J. Math. Biol. (2016) 73:947–976
DOI 10.1007/s00285-016-0976-6

Mathematical Biology

Backward bifurcations, turning points and rich dynamics in simple disease models

Wenjing Zhang[1] · **Lindi M. Wahl**[1] · **Pei Yu**[1]

Received: 5 September 2014 / Revised: 3 December 2015 / Published online: 26 February 2016

Abstract In this paper, dynamical systems theory and bifurcation theory are applied to investigate the rich dynamical behaviours observed in three simple disease models. The 2- and 3-dimensional models we investigate have arisen in previous investigations of epidemiology, in-host disease, and autoimmunity. These closely related models display interesting dynamical behaviors including bistability, recurrence, and regular oscillations, each of which has possible clinical or public health implications. In this contribution we elucidate the key role of backward bifurcations in the parameter regimes leading to the behaviors of interest. We demonstrate that backward bifurcations with varied positions of turning points facilitate the appearance of Hopf bifurcations, and the varied dynamical behaviors are then determined by the properties of the Hopf bifurcation(s), including their location and direction. A Maple program developed earlier is implemented to determine the stability of limit cycles bifurcating from the Hopf bifurcation. Numerical simulations are presented to illustrate phenomena of interest such as bistability, recurrence and oscillation. We also discuss the physical motivations for the models and the clinical implications of the resulting dynamics.

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) (No. 238389-RGPIN and No. R2686A02).
The first draft of this paper has been posted on arXiv.org since April 20, 2015 (No. arxiv:1504.05260).

✉ Wenjing Zhang
zhang.wenjing14@gmail.com

Lindi M. Wahl
lwahl@uwo.ca

Pei Yu
pyu@uwo.ca

[1] Applied Mathematics, Western University, London, ON N6A 5B7, Canada

Keywords Backward bifurcation · Positive backward bifurcation · Negative backward bifurcation · Forward bifurcation · Turning point · Disease model · Concave incidence rate · Convex incidence rate · Hopf bifurcation · Bistability · Recurrence

Mathematics Subject Classification 34C · 34D · 92B · 92D

1 Introduction

In the mathematical modelling of epidemic diseases, the fate of the disease can be predicted through the uninfected and infected equilibria and their stability. The basic reproduction number, R_0, represents the average number of new infectives introduced into an otherwise disease-free system by a single infective, and is usually chosen as the bifurcation parameter. If the model involves a *forward bifurcation* in a biologically meaningful domain, the uninfected equilibrium is in general globally asymptotically stable (Korobeinikov and Maini 2005), characterized by $R_0 < 1$, and infection fails to invade in this parameter regime. The threshold $R_0 = 1$ defines a bifurcation (or critical) point, and when $R_0 > 1$, a stable infected equilibrium emerges. If the infected equilibrium is globally asymptotically stable, then no complex dynamics can occur.

In contrast, *backward bifurcations* describe a scenario in which a *turning point* of the infected equilibrium exists. If the turning point is located in a region where all state variables are positive (biologically meaningful), we call this type of backward bifurcation a *positive backward bifurcation*. Otherwise, if the turning point is located in a region where some state variables are negative (mathematically meaningful, but the solutions do not exist biologically), we call this type of backward bifurcation a *negative backward bifurcation*. In the case of a negative backward bifurcation, the region for which all state variables are positive (i.e., restricted to the biologically meaningful regime) shows a seemingly *forward bifurcation*, which can, however, exhibit complex dynamical behaviours. Moreover, when the positive backward bifurcation also satisfies $R_0 < 1$, it will induce multiple infected equilibria, disrupting the global stability of the uninfected equilibrium, and multiple stable states (e.g., bistability) may likewise appear (Dushoff et al. 1998; Blayneh et al. 2010; Arino et al. 2003; Yu et al. 2015). In this case, instead of converging globally to the uninfected equilibrium, the solution may approach an infected equilibrium, depending on initial conditions. Further, when $R_0 > 1$, Hopf bifurcation(s) may occur from the infected equilibrium, leading to oscillations or even more complex behaviours.

In practice, the phenomenon of backward bifurcation gives rise to new challenges in disease control, since reducing R_0 such that $R_0 < 1$ is not sufficient to eliminate the disease (Hadeler and van den Driessche 1997; Brauer 2004). Instead, R_0 needs to be reduced past the critical value given by the turning point (Hadeler and van den Driessche 1997), since the result given in Yu et al. (2015) shows that the uninfected equilibrium in positive backward bifurcation is globally stable if R_0 is smaller than the turning point. Furthermore, an infective outbreak or catastrophe may occur if R_0 increases and crosses unity, while the upper branch of the infected equilibrium remains stable (Dushoff et al. 1998; Gomez-Acevedo and Li 2005; Zhang et al. 2013, 2014a). In addition, oscillation or even recurrent phenomena may occur if uninfected and

infected equilibria coexist in a parameter range, and both are unstable (Zhang et al. 2013, 2014a). Hadeler and van den Driessche (1997) predicted oscillations arising from backward bifurcation, and Brauer (2004) pointed out that the unstable infected equilibrium "commonly arises from Hopf bifurcation", but did not demonstrate oscillations.

Several mechanisms leading to backward bifurcations have been proposed, such as partially effective vaccination programs (Brauer 2004; Arino et al. 2003), educational influence on infectives' behavior (Hadeler and van den Driessche 1997), the interaction among multi-group models (Castillo-Chavez et al. 1989a, b; Huang et al. 1992) and multiple stages of infection (Simon and Jacquez 1992). In this study, we will investigate the emergence of positive/negative backward bifurcations in three simple disease models which have arisen in the study of epidemiology, in-host disease and autoimmunity. In each case, we find that backward bifurcation facilitates the emergence of Hopf bifurcation(s), and Hopf bifurcation in turn underlies a range of complex and clinically relevant dynamical behaviors.

Our investigation in the central theme of backward bifurcation starts with the role of incidence rate in the epidemiological and in-host disease models. The incidence rate describes the speed at which an infection spreads; it denotes the rate at which susceptibles become infectives. Under the assumptions of mass action, incidence is written as the product of the infection force and the number of susceptibles. For example, if S and I denote the susceptible and infective population sizes respectively, a bilinear incidence rate, $f(S, I) = \beta SI$ (where β is a positive constant), is linear in each of the state variables: S and I.

The possibility of saturation effects (Capasso and Serio 1978; Brown and Hasibuan 1995) has motivated the modification of the incidence rate from bilinear to nonlinear. Saturation occurs when the number of susceptible contacts per infective drops off as the proportion of infectives increases. A nonlinear incidence rate, therefore, typically increases sublinearly with respect to the growth of the infective population, and may finally reach an upper bound. The development of nonlinear incidence was first investigated in the form $\beta I^p S^q$, where β, p, and q are positive constants, (see, Liu et al. 1986, 1987; Hethcote et al. 1989; Hethcote and van den Driessche 1991; Derrick and van den Driessche 1993; Li and Muldowney 1995). Other forms of nonlinear incidence have also been analysed, such as $kI^p S/(1 + \alpha I^l)$ (Liu et al. 1986), and $kS\ln(1 + vP/k)$ (Briggs and Godfray 1995).

Since the nonlinear incidence functions described above were often developed to incorporate saturation effects, these functions are typically *concave* at realistic parameter values. Korobeinikov and Maini (2005) used this feature to derive general results for disease models with concave incidence. They proved that standard epidemiological models with concave incidence functions will have globally asymptotically stable uninfected and infected equilibria for $R_0 < 1$ and $R_0 > 1$, respectively.

More specifically, denoting the incidence rate function as $f(S, I, N)$, where N is the population size, the classical SIRS model considered in Korobeinikov and Maini (2005) takes the form

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \mu N - f(S,\ I,\ N) - \mu S + \alpha R, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = f(S,\ I,\ N) - (\delta + \mu)I,$$
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \delta I - \alpha R - \mu R, \tag{1}$$

where μ, δ, and α represent the birth/death rate, the recovery rate and the loss of immunity rate, respectively. When $\alpha = 0$, system (1) becomes an SIR model. Assuming that the total population size is constant, that is, $N = S + I + R$, the above system can be reduced to a 2-dimensional model:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = (\alpha+\mu)N - f(S,\ I,\ N) - \alpha I - (\alpha+\mu)S, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = f(S,\ I,\ N) - (\delta+\mu)I. \tag{2}$$

Moreover, it is assumed in Korobeinikov and Maini (2005) that the function $f(S,\ I,\ N)$, denoting the incidence rate, satisfies the following three conditions:

$$f(S,\ 0,\ N) = f(0,\ I,\ N) = 0, \tag{3a}$$
$$\frac{\partial f(S,\ I,\ N)}{\partial I} > 0, \quad \frac{\partial f(S,\ I,\ N)}{\partial S} > 0, \quad \forall\ S,\ I > 0 \tag{3b}$$
$$\frac{\partial^2 f(S,\ I,\ N)}{\partial I^2} \leq 0, \quad \forall\ S,\ I > 0. \tag{3c}$$

The first two conditions (3a) and (3b) are necessary to ensure that the model is biologically meaningful. The third condition (3c) implies that the incidence rate $f(S,\ I,\ N)$, is concave with respect to the number of infectives. It is also assumed that $\frac{\partial f(S,\ I,\ N)}{\partial I}$ evaluated at the uninfected equilibrium is proportional to the basic reproduction number R_0 (van den Driessche and Watmough 2002), and thus should be a positive finite number (Korobeinikov and Maini 2005). Korobeinikov and Maini first considered $\frac{dI}{dt} = 0$, or $f(S,\ I,\ N) - (\delta + \mu)I = 0$, and showed that forward bifurcation occurs in model (2) with a concave incidence function. They further proved that the uninfected equilibrium $Q_0 = (S_0,\ I_0) = (N,\ 0)$ and the infected equilibrium $\bar{Q} = (\bar{S},\ \bar{I})$ are globally asymptotically stable, when $R_0 = \frac{1}{\delta+\mu}\frac{\partial f(S_0,\ I_0,\ N)}{\partial I} < 1$ and $R_0 > 1$, respectively.

In the sections to follow, for an incidence rate function $f(S,\ I)$, satisfying (3a) and (3b), we define $f(S,\ I)$ as concave, if it satisfies (3c); as *convex*, if $\frac{\partial^2 f(S,\ I)}{\partial I^2} > 0$, $\forall\ I > 0$; and as *convex-concave*, if there exist $0 < I_1 < I_2 \leq +\infty$, such that $\frac{\partial f(S,\ I)}{\partial I} > 0, \forall\ I \in (0,\ I_2)$, and $\frac{\partial^2 f(S,\ I)}{\partial I^2} > 0, \forall\ I \in (0,\ I_1)$, $\frac{\partial^2 f(S,\ I)}{\partial I^2} = 0$, for $I = I_1$, $\frac{\partial^2 f(S,\ I)}{\partial I^2} < 0, \forall\ I \in (I_1,\ I_2)$.

Several models closely related to system (2) have been previously studied. For example, by adding a saturating treatment term to model (2) with a concave incidence rate, Zhou and Fan (2012) showed that this model may yield backward bifurcation and Hopf bifurcation. With an even more sophisticated nonlinear incidence rate function: $kI^pS/(1 + \alpha I^l)$, where $p = l = 2$, Ruan and Wang (2003) proved that a reduced 2-dimensional SIRS model could exhibit backward bifurcation, Hopf bifurcation, and even Bogdanov–Takens bifurcation and homoclinic bifurcation. Although the choice of $p = l = 2$ was not motivated by a specific physical process, this important result

demonstrates that a nonlinear incidence rate can induce backward bifurcation, and further generate complex dynamics in a simple disease model.

One of the focal points of our study will be a convex incidence function which arose in a 4-dimensional HIV antioxidant therapy model (van Gaalen and Wahl 2009). In this model, the infectivity of infected cells was proposed to be an increasing function of the density of reactive oxygen species, which themselves increase as the infection progresses. In van Gaalen and Wahl (2009), meaningful parameter values were carefully chosen by data fitting to both experimental and clinical results. In this parameter regime, the model was observed to capture the phenomenon of viral blips, that is, long periods of undetectable viral load punctuated by brief episodes of high viral load. Viral blips have been observed clinically in HIV patients under highly active antiretroviral therapy (Collier et al. 1996; Dornadula et al. 1999; Palmer et al. 2003, 2008), and have received much attention in the research literature, both by experimentalists (Fung et al. 2012; Garretta et al. 2012; Grennan et al. 2012) and mathematicians (Fraser et al. 2001; Jones and Perelson 2005; Conway and Coombs 2011; Rong and Perelson 2009a, b). Nonetheless, the mechanisms underlying this phenomenon are still not thoroughly understood (Grennan et al. 2012; Rong and Perelson 2009a).

We recently re-examined the model developed in van Gaalen and Wahl (2009), with the aim of providing new insight into the mechanism of HIV viral blips (Zhang et al. 2013, 2014a). Focusing on the dynamics of the slow manifold of this model, we reduced the dimension of the 4-dimensional model by using quasi-steady state assumptions. After a further generalization and parameter rescaling process, a 2-dimensional in-host HIV model (Zhang et al. 2013, 2014a) was obtained, given by

$$\frac{\mathrm{d}X}{\mathrm{d}\tau} = 1 - DX - \left(B + \frac{AY}{Y+C}\right)XY, \qquad \frac{\mathrm{d}Y}{\mathrm{d}\tau} = \left(B + \frac{AY}{Y+C}\right)XY - Y, \tag{4}$$

where X and Y denote the dimensionless concentrations of the uninfected and infected cells respectively. The constant influx rate of X and the death rate of Y have been scaled to 1. The death rate of X is D. This 2-dimensional infection model (4), reduced from the 4-dimensional HIV model (van Gaalen and Wahl 2009), preserves the viral blips observed in the 4-dimensional HIV model.

Importantly, system (4) is equivalent to the SIR model (2), except that the incidence function is convex, as we will show in Sect. 2.2. This equivalence can be demonstrated if we set $S = e_1 x$, $I = e_2 y$, and $t = e_3 \tau$ with $e_1 = e_2 = \frac{\mu N}{\delta+\mu}$ and $e_3 = \frac{1}{\delta+\mu}$. In this case, system (2) is rescaled to

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = 1 - \frac{\mu}{\delta+\mu}x - \frac{1}{\mu N}f(x,\ y), \qquad \frac{\mathrm{d}y}{\mathrm{d}\tau} = \frac{1}{\mu N}f(x,\ y) - y,$$

which takes the same form as system (4). Therefore, although system (2) arises in epidemiology and system (4) was derived as an in-host model, they are mathematically equivalent in this sense. We will refer to both systems (2) and (4) as infection models.

In previous work (Zhang et al. 2013, 2014a), we analyzed the recurrent behavior which emerges in system (4) in some detail. Recurrence is a particular form of oscillatory behavior characterized by long periods of time close to the uninfected

equilibrium, punctuated by brief episodes of high infection (Yao et al. 2006). Thus HIV viral blips are an example of recurrent behavior, but recurrence is a more general feature of many diseases (Yao et al. 2006; Zhang et al. 2014a). We have demonstrated that the increasing and saturating infectivity function of system (4) is critical to the emergence of recurrent behaviour. This form of an infectivity function corresponds to a convex incidence rate function in the associated 2-dimensional infection model (4), and can likewise induce recurrence in this model. Convex incidence has been previously suggested to model 'cooperation effects' in epidemiology (Korobeinikov and Maini 2005), or cooperative phenomena in reactions between enzyme and substrate, as proposed by Murray (2002).

The rest of this paper is organized as follows. In Sect. 2, we study two 2-dimensional infection models, both closely related to system (2). We show that system (2) with either (a) a concave incidence rate and saturating treatment term or (b) a convex incidence rate as shown in system (4), can exhibit backward bifurcations; we then identify the necessary terms in the system equations which cause this phenomenon. In Sect. 3, we demonstrate that in both models, backward bifurcation increases the likelihood of a Hopf bifurcation on the upper branch of the infected equilibrium. Studying system (4) in greater detail, we illustrate how the location of the turning points, and the location of the Hopf bifurcations and their directions (supercritical or subcritical), determine the possible dynamical behaviors, concluding that backward bifurcations with varied location of turning points facilitate Hopf bifurcation(s), which then underly the rich behaviours observed in these models. In Sect. 4, we explore backward bifurcation further, presenting an autoimmune disease model which exhibits negative backward bifurcation only. Although this bifurcation introduces two branches of the infected equilibrium, we demonstrate that, in the biologically feasible area, only forward bifurcation exists in this model and Hopf bifurcation does not occur. We then present a modification to this autoimmune model, motivated by the recent discovery of a new cell type, which generates a negative backward bifurcation and Hopf bifurcation, and allows recurrent behavior to emerge. A conclusion is drawn in Sect. 5.

2 Backward bifurcations

In this section, we study backward bifurcation in two 2-dimensional infection models. In particular, we explore the essential terms and parameter relations which are needed to generate positive backward bifurcation and negative backward bifurcation. Furthermore, we examine the convex incidence rate, and reveal its underlying role in determining the emergence of backward bifurcations.

2.1 Backward bifurcation in an infection model with concave incidence

First, we consider the SIR model with concave incidence, described by the following equations (Zhou and Fan 2012):

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta SI}{1+kI} - (d+\gamma+\epsilon)I, \quad \frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - dR, \quad (5)$$

where S, I and R denote the number of susceptible, infective, and recovered individuals, respectively; Λ is the constant recruitment rate of susceptibles; d, γ, and ϵ represent the rates of natural death, recovery, and the disease-induced mortality, respectively. Note that the function $\frac{\beta SI}{1+kI}$ is an incidence rate of the form $\frac{kI^l S}{1+\alpha I^h}$ (Liu et al. 1986), when $l = h = 1$. Here, β is the infection rate, and k measures the inhibition effect. Since the variable R is not involved in the first two equations, system (5) can be reduced to a 2-dimensional model as

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta SI}{1+kI} - (d+\gamma+\epsilon)I. \tag{6}$$

In Zhou and Fan (2012), an additional assumption regarding limited medical treatment resources is introduced to the above model, leading to a model with a saturating treatment term, given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = f_1(S,\ I) = \Lambda - \frac{\beta SI}{1+kI} - dS, \quad \frac{\mathrm{d}I}{\mathrm{d}t} = f_2(S,\ I) = \frac{\beta SI}{1+kI} - (d+\gamma+\epsilon)I - \frac{\alpha I}{\omega+I}, \tag{7}$$

where the real, positive parameter α represents the maximal medical resources per unit time, and the real, positive parameter ω is the half-saturation constant. For convenience, let

$$f_3(S,\ I) = \frac{\beta SI}{1+kI} \quad \text{and} \quad f_4(S,\ I) = \frac{\beta SI}{1+kI} - \frac{\alpha I}{\omega+I}. \tag{8}$$

Further, define the reproduction numbers for systems (6) and (7) as

$$R_0^{(6)} = \frac{\beta\Lambda}{d(d+\gamma+\epsilon)} \quad \text{and} \quad R_0^{(7)} = \frac{\beta\Lambda}{d(d+\gamma+\epsilon+\alpha/\omega)}, \tag{9}$$

respectively. For convenience, let

$$g_1(I) = (d+\gamma+\epsilon)I. \tag{10}$$

Now, if we do not consider the medical treatment term $\frac{\alpha I}{\omega+I}$ and remove it from system (7), that leads to system (6), which is a typical example of an SIR model studied in system (2). Actually, it is easy to see that $f_3(0,\ I) = f_3(S,\ 0) = 0$; $\frac{\partial f_3(S,\ I)}{\partial S} = \frac{\beta I}{1+kI} > 0$ and $\frac{\partial f_3(S,\ I)}{\partial I} = \frac{\beta S}{(1+kI)^2} > 0$ for all $S,\ I > 0$; and $\frac{\partial^2 f_3(S,\ I)}{\partial I^2} = -2\beta k S(1+kI)^{-3} < 0$ for all $S,\ I > 0$. Therefore, the incidence function $f_3(S,\ I)$, satisfies the conditions given in (3). In particular, the function is concave, and can only have one intersection point with the line $(d+\gamma+\epsilon)I$ in the I-S plane, as shown in Fig. 1a. Thus, the uniqueness of the positive infected equilibrium implies that backward bifurcation cannot occur in this case. Moreover, according to the result in Korobeinikov and Maini (2005), the uninfected and infected equilibria are globally asymptotically stable for $R_0^{(6)} < 1$ and $R_0^{(6)} > 1$, respectively. No complex dynamical behavior happens in system (6).

For model (7), we can show that $f_4(\tilde{S},\ I)$ actually has a convex-concave 'S' shape, and may have two positive intersection points with the ray line, $g_1(I)$, in the first

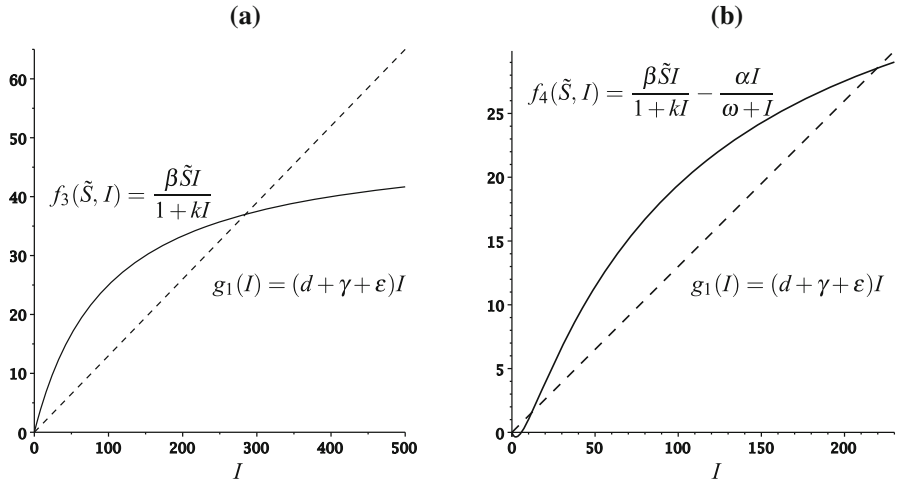

Fig. 1 Graphs of the incidence function f_3 in system (5), (6) and function f_4 in system (7) with respect to I, for which $\tilde{S} = 50$ has been used. The parameter values are chosen as $\beta = 0.01, k = 0.01, \alpha = 6, \omega = 7, d = 0.1, \gamma = 0.01, \epsilon = 0.02$, according to (Zhou and Fan 2012). The *solid lines* denote f_3 in **a** and f_4 in **b**, while the *dashed ray lines* in both graphs represent $g_1(I) = (d + \gamma + \varepsilon)I$. **a** The incidence function $f_3(S, I) = \frac{\beta SI}{1+kI}$, showing one intersection point with g_1; and **b** the function $f_4(S, I) = \frac{\beta SI}{1+kI} - \frac{\alpha I}{\omega+I}$, showing two intersection points with g_1

quadrant; see Fig. 1b. These intersections contribute the two positive equilibrium solutions that are a necessary feature of backward bifurcation. The detailed analysis is given in Appendix A.

In summary we may conclude that the necessary terms which should be contained in system (7) in order to have backward bifurcation are the constant influx Λ, the infection force β, and the saturating medical treatment $\frac{\alpha I}{\omega+I}$.

2.2 Backward bifurcation in the infection model (4) with convex incidence

Now we consider the 2-dimensional infection model (4) which exhibits viral blips, studied in Zhang et al. (2013, 2014a). The motivation for this model was a series of clinical discoveries indicating that viral infection can increase the density of a harmful chemical substance (Gil et al. 2003; Li and Karin 1999; Schwarz 1996; Israel and Gougerot-Pocidalo 1997), thereby amplifying an associated biochemical reaction (Stephenson et al. 2005), and thus accelerating the infection rate (Gil et al. 2003). This cooperative phenomenon in viral infection is expressed by an increasing, saturating infectivity function: $(B + \frac{AY}{Y+C})$. According to the principle of mass action, the incidence function is then denoted as $(B + \frac{AY}{Y+C})XY$, which is a convex function with respect to the infectives' density Y.

To analyze the occurrence of possible backward bifurcation, we first examine the two equilibrium solutions from the following equations:

$$f_5(X,\ Y) = 1 - DX - \left(B + \frac{AY}{Y + C}\right) XY = 0,$$
$$f_6(X,\ Y) = \left(B + \frac{AY}{Y + C}\right) XY - Y = 0, \tag{11}$$

where all parameters A, B, C and D are positive constants. Note that the incidence function of system (4) is given by

$$f_7(X,\ Y) = \left(B + \frac{AY}{Y + C}\right) XY. \tag{12}$$

Similarly, as shown in Appendix B, with appropriate parameter values, $f_7(Y)$ can have a convex-concave 'S' shape, yielding two intersection points with the ray line, $g_2(Y) = Y$, in the first quadrant of the X–Y plane, as shown in Fig. 2b. But if we only consider the second equation in (11), as was done for model (2) in Korobeinikov and Maini (2005), it results in the graph as shown Fig. 2a, implying that only one equilibrium solution would exist. The above discussion, as illustrated in Fig. 2, implies that system (4) can have two positive equilibrium solutions when $R_0 = \frac{B}{D} < 1$, and thus backward bifurcation may occur.

Remark 1 Summarizing the discussions and results given in this section indicates that a disease model with a convex-concave incidence function may lead to backward bifurcation, which in turn implies: (a) the system has at least two equilibrium solutions, and the two equilibrium solutions intersect at a transcritical bifurcation point; and (b) at least one of the equilibrium solutions is determined by a nonlinear equation.

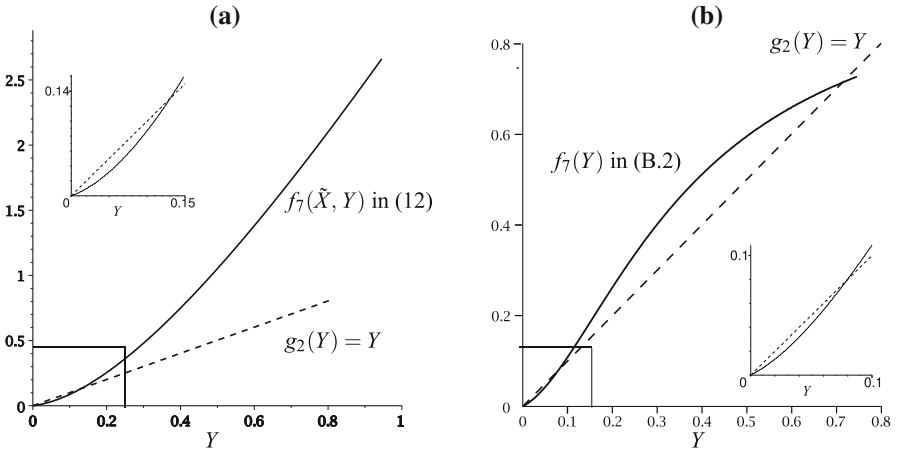

Fig. 2 Graphs of the incidence functions $f_7(\tilde{X}, Y)$ and $f_7(Y)$ for the parameter values $A = 0.364$, $B = 0.03$, $C = 0.823$, and $D = 0.057$. The incidence functions are denoted by the *solid lines*, while the ray lines, determined by $g_2(Y) = Y$, are denoted by *dotted lines*: **a** the incidence function $f_7(\tilde{X}, Y)$, showing one intersection point with g_2 with an *inset*, with a fixed value $\tilde{X} = 12.54$; and **b** the incidence function $f_7(Y)$, showing two intersection points with an *inset*

3 Hopf bifurcations

In the previous section, we studied backward bifurcation and established the necessary conditions for the occurrence of backward bifurcation in two models. In this section, we turn to Hopf bifurcation, since it typically underlies the change of stability in the upper branch of the infected equilibrium, the key condition in determining whether a model can exhibit oscillation or even recurrence. Again, we will present detailed studies for the two models.

3.1 Hopf bifurcation in the infection model (7) with concave incidence

In this subsection, we study two cases of an infection model with concave incidence: system (6) and (7) and show that Hopf bifurcation only appears in model (7). First, we discuss the equilibrium solutions and their stability by using the Jacobian matrix, denoted by J, and examining the corresponding characteristic polynomial,

$$P|_J(L) = L^2 + \mathrm{Tr}(J)L + \mathrm{Det}(J). \tag{13}$$

Bifurcation analysis is conducted by choosing Λ as the bifurcation parameter.

First, we consider the case without saturating medical treatment, system (6). This system satisfies the three conditions in (3), and consequently, its uninfected equilibrium $\bar{\mathrm{E}}_0 = (\frac{\Lambda}{d}, 0)$ is globally asymptotically stable if $R_0^{(6)} \le 1$, while the infected equilibrium $\bar{\mathrm{E}}_1 = \left(\frac{k\Lambda+d+\gamma+\epsilon}{dk+\beta}, \frac{\beta\Lambda-(d+\gamma+\epsilon)d}{(dk+\beta)(d+\gamma+\epsilon)}\right)$ emerges and is globally asymptotically stable if $R_0^{(6)} > 1$. Therefore, for this case the system has only one transcritical bifurcation point at $R_0^{(6)} = 1$ and no complex dynamics can occur. This can also be easily seen from the solution of the infected equilibrium $\bar{\mathrm{E}}_1$ for which $S = \frac{1}{d}[\Lambda - (d+\gamma+\varepsilon)I]$ and I is determined from a linear equation $(d+\gamma+\varepsilon)[d+(dk+\beta)I] - \beta\Lambda = 0$, which implies that no turning point exists.

Next, with the saturating treatment term, system (7) violates the conditions established for model (3), but leads to the possibility of complex dynamical behaviors. In fact, evaluating the Jacobian matrix $J_1 = J|_{(7)}(\bar{\mathrm{E}}_0)$ at the uninfected equilibrium, $\bar{\mathrm{E}}_0 = \left(\frac{\Lambda}{d}, 0\right)$, yields the characteristic polynomial in the form of (13), denoted by $P|_{J_1}(L)$, with $\mathrm{Tr}(J_1) = \left(-\frac{\beta\Lambda}{d}+\epsilon+\frac{\alpha}{\omega}+2d\right)$, and $\mathrm{Det}(J_1) = \left(-\beta\Lambda+d^2+d\epsilon+\frac{\alpha d}{\omega}\right) = \mathrm{Tr}(J_1)\,d - d^2$. This indicates that $\mathrm{Det}(J_1) < 0$ when $\mathrm{Tr}(J_1) = 0$, and thus Hopf bifurcation cannot occur from $\bar{\mathrm{E}}_0$. On the other hand, a static bifurcation can occur when $\mathrm{Det}(J_1) = 0$, that is, $\Lambda_S = \frac{1}{\beta}\left(d^2 + d\epsilon + \frac{\alpha d}{\omega}\right)$, where the subscript '$S$' refers to *static bifurcation*. Therefore, $\bar{\mathrm{E}}_0$ is stable (unstable) for $\Lambda < \Lambda_S$ ($> \Lambda_S$), or $R_0 < 1 (> 1)$, with $R_0 = \beta\Lambda d^{-1}(d+\gamma+\epsilon+\frac{\alpha}{\omega})^{-1}$ (Zhou and Fan 2012).

Next, we show that complex dynamical behaviors can emerge in system (7) from the infected equilibrium $\bar{\mathrm{E}}_1 = (\bar{S}, \bar{I})$. In the Λ-I plane, the bifurcation diagram as shown in Fig. 3(1–4), indicates a turning point on the curve with appropriate parameter values, given by

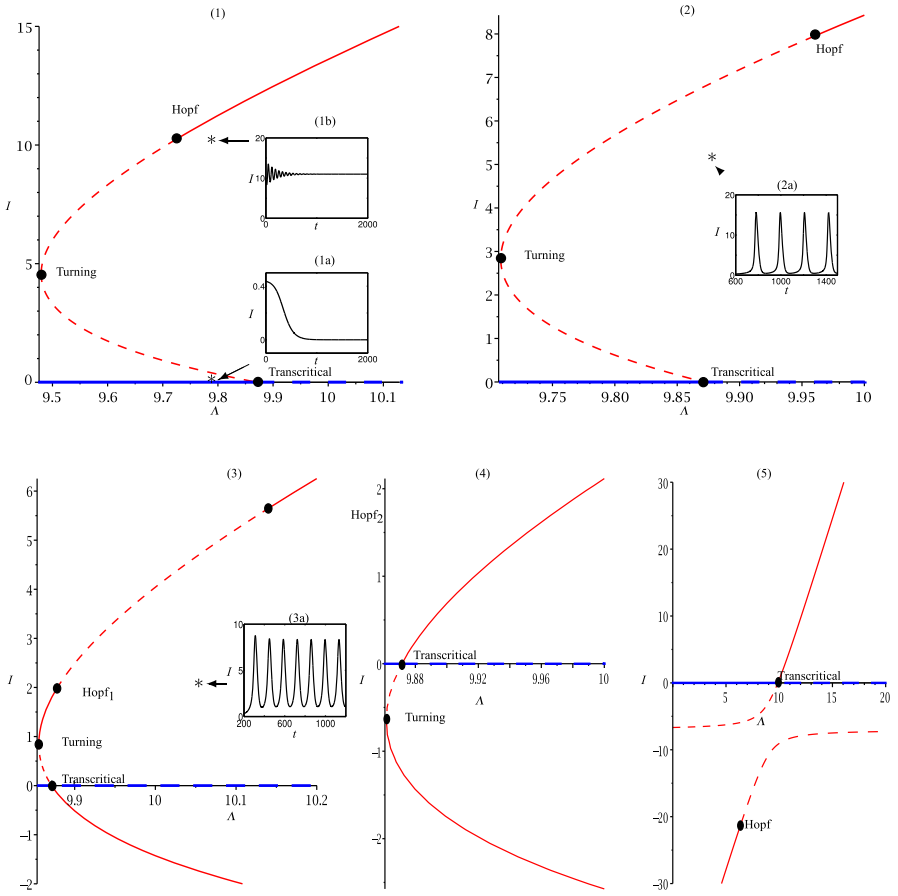

Fig. 3 Bifurcation diagrams and simulations associated with the five cases given in Table 1, demonstrating various dynamical behaviors. All insets are simulated time histories of I vs. t where the values of the initial condition for I and the bifurcation parameter Λ are indicated by the respective stars

$$I_T = \frac{1}{2}\left[\frac{\beta\Lambda_T}{(dk+\beta)(d+\epsilon)} - \omega - \frac{d}{dk+\beta} - \frac{\alpha}{d+\epsilon}\right]. \tag{14}$$

Thus, when $I_T > 0$ (<0), the turning point of the quadratic curve appears above (below) the Λ-axis, meaning that a positive (a negative) backward bifurcation occurs for $I > 0$ (<0). More detailed analysis can be seen in Appendix A. Further, using the Jacobian matrix evaluated at the infected equilibrium $\bar{E}_1$ shows that the necessary condition for system (7) to have a Hopf bifurcation from the infected equilibrium $\bar{E}_1$ is that $h_1(I)$ (whose expression, given in (22), and its analysis are shown in Appendix A) is negative when $\mathrm{Tr}(J_2) = 0$.

In the remaining part of this subsection, we demonstrate various dynamics which may happen in system (7) with different parameter values of k, as shown in Table 1. Taking other parameter values as $\alpha = 6$, $\omega = 7$, $\epsilon = 0.02$, $\gamma = 0.01$, $\beta = 0.01$,

Table 1 Dynamics of system (7) for different values of k, with $\alpha = 6$, $\omega = 7$, $\epsilon = 0.02$, $\gamma = 0.01$, $\beta = 0.01$, $d = 0.1$, and a fixed transcritical bifurcation point $(\Lambda_S, I_S) = (9.87, 0)$

Case	k	(Λ_T, I_T)	$h_1(I) < 0$	(Λ_H, I_H)	Dynamics	Notes
1	0.001	(9.48, 4.57)	$I \in [1.72, \infty]$	(9.73, 10.28)	Bistability	$\Lambda_H < \Lambda_S$
2	0.01	(9.71, 2.82)	$I \in [1.76, \infty]$	(9.96, 8.00)	Recurrence	$\Lambda_H > \Lambda_S$
3	0.02	(9.85, 0.84)	$I \in [1.82, \infty]$	(9.88, 2.09), (10.14, 5.62)	Oscillation	Two Hopf critical points
4	0.027	(9.86, −0.65)	$I \in [1.85, 30.65]$	No Hopf	No oscillation	Negative backward bifurcation
5	0.05	No turning	$I \in [2.01, 15.03]$	(6.18, −22.15)	No oscillation	No backward bifurcation

and $d = 0.1$, and solving the two equations $\mathrm{Tr}(J_2) = 0$ and $\mathscr{F}(I) = 0$ in (21) gives the Hopf bifurcation point candidates, $(\Lambda_H,\ I_H)$, for which $h_1(I_H) < 0$. Since the formula for the transcritical bifurcation point Λ_S has no relation with k, $(\Lambda_S,\ I_S) = (9.87,\ 0)$ is a fixed value pair in Table 1. Bifurcation diagrams and associated numerical simulations are shown in Fig. 3 corresponding to the five cases given in Table 1. The blue lines and red curves represent the uninfected equilibrium $\bar{\mathrm{E}}_0$ and infected equilibrium $\bar{\mathrm{E}}_1$, respectively. The stable and unstable equilibrium solutions are shown by solid and dashed lines/curves, respectively. Positive backward bifurcations occur in Cases 1, 2, and 3, a negative backward bifurcation appears in Case 4, and no backward bifurcation occurs in Case 5 (see Table 1), which are illustrated by the corresponding bifurcation diagrams in Fig. 3(1–5), respectively. For Cases 1 and 2, only one Hopf bifurcation occurs on the upper branch of the infected equilibrium $\bar{\mathrm{E}}_1$, and this bifurcation point exists at the critical point $\Lambda_H < \Lambda_S$ for Case 1 and $\Lambda_H > \Lambda_S$, for Case 2. For Case 1 with $\Lambda = 9.78$, the simulated time history converges to $\bar{\mathrm{E}}_0$ with initial condition IC $= [93.6,\ 0.44]$, shown in Fig. 3(1a), but converges to $\bar{\mathrm{E}}_1$ with initial condition IC $= [46.8,\ 10]$, shown in Fig. 3(1b). This clearly indicates the bistable behavior when $\Lambda_H < \Lambda_S$, and an overlapping stable region for both $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$ exist (see Fig. 3(1). The recurrent behavior for Case 2 is simulated at $\Lambda = 9.87$ with IC $= [50,\ 5]$, shown in Fig. 3(2a). For Case 2, $\Lambda_H > \Lambda_S$, and an overlapping unstable parameter region for both $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$ occurs between Λ_S and Λ_H (see Fig. 3(2). For Case 3, two Hopf bifurcations occur on the left side of Λ_S, and a stable part in the upper branch of $\bar{\mathrm{E}}_1$ exists when Λ passes through the critical value $\Lambda = \Lambda_S$. In this case, although backward bifurcation still exists and the turning point is also located above the Λ-axis, giving two branches of biologically feasible $\bar{\mathrm{E}}_1$, only regular oscillating behavior is observed. The simulated time history is conducted at $\Lambda = 10$, with initial condition IC $= [50,\ 2]$, shown in Fig. 3(3a). For Case 4, only forward bifurcation occurs in the biologically feasible region, and the turning point for backward bifurcation moves down to the fourth quadrant, that is, negative backward bifurcation occurs in this case. The whole upper branch of $\bar{\mathrm{E}}_1$ in the first quadrant is stable, therefore, no oscillations (or recurrence) can happen. Finally, further increases to the value of k change the shape of the red curves, as shown in Fig. 3(5), which again indicates that no biologically meaningful backward bifurcation or oscillations can occur. Note that in Fig. 3(5) a Hopf bifurcation point exists on the lower branch of the equilibrium solution, which is biologically unfeasible since it is entirely below the horizontal axis. In conclusion, interesting dynamical behaviors can emerge in system (7) if backward bifurcation occurs.

3.2 Hopf bifurcation in the infection model (7) with convex incidence

In this subsection, we return to system (4), that is, the 2-dimensional HIV model with convex incidence derived in Zhang et al. (2013, 2014a), and analyze the various dynamical phenomena which system (4) could possibly exhibit. To achieve this, we set B as the bifurcation parameter, and A as a control parameter; the bifurcation analysis will be carried out for various values of A. Also, simulated time histories are provided to illustrate the dynamical behavior predicted in the analysis.

As shown in Appendix B, the uninfected equilibrium $\bar{\mathrm{E}}_0 = (\frac{1}{D},\ 0)$ is stable when $R_0 = \frac{B}{D} < 1$, loses its stability and becomes unstable when B increases to pass through $B_S = D$ (where the subscript S stands for "static bifurcation"), that is $R_0 > 1$, and no other bifurcations can happen.

Next, we examine the infected equilibrium $\bar{\mathrm{E}}_1 = (\bar{X},\ \bar{Y})$ and can find the turning point $(B_T,\ Y_T)$ defined as (see Appendix B)

$$B_T = \frac{-A + D + 2\sqrt{ACD}}{C+1}, \quad Y_T = \frac{A + B - BC - D}{A+B}, \tag{15}$$

where 'T' in the subscript stands for *turning point*. Further, using the function $h_2(Y)$ given in (26) in Appendix B, we know that Hopf bifurcation can occur from the infected equilibrium $\bar{\mathrm{E}}_1$ if $h_2(Y) < 0$. It can be shown (see Appendix B) that a Hopf bifurcation, denoted by (B_H, Y_H), can happen only from the upper branch of the infected equilibrium $\bar{\mathrm{E}}_1$.

The various dynamical behaviors which may appear in system (4) have been classified in Table 2 for different values of the parameter A, with fixed values of $C = 0.823$ and $D = 0.057$. Thus, the transcritical bifurcation point is fixed for all cases: $B_S = D = 0.057$ and $Y_S = 0$. The two solutions B_{h1} and B_{h2} are solved from the two equations (25) $P|_{\bar{\mathrm{E}}_1}(\lambda, Y) = 0$ and (23) $\mathscr{F}_5(Y) = 0$, respectively. They become a Hopf bifurcation point only if their corresponding Y values (Y_{h1} and Y_{h2}, respectively) are in the range such that $h_2(Y) < 0$. Otherwise, system (4) has a pair of real eigenvalues with opposite signs at $(B_{h1},\ Y_{h1})$ or $(B_{h2},\ Y_{h2})$, which is denoted by the superscript '$*$' (which is actually a saddle point) in Table 2, while the Hopf bifurcation point is denoted by the superscript 'H' in Table 2.

Next, we further examine the direction of the Hopf bifurcation, that is, check whether it is a supercritical or subcritical Hopf bifurcation. Since the Jacobian matrix of the system evaluated at the Hopf bifurcation point has a pair of purely imaginary eigenvalues, the linearized system (4) does not determine the nonlinear behavior of the system. Therefore, we take advantage of normal form theory to study the existence of the limit cycles bifurcating from the Hopf bifurcation point as well as their stability. As mentioned earlier, Hopf bifurcation can only occur from the upper branch of the infected equilibrium $\bar{\mathrm{E}}_1$, therefore we first transform the fixed point $\bar{\mathrm{E}}_1$ to the origin by a shifting transformation, and, in addition, make the parameter transformation $B = B_H + \mu$; the Hopf bifurcation point is thus defined as $\mu = \mu_H = 0$. Then the normal form of system (4) near the critical point, $\mu = \mu_H = 0$, takes the form up to third-order approximation:

$$\dot{r} = d\,\mu\, r + a\, r^3 + \mathscr{O}(r^5), \qquad \dot{\theta} = \omega_c + c\,\mu + b\, r^2 + \mathscr{O}(r^4), \tag{16}$$

where r and θ represent the amplitude and phase of the motion, respectively. The first equation of (16) can be used for bifurcation and stability analysis, while the second equation of (16) can be used to determine the frequency of the bifurcating periodic motions. The positive ω_c in the second equation of (16) is the imaginary part of the eigenvalues at the Hopf bifurcation point. The parameters d and c can be

Table 2 Parameter values taken to illustrate various dynamics of system (4)

Case	A	$(B_T,\ Y_T)$	$h_2(Y) < 0,\ Y \in$	$(B_{h1},\ Y_{h1})$
1	0.80	$(-0.1950,\ 0.5850)$	$(0.0036,\ 0.9830)$	$(0.0355,\ 0.8725)^H$
2	0.71	$(-0.1580,\ 0.5660)$	$(0.0040,\ 0.9800)$	$(0.0539,\ 0.0038)^*$
3	0.60	$(-0.1140,\ 0.5380)$	$(0.0048,\ 0.9769)$	$(0.0540,\ 0.0045)^*$
4	0.07	$(0.0557,\ 0.0909)$	$(0.0476,\ 0.8030)$	$(0.0560,\ 0.0470)^*$
5	0.06	$(0.056558,\ 0.05581)$	$(0.0574,\ 0.7700)$	$(0.056559,\ 0.0574)^H$
6	0.05	$(0.05697,\ 0.01442)$	$(0.0724,\ 0.7232)$	$(0.0574,\ 0.0741)^H$
7	0.04	$(0.0569,\ -0.0358)$	$(0.0986,\ 0.6507)$	$(0.0592,\ 0.1071)^H$
8	0.03	$(0.0559,\ -0.0994)$	$(0.1611,\ 0.5149)$	–
Case	A	$(B_{h2},\ Y_{h2})$	Dynamics	Notes
1	0.80	$(0.054,\ 0.0034)^*$	Unstable limit cycle, bistable	$B_{h1} < B_S$
2	0.71	$(0.0574,\ 0.8650)^H$	Recurrence	$B_{h2} > B_S$
3	0.60	$(0.0819,\ 0.8530)^H$	Recurrence	$B_{h2} > B_S$
4	0.07	$(0.1015,\ 0.5612)^H$	Recurrence	$B_{h2} > B_S$
5	0.06	$(0.0961,\ 0.5225)^H$	Recurrence	$B_{h1} < B_S < B_{h2}$
6	0.05	$(0.0894,\ 0.4701)^H$	Recurrence	$B_{h1} < B_S < B_{h2}$
7	0.04	$(0.0806,\ 0.3897)^H$	Oscillation	$B_{h1} < B_S < B_{h2}, Y_T < 0$
8	0.03	–	$\bar{\mathrm{E}}_1$ stable	$Y_T < 0$

The fixed transcritical bifurcation point: $(B_S,\ Y_S) = (0.057,\ 0)$

easily obtained from a linear analysis, while a and b must be derived using a nonlinear analysis, with the Maple program available in, say, (Yu 1998).

Note that the infected equilibrium $\bar{\mathrm{E}}_1$ is represented by the fixed point $\bar{r} = 0$ of system (16), while the nonzero fixed point $\bar{r} > 0$ (satisfying $\bar{r}^2 = \frac{-d\mu}{a}$) is an approximate solution for a limit cycle or periodic orbit. The periodic orbit is asymptotically stable (unstable) if $a < 0$ ($a > 0$), and the corresponding Hopf bifurcation is called supercritical (subcritical). According to the Poincare–Andronov Hopf Bifurcation theorem (Wiggins 2003), for μ sufficiently small, there are four possibilities for the existence of periodic orbits and their stability, which are classified in Table 3, based on the four sets of the parameter values in the normal form (16). Then we use the results presented in Table 3 with a nonlinear analysis based on normal form theory to classify the Hopf bifurcations appearing in Table 2, and the results are shown in Table 4.

To illustrate the analytical results given in Tables 2 and 4, we provide the bifurcation diagrams in Fig. 4(1–8). These figures depict the uninfected equilibrium $\bar{\mathrm{E}}_0$ and the infected equilibrium $\bar{\mathrm{E}}_1$ in blue and red, respectively. The solid and dashed lines differentiate stable and unstable states of the equilibrium solutions. The bifurcation points on the equilibrium solutions are highlighted by solid black dots. Moreover, 'Transcritical', 'Turning', '$\mathrm{Hopf}_{\mathrm{sub}}$', and '$\mathrm{Hopf}_{\mathrm{super}}$', are used to denote *Transcritical bifurcation*, *Turning point*, *subcritical Hopf bifurcation*, and *supercritical Hopf bifur-*

Table 3 Classification of Hopf bifurcations based on the normal form (16)

Class	Stability of $\bar{r} = 0$		Stability of $\bar{r}^2 = -\frac{d\mu}{a}$		Hopf bifurcation
	$\mu < 0$	$\mu > 0$	$\mu < 0$	$\mu > 0$	
(a): $d > 0, a > 0$	Stable	Unstable	Unstable	–	Subcritical
(b): $d > 0, a < 0$	Stable	Unstable	–	Stable	Supercritical
(c): $d < 0, a > 0$	Unstable	Stable	–	Unstable	Subcritical
(d): $d < 0, a < 0$	Unstable	Stable	Stable	–	Supercritical

Table 4 Classification of Hopf bifurcations appearing in Table 2

Case	A	Hopf bifurcation point $(B_H,\ Y_H)$	d	a	Stability of limit cycles	Table 3 class
1	0.8	(0.0355, 0.8725)	−1.0722	0.2114×10^{-2}	Unstable	(c)
2	0.71	(0.0574, 0.8650)	−1.0726	0.1424×10^{-2}	Unstable	(c)
3	0.6	(0.0819, 0.8530)	−1.0733	0.6755×10^{-3}	Unstable	(c)
4	0.07	(0.1015, 0.5612)	−1.0307	-0.8791×10^{-3}	Stable	(d)
5	0.06	(0.056559, 0.0574)	884.27	−0.1019	Stable	(b)
		(0.0961, 0.5225)	−1.0079	-0.8613×10^{-3}	Stable	(d)
6	0.05	(0.0574, 0.0741)	18.232	-0.3145×10^{-2}	Stable	(b)
		(0.0894, 0.4701)	−0.9629	-0.8457×10^{-3}	Stable	(d)
7	0.04	(0.0592, 0.1071)	4.7242	-0.1577×10^{-2}	Stable	(b)
		(0.0805, 0.3897)	−0.8437	-0.8438×10^{-3}	Stable	(d)

cation, respectively. Simulated time histories are used to validate the analytical results, and to show different dynamical behaviors in each case listed in Tables 2 and 4. Subcritical Hopf bifurcation occurs in Cases 1–3, shown in Fig. 4(1–3). $A = 0.8$ is used in Fig. 4(1) for Case 1. Choosing $B = 0.036$, we have $\bar{\mathrm{E}}_0 = [17.1282566,\ 0.023689]$ and $\bar{\mathrm{E}}_1 = [2.233533,\ 0.8726886]$. The simulated solution converges to E_0 or E_1, with initial condition taken as $\mathrm{IC}_d = [17.13,\ 0.024]$ or $\mathrm{IC}_c = [2.233,\ 0.873]$, shown in Fig. 4(1d), (1c), respectively. Fig. 4(1a), (1b), on the other hand, show the unstable limit cycle bifurcating from the subcritical Hopf bifurcation with $\mathrm{IC}_c = [2.233,\ 0.873]$.

Figure 4(2) corresponds to Case 2 with $A = 0.71$. Choosing $B = 0.0572 \in [B_S,\ B_H]$ yields recurrence, independent of the initial conditions, see, for example, the result given in Fig. 4(2b) with $\mathrm{IC}_b = [2.4,\ 0.5]$. However, for $B = 0.06 > B_H$, the simulated time history converges to $\bar{\mathrm{E}}_1$, with an initial condition close to $\bar{\mathrm{E}}_1$, such as $\mathrm{IC}_a = [2.4,\ 0.6]$ as shown in Fig. 4(2a); or shows recurrence with an initial condition far away from $\bar{\mathrm{E}}_1$, such as $\mathrm{IC}_c = [2.4,\ 0.4]$, as shown in Fig. 4(2c).

Figure 4(3) plots the result for Case 3 with $A = 0.6$, and shows a broader region between the transcritical and Hopf bifurcation points, associated with a larger recurrent region. Recurrence occurs independent of the initial conditions for $B = 0.083 \in [B_S,\ B_H]$, giving $\bar{\mathrm{E}}_0 = [12.048,\ 0]$ and $\bar{\mathrm{E}}_1 = [2.576,\ 0.852]$, as shown in Fig. 4(3a),

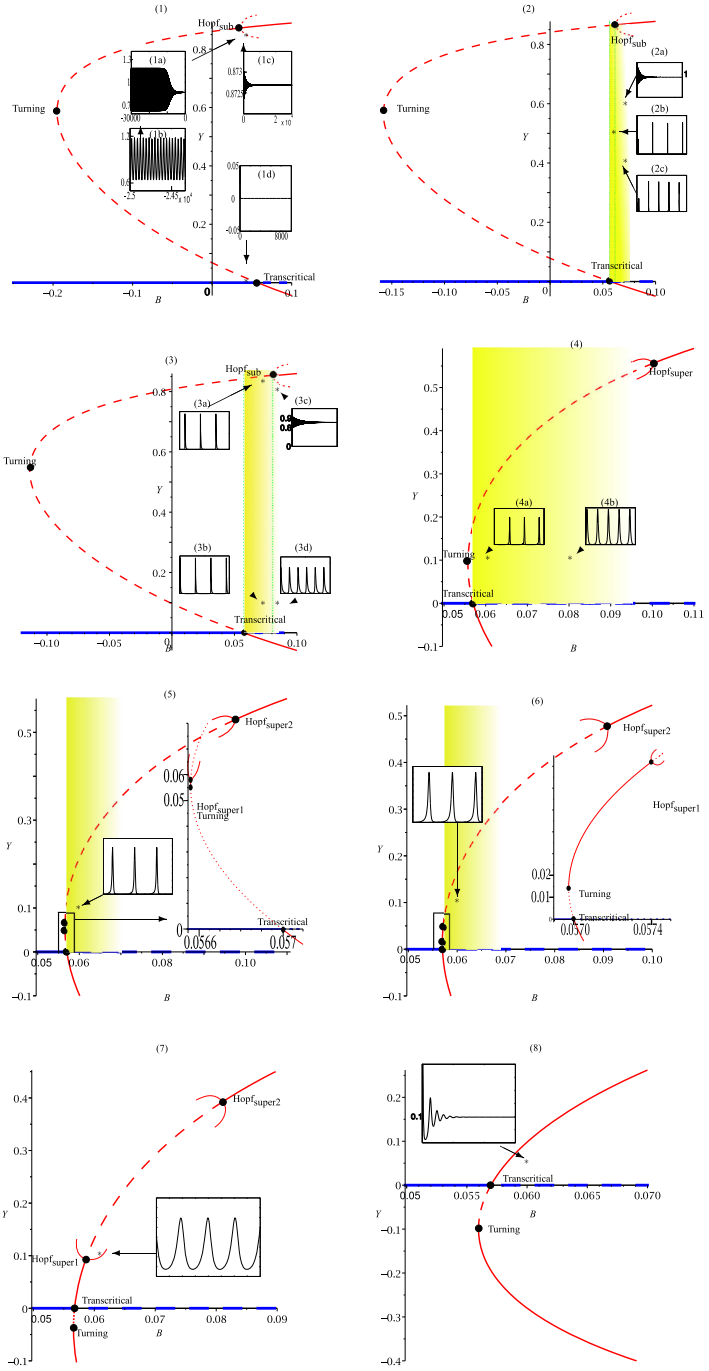

Fig. 4 Dynamical behaviors of system (4) corresponding to eight cases listed in Tables 2 and 4. All *insets* are simulated time histories of Y vs. t. The *yellow* areas fading to *white* show regions in which recurrent behavior occurs and fades to regular oscillations

(3b), with $\mathrm{IC}_a = [2.7,\ 0.84]$ and $\mathrm{IC}_b = [14,\ 0.1]$, respectively. But if we choose $B = 0.07 > B_H$, we have $\bar{\mathrm{E}}_0 = [14.286, 0]$ and $\bar{\mathrm{E}}_1 = [2.67,\ 0.8478]$. The time history converges to $\bar{\mathrm{E}}_1$ with $\mathrm{IC}_c = [2.6,\ 0.8]$, or shows recurrence with $\mathrm{IC}_d = [2.6,\ 0.1]$, as shown in Fig. 4(3c), (3d), respectively.

Supercritical Hopf bifurcations occur in Cases 4–7, as shown in Fig. 4(4–7). Figure 4(4) depicts the result for Case 4 with $A = 0.07$. Only one supercritical Hopf bifurcation happens in this case, and gives a large recurrent parameter region between the transcritical and Hopf bifurcation points. Although the simulated recurrent behavior does not depend on initial conditions, the recurrent pattern will fade out with the growth of the value of B from the transcritical point to the Hopf bifurcation point, see Fig. 4(4a), (4b) with the same $\mathrm{IC}_{a,\,b} = [8,\ 0.1]$, but different values of B: $B = 0.06$ and $B = 0.09$, respectively.

Figure 4(5) shows the result for Case 5 with $A = 0.06$. A transcritical bifurcation happens between two supercritical Hopf bifurcations. The recurrent region still starts from the transcritical point and independent of the initial conditions, but is narrower than that shown in Fig. 4(4). The simulated recurrent behavior for this case is conducted at IC$= [12,\ 0.1]$ and $B = 0.06$. Figure 4(6) corresponds to Case 6 with $A = 0.05$, and two supercritical Hopf bifurcations occur on the right side of the transcritical bifurcation point, which makes the recurrent region even narrower and the recurrent pattern less obvious, as shown in the simulated time history with IC $= [10,\ 0.1]$ and $B = 0.06$. Negative backward bifurcations occur in Cases 7 and 8, as shown in Fig. 4(7, 8). Although two Hopf bifurcations are still present in Case 7, see Fig. 4(7), only a regular oscillating pattern exists. For Case 8, no Hopf bifurcation happens in the biologically feasible part of $\bar{\mathrm{E}}_1$, and therefore no more interesting dynamics occur.

In general, backward bifurcations are much more likely to induce Hopf bifurcation. A Hopf bifurcation can only occur along the upper branch of $\bar{\mathrm{E}}_1$, since $\bar{\mathrm{E}}_0$ only changes its stability at a transcritical bifurcation point, and any point on the lower branch of $\bar{\mathrm{E}}_1$ is a saddle node (Yu et al. 2015). Moreover, Hopf bifurcation can lead to a change in the stability of the upper branch of the infected equilibrium $\bar{\mathrm{E}}_1$. Thus the system further develops bistable, recurrent, or regular oscillating behavior, corresponding to Cases 1–7 in Tables 2 and 4, and in Fig. 4(1–7). In particular, bistability happens when both equilibria $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$ share a stable parameter region, see Case 1 in Table 2 and Fig. 4(1).

As for recurrent behavior, we observe that recurrence is more likely to happen if the following two conditions are satisfied for the upper branch of $\bar{\mathrm{E}}_1$: (1) the equilibrium remains unstable as the bifurcation parameter increases and crosses the transcritical point, where $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$ intersect, such that the two equilibria share an unstable parameter range; and (2) at least one Hopf bifurcation occurs from $\bar{\mathrm{E}}_1$. As shown in Cases 2–6 in Table 2, and the corresponding Fig. 4(2–8), the common recurrent parameter region for both subcritical and supercritical Hopf bifurcations starts beside the transcritical point, and is located entirely in the unstable parameter region of $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$. The simulated recurrent pattern becomes more pronounced if the value of the bifurcation parameter is close to the transcritical point, but approaches an oscillatory pattern as the parameter diverges from the transcritical point, as shown in Fig. 4(4a), (4b). In this common recurrent parameter region, recurrence occurs independent of initial conditions; see Fig. 4(3a), (3b). In addition to the common recurrent region, for subcritical bifurcation, seen in Table 2 for Cases (2) and (3) and Fig. 4(2), (3), recurrence may

also appear on the stable side of the subcritical Hopf bifurcation point with an initial condition close to $\bar{E}_1$. Moreover, the subcritical Hopf bifurcation and the transcritical point should be close to each other for a clear recurrent pattern. When this is not the case, the periodic solutions show a more regular oscillating pattern, as compared in Fig. 4(2c), (3d). Although two Hopf bifurcation points occur in Table 2 for Case 5, see Fig. 4(5), the transcritical point is located inside the unstable range of the upper branch of $\bar{E}_1$, between the two Hopf bifurcation points. A recurrent pattern still characterizes the dynamical behavior in this case. However, if the unstable range of $\bar{E}_1$, between the two Hopf bifurcation points, is located entirely in the unstable range of $\bar{E}_0$, and moves further away from the transcritical point, the recurrent motion gradually becomes a regular oscillation, as shown in Fig. 4(6, 7).

Remark 2 It has been proved in Zhang et al. (2014a) that the recurrence appearing when $B > D$ (i.e., $R_0 > 1$) is not due to homoclinic bifurcation since no homoclinic orbits can exist for system (4) when $R_0 > 1$. However, it has also been shown in Yu et al. (2015) that when $B < D$ (i.e., $R_0 < 1$), the system could have Bogdanov–Takens bifurcation if proper parameter values of A and C are chosen, which may yield homoclinic bifurcation, leading to another type of mechanism for generating viral blips. This is not discussed in this paper. Therefore, for the HIV model (4) when $B > D$, numerical continuation of the periodic orbits produced by the Hopf bifurcation does not lead to a homoclinic orbit.

Summarizing the results and discussions presented in the previous two sections, we have the following observations.

1. Due to the fact that $\bar{E}_0$ only changes its stability at the transcritical bifurcation point, and the fact that any point on the lower branch of $\bar{E}_1$ is a saddle node, Hopf bifurcation can only occur from the upper branch of $\bar{E}_1$. A Hopf bifurcation may result in convergent, recurrent, bistable, or regular oscillating behaviors.
2. Backward bifurcation gives rise to two branches in the infected equilibrium $\bar{E}_1$. Hopf bifurcation is more likely to happen when the turning point of the backward bifurcation is located on the positive part of the equilibrium solution in the bifurcation diagram, as shown in Fig. 4(2–6). This means that we have two biologically feasible infected equilibria, which is essential to observe bistability, as shown in Fig. 4(1).
3. However, if the turning point on the infected equilibrium $\bar{E}_1$, or the backward bifurcation moves down to the negative part of a state variable in the bifurcation diagram, that is, negative backward bifurcation occurs, then Hopf bifurcation is very unlikely to happen. Although Fig. 4(7) shows an exceptional case, the parameter range for such a Hopf bifurcation is very narrow.
4. The bifurcation diagram for system (4) with $A = 0.03$, shown in Fig. 4(8), is a typical model with negative backward bifurcation. Such negative backward bifurcation may occur in higher-dimensional systems. However, by considering more state variables, which make the system more complicated, Hopf bifurcation can happen in the upper branch of the negative backward bifurcation. We will discuss this possibility in more detail in the next section by examining an autoimmune disease model.

The results obtained in this section suggest the following summary.

Remark 3 If a disease model contains a backward bifurcation on an equilibrium solution, then as the system parameters are varied, there may exist none, one or two Hopf bifurcations from the equilibrium solution, which may be supercritical or subcritical. If further this equilibrium has a transcritical bifurcation point at which it exchanges its stability with another equilibrium, then recurrence can occur between the transcritical and Hopf bifurcation points and near the transcritical point, where both equilibrium solutions are unstable, and bistability happens when Hopf bifurcation makes a shared stable parameter region for both equilibria.

4 An autoimmune disease model with negative backward bifurcation

In the previous section, we examined three cases of negative backward bifurcation: Table 1 Case 4 for system (7) and Table 2 Case (7) and (8) for system (4). The analytical and numerical results showed that solutions typically converge to the infected equilibrium in these cases, and the parameter range for Hopf bifurcation is very limited. As a result, negative backward bifurcation tends to give no interesting behavior. In this section, however, we shall explore an established autoimmune model (Alexander and Wahl 2011) in which negative backward bifurcation occurs. We demonstrate that after modification, the autoimmune model can also exhibit recurrence.

The autoimmune model (Alexander and Wahl 2011) takes the form

$$\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}t} &= f\tilde{v}G - (\sigma_1 R_n + b_1)A - \mu_A A \\
\frac{\mathrm{d}R_n}{\mathrm{d}t} &= (\pi_1 E + \beta)A - \mu_n R_n \\
\frac{\mathrm{d}E}{\mathrm{d}t} &= \lambda_E A - \mu_E E \\
\frac{\mathrm{d}G}{\mathrm{d}t} &= \gamma E - \tilde{v}G - \mu_G G,
\end{aligned} \tag{17}$$

where mature pAPCs (A) undergo maturation by intaking self-antigen (G), at rate $f\tilde{v}$, and are suppressed by specific regulatory T cells, T_{Reg} cells (R_n), at rate σ_1; b_1 represents additional non-specific background suppression. The T_{Reg} cells are activated by mature pAPCs at a rate proportional to the number of auto-reactive effector T cells (E) at rate π_1, and by other sources at rate β. Active auto-reactive effector T cells (E) come from the activation process initiated by mature pAPCs, at rate λ_E, then attack healthy body tissue and release free self-antigen (G) at rate γ, which is ready for mature pAPCs to engulf; the antigen engulfing rate is $\tilde{v}$. The death rates of the populations A, R_n, E, and G are denoted by μ_A, μ_n, μ_E, and μ_G, respectively.

Following the steps described by Zhang et al. (2014b), system (17), can be reduced via quasi-steady state analysis to a 2-dimensional system:

$$\begin{aligned}\frac{\mathrm{d}A}{\mathrm{d}t} &= \left[\frac{f\tilde{v}\gamma\lambda_E}{\mu_E(\tilde{v}+\mu_G)} - b_1 - \mu_A\right]A - \sigma_1 R_n A,\\ \frac{\mathrm{d}R_n}{\mathrm{d}t} &= \left(\frac{\pi_1\lambda_E}{\mu_E}A + \beta\right)A - \mu_n R_n.\end{aligned} \tag{18}$$

For simplicity, we set $a = \frac{f\tilde{v}\gamma\lambda_E}{\mu_E(\tilde{v}+\mu_G)} - b_1 - \mu_A$ and $b = \frac{\pi_1\lambda_E}{\mu_E}$. For the stability and bifurcation analysis, we define the reproduction number $R_0 = \frac{(b_1+\mu_A)(\tilde{v}+\mu_G)\mu_E}{\lambda_E f\tilde{v}\gamma}$. System (18) has a disease-free equilibrium $\bar{\mathrm{E}}_0 = (0,\ 0)$, which is stable if $R_0 < 1$ and unstable if $R_0 > 1$. Thus a static bifurcation occurs on $\bar{\mathrm{E}}_0$ when $R_0 = 1$. The disease equilibrium is given by $\bar{\mathrm{E}}_1 = (\bar{A},\ \bar{R}_n)$, in which $\bar{R}_n = \frac{(b\bar{A}+\beta)\bar{A}}{\mu_n}$, and $\bar{A}$ is given by the roots of the following equation,

$$f_8(A) = b\sigma_1 A^2 + \beta\sigma_1 A - \mu_n a. \tag{19}$$

It follows from (19) that the turning point is given by $A_T = -\frac{\beta}{2b} < 0$, clearly showing that this is a negative backward bifurcation. We further examine the characteristic equation at $\bar{\mathrm{E}}_1$, which shares the same form as equation (13), with $\mathrm{Tr}(J|_{\bar{\mathrm{E}}_1}) = \frac{1}{\mu_n}(b\sigma_1 A^2 + \beta\sigma_1 A + \mu_n^2 - a\mu_n) := a_{11}$ and $\mathrm{Det}(J|_{\bar{\mathrm{E}}_1}) = 3b\sigma_1 A^2 + 2\beta\sigma_1 A - a\mu_n := a_{12}$. Solving $f_8(A) = 0$ and $a_{12} = \mathrm{Det}(J|_{\bar{\mathrm{E}}_1}) = 0$, gives the static bifurcation point of $\bar{\mathrm{E}}_1$ at $(\bar{A},\ a) = (0,\ 0)$, which is a transcritical bifurcation point between $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$. Moreover, Hopf bifurcation can happen if and only if $f_8(A) = 0$ and $a_{11} = \mathrm{Tr}(J|_{\bar{\mathrm{E}}}) = 0$, which can be satisfied only if $\mu_n = 0$. This implies that the positive branch of $\bar{\mathrm{E}}_1$ is stable for any positive values of μ_n. Thus, this model cannot exhibit recurrence, bistability, or even regular oscillation. The same conclusion was obtained in Zhang et al. (2014b) for the original 4-dimensional model (17).

However, a recent experimental discovery (Baecher-Allan et al. 2006) has revealed a new class of terminally differentiated $\mathrm{T}_{\mathrm{Reg}}$ cells. As described in detail in Zhang et al. (2014b), introducing this cell population, denoted R_d, into the model yields the full system

$$\begin{aligned}\frac{\mathrm{d}A}{\mathrm{d}t} &= f\tilde{v}G - \sigma_1(R_n + dR_d)A - (b_1+\mu_A)A\\ \frac{\mathrm{d}R_n}{\mathrm{d}t} &= (\pi_1 E + \beta)A - \mu_n R_n - \xi R_n\\ \frac{\mathrm{d}R_d}{\mathrm{d}t} &= c\xi R_n - \mu_d R_d\\ \frac{\mathrm{d}E}{\mathrm{d}t} &= \lambda_E A - \mu_E E\\ \frac{\mathrm{d}G}{\mathrm{d}t} &= \gamma E - \tilde{v}G - \mu_G G\end{aligned}$$

and quasi-steady state analysis then yields a reduced 3-dimensional model in the form

$$\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}t} &= \left[\frac{f\tilde{v}\gamma\lambda_E}{(\tilde{v}+\mu_G)\mu_E} - (b_1+\mu_A)\right]A - \sigma_1(R_n + dR_d)A,\\
\frac{\mathrm{d}R_n}{\mathrm{d}t} &= \left(\frac{\pi_1\lambda_E}{\mu_E}A + \beta\right)A - \mu_n R_n - \xi R_n,\\
\frac{\mathrm{d}R_d}{\mathrm{d}t} &= c\xi R_n - \mu_d R_d.
\end{aligned} \tag{20}$$

Again, here λ_E is chosen as the bifurcation parameter for stability and bifurcation analysis. It is easy to show that system (20) still has a disease-free equilibrium $\bar{\mathrm{E}}_0$ as $(A,\ R_n,\ R_d) = (0,\ 0,\ 0)$, and a disease equilibrium $\bar{\mathrm{E}}_1$ as $(\bar{A},\ \bar{R}_n,\ \bar{R}_d)$, where $\bar{R}_d = \frac{c\xi\bar{R}_n}{\mu_d}$, $\bar{R}_n = \frac{\beta\mu_E+\pi_1\lambda_E\bar{A}}{\mu_E(\mu_n+\xi)}\bar{A}$, and $\bar{A}$ is determined from the following quadratic equation:

$$\begin{aligned}
f_9(A) = {} & \pi_1\lambda_E A^2 + \beta\mu_E A\\
& + \frac{\mu_d(\mu_n+\xi)}{(\tilde{v}+\mu_G)(cd\xi+\mu_d)\sigma_1}[-f\gamma\tilde{v}\lambda_E + (b_1+\mu_A)(\mu_G+\tilde{v})\mu_E],
\end{aligned}$$

which gives two negative roots if $R_0 > 1$ and two roots with opposite signs when $R_0 < 1$. The critical point is determined by $R_0 = 1$, which is actually the intersection point of $\bar{\mathrm{E}}_0$ and $\bar{\mathrm{E}}_1$. The two equilibrium solutions exchange their stability at $R_0 = 1$, leading to a transcritical bifurcation at $(\bar{A},\ \lambda_E) = (0,\ \frac{(b_1+\mu_A)(\mu_G+\tilde{v})\mu_E}{f\gamma\tilde{v}})$. Note that the negative backward bifurcation still happens in system (20). Moreover, now a Hopf bifurcation occurs from the upper branch of $\bar{\mathrm{E}}_1$, giving rise to oscillation and recurrence.

Realistic parameter values have been obtained in Zhang et al. (2014b), and are given as follows:

$$\begin{aligned}
& f = 1\times10^{-4},\ \ \tilde{v} = 0.25\times10^{-2},\ \ \sigma_1 = 3\times10^{-6},\ \ b_1 = 0.25,\\
& \quad \mu_A = 0.2,\ \ \pi_1 = 0.016,\\
& \beta = 200,\ \ \mu_n = 0.1,\ \ \mu_E = 0.2,\ \ \gamma = 2000,\ \ \mu_G = 5,\ \ \mu_d = 0.2,\\
& \quad c = 8,\ \ d = 2,\ \ \xi = 0.025.
\end{aligned}$$

For the above parameter values, the Hopf critical point is obtained at $(A_H,\ \lambda_{EH}) = (5.6739,\ 1691.6414)$, while the turning point is at $(A_T,\ \lambda_{ET}) = (-1.4205,\ 879.9848)$, indeed showing a negative backward bifurcation, and the transcritical bifurcation point is at $(A_S,\ \lambda_{ES}) = (0,\ 900.45)$. These three bifurcation points and the stability of equilibrium solutions are shown in the bifurcation diagram given in Fig. 5a, and the simulated recurrent time history is plotted in Fig. 5b for $\lambda_E = \lambda_{EH} + 1000$.

In summary, when a negative backward bifurcation occurs, that is, when the turning point is located in the negative state variable space, less complex dynamical behavior will be present. Hopf bifurcation in a biologically feasible area does not happen in the reduced 2-dimensional system (18), nor in the original system (17) (Zhang et al. 2014b). However, if we increase the dimension of the system, Hopf bifurcation and complex dynamical phenomena can emerge, as shown in our results for system (20).

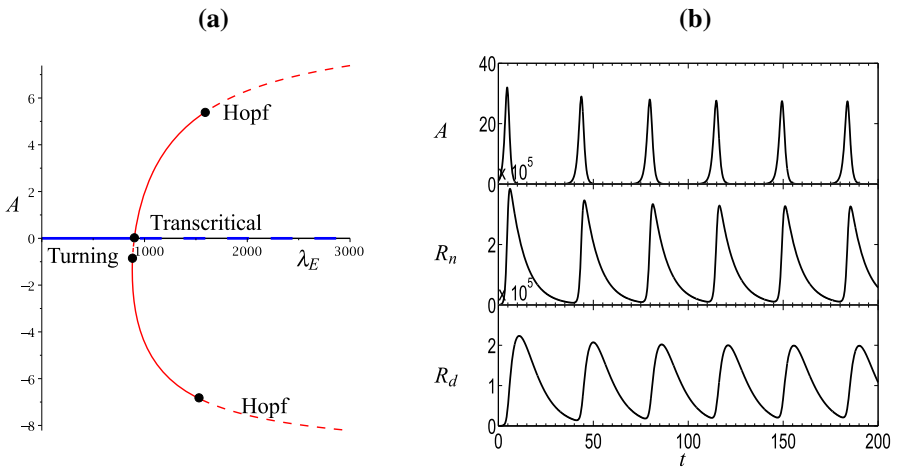

Fig. 5 Dynamics of system (20): **a** bifurcation diagram; and **b** simulated time history for $\lambda_E = \lambda_{EH} + 1000$

5 Conclusion and discussion

In this paper, we demonstrate that the occurrence of turning points in bifurcation diagrams may give rise to backward bifurcations, which induce rich dynamical behaviors. The existence of a turning point implies a nonlinear form (typically quadratic) of the equilibrium solution in terms of at least one state variable. We use the location of the turning point—above or below the horizontal axis in the bifurcation diagram—to define whether the backward bifurcation is "positive" or "negative". Both types of backward bifurcation are capable of yielding stability changes in the positive endemic equilibrium solution, for example through Hopf bifurcation. The emergence of one or two Hopf bifurcation(s) can contribute regular oscillations or even large-amplitude oscillations, namely recurrence, to the system behavior. In fact, a turning point actually defines a saddle-node bifurcation on a one-dimensional manifold embedded in the whole system. In the case of positive backward bifurcation, two endemic equilibrium solutions exist, one of which is a saddle point (Yu et al. 2015), and bistability can occur.

Three simple models are analyzed in this paper, describing problems in epidemiology, in-host viral dynamics and immunology. The in-host model is shown to be mathematically equivalent to a well-studied population level SIR model. In this general infectious disease model, different nonlinear incidence functions are examined. A concave incidence function which includes a saturation effect in terms of the number of infectives generates a forward bifurcation, and globally stable disease-free and endemic equilibria emerge for reproductive number $R_0 < 1$ and $R_0 > 1$, respectively (Korobeinikov and Maini 2005). However, this global stability is broken when an extra term denoting saturating treatment or hospital resources is added to a system which has the above property (Zhou and Fan 2012).

A convex incidence with respect to the number of infectives, on the other hand, enables backward bifurcation to occur on the positive branch of the disease equi-

librium solution. In general, the appearance of backward bifurcation increases the parameter range for Hopf bifurcation, leading to recurrent, bistable and regular oscillating behaviors. These dynamics are determined by the location and direction of the Hopf bifurcation(s), as determined by parameter values.

For the autoimmune disease model in Zhang et al. (2014b), conventional forward bifurcation appears in the first quadrant of the bifurcation diagram, however this branch of the solution is actually determined by a quadratic equation and has a turning point below the horizontal axis in the bifurcation diagram—a negative backward bifurcation. Moreover, a Hopf bifurcation could occur in the positive endemic equilibrium solution in the first quadrant and serve to switch stability of the equilibrium. Oscillation and recurrence are possible as well.

Biologically, a convex incidence function implies a cooperative effect in disease spread or disease progression, that is, the infection rate increases superlinearly with the number of infectives. In an SIR model, this effect occurs when health care resources saturate (Zhou and Fan 2012). In an in-host model, this occurs when existing infection makes the host more vulnerable to further infection. For example, HIV infection produces high levels of reactive oxygen species (ROS) (van Gaalen and Wahl 2009), which in turn accelerate viral production by infected cells. Similarly, the autoimmune model we investigate shows cooperative effects. In particular, the suppressors, natural T_{Reg} cells, are activated to generate the terminal T_{Reg} cells with higher suppressive capability. Both natural and terminal T_{Reg} cells work together to enhance immune regulation of the auto-reactive T cells, which are responsible for T cell mediated autoimmune diseases.

The main implication of our work from a disease perspective is that cooperative infections can yield systems capable of bistability and recurrence. We predict that these complex behaviours may be displayed by any disease in which, as with HIV or autoimmune disorders, the host becomes more vulnerable to further infection as the disease progresses. Such diseases are likely to display bistability and relapse-remission cycles. Thus infections which appear to be cleared may recur after a long interval, and patients for whom R_0 has been reduced below unity by drug therapy may nonetheless continue in a symptomatic state. Our work suggests that in many diseases which are known to display cycles of relapse and remission, a cooperative or convex infection term should be sought as the underlying cause. A more positive implication of this work is that it may be possible, in such diseases, to move the patient to a stable infection-free equilibrium state through a parameter change or through a perturbation to one of the state variables.

Backward bifurcation has also been observed in a multi-group compartmental AIDS model in Dushoff et al. (1998), which reveals the effect of a core group on disease transmission. Similarly, by considering multiple groups in a mosquito population, the properties of a traditional vector-born infectious disease model of malaria are altered and show backward bifurcation and multiple Hopf bifurcations (Ngonghala et al. 2014). The bifurcation diagrams of these models are similar to Fig. 4(6). Here, the positive endemic equilibrium becomes stable, oscillating, and then stable again as the bifurcation parameter increases. Thus, these system behaviors indicate that incomplete malaria treatment could suppress clinical symptoms, but a disease outbreak is still possible if certain conditions are satisfied. Rich dynamical behaviors are also shown in

a predator-prey food chain system including an inorganic resource as a state variable. The Nitrogen–Chlorella–Brachionus model (Fussmann et al. 2000) undergoes two Hopf bifurcations on the positive equilibrium and likewise exhibits stable, oscillating, and stable equilibria as the bifurcation parameter, indicating the nutrition enrichment rate, is varied.

Our results demonstrate that systems involving backward bifurcations, with turning points that occur either in positive or negative state variable space, can exhibit rich behaviors, including bistability and recurrence. Relevant systems cover a broad range of research problems including infectious disease models at both the population and in-host levels, immunological models, and ecological models. We show that convex incidence functions and cooperative effects contribute to these behaviors.

Appendix A

The equilibrium solutions of system (7) are obtained by solving the following algebraic equations: $f_1(S,\ I) = 0$ and $f_2(S,\ I) = 0$, from which the disease-free equilibrium can be easily obtained as $\bar{\mathrm{E}}_0 = (\Lambda/d,\ 0)$. For the infected equilibrium $\bar{\mathrm{E}} = (\bar{S},\ \bar{I})$, $\bar{S}$ is solved from $f_1 = 0$ as $\bar{S}(I) = \frac{\Lambda(1+kI)}{(dk+\beta)I+d}$. Then, substituting $S = \bar{S}(I)$ into $f_2 = 0$ yields a quadratic equation of the form

$$\mathscr{F}(I) = \mathscr{A}I^2 + \mathscr{B}I + \mathscr{C} = 0, \tag{21}$$

which in turn gives two roots: $\bar{I}_{1,\,2} = \frac{-\mathscr{B}\pm\sqrt{\mathscr{B}^2-4\mathscr{A}\mathscr{C}}}{2\mathscr{A}}$, where, $\mathscr{A} = (d+\gamma+\epsilon)(dk+\beta)$, $\mathscr{B} = [(dk+\beta)\omega+d](d+\gamma+\epsilon)+(dk+\beta)\alpha-\beta\Lambda$, $\mathscr{C} = [(d+\gamma+\epsilon)\omega+\alpha]d-\beta\Lambda\omega$ for system (7). Since all parameters take positive values, we have $\mathscr{A} > 0$. To get the two positive roots essential for backward bifurcation, it is required that $\mathscr{B} < 0$ and $\mathscr{C} > 0$. Noticing that $\beta,\ \Lambda,\ \omega > 0$, we can see that the infection force, β, the constant influx of the susceptibles, Λ, and the effect of medical treatment $\frac{\alpha I}{\omega+I}$ are indispensable terms for backward bifurcation. The number of positive infected equilibrium solutions changes from two to one when the value of $\mathscr{C}$ passes from negative to positive, which gives a critical point at $\mathscr{C} = 0$, that is, $[(d+\gamma+\epsilon)\omega+\alpha]d = \beta\Lambda\omega$, which is equivalent to $R_0^{(7)} = 1$.

On the other hand, we may infer the emergence of backward bifurcation without solving the equilibrium conditions. When we introduce the loss of the infectives due to medical treatment, the dynamics of system (7) differ greatly from system (6). In particular, backward bifurcation emerges and complex dynamical behaviors may occur. To clarify this effect, we obtain the function $f_4(S,\ I)$ from the equation $\frac{dI}{dt} = 0$ of (7). Note that $f_4(S,\ I)$ is not an incidence rate. But, if we fix $S = \tilde{S} > 0$, there exist $0 < I_1 < I_2 < +\infty$, such that $\frac{\partial f_4(\tilde{S},\ I)}{\partial I} = \frac{1}{(1+kI)^2(\omega+I)^2}[\beta\tilde{S}(\omega+I)^2 - \alpha\omega(1+kI)^2] > 0$, $\forall\ I \in (0,\ I_2)$; and $\frac{\partial^2 f_4(\tilde{S},\ I)}{\partial I^2} = -2k\beta\tilde{S}(1+kI)^{-3}+2\alpha\omega(\omega+I)^{-3} > 0, \forall\ I \in (0,\ I_1)$, $\frac{\partial^2 f_4(\tilde{S},\ I)}{\partial I^2} = 0$, for $I = I_1$, $\frac{\partial^2 f_4(\tilde{S},\ I)}{\partial I^2} < 0, \forall\ I \in (I_1,\ I_2)$. Thus, $f_4(\tilde{S},\ I)$ actually has a convex-concave 'S' shape, and may have two positive intersection points with the ray line, $g_1(I)$, in the first quadrant.

The infected equilibrium of (7) is denoted as $\bar{\mathrm{E}}_1 = (\bar{S},\ \bar{I})$, where $\bar{I}$ is solved from the equation $\mathscr{F}(I) = 0$ in (21). The turning point is determined by both the quadratic equation (21) and the relation $\frac{\mathrm{d}\Lambda}{\mathrm{d}I} = -\frac{\partial\mathscr{F}}{\partial I}/\frac{\partial\mathscr{F}}{\partial\Lambda} = 0$, which is equivalent to $\frac{\partial\mathscr{F}}{\partial I} = 0$. Solving $\frac{\partial\mathscr{F}}{\partial I} = 0$ yields the expression of the turning point of I, given in (14). To find the stability of the infected equilibrium $\bar{\mathrm{E}}_1$, evaluating the Jacobian matrix at $\bar{\mathrm{E}}_1$, and further denoting it as $J_2 = J|_{(7)}(\bar{\mathrm{E}}_1)$, we obtain the characteristic polynomial in the form of (13), with $\mathrm{Tr}(J_2) = a_{11}/[(\omega + I)^2(kI + 1)(dkI + \beta I + d)]$ and $\mathrm{Det}(J_2) = a_{21}/[(\omega + I)^2(kI + 1)(dkI + \beta I + d)]$, where $a_{11} = a_{1a} - a_{1b}$ and $a_{21} = a_{2a} - a_{2b}$, with $a_{1b} = \beta\Lambda(\omega + I)^2$ and $a_{2b} = da_{1b}$, and a_{1a} and a_{2a} only contain positive terms (their expressions are omitted here for brevity). Determining whether a Hopf bifurcation can occur from $\bar{\mathrm{E}}$ is equivalent to finding whether $\mathrm{Det}(J_2)$ remains positive when $\mathrm{Tr}(J_2) = 0$. Consider the following expression:

$$\begin{aligned} h_1(I) &= \left[\mathrm{Tr}(J_2) - \frac{1}{d}\mathrm{Det}(J_2)\right](\omega + I)^2(kI + 1)(dkI + \beta I + d) \\ &= a_{11} - \frac{1}{d}a_{21} = a_{1a} - \frac{1}{d}a_{2a} \\ &= \frac{1}{d}(dkI + \beta I + d)[(kI + 1)d^2(\omega + I)^2 - \beta\epsilon I(\omega + I)^2 - \alpha\beta\omega I], \end{aligned} \tag{22}$$

where the expressions of a_{1a} and a_{2a} have been used. This indicates that $h_1(I)$ may take negative values, for which $\mathrm{Det}(J_2) > 0$.

Appendix B

It is easy to find the uninfected equilibrium of model (4), $\bar{\mathrm{E}}_0 = (\bar{X}_0,\ \bar{Y}_0) = (\frac{1}{D},\ 0)$, whose characteristic polynomial has two roots: $\lambda_1 = -D < 0$, and $\lambda_2 = \frac{B}{D} - 1$, which gives $R_0 = \frac{B}{D}$. Consequently, $\bar{\mathrm{E}}_0$ is stable (unstable) for $R_0 < 1$ (>1). To find the infected equilibrium solution, setting $f_6(X,\ Y) = 0$ yields $\bar{X}_1(Y) = \frac{Y+C}{(A+B)Y+BC}$, which is then substituted into $f_5(X,\ Y) = 0$ to give the following quadratic equation:

$$\mathscr{F}_5(Y) = (A + B)Y^2 + (BC + D - A - B)Y + C(D - B) = 0. \tag{23}$$

In order to have two real, positive roots, two conditions must be satisfied, that is, $BC + D - A - B < 0$ and $D - B > 0$, or in compact form, $0 < D - B < A - BC$. The condition $D - B > 0$ is equivalent to $0 < R_0 = \frac{B}{D} < 1$, which is a necessary condition for backward bifurcation. Moreover, the positive influx constant, having been scaled to 1, is a necessary term for the positive equilibrium of Y. Therefore, the positive influx rate term and the increasing and saturating infectivity function are necessary for backward bifurcation.

We further examine the incidence function, $f_7(X,\ Y)$ defined in (12), without solving for the equilibrium solutions. The incidence function f_7 obviously satisfies the condition (3a), as well as the condition (3b) since $\frac{\partial}{\partial X}f_7(X,\ Y) = [B + AY(Y + C)^{-1}]Y > 0$ and $\frac{\partial}{\partial Y}f_7(X,\ Y) = ACXY(Y + C)^{-2} + [B + AY(Y + C)^{-1}]X > 0$

for all $X,\ Y > 0$. However, the second partial derivative of $f_7(X,\ Y)$ with respect to Y, $\frac{\partial^2}{\partial Y^2} f_7(X,\ Y) = 2AC^2X(X+C)^{-3} > 0$ for all $X,\ Y > 0$, showing that $f_7(X,\ Y)$ is a convex function with respect to the variable Y. Consequently, $f_7(X,\ Y)$ can only have one intersection with $g_2(Y) = Y$, implying that only one equilibrium solution would exist if we only consider the second equation in (11), as shown Fig. 2a. However, when considering both conditions given in (11) for equilibrium solutions, we will have two intersection points between f_7 and g_2. According to the first equation in (11), that is $f_5(X,\ Y) = 0$, we can use Y to express X in the equilibrium state as $\bar{X}(Y) = (Y+C)[(A+B)Y^2 + (BC+D)Y + DC]^{-1}$. Substituting $\bar{X}(Y)$ into $f_7(X,\ Y)$ in (12), we obtain

$$f_7(Y) = Y[(A+B)Y + BC][(A+B)Y^2 + (BC+D)Y + CD]^{-1}, \tag{24}$$

and $\frac{\partial}{\partial Y} f_7(Y) = D[(A+B)Y^2 + 2(A+B)CY + BC^2][(A+B)Y^2 + (BC+D)Y + CD]^{-2} > 0$ for all $X,\ Y > 0$. However, the sign of $\frac{\partial^2}{\partial Y^2} f_7(Y) = -2D[(A+B)^2Y^3 + 3C(A+B)^2Y^2 + 3(A+B)BC^2Y + (B^2C - AD)C^2][(A+B)Y^2 + (BC+D)Y + CD]^{-3}$, could alter at the inflection point from positive to negative as Y increases. Therefore, with appropriate parameter values, $f_7(Y)$ can have a convex-concave 'S' shape.

If we choose the parameter B as the bifurcation parameter, then $R_0 = \frac{B}{D} = 1$ defines $B_S = D$ where the 'S' in subscript stands for *static bifurcation*. Further, it can be proved that this is a transcritical bifurcation. Therefore, $\bar{\mathrm{E}}_0$ is stable when $B < D$ (or $R_0 < 1$), loses its stability and becomes unstable when B increases to pass through $B_S = D$, that is $B > D$ (or $R_0 > 1$), and no other bifurcations can happen.

Next, we examine the infected equilibrium $\bar{\mathrm{E}}_1 = (\bar{X},\ \bar{Y})$. Since $\bar{X}(Y) = \frac{Y+C}{(A+B)Y+BC}$, $\bar{Y}$ is determined by the quadratic equation (23), which gives the turning point $(B_T,\ Y_T)$, as given in (15). We perform a further bifurcation analysis on its corresponding characteristic polynomial (13), which takes the form

$$\begin{aligned} P|_{\bar{\mathrm{E}}_1}(\lambda,\ Y) &= \lambda^2 + \frac{a_{1a}}{[(A+B)Y+BC](Y+C)}\lambda + \frac{a_{2a}}{[(A+B)Y+BC](Y+C)}, \quad \text{where} \\ a_{1a} &= (A+B)^2Y^3 + (2BC+D)(A+B)Y^2 + (B^2C^2 + ACD + 2BCD \\ &\quad -AC)Y + BC^2D, \\ a_{2a} &= (A+B)^2Y^3 + 2(A+B)BCY^2 + (B^2C - AD)CY. \end{aligned} \tag{25}$$

Therefore, the sign of the subtraction between the trace and determinant is determined by $h_2(Y) = a_{1a} - a_{2a} = D(A+B)Y^2 + [2CD(A+B) - AC]Y + BC^2D$. Here the equilibrium solution of Y and other parameters satisfy the quadratic equation (23), which leads to an explicit expression, given by $\bar{B} = -\frac{AY^2+(D-A)Y+CD}{Y^2+(C-1)Y-C}$. Substituting $B = \bar{B}$ into $h_2(Y)$, we obtain

$$h_2(Y)|_{B=\bar{B}} = a_{1a} - a_{2a} = \frac{[AC(D-1) - D^2]Y^2 - [AC(D-1) + 2CD^2]Y - C^2D^2}{Y-1}. \tag{26}$$

Hopf bifurcation may occur when the trace is zero, while the determinant is still positive. This implies $h_2(Y) < 0$, which is possible with appropriately chosen parameter values. Hence, by solving $a_{1a} = 0$ in (25) together with the quadratic equation (23), we get two pairs of points denoted by $(B_{h1},\ Y_{h1})$ and $(B_{h2},\ Y_{h2})$, which are candidates for Hopf bifurcation. Then validating the above two points by substituting them back into the characteristic polynomial (25), respectively, we denote the Hopf bifurcation point as $(B_H,\ Y_H)$ if this validation confirms their existence. According to Yu et al. (2015), Hopf bifurcation can happen only from the upper branch of the infected equilibrium $\bar{E}_1$.

References

Alexander HK, Wahl LM (2011) Self-tolerance and autoimmunity in a regulatory T cell model. Bull Math Biol 73(1):33–71

Arino J, McCluskey CC, van den Driessche P (2003) Global results for an epidemic model with vaccination that exhibits backward bifurcation. SIAM J Appl Math 64(1):260–276

Baecher-Allan C, Wolf E, Hafler DA (2006) MHC class II expression identifies functionally distinct human regulatory T cells. J Immunol 176:4622–4631

Blayneh KW, Gumel AB, Lenhart S, Clayton T (2010) Backward bifurcation and optimal control in transmission dynamics of west nile virus. Bull Math Biol 72(4):1006–1028

Brauer F (2004) Backward bifurcations in simple vaccination models. J Math Anal Appl 298(2):418–431

Briggs CJ, Godfray HCJ (1995) The dynamics of insect-pathogen interactions in stage-structured populations. Am Nat 145(6):855–887

Brown GC, Hasibuan R (1995) Conidial discharge and transmission efficiency of Neozygites floridana, an entomopathogenic fungus infecting two-spotted spider mites under laboratory conditions. J Invertebr Pathol 65(1):10–16

Capasso V, Serio G (1978) A generalization of the Kermack–McKendrick deterministic epidemic model. Math Biosci 42(1):43–61

Castillo-Chavez C, Cooke K, Huang W, Levin S (1989a) On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS). Part 2: Multiple group models. In: Castillo-Chavez C (ed) Mathematical and statistical approaches to AIDS epidemiology, lecture notes in biomathematics, vol 83. Springer, Berlin, pp 200–217

Castillo-Chavez C, Cooke K, Huang W, Levin SA (1989b) Results on the dynamics for models for the sexual transmission of the human immunodeficiency virus. Appl Math Lett 2(4):327–331

Collier AC, Coombs RW, Schoenfeld DA, Bassett RL, Timpone J, Baruch A, Jones M, Facey K, Whitacre C, McAuliffe VJ, Friedman HM, Merigan TC, Reichman RC, Hooper C, Corey L (1996) Treatment of human immunodeficiency virus infection with Saquinavir, Zidovudine, and Zalcitabine. N Engl J Med 334(16):1011–1018

Conway JM, Coombs D (2011) A stochastic model of latently infected cell reactivation and viral blip generation in treated HIV patients. PLoS Comput Biol 7(4):e1002033

Derrick WR, van den Driessche P (1993) A disease transmission model in a nonconstant population. J Math Biol 31(5):495–512

Dornadula G, Zhang H, VanUitert B, Stern J, Livornese L Jr, Ingerman MJ, Witek J, Kedanis RJ, Natkin J, DeSimone J, Pomerantz RJ (1999) Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. JAMA 282(17):1627–1632

Dushoff J, Huang W, Castillo-Chavez C (1998) Backwards bifurcations and catastrophe in simple models of fatal diseases. J Math Biol 36(3):227–248

Fraser C, Ferguson NM, de Wolf F, Anderson RM (2001) The role of antigenic stimulation and cytotoxic T cell activity in regulating the long-term immunopathogenesis of HIV: mechanisms and clinical implications. Proc R Soc B Biol Sci 268(1481):2085–2095

Fung ICH, Gambhir M, van Sighem A, de Wolf F, Garnett GP (2012) The clinical interpretation of viral blips in HIV patients receiving antiviral treatment: are we ready to infer poor adherence? J Acquir Immune Defic Syndr 60(1):5–11

Fussmann GF, Ellner SP, Shertzer KW, Hairston NG Jr (2000) Crossing the Hopf bifurcation in a live predator–prey system. Science 290(5495):1358–1360

Garretta NJ, Apeaa V, Noria A, Ushiro-Lumbb I, Oliverb AR, Bailya G, Clarkb DA (2012) Comparison of the rate and size of HIV-1 viral load blips with Roche COBAS TaqMan HIV-1 versions 1.0 and 2.0 and implications for patient management. J Clin Virol 53(4):354–355

Gil L et al (2003) Contribution to characterization of oxidative stress in HIV/AIDS patients. Pharmacol Res 47(3):217–224

Grennan JT, Loutfy MR, Su D, Harrigan PR, Cooper C, Klein M, Machouf N, Montaner JSG, Rourke S, Tsoukas C, Hogg B, Raboud J, the CANOC Collaboration (2012) Magnitude of virologic blips is associated with a higher risk for virologic rebound in HIV-infected individuals: a recurrent events analysis. J Infect Dis 205(8):1230-1238

Gomez-Acevedo H, Li MY (2005) Backward bifurcation in a model for HTLV-I infection of $CD4^+$ T cells. Bull Math Biol 67(1):101-114

Hadeler KP, van den Driessche P (1997) Backward bifurcation in epidemic control. Math Biosci 146(1):15–35

Hethcote HW, van den Driessche P (1991) Some epidemiological models with nonlinear incidence. J Math Biol 29(3):271–287

Hethcote HW, Lewis MA, van den Driessche P (1989) An epidemiological model with a delay and a nonlinear incidence rate. J Math Biol 27(1):49–64

Huang W, Cooke KL, Castillo-Chavez C (1992) Stability and bifurcation for a multiple-group model for the dynamics of HIV/AIDS transmission. SIAM J Appl Math 52(3):835–854

Israel N, Gougerot-Pocidalo MA (1997) Oxidative stress in human immunodeficiency virus infection. Cell Mol Life Sci 53(11–12):864–870

Jones LE, Perelson AS (2005) Opportunistic infection as a cause of transient viremia in chronically infected HIV patients under treatment with HAART. Bull Math Biol 67(6):1227–1251

Korobeinikov A, Maini PK (2005) Non-linear incidence and stability of infectious disease models. Math Med Biol 22(2):113–128

Li MY, Muldowney JS (1995) Global stability for the SEIR model in epidemiology. Math Biosci 125(2):155–164

Li N, Karin M (1999) Is NF-κB the sensor of oxidative stress? FASEB J 13(10):1137–1143

Liu W, Levin S, Iwasa Y (1986) Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. J Math Biol 23(2):187–204

Liu W, Hethcote HW, Levin SA (1987) Dynamical behavior of epidemiological models with nonlinear incidence rates. J Math Biol 25(4):359–380

Murray JD (2002) Mathematical biology: I. An introduction. Interdisciplinary applied mathematics. Springer, Berlin

Ngonghala CN, Miranda IT, Ngwa GA (2014) Persistent oscillations and backward bifurcation in a malaria model with varying human and mosquito populations: implications for control. J Math Biol 70(7):1581–1622

Palmer S, Wiegand AP, Maldarelli F, Bazmi H, Mican JM, Polis M, Dewar RL, Planta A, Liu S, Metcalf JA, Mellors JW, Coffin JM (2003) New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. J Clin Microbiol 41(10):4531–4536

Palmer S, Maldarelli F, Wiegand A, Bernstein B, Hanna GJ, Brun SC, Kempf DJ, Mellors JW, Coffin JM, King MS (2008) Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. Proc Natl Acad Sci 105(10):3879–3884

Rong L, Perelson AS (2009a) Asymmetric division of activated latently infected cells may explain the decay kinetics of the HIV-1 latent reservoir and intermittent viral blips. Math Biosci 217(1):77–87

Rong L, Perelson AS (2009b) Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. PLoS Comput Biol 5(10):e1000533

Ruan S, Wang W (2003) Dynamical behavior of an epidemic model with a nonlinear incidence rate. J Differ Equ 188(1):135–163

Schwarz KB (1996) Oxidative stress during viral infection: a review. Free Radic Biol Med 21(5):641–649

Simon CP, Jacquez JA (1992) Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations. SIAM J Appl Math 52(2):541–576

Stephenson CB, Marquis GS, Douglas SD, Wilson CM (2005) Immune activation and oxidative damage in HIV-positive and HIV negative adolescents. J Acquir Immune Defic Syndr 38(2):180–190

van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180(12):29–48

van Gaalen RD, Wahl LM (2009) Reconciling conflicting clinical studies of antioxidant supplementation as HIV therapy: a mathematical approach. BMC Public Health 9(Suppl. 1):1–18
Wiggins S (2003) Introduction to applied nonlinear dynamical systems and chaos, vol 2. Springer Science and Business Media, Berlin
Yao W, Hertel L, Wahl LM (2006) Dynamics of recurrent viral infection. Proc R Soc B Biol Sci 273(1598):2193–2199
Yu P (1998) Computation of normal forms via a perturbation technique. J Sound Vib 211(1):19–38
Yu P, Zhang W, Wahl LM (2016) Dynamical analysis and simulation of a 2-dimensional disease model with convex incidence. Commun Nonlinear Sci Numer Simul 37:163–192
Zhang W, Wahl LM, Yu P (2013) Conditions for transient viremia in deterministic in-host models: viral blips need no exogenous trigger. SIAM J Appl Math 73(2):853–881
Zhang W, Wahl LM, Yu P (2014) Viral blips may not need a trigger: How transient viremia can arise in deterministic in-host models. SIAM Rev 56(1):127–155
Zhang W, Wahl LM, Yu P (2014) Modelling and analysis of recurrent autoimmune disease. SIAM J Appl Math 74(6):1998–2025
Zhou L, Fan M (2012) Dynamics of an SIR epidemic model with limited medical resources revisited. Nonlinear Anal Real 13(1):312–324