Available online at www.sciencedirect.com
ScienceDirect
Mathematical Biosciences 207 (2007) 89–103
Mathematical Biosciences
www.elsevier.com/locate/mbs

Modeling relapse in infectious diseases

P. van den Driessche [a,1], Xingfu Zou [b,*,2]

[a] Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4
[b] Department of Applied Mathematics, University of Western Ontario, London, Ont., Canada N6A 5B7

Received 10 May 2005; received in revised form 10 June 2006; accepted 18 September 2006
Available online 7 October 2006

Abstract

An integro-differential equation is proposed to model a general relapse phenomenon in infectious diseases including herpes. The basic reproduction number $\mathscr{R}_0$ for the model is identified and the threshold property of $\mathscr{R}_0$ established. For the case of a constant relapse period (giving a delay differential equation), this is achieved by conducting a linear stability analysis of the model, and employing the Lyapunov–Razumikhin technique and monotone dynamical systems theory for global results. Numerical simulations, with parameters relevant for herpes, are presented to complement the theoretical results, and no evidence of sustained oscillatory solutions is found.

Keywords: Infectious disease; Relapse; Basic reproduction number; Integro-differential equation; Delay

1. Introduction

Herpes simplex virus type 2 (herpes) is a human disease that is transmitted by close physical or sexual contact, and the incidence of this disease has risen over the last three decades [10]. Important features of herpes are that an individual once infected remains infected for life, and the virus

* Corresponding author. Tel.: +1 519 661 2111x88781; fax: +1 519 661 3523.
E-mail addresses: pvdd@math.uvic.ca (P. van den Driessche), xzou@uwo.ca (X. Zou).
[1] Research supported by NCE-MITACS of Canada and NSERC of Canada.
[2] Research supported by NCE-MITACS of Canada, NSERC of Canada, and by a Premier's Research Excellence Award of Ontario.

0025-5564/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2006.09.017

reactivates regularly with reactivation producing a relapse period of infectiousness (see, e.g., Blower et al. [5] and the references therein, and Hart [10]).

An ordinary differential equation (ODE) compartmental model for herpes was formulated by Tudor [18], who also noted that such a model is appropriate for pseudorabies in swine (see also Smith and Grenfell [15]). In this model the constant population is divided into three compartments depending on disease status. Individuals not previously exposed to the virus are *susceptible*, individuals infected and shedding virus are *infective* (infectious), and individuals previously infected with the virus but not currently shedding virus are *recovered* (latent). At time t, the numbers in each of these compartments are denoted by $S(t)$, $I(t)$, and $R(t)$, respectively, giving an SIRI model. Assuming standard incidence, a basic reproduction number $\mathscr{R}_0$ is identified, and it is shown to be a sharp threshold determining whether or not the disease dies out or approaches an endemic value.

This ODE model was extended to include more general incidence functions by Moreira and Wang [13] and a similar threshold result identified. Blower [4] summarized four different compartmental models for herpes. One model [5] contains an ODE model with six compartments to predict how much drug resistance would emerge if antiviral treatment rates of herpes were increased.

Our aim is to formulate a more general three compartmental model for a disease with relapse, and in particular to investigate the consequences of different assumptions about the relapse period. For the ODE models cited above, the infectious and relapse periods are assumed to have distributions that are negative exponentials. We allow for a more general relapse distribution, and in particular consider a case in which the relapse time is a constant. Mathematically this arises from taking a step function distribution for the relapse period, and leads to a delay differential equation. Such equations can have a Hopf bifurcation leading to sustained oscillatory solutions, but we find no evidence of this in our model.

In Section 2, we formulate our general SIRI model that can be applied to a disease with relapse. This is given in terms of $P(t)$, the fraction of recovered individuals remaining in the recovered class t units after recovery. Some basic results, including calculation of $\mathscr{R}_0$, are given in Section 3. For $P(t)$ a negative exponential, the ODE model dynamics are briefly summarized in Section 4. In Section 5, $P(t)$ is assumed to have compact support. The disease-free equilibrium is shown to be globally asymptotically stable if $\mathscr{R}_0 < 1$, and a Lyapunov–Razumikhin type theorem is used to determine a condition under which the endemic equilibrium is globally asymptotically stable if $\mathscr{R}_0 > 1$. For $P(t)$ a step function (Section 6), the endemic equilibrium is proved to be locally asymptotically stable if $\mathscr{R}_0 > 1$, and global asymptotically stable if, in addition, the relapse time is short. Finally in Section 6, numerical simulations using parameters appropriate for herpes [5] are presented that complement the theoretical results and indicate that $\mathscr{R}_0$ is a sharp threshold also for the step function case.

2. Model formulation

Let $S(t)$, $I(t)$ and $R(t)$ be the numbers of individuals in the susceptible, infective and the recovered classes, respectively, with the total population $N(t) = S(t) + I(t) + R(t)$. Assuming standard incidence for the disease transmission, the rate of change of $S(t)$ with time is

$$S'(t) = bN(t) - \lambda \frac{S(t)I(t)}{N(t)} - dS(t), \tag{2.1}$$

where the parameters $b > 0$ and $d > 0$ are the birth rate and death rate constants, respectively, and $\lambda > 0$ is the average number of effective contacts of an infectious individual per unit time (a fraction S/N are with susceptibles). For simplicity, we only consider a closed community in which the birth rate and death rate constants are equal, thus $b = d$. For genital herpes, these parameters can be considered as the rate of entering and leaving the sexually active population. We assume that the disease is not fatal, thus the death rate is the same for all individuals. In such a case, the total population remains a constant (since $N'(t) = 0$).

Rescaling $S(t)/N \to S(t)$, $I(t)/N \to I(t)$, and $R(t)/N \to R(t)$ with $S + I + R = 1$, we work with proportions in each class, and (2.1) becomes

$$S'(t) = d - \lambda S(t)I(t) - dS(t). \tag{2.2}$$

As outlined in Section 1, one important feature of herpes is that recovered individuals may relapse. Denote by $P(t)$ the fraction of recovered individuals remaining in the recovered class t time units after recovery. By the meaning of $P(t)$, it is reasonable to assume the following properties.

(A) $P : [0, \infty) \to [0, \infty)$ is differentiable (hence continuous) except at possibly finitely many points where it may have jump discontinuities, non-increasing and satisfies $P(0) = 1$, $\lim_{t\to\infty} P(t) = 0$ and $\int_0^\infty P(u)\,\mathrm{d}u$ is positive and finite.

The proportion of recovered individuals can be expressed by the integral

$$R(t) = \int_0^t \gamma I(\xi)\,\mathrm{e}^{-d(t-\xi)}P(t-\xi)\,\mathrm{d}\xi, \tag{2.3}$$

where $\gamma > 0$ is the recovery rate constant assuming that the infective period is exponentially distributed. The term $\mathrm{e}^{-d(t-\xi)}$ in the above integral accounts for the death of infectives. It is assumed that no individuals are initially in the recovered class, i.e., $R(0) = 0$. Differentiating (2.3) gives

$$R'(t) = -dR(t) + \gamma I(t) + \int_0^t \gamma I(\xi)\,\mathrm{e}^{-d(t-\xi)}d_t P(t-\xi)\,\mathrm{d}\xi. \tag{2.4}$$

Here, the integral is in the Riemann–Stieltjes sense to allow for possible jump discontinuities of $P(t)$. Substituting (2.3) and (2.4) into $I'(t) = -S' - R'$ leads to

$$\begin{aligned} I'(t) = &-(d+\gamma)I(t) + \lambda I(t)\left[1 - I(t) - \int_0^t \gamma I(\xi)\,\mathrm{e}^{-d(t-\xi)}P(t-\xi)\,\mathrm{d}\xi\right] \\ &- \int_0^t \gamma I(\xi)\,\mathrm{e}^{-d(t-\xi)}d_t P(t-\xi)\,\mathrm{d}\xi. \end{aligned} \tag{2.5}$$

Fig. 1 gives a digram of interactions between $S(t)$, $I(t)$ and $R(t)$.

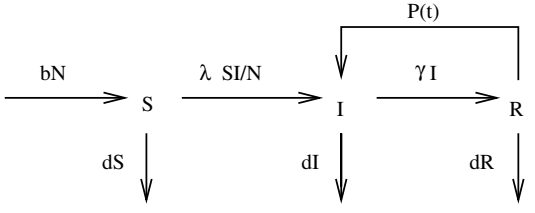

Fig. 1. Diagram of the model.

From (2.2), (2.3), (2.5) and the relation $S(t) + I(t) + R(t) = 1$, it is obvious that in order to determine the dynamics of each class, we only need to study (2.5) with assumption (A) and initial condition taken as $0 < I(0) \leqslant 1$. For example, if $I(t)$ approaches a constant $\bar{I}$ as $t \to \infty$, then by (2.2) and the theory of asymptotically autonomous systems, it follows that $S(t) \to \bar{S} = d/(\lambda\bar{I} + d)$, and therefore, $R(t) \to \bar{R} = 1 - \bar{I} - \bar{S}$.

Since (2.5) is a differential equation with possible unbounded delay, the existence and uniqueness of a solution to the initial value problem of (2.5) can be obtained by the fundamental theory established in Schumacher [14]. The focus of our study is on the dynamic behavior of solutions of (2.5). Note that due to the presence of the last term in (2.5), this scalar equation cannot be written in the form $I'(t) = I(t)g(I_t)$ (here $g(I_t)$ is the standard notation for general functional differential equations including integro-differential equations) for general $P(t)$, thus the analysis of the model is non-trivial.

3. Some basic results

For biological reasons, we only need to consider $I(t) \in [0, 1]$. The following result establishes this precisely.

Theorem 3.1. *Assume that $P(t)$ satisfies* (A). *If $0 < I(0) \leqslant 1$, then it follows that $0 < I(t) < 1$ for all finite $t > 0$.*

Proof. Assume there is a $t_1 > 0$ such that $I(t) > 0$ for $t \in [0, t_1)$ and $I(t_1) = 0$.

Let

$$h(t) := -\int_0^t \gamma I(\xi)\, \mathrm{e}^{-d(t-\xi)} d_t P(t-\xi)\, \mathrm{d}\xi = -\int_0^t \gamma I(t-\xi)\, \mathrm{e}^{-d\xi} d_\xi P(\xi)\, \mathrm{d}\xi.$$

By (A) and the choice of t_1, $h(t_1) \geqslant 0$. Denote

$$a(t) = -(d+\gamma) + \lambda\left[1 - I(t) - \int_0^t \gamma I(\xi)\, \mathrm{e}^{-d(t-\xi)} P(t-\xi)\, \mathrm{d}\xi\right].$$

Using an integrating factor for (2.5) gives

$$I(t_1) = \mathrm{e}^{\int_0^{t_1} a(s)\,\mathrm{d}s}\left[I(0) + \int_0^{t_1} h(s)\, \mathrm{e}^{-\int_0^s a(\xi)\,\mathrm{d}\xi}\,\mathrm{d}s\right] > 0.$$

This implies that such a $t_1 > 0$ cannot exist, and thus, $I(t) > 0$ for all finite $t > 0$.

To prove $I(t) < 1$ for $t > 0$, we consider two cases: $I(0) < 1$ and $I(0) = 1$. For the first case, assume that there is a $t_2 > 0$ such that $0 < I(t) < 1$ for $t \in [0, t_2)$ and $I(t_2) = 1$. Then from (2.5),

$$\begin{aligned} I'(t_2) &= -(d+\gamma) - \lambda\gamma \int_0^{t_2} I(\xi)\, \mathrm{e}^{-d(t_2-\xi)} P(t_2-\xi)\, \mathrm{d}\xi - \int_0^{t_2} \gamma I(\xi)\, \mathrm{e}^{-d(t_2-\xi)} d_t P(t_2-\xi)\, \mathrm{d}\xi \\ &\leqslant -(d+\gamma) - \gamma[P(t_2) - P(0)] \leqslant -(d+\gamma) + \gamma = -d < 0. \end{aligned}$$

This implies that such a t_2 cannot exist, and therefore, $I(t) < 1$ for all $t > 0$. For the second case, (2.5) gives $I'(0) = -(d+\gamma) < 0$, also implying that $I(t) < 1$ for all $t > 0$. □

Remark 3.1. If $I(0) = 0$, then $I(t) = 0$ for $t \geqslant 0$. Otherwise, there would exist a $t_0 \geqslant 0$ such that $I(t_0) \geqslant 0$, $I'(t_0) > 0$ and $I(t) = 0$ for $t \in [0, t_0)$. But then (2.5) implies that $I'(t_0) = 0$, a contradiction. Since we are interested in disease dynamics that start with some infectives at $t = 0$, and $I(t)$ is the proportion of infectives, we assume that $0 < I(0) \leqslant 1$.

Note that once the dynamics of $I(t)$ is known, (2.3) and the relation $S + I + R = 1$ will determine the behavior of $R(t)$ and $S(t)$. It is clear that the model has the disease-free equilibrium (DFE) with $(S, I, R) = (1, 0, 0)$.

Let

$$\hat{P} = \lim_{t\to\infty} \int_0^t \mathrm{e}^{-dv} P(v)\,\mathrm{d}v \leqslant \frac{1}{d} \tag{3.1}$$

and define

$$\mathscr{R}_0 = \frac{\lambda}{d(1+\gamma\hat{P})}. \tag{3.2}$$

Then $\hat{P}$ is the average time that an individual remains recovered before relapsing or dying. To interpret formula (3.2) for $\mathscr{R}_0$, note that the average time in the infective class on the first pass is $1/(d+\gamma)$ and the probability of surviving this class is $\gamma/(d+\gamma)$. Since $\hat{P}$ is the average time in the recovered class on the first pass, the probability of surviving the recovered class is $1 - d\hat{P}$. Thus, the total average time in the infective class (on multiple passes) is

$$\frac{1}{d+\gamma}\left[1 + \frac{\gamma(1-d\hat{P})}{d+\gamma} + \frac{\gamma^2(1-d\hat{P})^2}{(d+\gamma)^2} + \cdots\right] = \frac{1}{d+\gamma}\left[\frac{1}{1-\frac{\gamma(1-d\hat{P})}{d+\gamma}}\right] = \frac{1}{d(1+\gamma\hat{P})}.$$

Multiplying this by λ gives $\mathscr{R}_0$, which is the average number of new infectives produced by one infective introduced into a susceptible population [1]. Thus, $\mathscr{R}_0$ is the *basic reproduction number*, and acts as a threshold as is shown in the following result.

Theorem 3.2. *Consider* (2.5) *with* $0 < I(0) \leqslant 1$ *and assumption* (A). *If* $\mathscr{R}_0 < 1$, *then the DFE is locally asymptotically stable. If* $\mathscr{R}_0 > 1$, *then the DFE is unstable.*

Proof. To consider the stability of the DFE, set $I(t)$ proportional to e^{zt} in the linearization of (2.5) at the DFE to obtain the characteristic equation

$$z + (d+\gamma) - \lambda = -\gamma \lim_{t\to\infty} \int_0^t \mathrm{e}^{-(d+z)(t-\xi)} d_t P(t-\xi)\,\mathrm{d}\xi. \tag{3.3}$$

It is well known that the zero solution of (2.5) is asymptotically stable if all roots of (3.3) have negative real parts, and is unstable if (3.3) has a root with a positive real part (see, e.g., Cooke and Grossman [7], or Beretta and Kuang [3]). Now taking $z = x + \mathrm{i}y$ in (3.3) and assuming that $x \geqslant 0$ gives

$$|x + (d+\gamma) - \lambda + \mathrm{i}y| \leqslant \gamma \lim_{t\to\infty} \int_0^t |\mathrm{e}^{-(d+z)(t-\xi)} d_t P(t-\xi)|\,\mathrm{d}\xi \leqslant \gamma(1-d\hat{P})$$

since $\hat{P} \leqslant 1/d$ from (3.1). Thus, $(x-\lambda+d+\gamma)^2+y^2 \leqslant \gamma^2(1-d\hat{P})^2$ implying that $(x-\lambda+d+\gamma)^2-\gamma^2(1-d\hat{P})^2 \leqslant -y^2$ which is impossible if $-\lambda+d+\gamma d\hat{P}>0$ (equivalently $\mathscr{R}_0<1$). Thus, if $\mathscr{R}_0<1$, then $x<0$ and the DFE is locally asymptotically stable.

When $\mathscr{R}_0>1$, the characteristic equation (3.3) has a positive real root. To verify this, consider real z. The left side of (3.3) is an increasing function of z with value $d+\gamma-\lambda$ at $z=0$ and tends to ∞ as $z\to\infty$; whereas the right side is a non-increasing function of z with value $\gamma(1-d\hat{P})$ at $z=0$. Thus, if $d+\gamma-\lambda<\gamma(1-d\hat{P})$, equivalently $\mathscr{R}_0>1$, then (3.3) has a positive real root, and therefore the DFE is unstable. □

We point out that the above result is on local stability of the zero solution of (2.5) for $\mathscr{R}_0<1$. For some particular forms of the function $P(t)$, we are able to obtain global stability of this solution. In addition, for some special forms of $P(t)$, we can determine, fully or partially, the dynamics of (2.5) for $\mathscr{R}_0>1$. In the rest of the paper, we consider three particular forms for $P(t)$, beginning with a form for which the complete qualitative behavior of solutions of (2.5) can be determined.

4. $P(t)=e^{-\alpha t}$: an ODE system

For $P(t)=e^{-\alpha t}$, i.e., a negative exponential relapse distribution with relapse rate constant α, the system (2.4) and (2.5) can be written as a two-dimensional ODE system

$$\begin{aligned} R'(t) &= -(d+\alpha)R(t)+\gamma I(t),\\ I'(t) &= -(d+\gamma)I(t)+\lambda I(t)[1-I(t)-R(t)]+\alpha R(t). \end{aligned} \tag{4.1}$$

In this case, from (3.2), the basic reproduction number is

$$\mathscr{R}_0=\frac{\lambda(d+\alpha)}{d(d+\alpha+\gamma)}.$$

This model is given in Diekmann and Heesterbeek [9, p. 33], where it is shown that $\mathscr{R}_0$ can be calculated from the integral over T of the expected infectivity at time T after infection takes place. For this ODE model, $\mathscr{R}_0$ can also be calculated as the spectral radius of the next generation matrix, see Diekmann and Heesterbeek [9]; van den Driessche and Watmough [19]. Here, $\mathscr{R}_0$ can be interpreted as the product of λ and the total average time in the infective class, namely,

$$\frac{\lambda}{d+\gamma}\left[1+\frac{\gamma}{d+\gamma}\frac{\alpha}{d+\alpha}+\frac{\gamma^2}{(d+\gamma)^2}\frac{\alpha^2}{(d+\alpha)^2}+\cdots\right]=\frac{\lambda(d+\alpha)}{d(d+\alpha+\gamma)}.$$

For $\mathscr{R}_0>1$, (4.1) has a unique endemic equilibrium (EE) given by

$$R^*=\frac{\gamma I^*}{d+\alpha},\quad I^*=\frac{d(\mathscr{R}_0-1)}{\lambda}.$$

The dynamics of (4.1) are summarized in the following theorem.

Theorem 4.1. *Consider system (4.1) subject to the initial conditions* $0<I(0)\leqslant 1$, $R(0)=0$ *with* $D=\{(I,R): I\geqslant 0,\ R\geqslant 0,\ I+R\leqslant 1\}$. *If* $\mathscr{R}_0<1$, *then* D *is an asymptotically stable region for the DFE* $(I,R)=(0,0)$. *If* $\mathscr{R}_0>1$, *then* $D-\{(0,0)\}$ *is an asymptotically stable region for the EE* $(I,R)=(I^*,R^*)$.

The above result, showing that $\mathscr{R}_0 = 1$ is a sharp threshold in the global sense, is given in Tudor [18]. Moreira and Wang [13] study a generalization of (4.1) in which more general incidence $I\phi(S)$ or $S\psi(I)$ is assumed. By transforming the system to a Liénard system and using a Lyapunov function, they also obtain a similar threshold result for the generalized model.

5. General $P(t)$ with compact support: a delay differential equation with finite distributed delay

In this section, we consider $P(t)$ with compact support. That is, we assume in addition to (A) that $P(t) = 0$ for $t \geqslant w$, where w is the maximum relapse time. For such a $P(t)$, $\hat{P} = \int_0^w e^{-dv} P(v)\,dv$ and Eq. (2.5) for $t \geqslant w$ becomes the following delay differential equation with distributed finite delay:

$$I'(t) = -(d+\gamma)I(t) + \lambda I(t)\left[1 - I(t) - \int_{t-w}^{t} \gamma I(\xi)\, e^{-d(t-\xi)} P(t-\xi)\,d\xi\right] - \int_{t-w}^{t} \gamma I(\xi)\, e^{-d(t-\xi)} d_t P(t-\xi)\,d\xi. \tag{5.1}$$

We can apply the Lyapunov–Razumikhin technique to obtain the global stability of the DFE.

Theorem 5.1. *Assume that $P(t)$ satisfies* (A) *and $P(t) = 0$ for $t \geqslant w$. If $\mathscr{R}_0 < 1$, then the DFE of* (5.1) *with $0 < I(0) \leqslant 1$ is globally asymptotically stable.*

Proof. Consider the Lyapunov function $V(I(t)) = \frac{1}{2} I^2(t)$. Then, for those values of $t \geqslant w$ such that $V(I(t+s)) \leqslant V(I(t))$ for $s \in [-w, 0]$, the derivative of $V(I(t))$ along (5.1) is estimated as below:

$$\begin{aligned}
V' &= \lambda S(t) I^2(t) - (d+\gamma) I^2(t) - \int_{t-w}^{t} \gamma I(t) I(\xi)\, e^{-d(t-\xi)} d_t P(t-\xi)\,d\xi \\
&\leqslant \lambda I^2(t) - (d+\gamma) I^2(t) \\
&\quad - \frac{\gamma}{2} \int_{t-w}^{t} [I^2(t) + I^2(\xi)]\, e^{-d(t-\xi)} d_t P(t-\xi)\,d\xi \\
&\leqslant \lambda I^2(t) - (d+\gamma) I^2(t) - \gamma I^2(t) \int_{t-w}^{t} e^{-d(t-\xi)} d_t P(t-\xi)\,d\xi \\
&= \lambda I^2(t) - (d+\gamma) I^2(t) - \gamma I^2(t) \int_{0}^{w} e^{-dv} d_v P(v)\,dv \\
&= \lambda I^2(t) - (d+\gamma) I^2(t) + \gamma I^2(t)[1 - d\hat{P}] \\
&= [\lambda - d(1 + \gamma \hat{P})] I^2(t) \\
&= -\lambda \left(\frac{1}{\mathscr{R}_0} - 1\right) I^2(t).
\end{aligned} \tag{5.2}$$

Here, we have used the assumption that $P(t)$ is non-increasing and the fact that $0 \leqslant S(t) = 1 - I(t) - R(t) \leqslant 1$. Now by (5.2), the assumption that $\mathscr{R}_0 < 1$ and a Lyapunov–Razumikhin type theorem (see, e.g., Bélair [2]), we conclude that $I = 0$ is globally asymptotically stable. □

When $\mathscr{R}_0 > 1$, the DFE becomes unstable (from Theorem 3.2) and there exists an EE given by

$$I^* = \frac{d(\mathscr{R}_0 - 1)}{\lambda}. \tag{5.3}$$

We now explore the global stability of this EE for the form of $P(t)$ specified at the beginning of this section.

Theorem 5.2. *Assume that $P(t)$ satisfies* (A) *and $P(t) = 0$ for $t \geqslant w$. If $\mathscr{R}_0 > 1$, then the EE of* (5.1) *with $0 < I(0) \leqslant 1$ is globally asymptotically stable, provided that $\gamma\hat{P} < 1$.*

Proof. Let $\mathscr{R}_0 > 1$ and $u(t) = I(t) - I^*$. Then, from (5.1) and (5.3), $x(t)$ satisfies the following equation:

$$\begin{aligned} u'(t) = &-\gamma(1 - d\hat{P})u(t) - \lambda I(t)u(t) - \lambda\gamma I(t)\int_{t-w}^{t} u(\xi)\,\mathrm{e}^{-d(t-\xi)}P(t-\xi)\,\mathrm{d}\xi \\ &- \gamma\int_{t-w}^{t} u(\xi)\,\mathrm{e}^{-d(t-\xi)}d_tP(t-\xi)\,\mathrm{d}\xi. \end{aligned} \tag{5.4}$$

Let $V(u(t)) = \frac{1}{2}u^2(t)$. From Theorem 3.1, $I(t) > 0$ when $I(0) > 0$. Thus, at those $t \geqslant 0$ such that $V(u(t+s)) \leqslant V(u(t))$ for $s \in [-w, 0]$, the derivative of V along (5.4) can be estimated as below

$$\begin{aligned} \frac{\mathrm{d}V(u(t))}{\mathrm{d}t} &= -\gamma(1 - d\hat{P})u^2(t) - \lambda I(t)u^2(t) \\ &\quad - \lambda\gamma I(t)\int_{t-w}^{t} u(t)u(\xi)\,\mathrm{e}^{-d(t-\xi)}P(t-\xi)\,\mathrm{d}\xi \\ &\quad - \gamma\int_{t-w}^{t} u(t)u(\xi)\,\mathrm{e}^{-d(t-\xi)}d_tP(t-\xi)\,\mathrm{d}\xi \\ &\leqslant -\gamma(1 - d\hat{P})u^2(t) - \lambda I(t)u^2(t) + \lambda\gamma I(t)u^2(t)\hat{P} \\ &\quad - \gamma u^2(t)\int_{t-w}^{t} \mathrm{e}^{-d(t-\xi)}d_tP(t-\xi)\,\mathrm{d}\xi \\ &= u^2(t)[-\gamma(1 - d\hat{P}) - \lambda I(t) + \lambda\gamma I(t)\hat{P} + \gamma(1 - d\hat{P})] \\ &= -\lambda I(t)[1 - \gamma\hat{P}]u^2(t) < 0. \end{aligned} \tag{5.5}$$

By a Lyapunov–Razumikhin type theorem (see, e.g., Bélair [2]), we conclude that $u(t) \to 0$ as $t \to \infty$, implying that $I(t) \to I^*$ as $t \to \infty$ if $I(0) > 0$. □

Remark 5.1. The condition $\mathscr{R}_0 > 1$ is equivalent to

$$\gamma\hat{P} < \frac{\lambda - d}{d}. \tag{5.6}$$

Thus, $\lambda > d$ is necessary for $\mathscr{R}_0 > 1$. Moreover, the following implications are obvious:

(i) if $d < \lambda \leqslant 2d$, then the condition $\gamma\hat{P} < 1$ is implied by $\mathscr{R}_0 > 1$;
(ii) if $\lambda > 2d$, then $\mathscr{R}_0 > 1$ is implied by the condition $\gamma\hat{P} < 1$.

The condition $\gamma\hat{P} < 1$ is equivalent to $\hat{P} < 1/\gamma$, and thus holds when the average time that an individual remains recovered before relapse or death is less than the average infective time. For herpes, this does not seem to be realistic. Thus, from a practical point of view, we only need to be concerned with the situation where $\lambda > 2d$, $\mathscr{R}_0 > 1$ and yet $\gamma\hat{P} > 1$. Unfortunately, we are unable to prove the global stability of the endemic equilibrium in this case. However, if $P(t)$ is further assumed to be a step function, then we can prove that the endemic equilibrium is (locally) asymptotically stable, as will be shown in the next section.

6. Step function $P(t)$: a delay differential equations with a single delay

In this section we further restrict to the case in which all individuals remain in the recovered class w time units before relapsing, thus $P(t)$ is the step function given by

$$P(t) = \begin{cases} 1 & t \in [0, w), \\ 0 & t \geqslant w. \end{cases} \tag{6.1}$$

For this $P(t)$, from (3.2) the basic reproduction number becomes

$$\mathscr{R}_0 = \frac{\lambda}{d + \gamma(1 - \mathrm{e}^{-dw})}. \tag{6.2}$$

When $t \geqslant w$,

$$\begin{aligned} \int_{t-w}^{t} I(\xi)\,\mathrm{e}^{-d(t-\xi)} d_t P(t-\xi)\,\mathrm{d}\xi &= \int_0^w I(t-\xi)\,\mathrm{e}^{-d\xi} d_\xi P(\xi)\,\mathrm{d}\xi = I(t-w)\,\mathrm{e}^{-dw}[P(w^+) - P(w^-)] \\ &= I(t-w)\,\mathrm{e}^{-dw}[0-1] = -I(t-w)\,\mathrm{e}^{-dw}. \end{aligned}$$

Therefore, (5.1) is further reduced to the following delay differential equation

$$I'(t) = -(d+\gamma)I(t) + \gamma I(t-w)\,\mathrm{e}^{-dw} + \lambda I(t)[1 - I(t) - \gamma \int_{t-w}^{t} I(\xi)\,\mathrm{e}^{-d(t-\xi)}\,\mathrm{d}\xi], \tag{6.3}$$

for $t \geqslant w$. For $t \in [0, w)$, the dynamics of $I(t)$ in (5.1) is governed by

$$I'(t) = -(d+\gamma)I(t) + \lambda I(t)\left[1 - I(t) - \gamma \int_0^t I(\xi)\,\mathrm{e}^{-d(t-\xi)}\,\mathrm{d}\xi\right]. \tag{6.4}$$

Obviously, the long-term behavior of $I(t)$ is determined by (6.3), and thus, we will mainly focus on (6.3) in the rest of this section.

When $\mathscr{R}_0 > 1$, (6.3) has the unique endemic equilibrium I^* as in (5.3) with $\mathscr{R}_0$ given by (6.2). For Eq. (6.3), in addition to the results established in the previous sections, we can obtain further stability results. First, we consider the local asymptotic stability of the EE. To this end, we study the linearization of (6.3) at I^* given by

$$u'(t) = -\gamma\,\mathrm{e}^{-dw}[u(t) - u(t-w)] - \lambda I^* u(t) - \lambda\gamma I^* \int_{t-w}^{t} u(\xi)\,\mathrm{e}^{-d(t-\xi)}\,\mathrm{d}\xi. \tag{6.5}$$

Assuming that $u(t)$ in (6.5) is proportional to e^{zt} gives the following characteristic equation

$$z^2 + (\gamma e^{-dw} + d\mathscr{R}_0)z + d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw} - \gamma e^{-dw} z e^{-zw} - \gamma d \mathscr{R}_0 e^{-dw} e^{-zw} = 0. \quad (6.6)$$

Denote

$$P(z, w) = z^2 + (\gamma e^{-dw} + d\mathscr{R}_0)z + d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw},$$

$$Q(z, w) = -\gamma e^{-dw} z - \gamma d \mathscr{R}_0 e^{-dw}.$$

Then, (6.6) can be rewritten as

$$P(z, w) + Q(z, w) e^{-zw} = 0. \quad (6.7)$$

This is a transcendental equations with delay-dependent coefficients. Stability analysis of many other models also leads to such equations; for example [8]. Recently, Beretta and Kuang [3] provide some criteria on stability switches for such equations, and one important feature is that such equations may allow multiple switches of the stability as delay increases [8,3]. It is also interesting to note that while (6.3) is a first-order scalar delay differential equation, its characteristic equation at I^* is a transcendental equation of degree 2. By analyzing (6.7) under the assumption $\mathscr{R}_0 > 1$, the local stability of I^* can be proved.

Theorem 6.1. *If $\mathscr{R}_0 > 1$ and $0 < I(0) \leqslant 1$, then the endemic equilibrium I^* of (6.3) is locally asymptotically stable.*

Proof. First, observe that $z = 0$ is not a root of (6.7), for

$$\begin{aligned} P(0, w) + Q(0, w) &= d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw} - \gamma d \mathscr{R}_0 e^{-dw} \\ &= d(\mathscr{R}_0 - 1)(\gamma + d - \gamma e^{-dw}) > 0. \end{aligned}$$

Second, for $w = 0$ (zero relapse time), then all roots of (6.7) have negative real parts since

$$\begin{aligned} P(z, 0) + Q(z, 0) &= z^2 + (\gamma + d\mathscr{R}_0)z + d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d - \gamma z - \gamma d \mathscr{R}_0 \\ &= z^2 + d\mathscr{R}_0 z + d^2(\mathscr{R}_0 - 1) \end{aligned}$$

has positive coefficients. By the above results and the standard theory of delay differential equations [7], as the delay w increases, roots of (6.7) can cross the imaginary axis *only* through a pair of purely imaginary roots. Let $z = iy$ $(y > 0)$ be a root of (6.7). Then

$$\begin{aligned} &-y^2 + iy(\gamma e^{-dw} + d\mathscr{R}_0) + d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw} \\ &\quad + (\cos yw - i \sin yw)(-\gamma e^{-dw} iy - \gamma d \mathscr{R}_0 e^{-dw}) = 0. \end{aligned}$$

Separating the real and the imaginary parts in the above equation gives

$$\begin{aligned} -y^2 + d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw} &= \gamma d \mathscr{R}_0 e^{-dw} \cos yw + \gamma e^{-dw} y \sin yw, \\ y(\gamma e^{-dw} + d\mathscr{R}_0) &= \gamma e^{-dw} y \cos yw - \gamma d \mathscr{R}_0 e^{-dw} \sin yw. \end{aligned} \quad (6.8)$$

Squaring the two equations in (6.8) and adding the resulting equations yields

$$\begin{aligned} &y^4 - 2y^2[d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d^{-dw}] + [d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw}]^2 \\ &\quad + y^2(\gamma e^{-dw} + d\mathscr{R}_0)^2 = \gamma^2 d^2 \mathscr{R}_0^2 e^{-2dw} + \gamma^2 e^{-2dw} y^2, \end{aligned}$$

which is equivalent to

$$y^4 + by^2 + c = 0, \tag{6.9}$$

in which

$$b = d^2\mathscr{R}_0^2 - 2d(\mathscr{R}_0 - 1)(d + \gamma - \gamma e^{-dw}) = d^2\mathscr{R}_0^2 - 2\lambda d(1 - 1/\mathscr{R}_0),$$
$$c = [d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw}]^2 - \gamma^2 d^2 \mathscr{R}_0^2 e^{-2dw}.$$

Since

$$d(\mathscr{R}_0 - 1)(\gamma + d) + \gamma d e^{-dw} - \gamma d\mathscr{R}_0 e^{-dw} = \lambda d(1 - 1/\mathscr{R}_0),$$

it follows that

$$c = \lambda^2 d^2 (1 - 1/\mathscr{R}_0)^2 + 2\lambda\gamma d^2 e^{-dw}(\mathscr{R}_0 - 1).$$

Thus, (6.9) can be written as

$$[y^2 - \lambda d(1 - 1/\mathscr{R}_0)]^2 + d^2\mathscr{R}_0^2 y^2 + 2\lambda\gamma d^2 e^{-dw}(\mathscr{R}_0 - 1) = 0.$$

Since the left side of the above is positive, this gives a contradiction, showing that no root $z = iy$, $y > 0$ exists. Therefore, for all $w \geqslant 0$, all roots of (6.7) remain in the left half of the complex plane, completing the proof of the local stability. □

Since the form of $P(t)$ given by (6.1) is a special case of the forms of $P(t)$ specified in Section 5, the result of Theorem 5.2 may be applied to (6.3), giving the global stability of I^* under the additional condition

$$w < \frac{1}{d} \ln\left(1 + \frac{d}{\gamma - d}\right), \tag{6.10}$$

provided that $\gamma > d$. This condition requires that the delay be sufficiently small. On the other hand, the fact that (6.3) has only a single delay allows us to employ the monotone dynamical systems theory to establish new criteria. The idea is to introduce an exponential ordering in the phase space of (6.3). For details, see, e.g., Thieme and Smith [17] and Smith [16]. Using this technique, we can prove that when $P(t)$ is the step function given by (6.3), the condition (6.10) can by replaced by

$$w < \sup_{\mu > d+\lambda+\gamma+\lambda\gamma/d} \frac{1}{\mu - d} \ln\left[1 + \frac{(\mu - d)(\mu - d - \lambda - \gamma - \frac{\lambda\gamma}{d})}{\lambda\gamma}\right]. \tag{6.11}$$

See Appendix A for a proof of this. This inequality, although not identical to (6.10), is again only true for small relapse time. We point out that (6.10) and (6.11) are obtained from different approaches for stability, and they do not seem to be related to each other.

7. Numerical simulations and discussion

To complement the analytical results of the previous sections, we now show some numerical simulations for Eqs. (6.3) and (6.4). To this end, we adopt the following values for the parameters

$1/d = 20$ years $=7300$ days, $1/\gamma = 3.5$ days and $w = 27$ days taken from the model parameters for herpes simplex virus type 2 presented by Blower et al. [5]. For these values, the reproduction number $\mathscr{R}_0$ and the endemic equilibrium I^* can be computed from (6.2) and (5.3) as

$$\mathscr{R}_0 = \frac{\lambda}{0.001917}, \quad I^* = 0.114948\left(1 - \frac{0.001917}{\lambda}\right). \tag{7.1}$$

Recalling that λ is the average number of effective contacts of an infectious individual per unit time, it can be evaluated as a product $\lambda = c\beta$ where c is the average number of contacts of an infectious individual per unit time, and β is the probability of transmission on each contact between an

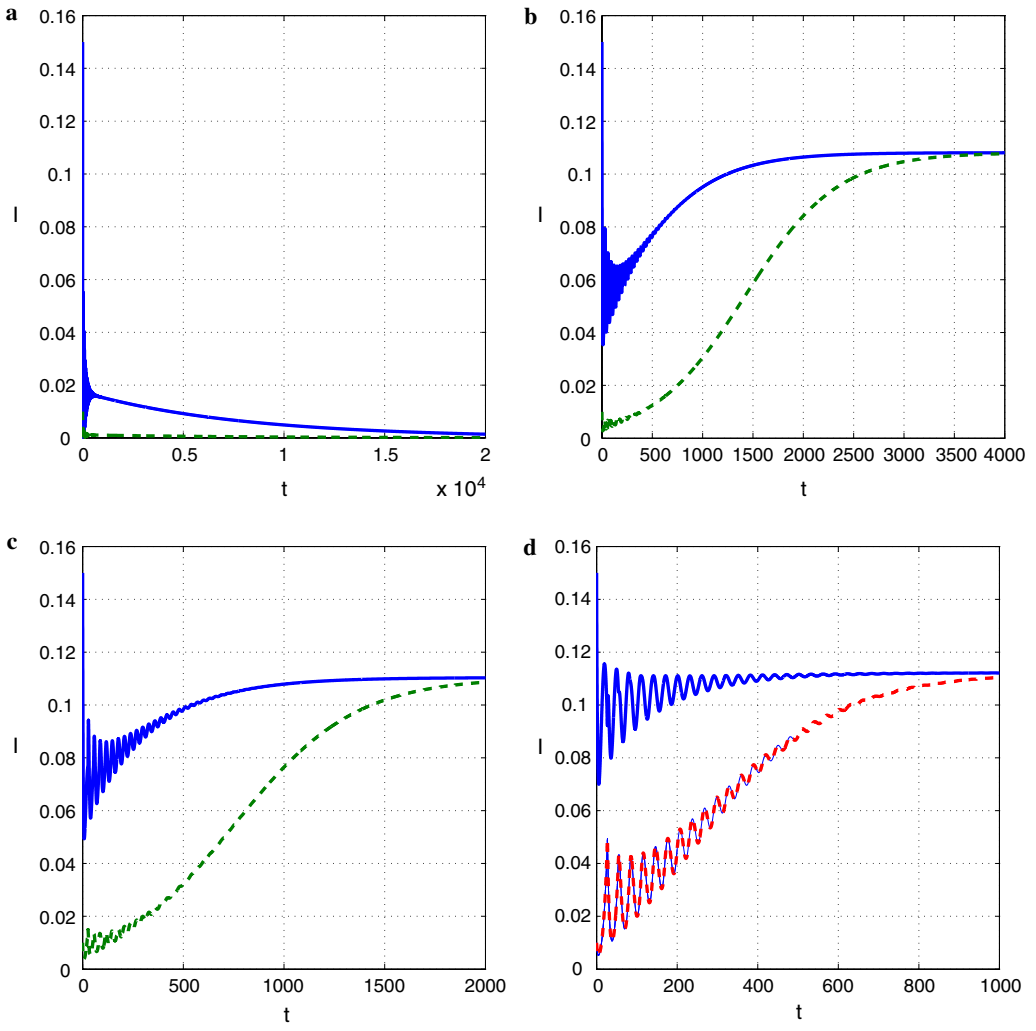

Fig. 2. Numerical simulations for (6.3) and (6.4). The initial values are taken as $I(0) = 0.01$ (dashed curves), and 0.15 (solid curves): (a) $\lambda = 0.0001$; (b) $\lambda = 0.02$; (c) $\lambda = 0.03$; (d) $\lambda = 0.05$. Others parameter values as in text. Note different scales on the time axes.

infectious and susceptible individual. As given in Blower et al. [5], $\beta \in [0.1, 0.5]$. Our simulations are done by varying the values of λ corresponding to various values of β and c within their biologically meaningful ranges. From (7.1), we see that for the above values of d,γ and w, the parameter λ has a critical value $\lambda_0 = 0.001917$. When λ is close to λ_0, the simulations also give convergence either to the DFE (when $\lambda < \lambda_0$) or to the EE (when $\lambda > \lambda_0$), but the convergence is so slow that we have to extend the time axis to a very large scale in order to observe the convergence. Here, we give simulation results for values of λ that are further away from λ_0 so that the convergence is sufficiently fast.

In Fig. 2, we fix the parameters as above, and choose two different values for $I(0) > 0$, namely $I(0) = 0.01$ and $I(0) = 0.15$. For Fig. 2(a), $\lambda = 0.0001$ (giving $\mathscr{R}_0$ approximately equal to 0.05); for Fig. 2(b), $\lambda = 0.02$ (giving $\mathscr{R}_0$ approximately equal to 10); for Fig. 2(c), $\lambda = 0.03$ (giving $\mathscr{R}_0$ approximately equal to 16); for Fig. 2(a), $\lambda = 0.05$ (giving $\mathscr{R}_0$ approximately equal to 25). From the figure, the disease dies out in the case $\mathscr{R}_0 < 1$ (in accordance with Theorem 5.1), whereas in all cases with $\mathscr{R}_0 > 1$, solutions converge to the endemic equilibrium, even though w is sufficiently large so that neither (6.10) nor (6.11) holds. In all simulations, transitory oscillations are observed for about the first 400 days, but no sustained oscillations are observed.

As noted in Section 1, several disease transmission models with a constant period in one compartment can exhibit sustained oscillatory solutions for certain parameters values with the basic reproduction number above a threshold. Such examples include an SIRS model in a population of fixed size with a constant period of temporary immunity [12], and a model with standard incidence and a distributed removed period [6]. In the light of these sustained oscillations and the fact that this model with relapse shows no such behavior, it remains a general open question as to what disease phenomena can generate sustained oscillations of the infectives; see Hethcote and Levin [11] for a survey of some such mechanisms.

Acknowledgments

We are grateful to the two anonymous referees for their valuable comments and suggestions that led to an improvement of the presentation. We also thank Dr. Lin Wang for helpful discussions in revising this paper and for his help with the numerical simulations.

Appendix A. Derivation of condition (6.11)

Theorem A1. *Let $\mathscr{R}_0$ be given by (6.2) and assume $\mathscr{R}_0 > 1$. If (6.11) holds, then the endemic equilibrium I^* of (6.3) (given by (5.3)) is globally asymptotically stable in the sense of Theorem 5.2.*

Proof. By Theorem 3.1, we only need consider $I \in (0, 1]$. Let $C = C([-w, 0], (0, 1])$ be equipped with the sup norm. For any $\mu \geqslant 0$, introduce the so-called exponential partial ordering $\leqslant_\mu$ in C as follows:

$$\phi_1 \leqslant_\mu \phi_2 \quad \text{if and only if } \phi_1 \leqslant \phi_2 \text{ and } [\phi_2(s) - \phi_1(s)]\, e^{\mu s} \text{ is non-decreasing.}$$

Denote by $f(\phi)$ the functional on the right-hand side of (6.3), i.e.,

$$f(\phi) = [\lambda - (d+\gamma)]\phi(0) + \gamma e^{-dw}\phi(-w) - \lambda\phi^2(0) - \lambda\gamma\phi(0)\int_{-w}^{0}\phi(s)e^{ds}\,ds.$$

Then for ϕ_1, $\phi_2 \in C$ with $\phi_1 \leqslant_\mu \phi_2$, calculation shows that for $\mu > d$

$$\begin{aligned}
&f(\phi_2) - f(\phi_1)\\
&\quad= [\lambda - (d+\gamma)][\phi_2(0) - \phi_1(0)] + \gamma e^{-dw}[\phi_2(-w) - \phi_1(-w)] - \lambda[\phi_2^2(0) - \phi_1^2(0)]\\
&\qquad - \lambda\gamma[\phi_2(0) - \phi_1(0)]\int_{-w}^{0}\phi_2(s)e^{ds}\,ds - \lambda\gamma\phi_1(0)\int_{-w}^{0}[\phi_2(s) - \phi_1(s)]e^{ds}\,ds\\
&\quad\geqslant [\lambda - (d+\gamma)][\phi_2(0) - \phi_1(0)] - \lambda[\phi_2(0) - \phi_1(0)][\phi_2(0) + \phi_1(0)]\\
&\qquad - \lambda\gamma[\phi_2(0) - \phi_1(0)]\int_{-w}^{0}e^{ds}\,ds - \lambda\gamma\int_{-w}^{0}[\phi_2(s) - \phi_1(s)]e^{ds}\,ds\\
&\quad\geqslant [\lambda - (d+\gamma)][\phi_2(0) - \phi_1(0)] - 2\lambda[\phi_2(0) - \phi_1(0)]\\
&\qquad - \lambda\gamma[\phi_2(0) - \phi_1(0)]\int_{-w}^{0}e^{ds}\,ds - \lambda\gamma[\phi_2(0) - \phi_1(0)]\int_{-w}^{0}e^{(d-\mu)s}\,ds\\
&\quad= \left[-\lambda - d - \gamma - \lambda\gamma\frac{1 - e^{-dw}}{d} - \lambda\gamma\frac{1 - e^{-(d-\mu)w}}{d-\mu}\right][\phi_2(0) - \phi_1(0)]\\
&\quad\geqslant -\left[\lambda + d + \gamma + \frac{\lambda\gamma}{d} + \frac{\lambda\gamma(e^{(\mu-d)w} - 1)}{\mu - d}\right][\phi_2(0) - \phi_1(0)].
\end{aligned} \tag{A.1}$$

By (6.11), there exists $\mu > d + \lambda + \gamma + \lambda\gamma/d$ such that

$$w < \frac{1}{\mu - d}\ln\left[1 + \frac{(\mu - d)(\mu - d - \lambda - \gamma - \frac{\lambda\gamma}{d})}{\lambda\gamma}\right],$$

which is equivalent to

$$\lambda + d + \gamma + \frac{\lambda\gamma}{d} + \frac{\lambda\gamma(e^{(\mu-d)w} - 1)}{\mu - d} < \mu. \tag{A.2}$$

For such a $\mu > d$, (A.1) and (A.2) lead to

$$f(\phi_2) - f(\phi_1) + \mu[\phi_2(0) - \phi_1(0)] > 0. \tag{A.3}$$

By Thieme and Smith [17], the semiflow generated by (6.3) in C is strongly order preserving in C in terms of the ordering $\leqslant_\mu$. Now, the global convergence theorem in Smith [16, p. 18, Theorem 3.1], the instability of $I = 0$ (from Theorem 3.1) and the uniqueness of the positive equilibrium I^* imply that all positive solutions of (6.3) converge to I^*. Global convergence to I^* and the local stability of I^* (Theorem 6.1) gives the global asymptotic stability of I^*, completing the proof. □

Taking $\mu = \lambda + 2d + \gamma + \lambda\gamma/d$ in (6.11), the following more explicit sufficient condition is obtained.

Corollary A2. *Assume* $\mathcal{R}_0 > 1$. *If*

$$w \leqslant \frac{1}{\lambda + d + \gamma + \frac{\lambda\gamma}{d}} \ln\left(1 + \frac{(\lambda + d + \gamma + \frac{\lambda\gamma}{d})d}{\lambda\gamma}\right), \tag{A.4}$$

then the endemic equilibrium I^* of (6.3) is globally asymptotically stable in the sense of Theorem 5.2.

References

[1] R.M. Anderson, R.M. May, Infectious Diseases of Humans, Dynamics and Control, Oxford University, Oxford, 1991.
[2] J. Bélair, Stability in a model of a delayed neural network, J. Dynam. Differ. Equat. 5 (1993) 607.
[3] E. Beretta, Y. Kuang, Geometric stability switch criteria in delay differential systems with delay dependent parameters, SIAM J. Math. Anal. 33 (2002) 1144.
[4] S. Blower, Modelling the genital herpes epidemic, Herpes 11 (Suppl. 3) (2004) 138A.
[5] S. Blower, T.C. Porco, G. Darby, Predicting and preventing the emergence of antiviral drug resistance in HSV-2, Nat. Med. 4 (1998) 673.
[6] C. Castillo-Chávez, H.R. Thieme, Asymptotically autonomous epidemic models, in: O. Arino, D. Axelrod, M. Kimmel, M. Langlais (Eds.), Mathematical Population Dynamics: Analysis of Heterogenity, Theory of Epidemics, vol. 1, Wuertz, 1995, p. 33.
[7] K. Cooke, Z. Grossman, Discrete delay, distributed delay and stability switches, J. Math. Anal. Appl. 86 (1982) 592.
[8] K. Cooke, P. van den Driessche, X. Zou, Interaction of maturation delay and nonlinear birth in population and epidemic models, J. Math. Biol. 39 (1999) 332.
[9] O. Diekmann, J.A.P. Heesterbeek, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Wiley, New York, 2000.
[10] T. Hart, Microterrors, Firefly Books, 2004.
[11] H.W. Hethcote, S.A. Levin, Periodicity in epidemiological models, in: S.A. Levin, T.G. Hallam, L.J. Gross (Eds.), Applied Mathematical Ecology, Springer, Berlin, 1989, p. 193.
[12] H.W. Hethcote, H.W. Stech, P. van den Driessche, Nonlinear oscillations in epidemic models, SIAM J. Appl. Math. 40 (1981) 1.
[13] H.N. Moreira, Y. Wang, Global stability in an $S \to I \to R \to I$ model, SIAM Rev. 39 (1997) 496.
[14] K. Schumacher, Existence and continuous dependence for functional-differential equations with unbounded delay, Arch. Rational Mech. Anal. 67 (1978) 315.
[15] G. Smith, B.T. Grenfell, Population biology of pseudorabies in swine, Am. J. Vet. Res. 51 (1990) 148.
[16] H.L. Smith, Monotone Dynamical Systems, An Introduction to the Theory of Competitive and Cooperative Systems, Mathematical Surveys and Monographs, vol. 40, American Mathematical Society, Providence, RI, 1995.
[17] H.R. Thieme, H.L. Smith, Monotone semiflows in scalar non-quasi-monotone functional-differential equations, J. Math. Anal. Appl. 150 (1990) 289.
[18] D. Tudor, A deterministic model for herpes infections in human and animal populations, SIAM Rev. 32 (1990) 136.
[19] P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (2002) 29.