $\sigma \leq 1$ hold. Then I is monotonically decreasing for $t \geq 0$. It holds $\lim_{t\to\infty} I(t) = 0$.*

4.2. $\sigma > 1$

In this subsection we consider the case that

$$\mathcal{R}_0 \leq 1, \ \sigma > 1 \tag{4.2}$$

hold. We show the following results for the graph of $\hat{\mathcal{R}}$.

Proposition 5. *Let us assume that $\mathcal{R}_0 \leq 1$ and $\sigma > 1$ hold.*

1. *If*

$$b\left(S_0 + \alpha\sigma^2 R_0\right) \geq \alpha\sigma \tag{4.3}$$

then $\hat{\mathcal{R}}(S) \leq 1$ for $0 \leq S \leq S(0)$.

2. *If*

$$b\left(S_0 + \alpha\sigma^2 R_0\right) < \alpha\sigma \tag{4.4}$$

then there is a unique maximum for $0 < S < S_0$ *at* $S = \tilde{S}$*, where*

$$\tilde{S} := \left(\frac{bS_0}{\sigma\alpha\,(1-\sigma b R_0)}\right)^{\frac{1}{\sigma-1}} S_0 < S_0. \tag{4.5}$$

Then

(a) If $\hat{\mathcal{R}}(\tilde{S}) > 1$ *then there are two roots for* $\hat{\mathcal{R}}(S) = 1$ *for* $0 < S < S_0$*. Denote the roots by* $\hat{S}_1$ *and* $\hat{S}_2$ *such that*

$$0 < \hat{S}_1 < \tilde{S} < \hat{S}_2 < S_0,$$

then

$$\hat{\mathcal{R}}(S)\begin{cases} < 1, & \hat{S}_2 < S \le S_0, \\ > 1, & \hat{S}_1 < S < \hat{S}_2, \\ < 1, & 0 < S < \hat{S}_1. \end{cases}$$

(b) If $\hat{\mathcal{R}}(\tilde{S}) \le 1$ *then* $\hat{\mathcal{R}}(S) \le 1$ *for* $0 \le S \le S_0$*.*

Proof. For $\sigma > 1$ one sees that

$$\lim_{S\downarrow 0} r'(S) = 0 \implies \lim_{S\downarrow 0} \hat{\mathcal{R}}'(S) = b > 0.$$

From the monotonicity of $\hat{\mathcal{R}}'$, if $\lim_{S\uparrow S_0} \hat{\mathcal{R}}'(S) > 0$ then $\hat{\mathcal{R}}(S) \le 1$ for $0 \le S \le S_0$ follows. Computing

$$\begin{aligned}\lim_{S\uparrow S_0} \hat{\mathcal{R}}'(S) &= b\left(1 + \alpha\sigma r'(S_0)\right) \\ &= \frac{1}{S(0)}\left(b\left(S_0 + \alpha\sigma^2 R_0\right) - \alpha\sigma\right),\end{aligned}$$

one can see that (4.3) is equivalent to that $\lim_{S\uparrow S_0} \hat{\mathcal{R}}'(S) > 0$ holds.

Next assume that (4.4) holds. Then

$$\lim_{S\uparrow S_0} \hat{\mathcal{R}}'(S) < 0 < \lim_{S\downarrow 0} \hat{\mathcal{R}}'(S).$$

From the monotonicity of $\hat{\mathcal{R}}'$, there is a unique maximum for $0 < S < S_0$. Solving $\hat{\mathcal{R}}'(S) = 0$, we obtain $\tilde{S}$ given as in (4.5). It is now straightforward to obtain the statements (a) and (b). □

In Figure 1, we plot the graph of the function $\hat{\mathcal{R}}(S) - 1$ for $0 \le S \le S_0$, where $\mathcal{R}_0 \le 1$ and $\sigma > 1$. Parameters are fixed so that (4.4) and $\hat{\mathcal{R}}(\tilde{S}) > 1$ hold.

From Proposition 5, Lemma 1 and 3.2, we first obtain the result for the extinction of the disease.

Theorem 3. *Let us assume that* $\mathcal{R}_0 \le 1$ *and* $\sigma > 1$ *holds. If either*

1. *(4.3) holds, or*
2. *(4.4) and* $\hat{\mathcal{R}}(\tilde{S}) \le 1$ *hold,*

then I is monotonically decreasing for $t \geq 0$. It follows that $\lim_{t\to\infty} I(t) = 0$.

Now it is assumed that (4.2) holds. If

$$\hat{\mathcal{R}}(\tilde{S}) > 1$$

holds, where $\tilde{S}$ is given by (4.5), then $I(t)$ may attain a minimum and a maximum (see Figure 1). This implies that even if $\mathcal{R}_0 \leq 1$, the epidemic curve may grow for a certain time interval, which we call *delayed epidemic.*

To determine if the delayed epidemic indeed occurs, we evaluate the minimum of $I(t)$ using (3.8) in Proposition 2. If $S(t) = \hat{S}_2$ for some t, $I(t)$ is given by $\underline{I}$, where

$$\underline{I} := I_0 + \left(S_0 - \hat{S}_2\right) + (1-\alpha)\frac{1}{b}\ln\left(\frac{\hat{S}_2}{S_0}\right) + \alpha\left(R_0 - r(\hat{S}_2)\right). \tag{4.6}$$

See also Figure 1 (B) for the phase portrait in the (I, S)-plane.

Theorem 4. *Let us assume that $\mathcal{R}_0 \leq 1$ and $\sigma > 1$ holds. Furthermore, assume that (4.4) and*

$$\hat{\mathcal{R}}(\tilde{S}) > 1$$

hold.

1. *If $\underline{I} \leq 0$, then I is monotonically decreasing for $t \geq 0$.*
2. *If $\underline{I} > 0$, then there is an interval $[t_1, t_2]$ such that I increases for $t_1 \leq t \leq t_2$ and decreases for $0 \leq t \leq t_1$ and $t_2 \leq t$.*

It follows that $\lim_{t\to\infty} I(t) = 0$.

Proof. One sees that I has a local maximum and minimum with respect to t and S, where $\hat{\mathcal{R}}(S) = 1$ holds (see (3.1) and (3.2)). I has a local minimum at $S = \hat{S}_2 \in \left(\tilde{S}, S_0\right)$ and I is increasing for $\hat{S}_2 \leq S \leq S_0$ (see Figure 1 (B)). Noting that $I(t) > 0$ for $t \geq 0$ and that S is a decreasing function with respect to t, $\underline{I} \leq 0$ implies that I is monotonically decreasing for $t \geq 0$. On the other hand, if $\underline{I} > 0$ then I is monotonically increasing for $S < \hat{S}_1$, decreasing for $\hat{S}_1 < S < \hat{S}_2$ and then increasing for $\hat{S}_2 < S$. There exist t_1 and t_2 such that $S(t_1) = \hat{S}_2$ and $S(t_2) = \hat{S}_1$. Thus we obtain the conclusion. From Lemma 1, it follows that $\lim_{t\to\infty} I(t) = 0$. □

Thus the model has three different transmission dynamics: no epidemic, normal epidemic and delayed epidemic as illustrated in Figure 2. Figure 3 shows parameter regions for the three different disease transmission dynamics. The region for the delayed epidemic become larger with respect to the initial condition of I and the susceptibility σ. Since the initial condition of I is involved in the condition of Theorem 4, the initial condition qualitatively changes the epidemic curve, see Figure 4: delayed epidemic is induced by a large initial condition.

Due to the nonlinearity, it may not be possible to explicitly express the condition under which a delayed epidemic occurs in terms of the the initial conditions. We here consider a special case $\sigma = 2$ in order to illustrate that a certain amount of the initial infectives is required for a delayed epidemic. We assume that $R_0 = B_0 = 0$ to keep the presentation simple, although one may follow the approach for $R_0 > 0$ or $B_0 > 0$. We let

$$x(t) = \frac{S(t)}{S_0},\ y(t) = \frac{1-\sigma b R(t)}{1-\sigma b R_0}.$$

Then, from (3.5) in the proof of Lemma 1, we get $y(t) = x(t)^\sigma$. Assume that $\sigma = 2$ and $R_0 = B_0 = 0$ hold. One easily sees that the condition $\mathcal{R}_0 \le 1$ is equivalent to

$$bS_0 \le 1. \tag{4.7}$$

It can be seen that $\hat{\mathcal{R}}(S) = 1$ has two roots if and only the following equation has two roots

$$f(x) = 0,\ x \in [0,1], \tag{4.8}$$

where

$$f(x) := bS_0 x - \alpha x^2 - (1-\alpha) = 0.$$

Note $f(0) = -(1-\alpha) < 0$ and $f(1) = bS_0 - 1 \le 0$. If $(bS_0)^2 > 4\alpha(1-\alpha)$ then the equation (4.8) has two roots. Together with the condition (4.7) we now assume that

$$2\sqrt{\alpha(1-\alpha)} < bS_0 \le 1 \tag{4.9}$$

holds. If (4.8) holds, we have

$$bS_0 = \alpha x + \frac{1-\alpha}{x}, \tag{4.10}$$

where we regard x as a free parameter. Under the condition (4.9), equation (4.10) has two roots in $[0,1]$ for $\frac{1}{2} < \alpha < 1$ and the larger root exists in the interval $\left(\sqrt{(1-\alpha)/\alpha}, 1\right]$, thus we vary x in $\left(\sqrt{(1-\alpha)/\alpha}, 1\right]$.

From (3.8) one can express $I(t)$ in terms of $x(t)$:

$$I(t) = \frac{1}{b}\left(b(1 - S_0 x(t)) + \alpha\frac{x(t)^2 - 1}{2} + (1-\alpha)\ln x(t)\right). \tag{4.11}$$

Assume $\underline{I} = 0$, which is the threshold condition for a delayed epidemic. Using (4.11) and (4.10), b and S_0 can be expressed in terms of x as

$$b = \left(1 - \frac{\alpha}{2}\right) + \frac{1}{2}\alpha x^2 - (1-\alpha)\ln x,$$
$$S_0 = \frac{\alpha x + \frac{1-\alpha}{x}}{\left(1 - \frac{\alpha}{2}\right) + \frac{1}{2}\alpha x^2 - (1-\alpha)\ln x}$$

for $x \in \left(\sqrt{(1-\alpha)/\alpha}, 1\right]$. Then we can plot a parametrized curve in (I_0, b) plane, using $I_0 = 1 - S_0$, which shows the threshold condition for a delayed epidemic, see Figure 5. In the Figure 5, we also plot two curves $bS_0 = b(1 - I_0) = 1$ (corresponding to the condition $\mathcal{R}_0 = 1$) and $bS_0 = b(1 - I_0) = 2\sqrt{\alpha(1-\alpha)}$. A delayed epidemic occurs in the region enclosed by the three curves, showing that delayed epidemic requires sufficiently large initial infectives. Figure 5 also shows that if I_0 decreases, then either the condition $\mathcal{R}_0 > 1$ holds (i.e., normal epidemic) or $\underline{I} \le 0$ holds (i.e., no epidemic), depending on the value of b.

Consider a special case that $R_0 \to 0$. The basic reproduction number is given as $\mathcal{R}_0 = bS_0$. The conditions (4.4) becomes

$$\mathcal{R}_0 < \sigma\alpha$$

and $\tilde{S} = \left(\frac{\mathcal{R}_0}{\sigma\alpha}\right)^{\frac{1}{\sigma-1}} S_0$. If $\hat{\mathcal{R}}(\tilde{S}) > 1$ holds, then the delayed epidemic may occur.

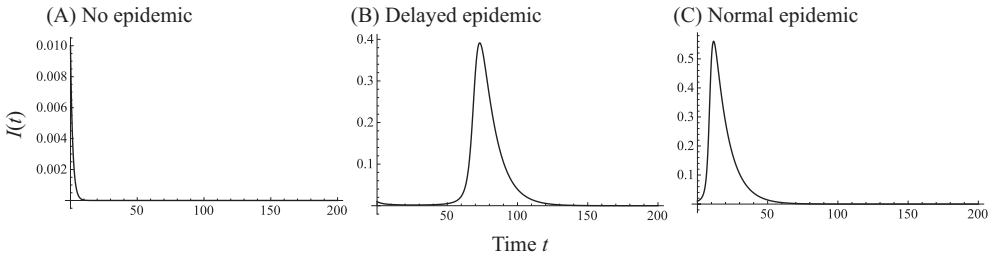


Figure 2. Examples of three types of epidemic curve. (A) shows no epidemic case, (B) shows delayed epidemic, and (C) shows normal epidemic, respectively. Parameters were set as $\mathcal{R}_0 = 0.4$ for (A), 0.8 for (B), and 1.2 for (C), other parameter values are identical between (A), (B) and (C): $\alpha = 0.9,\ \sigma = 5,\ S_0 = 0.99,\ I_0 = 0.01$, and $R_0 = B_0 = 0$.

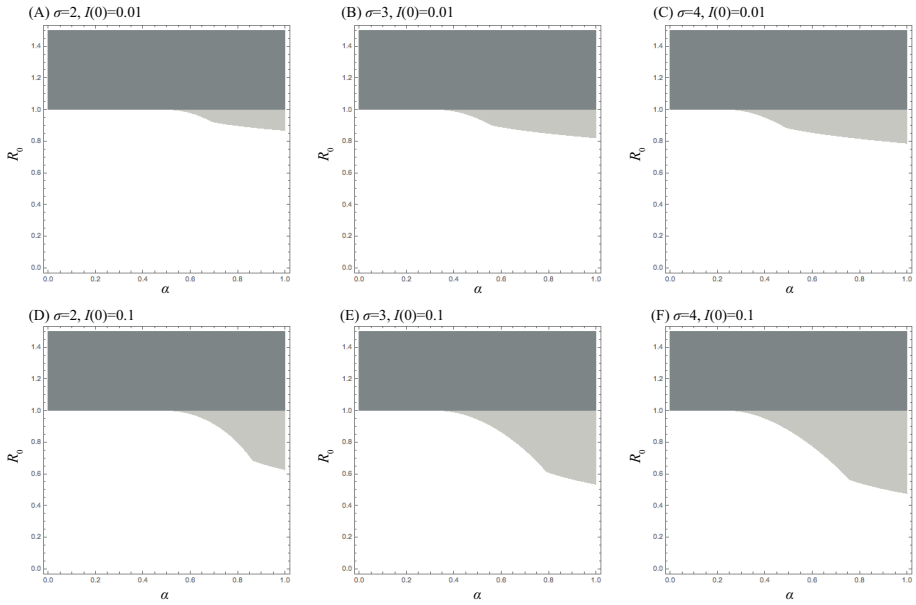


Figure 3. The dependency of epidemic type on $\mathcal{R}_0$ and α for several σ and I_0. White are denotes "no epidemic", light gray area denotes "delayed epidemic", and gray area denotes "normal epidemic", respectively.

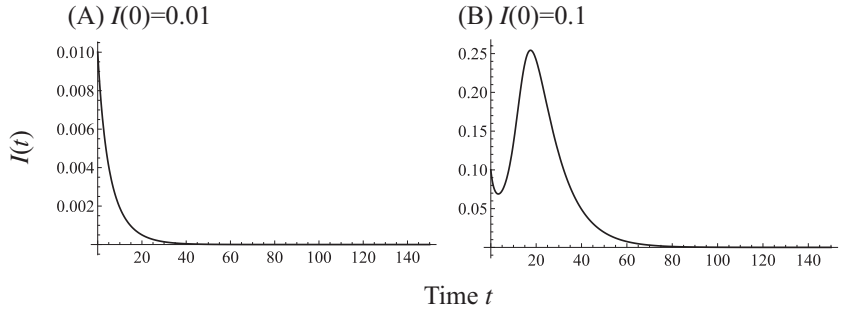


Figure 4. The initial condition for I changes the disease transmission dynamics. Here parameters are chosen as $\beta = 0.8,\ \gamma = 1,\ \alpha = 0.9,\ \sigma = 3$. $R_0 = B_0 = 0$. (A) shows no epidemic for $I_0 = 0.01$ while (B) shows delayed epidemic for $I_0 = 0.1$.

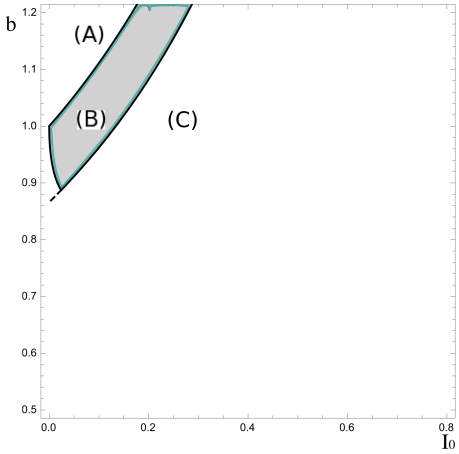


Figure 5. The disease transmission dynamics in a parameter plane (I_0, b). Here parameters are chosen as $\sigma = 2,\ \alpha = 0.75$ and $R_0 = B_0 = 0$. In the region (A) $\mathcal{R}_0 = bS_0 > 1$ holds, thus an epidemic occurs. In the regions (B) and (C), $\mathcal{R}_0 < 1$ holds. In the region (B) $\underline{I} > 0$, thus a delayed epidemic occurs and in the region (C) $\underline{I} \leq 0$, thus no epidemic occurs.

5. Final epidemic size

Let

$$(S(\infty), I(\infty), R(\infty), B(\infty)) = \lim_{t\to\infty} (S(t), I(t), R(t), B(t)) .$$

It follows that $I(\infty) = 0$. From the relations (3.8), (3.3) and (3.9), one sees that $(S(\infty), R(\infty), B(\infty))$ satisfy the following equations

$$0 = I_0 + (S_0 - S(\infty)) + (1-\alpha)\frac{1}{b}\ln\left(\frac{S(\infty)}{S_0}\right) + \alpha\,(R_0 - R(\infty)), \tag{5.1}$$

$$R(\infty) = r(S(\infty)), \tag{5.2}$$

$$B(\infty) = B_0 + (1-\alpha)\left\{-\frac{1}{b}\ln\left(\frac{S(\infty)}{S_0}\right) + (R_0 - R(\infty))\right\}. \tag{5.3}$$

The final epidemic size is given by $R(\infty) + B(\infty)$, the number of individuals who are infected at least once. From (5.1) and (5.2) we get the following equation

$$0 = I_0 + (S_0 - S(\infty)) + (1-\alpha)\frac{1}{b}\ln\left(\frac{S(\infty)}{S_0}\right) + \alpha\,(R_0 - rS\,((\infty))) . \tag{5.4}$$

In Figure 6, we plot $R(\infty) + B(\infty) = 1 - S(\infty),\ R(\infty)$ and $B(\infty)$ with respect to $\mathcal{R}_0$.

Numerically we observe that $R(\infty)$ is not monotone with respect to $\mathcal{R}_0$. Small $\mathcal{R}_0$ allows the increase of $R(\infty)$, on the other hand, does not contribute to the increase of $B(t)$, the outbreak ends before the transition from $R(t)$ to $B(t)$ via $I(t)$ occurs among most $R(t)$. Increase of $\mathcal{R}_0$ contributes the transition from $R(t)$ to $B(t)$, consequently, $R(\infty)$ decreases. Despite of non-monotnic relation of $R(\infty)$ with respect to $\mathcal{R}_0$, $R(\infty) + B(\infty)$ is likely to increase monotonically with the increase of $\mathcal{R}_0$ as shown in Figure 6.

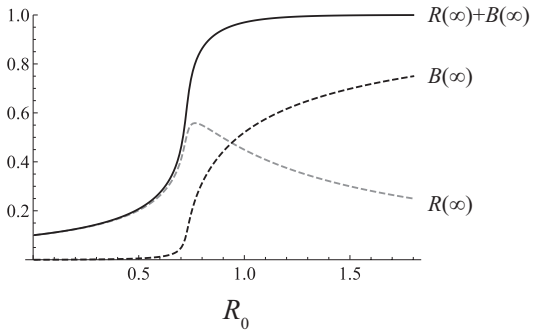


Figure 6. Final epidemic size with respect to $\mathcal{R}_0$. Initial conditions are fixed as $S(0) = 0.9,\ I(0) = 0.1,\ R(0) = B(0) = 0$. Parameters are fixed as $\sigma = 2,\ \alpha = 0.8$.

When $\alpha = 0$ we obtain the standard SIR setting. Letting $I(0) \to 0$ and $S(0) \to 1$, the basic reproduction number is given as $\mathcal{R}_0 = b$. In this case, from (5.1), we obtain the well known final size relation

$$0 = (1 - S(\infty)) + \frac{1}{\mathcal{R}_0} \ln(S(\infty)),$$

see e.g. [10, 19].

6. Discussion and conclusion

In this paper, we study a disease transmission dynamics model after incorporating changes in immune protection level following an infection. The proposed modelling approach describing waning and enhancement of susceptibility following an infection captures the delayed epidemics in addition to the standard transmission dynamics. A delayed epidemic shows a negative slope in the initial phase of the epidemic. Therefore, the estimation of $\mathcal{R}_0$ using the initial slope of the epidemic makes it difficult to capture the actual epidemic coming later. The condition for a delayed epidemic is derived using the analytic transient solution of $I(t)$.

A delayed epidemic, illustrated in Figures 2 and 4, occurs due to the enhancement of susceptibility of the recovered population (i.e., $\sigma > 1$). For example, antibody dependent enhancement can boost the viral replication within the host body. Consequently, the host susceptibility can be enhanced at the time of reinfection [33]. Theorem 4 formulates a condition for the delayed epidemic. One of the necessary condition for a delayed epidemic is (4.4) in Proposition 5. The condition (4.4) is necessary for the epidemic curve increase; it relates the effective susceptible population increase, which is defined as

$$L(t) := S(t) + \alpha\sigma R(t).$$

Since it holds that

$$\begin{aligned} L'(0) &= S'(0) + \alpha\sigma R'(0) \\ &= \gamma I_0 \left(-bS_0 + \alpha\sigma (1 - b\sigma R_0)\right), \end{aligned}$$

one can see that

$$L'(0) > 0 \Leftrightarrow bS_0 < \sigma\alpha (1 - b\sigma R_0).$$

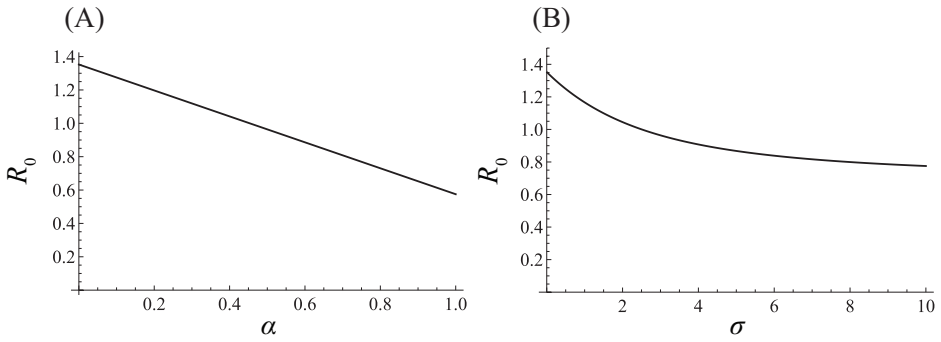


Figure 7. $\mathcal{R}_0$ derived from a given final epidemic size, $R(\infty) + B(\infty) = 0.5$ with varied α and σ. When $\alpha = 0$ or $\sigma = 0$ the boosting and waning immunity do not occur (the standard SIR model). The initial conditions are fixed as $S_0 = 0.99$, $I_0 = 0.01$, $R_0 = B_0 = 0$. Parameters are fixed as $\sigma = 3$ for (A) and $\alpha = 0.5$ for (B).

Therefore, the effective susceptible population increase at the initial time is necessary for a delayed epidemic and may induce the delayed epidemic even if $\mathcal{R}_0 \leq 1$.

We remark that $\mathcal{R}_0$ cannot measure the outbreak potential of a delayed epidemic. In principle, $\mathcal{R}_0$ is derived based on the linearized system of the the initial disease transmission dynamics. The linearized system at the initial phase does not provide enough information to predict a delayed epidemic. Similar phenomena can be observed in epidemic models that show backward bifurcation of the endemic equilibrium [2, 6, 14, 19, 26]. These studies show that there is a stable endemic equilibrium even if the basic reproduction number is less than unity. Unlike these models, the short-term disease transmission dynamics model has many equilibria, which are associated with the zero eigenvalue. Our study illustrates that the outbreak potential should be carefully examined in the short-term disease transmission dynamics using the transient solution.

$\mathcal{R}_0$ can be estimated from the final epidemic size. It should be noted that $\mathcal{R}_0$ can be overestimated if the model neglects the enhancement of susceptibility following an infection. Figure 7 shows the estimated $\mathcal{R}_0$ using a fixed final epidemic size of 0.5 with varied α and σ. In this case, our model is equivalent to the standard SIR model when $\alpha = 0$ or $\sigma = 0$. If waning and boosting of immunity are introduced, i.e., $\alpha > 0$ or $\sigma > 0$, the estimated $\mathcal{R}_0$ is always lower than the estimate obtained using the standard epidemic model with $\alpha = 0$ or $\sigma = 0$. To estimate the precise $\mathcal{R}_0$ from the final epidemic size, an appropriate modelling with respect to waning and boosting of immunity is required.

$\mathcal{R}_0$ can also be estimated from the time-series of $I(t)$. Since $I(t)$ in the proposed model can be expressed as the closed-form function of $S(t)$, if $I(t) + S(t)$ is given, e.g., $B(t) + R(t)$ is estimated from a serological survey, $I(t)$ can be obtained as a function of time, which can be used to estimate $\mathcal{R}_0$.

The proposed mathematical model describing waning and enhancing susceptibility following an infection has a limitation. In particular, it describes a step-wise level of boosting immunity i.e., R has susceptibility σ to an infectious disease while B has a complete protection against reinfection. This setting is suitable for infectious diseases where multiple infections can establish drastic increase of the immunity level. On the other hand, several classes of R with varied immune protection level are required to describe gradual change of the immunity level resulted from waning and boosting of immunity.

Acknowledgments

The authors thank the reviewer for the careful reading of the manuscript. The original manuscript has been improved by the reviewer's helpful comments. The first author was supported by JSPS Grant-in-Aid for Young Scientists (B) 16K20976 of Japan Society for the Promotion of Science. The second author was supported by CREST, Japan Science and Technology Agency.

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. R.M. Anderson and R.M. May, Infectious diseases of humans: dynamics and control, Vol. 28. Oxford: Oxford University Press, 1992.
2. J. Arino, C. C. McCluskey and P. van den Driessche, Global results for an epidemic model with vaccination that exhibits backward bifurcation, *SIAM J. Appl. Math.*, **64** (2003), 260–276.
3. N. Arinaminpathy, J. S. Lavine and B. T. Grenfell, Self-boosting vaccines and their implications for herd immunity, *PNAS*, **109** (2012), 20154-20159.
4. M.V. Barbarossa, A. Denes, G. Kiss, Y. Nakata and G. R$\acute{}$ost, Zs. Vizi, Transmission dynamics and final epidemic size of Ebola Virus Disease outbreaks with varying interventions, *PLoS One*, **10** (2015), e0131398.
5. W. Dejnirattisai, A. Jumnainsong, N. Onsirisakul, P. Fitton, S. Vasanawathana, W. Limpitikul, C. Puttikhunt, C. Edwards, T, Duangchinda , S. Supasa, K. Chawansuntati, P. Malasit, J. Mongkolsapaya and G. Screaton, Cross-reacting antibodies enhance dengue virus infection in humans, *Science*, **328** (2010), 745-748.
6. F. Brauer, Backward bifurcations in simple vaccination models, *J. Math. Anal. Appl.*, **29** (2004), 418–431.
7. O. Diekmann, J. A. P. Heesterbeek and J. A. J. Metz, On the definition and the computation of the basic reproduction ratio $\mathcal{R}_0$ in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382.
8. O. Diekmann, J. A. P. Heesterbeek and J. A. J. Metz, The legacy of Kermack and McKendrick, Epidemic Models: Their Structure and Relation to Data, (D. Mollison, ed.) (1995), 95–115.
9. O. Diekmann and R. Montijn, Prelude to Hopf bifurcation in an epidemic model: analysis of a characteristic equation associated with a nonlinear Volterra integral equation, *J. Math. Biol.*, **14** (1982), 117–127.
10. O. Diekmann, H. Heesterbeek and T. Britton, Mathematical tools for understanding infectious disease dynamics, Princeton University Press, 2012.
11. W.J. Edmunds, O.G. Van de Heijden, M. Eerola and N. J. Gay, Modelling rubella in Europe, *Epidemiol. Infect.*, **125** (2000), 617–634.

12. P. van den Driessche and J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48.
13. M.G.M. Gomes, L.J. White and G.F. Medley, Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives, *J. Theor. Biol.*, **228** (2004), 539–549.
14. D. Greenhalgh, O. Diekmann and M. C. M de Jong, Subcritical endemic steady states in mathematical models for animal infections with incomplete immunity, *Math. Biosci.*, **165** (2000), 1–25.
15. C. Hall, E. Walsh, C. Long and K. Schnabel, Immunity to and frequency of reinfection with respiratory syncytial virus, *J. Infect. Dis.*, **163**(1991), 693–698.
16. H.W. Hethcote, H.W. Stech and P. van den Driessche, Nonlinear oscillations in epidemic models, *SIAM J. Appl. Math.*, **40** (1981), 1–9.
17. H. Inaba, Kermack and McKendrick revisited: the variable susceptibility model for infectious diseases, *J. Ind. Appl. Math.*, **18** (2001), 273–292.
18. H. Inaba, Endemic threshold analysis for the Kermack-McKendrick reinfection model, *Josai Math. Mono.*, **9** (2016), 105–133.
19. H. Inaba, Age-structured population dynamics in demography and epidemiology, Springer, 2017.
20. D. Isaacs, D. Flowers, J. R. Clarke, H. B. Valman and M. R. MacNaughton, Epidemiology of coronavirus respiratory infections, *Arch. Dis. Child.*, **58** (1983), 500–503.
21. G. Katriel, Epidemics with partial immunity to reinfection, *Math. Biosci.*, **228** (2010), 153–159.
22. W.O. Kermack and A.G. McKendrick, Contributions to the mathematical theory of epidemics-II. The problem of endemicity. Proceedings of the Royal Society, 138A (1932), 55–83; Reprinted in Bull, *Math. Biol.*, **53** (1991), 57–87.
23. W.O. Kermack and A.G. McKendrick, Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity. Proceedings of the Royal Society, 141A (1933), 94–122; Reprinted in Bull, *Math. Biol.*, **53** (1991), 89–118.
24. R. Kohlmann, A. Salmen and A. Chan, Serological evidence of increased susceptibility to varicella-zoster virus reactivation or reinfection in natalizumab-treated patients with multiple sclerosis, *Mult. Scler. J.*, **21** (2015), 1823–1832.
25. J.R. Kremer, F. Schneider and C.P. Muller, Waning antibodies in measles and rubella vaccinees-a longitudinal study, *Vaccine*, **24** (2006), 2594–2601.
26. C.M. Kribs-Zaleta and J.X. Velasco-Hernandez, A simple vaccination model with multiple endemic states, *Math. Biosci.*, **164** (2000), 183–201.
27. J.S. Lavine, A. A. King and O. N. Bjørnstad, Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure, *PNAS*, **108** (2011), 7259–7264.
28. Y. Nakata, Y. Enatsu, H. Inaba, T. Kuniya, Y. Muroya and Y. Takeuchi, Stability of epidemic models with waning immunity, *SUT J. Math.*, **50** (2014), 205–245.
29. C.A. Siegrist, Vaccine immunology, *Vaccines*, **5** (2008), 1725.
30. T.J.W. van de Laar, R. Molenkamp, C. V. D. Berg, Frequent HCV reinfection and superinfection in a cohort of injecting drug users in Amsterdam, *J. Hepatol.*, **51** (2009), 667–674.

31. S. Verver, Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis, *Am. J. Resp. Crit. Care*, **171** (2005), 1430–1435.

32. G. Wei, H. Bull, X. Zhou and H. Tabel, Intradermal infections of mice by low numbers of African trypanosomes are controlled by innate resistance but enhance susceptibility to reinfection, *J.Infect. Dis.*, **203** (2011), 418–429.

33. S.S. Whitehead, J.E. Blaney, A.P. Durbin and B.R. Murphy, Prospects for a dengue virus vaccine, *Nat. Rev. Microbiol.*, **5** (2007), 518–528.