J. Math. Biol. (1993) 31: 495–512

Journal of
Mathematical
Biology

A disease transmission model in a nonconstant population

W. R. Derrick[1], P. van den Driessche[2,3]*

[1] Department of Mathematics, University of Montana, Missoula, MT 59812, USA

[2] Department of Mathematics, University of Victoria, Victoria, B.C. V8W 3P4, Canada

[3] Mathematisches Institut, Universität München, W-8000 München 2, Germany

Received November 15, 1991; received in revised form March 25, 1992

Abstract. A general SIRS disease transmission model is formulated under assumptions that the size of the population varies, the incidence rate is nonlinear, and the recovered (removed) class may also be directly reinfected. For a class of incidence functions it is shown that the model has no periodic solutions. By contrast, for a particular incidence function, a combination of analytical and numerical techniques are used to show that (for some parameters) periodic solutions can arise through homoclinic loops or saddle connections and disappear through Hopf bifurcations.

Key words: Epidemiological model – Nonlinear incidence function – Hopf bifurcation – Homoclinic loop – Saddle connection

1 Introduction

The behavior of a general SIRS epidemiological model in a population of varying size is modeled by a system of differential equations in $\mathbb{R}^3_+$. Here S is the number of susceptible individuals, I is the number of infective individuals, R is the number of recovered (or removed) individuals in the population, and $N = S + I + R$ is the total population number. Natural births and deaths are assumed to be proportional to the class numbers, with all newborns susceptible. Excess death due to the disease may occur in both the infective and recovered classes. The rate at which susceptibles become infective is taken as a nonlinear incidence function $I\Phi(S, I, N)$. Recovered individuals lose immunity and cycle back into the susceptible class and may also be directly reinfected at a rate $I\Psi(R, I, N)$.

The assumption that the recovered class may also be susceptible applies to diseases such as influenza, where infection by one strain provides some perma-

* Supported in part by NSERC grant A-8965, the University of Victoria Committee on Faculty Research & Travel, and the Institute for Mathematics and its Applications, Minneapolis, MN, with funds provided by NSF

nent protection, or cross-immunity, against infection by a related strain. Castillo-Chavez et al. (1989) have studied models involving two viral strains and cross-immunity. They divided the infectious and recovered individuals into separate classes depending on the viral strain with which they have been infected. We will not make such a separation; instead individuals who have recovered from one or more bouts of the disease will be assumed to be susceptible to infection at a different rate than individuals who have never had the disease. In some sense, our infective and recovered classes will be a mass average of the separate classes considered by Castillo-Chavez et al. In mathematical models for tuberculosis, Revelle et al. (1967) identified classes of recovered inactive individuals who had been previously infected but may become active again. They considered a model with nine classes which included forms of control. In our simpler model, we can assume such individuals are in the R class, and they may become reinfected at a rate $I\Psi(R, I, N)$. Tudor (1990) developed a constant population SIRI model for herpes viral infections in which R is a latent class.

Our model is a generalization of one considered by Busenberg and van den Driessche (1990) in which $I\Phi$ is taken proportional to SI/N, thus the force of infection is of the proportionate mixing type, and Ψ is taken to be zero. This latter model, as well as a model with $I\Phi$ proportional to SI has also been analyzed by Mena-Lorca and Hethcote (1991). Busenberg and van den Driessche (1990, 1991) proved a generalization of the Bendixson-DuLac criterion which rules out periodic solutions when the equations are transformed to proportional variables S/N, I/N, and R/N. Here we give a simpler proof of this result, which we extend to more general Φ and Ψ.

For models with constant population size, it has been shown that certain forms of nonlinear incidence functions can lead to periodic solutions. An excellent review of disease transmission models that exhibit periodic solutions is given by Hethcote and Levin (1989). In Sect. 5, we take the nonlinear incidence function that was used by Liu et al. (1986), namely $I\Phi$ proportional to $S^q I^p / N^{p+q-1}$. (Note that Liu et al. assumed that N was constant, but we allow N to vary.) For the case $q = 1, p = 2$, we show that it is possible for the proportional variables to exhibit stable periodic solutions. We show that as we increase the per capita loss of immunity rate of recovered individuals, for certain values of the parameters, a saddle connection or a homoclinic loop occurs giving rise to stable periodic solutions. As that parameter is increased, the period and amplitude of the periodic solutions decrease and vanish through a Hopf bifurcation.

2 A general SIRS model

We consider a model of disease transmission in a nonconstant population of size N divided into three classes: susceptibles, infectives, and recovereds, the numbers in each class given by S, I, R, respectively. We set $N = S + I + R$, and use the following parameters, which are assumed non-negative unless otherwise specified:

b = per capita birth rate, assumed to be positive (except as indicated in Sects. 5 and 6),
d = per capita disease free death rate, assumed to be positive,
ϵ = excess per capita death rate of infectives,
δ = excess per capita death rate of recovereds,
γ = per capita recovery rate of infectives,
ϱ = per capita loss of immunity rate of recovereds.

We make the following assumptions in this model:

(i) The incidence of disease in the susceptible class is given by the function $I\Phi(S, I, N)$, where $\Phi(S, I, N)$ is a non-negative, bounded, C^1, homogeneous function of degree zero on $\mathbb{R}^3_+$. We also assume that

$$\lim_{S \to 0+} \Phi(S, I, N) = 0$$

and that the values

$$\overline{\lim_{I \to 0+}} \Phi(S, I, N) = 0, \quad \text{and} \quad \Phi_0 = \overline{\lim_{N \to 0+}} \Phi(S, I, N),$$

where $N = S + I + R$, are nonnegative and bounded. For example, if we let $\Phi(S, I, N) = \phi S/N$, then the incidence of disease is of the *proportionate mixing* type introduced by Nold (1980), with "force of infection" given by $\phi I/N$, where the non-negative parameter ϕ is the effective per capita contact rate of infectives with susceptible individuals. This incidence function has since been used in many models; see, for example, Mena-Lorca and Hethcote (1991) where it is called the standard incidence. Other bilinear and nonlinear expressions have been used for the incidence function, and other choices we can take for $\Phi(S, I, N)$ include the following. The function $\Phi(S, I, N) = \phi S^q I^{p-1}/N^{p+q-1}$ with $p \geqslant 1, q \geqslant 1$, was used for constant populations by Liu et al. (1986); see also Liu et al. (1987). In a study of AIDS, Jacquez et al. (1988) used $\Phi(S, I, N) = \phi S/(S + I)$. We could also take a more general function displaying a saturation effect, such as $\Phi(S, I, N) = \phi S/(S + I + cN)$, for constant $c \geqslant 0$. Diekmann and Kretzschmar (1991) have considered functions of this type.

(ii) The recovered class may also be directly reinfected by the disease with transfer rate $I\Psi(R, I, N)$, where $\Psi(R, I, N)$, is also a non-negative, bounded, C^1, homogeneous function of degree zero on $\mathbb{R}^3_+$. We also assume that

$$\lim_{R \to 0+} \Psi(R, I, N) = 0$$

and that the values

$$\overline{\lim_{I \to 0+}} \Psi(R, I, N), \quad \text{and} \quad \Psi_0 = \overline{\lim_{N \to 0+}} \Psi(R, I, N),$$

where $N = S + I + R$, are nonnegative and bounded. A typical example of the function $\Psi(R, I, N)$ might be a proportionate mixing type $\Psi(R, I, N) = \psi R/N$, where the non-negative parameter ψ measures the average effective per capita contact rate of infectives with recovered individuals.

Our model is given schematically in Fig. 1. Observe that the flow diagram in Fig. 1 is a special case of Fig. 1 in Busenberg and van den Driessche (1991).

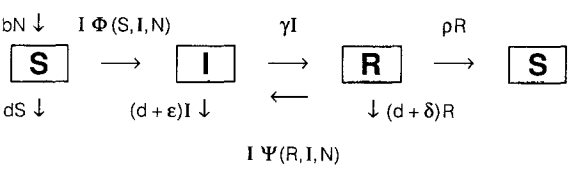

Fig. 1 Flow diagram for the SIRS model

The above hypotheses lead to the following system of differential equations, where the derivative d/dt is denoted by $'$:

$$S' = bN - dS + \varrho R - I\Phi(S, I, N), \tag{2.1}$$

$$I' = I[\Phi(S, I, N) + \Psi(R, I, N) - (d + \epsilon + \gamma)], \tag{2.2}$$

$$R' = \gamma I - (d + \delta + \varrho)R - I\Psi(R, I, N), \tag{2.3}$$

or, in matrix form, by

$$\begin{bmatrix} S \\ I \\ R \end{bmatrix}' = \begin{bmatrix} b-d & b & b+\varrho \\ 0 & -(d+\epsilon+\gamma) & 0 \\ 0 & \gamma & -(d+\delta+\varrho) \end{bmatrix} \begin{bmatrix} S \\ I \\ R \end{bmatrix} + \begin{bmatrix} -I\Phi(S,I,N) \\ I\Phi(S,I,N) + I\Psi(R,I,N) \\ -I\Psi(R,I,N) \end{bmatrix}.$$

Observe that this matrix equation is in the form $\boldsymbol{x}' = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{f}(\boldsymbol{x})$, where $\boldsymbol{x} = \mathrm{col}(S, I, R)$, $\boldsymbol{M}$ is a 3×3 *essentially nonnegative* matrix, that is, one with off-diagonal entries nonnegative and at least one such entry positive, and $\boldsymbol{f}: \mathbb{R}^3_+ \to \mathbb{R}^3_+$ is a C^1 function that is homogeneous of degree one, that is $\boldsymbol{f}(a\boldsymbol{x}) = a\boldsymbol{f}(\boldsymbol{x})$, for $a > 0$, and has entries adding to zero.

The equation for the total population N is obtained by adding (2.1)–(2.3):

$$N' = (b - d)N - \epsilon I - \delta R. \tag{2.4}$$

The model Eqs. (2.1)–(2.4) are *well-posed*, in the sense that non-negative initial data (S, I, R) leads to non-negative solutions which are defined for all time $t \geqslant 0$, because, by our assumptions on $\Phi(S, I, N)$ and $\Psi(R, I, N)$

$$bN + \varrho R \geqslant I \lim_{S \to 0+} \Phi(S, I, N) = 0 \quad \text{and} \quad \gamma \geqslant \lim_{R \to 0+} \Psi(R, I, N) = 0. \tag{2.5}$$

Letting $S \to 0+$ in (2.1) and using the first inequality in (2.5) yields $S' \geqslant 0$, so that the IR-plane is a barrier for trajectories. Similarly, letting $I \to 0+$ in (2.2) yields $I' = 0$, while letting $R \to 0+$ in (2.3) and applying the second inequality in (2.5) yields $R' \geqslant 0$. Hence, all trajectories stay in $\mathbb{R}^3_+$.

If we seek the equilibrium points of (2.1)–(2.4), we observe that $I = 0$ implies that $R = 0$, so that $bN = dS$. If $b = d$, then *any* $(N, 0, 0)$ is an equilibrium point, but if $b \neq d$, then the origin is the only equilibrium point with $I = 0$. If $I \neq 0$, $b = d$, and $\epsilon = d = 0$, then the curve (if any) satisfying

$$\Phi(S, I, N) + \Psi(R, I, N) = b + \gamma \quad \text{and} \quad (\gamma - \Psi(R, I, N))I = (b + \varrho)R, \tag{2.6}$$

consists of equilibrium points, because (2.6) forces the right hand sides of (2.1)–(2.3) to be zero. This case models disease transmission with no excess death due to disease and with equal population birth and death rates. Thus, the total population remains constant. As we are interested in varying population size, we assume in what follows that $b \neq d$.

Calculating the Jacobian matrix of (2.1)–(2.3) at the origin, we obtain the eigenvalues $b - d$, $-(\varrho + d + \delta)$, and $(\Phi_0 + \Psi_0) - (d + \epsilon + \gamma)$. If $I = R = 0$ and $b > d$, then the origin is the only equilibrium point, and it is unstable. If $b < d$ and $(\Phi_0 + \Psi_0)/(d + \epsilon + \gamma) < 1$, the origin is locally asymptotically stable.

3 Nonexistence of periodic solutions

We are interested in studying this model in situations where the population $N(t)$ is not stationary. To do so we consider the proportions of individuals in the

three epidemiological classes,

$$s = S/N, \quad i = I/N, \quad \text{and} \quad r = R/N, \tag{3.1}$$

and examine the dynamics of the resulting system in the feasibility region

$$\mathscr{D} = \{s \geqslant 0, \quad i \geqslant 0, \quad r \geqslant 0 : s + i + r = 1\}.$$

With these variables, system (2.1)–(2.3) becomes

$$s' = b(1 - s) + \varrho r + \epsilon si + \delta sr - i\Phi(s, i), \tag{3.2}$$

$$i' = -(b + \epsilon + \gamma)i + \epsilon i^2 + \delta ir + i\Phi(s, i) + i\Psi(r, i), \tag{3.3}$$

$$r' = \gamma i - (b + \varrho + \delta)r + \epsilon ir + \delta r^2 - i\Psi(r, i), \tag{3.4}$$

where $\Phi(s, i) = \Phi(s, i, 1) = \Phi(S/N, I/N, N/N) = \Phi(S, I, N)$ and $\Psi(r, i) = \Psi(r, i, 1) = \Psi(R/N, I/N, N/N) = \Psi(R, I, N)$ by homogeneity. Rewriting (3.2)–(3.4) in matrix form, and letting $1 - s = i + r$

$$\begin{bmatrix} s \\ i \\ r \end{bmatrix}' = \begin{bmatrix} 0 & b & b + \varrho \\ 0 & -(b + \epsilon + \gamma) & 0 \\ 0 & \gamma & -(b + \varrho + \delta) \end{bmatrix} \begin{bmatrix} s \\ i \\ r \end{bmatrix} + (\epsilon i + \delta r) \begin{bmatrix} s \\ i \\ r \end{bmatrix} + \begin{bmatrix} -i\Phi(s, i) \\ i\Phi(s, i) + i\Psi(r, i) \\ -i\Psi(r, i) \end{bmatrix}. \tag{3.5}$$

We seek conditions for the existence and stability of equilibria and the existence or non-existence of periodic solutions and oriented phase polygons for this model. A closed curve connecting several equilibrium points whose segments between successive equilibria are solution trajectories of a system of differential equations is called a *phase polygon* (see Hahn 1967). A phase polygon whose sides are solutions traversed in the same time sense is called an *oriented phase polygon.*

It is not hard to show that system (3.5) has the form

$$\boldsymbol{y}' = \boldsymbol{F}(\boldsymbol{y}) = \boldsymbol{M}\boldsymbol{y} - (\mathbf{1} \cdot \boldsymbol{M}\boldsymbol{y})\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y}), \tag{3.6}$$

where $\boldsymbol{y} = \text{col}(s, i, r)$, $\mathbf{1} = \text{col}(1, 1, 1)$, $\boldsymbol{M}$ is the essentially nonnegative matrix given in Sect. 2, and $\mathbf{1} \cdot \boldsymbol{f}(\boldsymbol{y}) = 0$. The hyperplane $\mathbf{1} \cdot \boldsymbol{y} = 1$ is invariant under the flow (3.6) since

$$(\mathbf{1} \cdot \boldsymbol{y})' = (\mathbf{1} \cdot \boldsymbol{M}\boldsymbol{y})(1 - (\mathbf{1} \cdot \boldsymbol{y})) + (\mathbf{1} \cdot \boldsymbol{f}(\boldsymbol{y})) = 0.$$

Busenberg and van den Driessche (1991) proved in Theorem 2.2, given an additional condition, that systems such as (3.6) have no *periodic solutions* (including closed orbits, homoclinic loops, and oriented phase polygons). For convenience, we provide a simpler proof of this result.

Let $\boldsymbol{g} = \alpha \boldsymbol{y} \times \boldsymbol{F}(\boldsymbol{y})$, where $\alpha = (y_1 y_2 y_3)^{-1}$. Then, since $\boldsymbol{y} \times \boldsymbol{y} = 0$, we have

$$\boldsymbol{g} = \alpha \boldsymbol{y} \times [\boldsymbol{M}\boldsymbol{y} - (\mathbf{1} \cdot \boldsymbol{M}\boldsymbol{y})\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})] = \alpha \boldsymbol{y} \times \boldsymbol{M}\boldsymbol{y} + \alpha \boldsymbol{y} \times \boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{g}_M + \boldsymbol{g}_f,$$

$$\boldsymbol{g} \cdot \boldsymbol{F} = (\alpha \boldsymbol{y} \times \boldsymbol{F}) \cdot \boldsymbol{F} = \alpha \boldsymbol{y} \cdot (\boldsymbol{F} \times \boldsymbol{F}) = 0. \tag{3.7}$$

A direct calculation, without writing the terms of $\boldsymbol{F}$ in alternate forms as was done in Busenberg and van den Driessche (1991), gives

$$(\text{curl}\, \boldsymbol{g}_M) \cdot \mathbf{1} = -\left(\frac{m_{13}}{y_1^2 y_2} + \frac{m_{23}}{y_1 y_2^2} + \frac{m_{32}}{y_1 y_3^2} + \frac{m_{12}}{y_1^2 y_3} + \frac{m_{21}}{y_2^2 y_3} + \frac{m_{31}}{y_2 y_3^2}\right), \tag{3.8}$$

and

$$(\operatorname{curl} \boldsymbol{g}_f) \cdot \mathbf{1} = \alpha\left\{\left(\frac{f_1}{y_2} - \frac{f_1}{y_1} + \frac{\partial f_1}{\partial y_1} - \frac{\partial f_1}{\partial y_2}\right) + \left(\frac{f_3}{y_2} - \frac{f_3}{y_3} + \frac{\partial f_3}{\partial y_3} - \frac{\partial f_3}{\partial y_2}\right)\right\}, \quad (3.9)$$

where we have expressed (3.9) in terms of f_1 and f_3, instead of f_2 and f_3 as was done in Busenberg and van den Driessche (1991), because of the form of system (3.5). We now obtain the following generalization of the Bendixson-Dulac criterion to the 2-dimensional surface $\mathscr{D}$ in $\mathbb{R}^3$.

Theorem 3.1 (Busenberg and van den Driessche 1991) *Let $\boldsymbol{M} = (m_{ij})$ be a 3×3 essentially nonnegative constant matrix, let $\boldsymbol{y} = \operatorname{col}(y_1, y_2, y_3)$, where $\mathbf{1} \cdot \boldsymbol{y} = 1$, and let $\boldsymbol{f}: \mathbb{R}^3_+ \to \mathbb{R}^3_+$ be a continuously differentiable function satisfying $\mathbf{1} \cdot \boldsymbol{f}(\boldsymbol{y}) = 0$. Let $\boldsymbol{g} = \alpha \boldsymbol{y} \times \boldsymbol{F}(\boldsymbol{y})$, where $\alpha = (y_1 y_2 y_3)^{-1}$. If $(\operatorname{curl} \boldsymbol{g}_f) \cdot \mathbf{1} \leqslant 0$ in $\mathscr{D}^0 = \mathscr{D} - \partial\mathscr{D}$, where $\mathscr{D} = \{y_1 \geqslant 0, y_2 \geqslant 0, y_3 \geqslant 0 : y_1 + y_2 + y_3 = 1\}$, then there are no periodic solutions in $\mathscr{D}^0$ of*

$$\boldsymbol{y}' = \boldsymbol{F}(\boldsymbol{y}) = \boldsymbol{M}\boldsymbol{y} - (\mathbf{1} \cdot \boldsymbol{M}\boldsymbol{y})\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y}).$$

Moreover, if $\boldsymbol{F}(\boldsymbol{y})_i$ evaluated at $y_i = 0$ is nonnegative for $i = 1, 2, 3$, and is positive for at least one i, then this result holds in the region $\mathscr{D}$.

Proof. Since $\boldsymbol{g} \cdot \boldsymbol{F} = 0$ by (3.7), and the off-diagonal entries of $\boldsymbol{M}$ are nonnegative with at least one of them positive, by (3.8) and applying the hypothesis to (3.9), it follows that $\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1} < 0$. The nonexistence of oriented phase polygons in $\mathscr{D}^0$ follows from either Theorem 4.1 or Corollary 4.2 of Busenberg and van den Driessche (1990). The condition $\boldsymbol{F}(\boldsymbol{y})_i \geqslant 0$ for $i = 1, 2, 3$, implies that the field $\boldsymbol{F}$ never points towards the exterior of $\mathscr{D}$ when evaluated on $\partial\mathscr{D}$, and since $\boldsymbol{F}(\boldsymbol{y})_i > 0$ for at least one i, it follows that $\partial\mathscr{D}$ is not a phase polygon for the system. Thus, there are no periodic solutions in $\mathscr{D}$. □

Using (3.8) and (3.9) we see that system (3.5) satisfies the conditions of Theorem 3.1), with $\boldsymbol{g} \cdot \boldsymbol{F} = 0$ and

$$\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1} = -\left(\frac{\gamma}{sr^2} + \frac{b}{s^2 r} + \frac{b + \varrho}{s^2 i}\right) + \ell_s \Phi + \ell_r \Psi, \quad (3.10)$$

on $\mathscr{D}^0$, where

$$\ell_s \Phi = \frac{1}{sr}\left(\frac{\Phi}{s} + \Phi_i - \Phi_s\right) \quad \text{and} \quad \ell_r \Psi = \frac{1}{rs}\left(\frac{\Psi}{r} + \Psi_i - \Psi_r\right) \quad (3.11)$$

and Φ_s, Φ_i, Ψ_i, Ψ_r are partial derivatives. Consequently, we have:

Corollary 3.2 *Let $\ell_s \Phi \leqslant 0$ and $\ell_r \Psi \leqslant 0$. Then the generalized* SIRS *model* (3.2)–(3.4) *has no periodic solutions in $\mathscr{D}$.*

Proof. Since $b > 0$, $\gamma \geqslant 0$, $\varrho \geqslant 0$ and the hypotheses, it follows that $\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1} < 0$, so Theorem 3.1 implies the nonexistence of periodic solutions in $\mathscr{D}^0$. As the problem is well-posed, the field $\boldsymbol{F}$ never points towards the exterior of $\mathscr{D}$ when evaluated on $\partial\mathscr{D}$, and $s'(0, i, r) = b + \varrho r > 0$, so that $\partial\mathscr{D}$ is not a phase polygon for the system. Thus, there are no periodic solutions in $\mathscr{D}$. □

Corollary 3.3 *If* $\Phi(s, i) = sh(s + i)$ *and* $\Psi(r, i) = rj(r + i)$, *where* h *and* j *are arbitrary nonnegative differentiable functions, then the model system given by* (3.2)–(3.4) *has no periodic solutions in the invariant region* $\mathscr{D}$.

Proof. Consider the partial differential equation $\ell_s \Phi = 0$ on $\mathscr{D}^0$, namely

$$\Phi_s - \Phi_i = \frac{\Phi}{s}. \tag{3.12}$$

By characteristics: $ds = -di = s\, d\Phi/\Phi$, so that $s + i = c_1$ and $\Phi = c_2 s$, implying that $\Phi(s, i) = sh(s + i)$ solves (3.12), where h is any differentiable function. Since $\Phi(s, i)$ is nonnegative in $\mathscr{D}$, h must also be nonnegative. Similarly, $\ell_r \Psi = 0$ has the solution $\Psi(r, i) = rj(i + r)$, for any nonnegative differentiable function j. The equality conditions of Corollary 3.2 are met, and the result follows. □

Many of the commonly used incidence functions $\Phi(S, I, N)$ satisfy $\ell_S \Phi = 0$. We now list some examples from the literature which were cited in Sects. 1 and 2.

(i) $\Phi(S, I, N) = \phi S/N$, giving $\phi SI/N$ as the rate at which susceptibles become infected. The force of infection is $\phi I/N$, the proportionate mixing type (Nold 1980). This yields $\Phi(s, i) = \phi s$, so that h is the constant function ϕ, which has been used in many models (see, for example, Busenberg and van den Driessche 1990, Mena-Lorca and Hethcote 1991).

(ii) $\Phi(S, I, N) = \phi S/(S + I)$, so that $\phi SI/(S + I)$ is the rate at which susceptibles become infected, so that the force of infection is $\phi I/(S + I)$. This incidence function was used by Jacquez et al. (1988) for an AIDS model with constant input term. Here $\Phi(s, i) = \phi s/(s + i)$, so that $h(s + i) = \phi/(s + i)$.

(iii) $\Phi(S, I, N) = \phi S/(S + I + cN)$, so that $\Phi(s, i) = \phi s/(s + i + c)$, with $h(s + i) = \phi/((s + i) + c)$. The function $\Phi(s, i)$ is a Holling type II response. Holling type II incidence functions have been considered by Diekmann and Kretzschmar (1991), but they consider numbers of susceptibles and infectives rather than proportions.

The functions (i)–(iii) satisfy $\ell_S \Phi = 0$. Thus, by Corollary 3.3, when $\Psi \equiv 0$, all such models have no periodic solutions. On the other hand,

(iv) $\Phi(S, I, N) = \phi S^q I^{p-1}/N^{p+q-1}$ yields $\Phi(s, i) = \phi s^q i^{p-1}$ which satisfies

$$\ell_S \Phi = \phi s^{q-2} i^{p-2}[(1 - q)i + (p - 1)s]/r,$$

which is nonzero for p or $q \neq 1$. Liu et al. (1986, 1987) studied nonlinear incidence functions of this type for constant populations, and found periodic oscillations for certain ranges of the parameters with $p > 1$. Liu et al. (1986) analyzed the case $p = 2$, $q = 1$ (for constant populations) in substantial detail; in Sects. 5 and 6 we will do a similar study for varying populations.

Observe that $\Psi(R, I, N)) = \psi R/N$ yields $\Psi(r, i) = \psi r$ which satisfies $\ell_r \Psi = 0$.

We now provide a characterization for the solutions of the partial differential inequality $\ell_s \Phi < 0$. Let $\Phi(s, i) = sH(s, i)$, then

$$\ell_s \Phi = \frac{1}{sr}\left(\frac{\Phi}{s} + \Phi_i - \Phi_s\right) = \frac{1}{sr}(H + sH_i - H - sH_s) = \frac{1}{r}(H_i - H_s),$$

implying that $\ell_s \Phi < 0$ whenever $H_i < H_s$. This inequality provides the following characterization adapted from a proof by Westphal (1947/49).

Theorem 3.4 *Let* $\Delta = \{(s, i) \in \mathbb{R}^2_+ : 0 < s + i < 1\}$, *and suppose* $H_s > H_i$ *and* $G_s = G_i$ *on* $\overline{\Delta}$, *and* $H > G$ *on* $\partial\Delta$. *Then* $H > G$ *on* $\overline{\Delta}$. *Exactly the same conclusion is obtained for the reverse inequality* $H_i > H_s$ *if all other conditions are unchanged.*

Proof. Define the continuous function $W = H - G$ and the closed set

$$A = \{s : W(s, i) \leqslant 0 \quad \text{for some } (s, i) \in \overline{\Delta}\}.$$

Assume $A \neq \emptyset$ and let s_0 be the smallest value in A. Then there is a sequence $\{(s_n, i_n)\}$ in Δ such that $W(s_n, i_n) \leqslant 0$ and $s_n \to s_0$. By compactness, some subsequence converges to (s_0, i_0) in Δ, and $W(s_0, i_0) \leqslant 0$. But if we approach (s_0, i_0) from the left, $W(s_0, i_0) \geqslant 0$. Thus, $W(s_0, i_0) = 0$, and

$$W_s(s_0, i_0) = \lim_{k \to 0} W(s_0 + k, i_0)/k \leqslant 0,$$

implying that $H_s(s_0, i_0) \leqslant G_s(s_0, i_0)$. But $W(s_0, i) \geqslant 0$ for all $0 \leqslant s_0 + i \leqslant 1$ in $\overline{\Delta}$, so that

$$W_i(s_0, i_0) = \lim_{k \to 0} W(s_0, i_0 + k)/k$$

is non-negative if $k > 0$, and is non-positive if $k < 0$. Thus, $W_i(s_0, i_0) = 0$, or $H_i(s_0, i_0) = G_i(s_0, i_0)$. But then $H_s \leqslant G_s = G_i = H_i < H_s$ at (s_0, i_0), a contradiction. The proof for $H_s < H_i$ follows by reversing the roles of s and i. □

It is trivial to show that $G_s = G_i$ has the solution $G = G(s + i)$, so whenever we can produce a function G such that $H > G$ on $\partial\Delta$, it follows that $H > G$ on $\overline{\Delta}$.

Example 3.5 Suppose $H(s, i)$ satisfies $H_s > H_i$ on $\overline{\Delta}$ and $H(s, 0) = 1/(s^2 + 1)$, $H(0, i) = 1/(2i + 1)$, $H(1 - i, i) = 1/(i^2 + 2)$ on $\partial\Delta$. First consider the constant function $G^* = \frac{1}{3} - \epsilon$, which clearly satisfies $G^*_s = G^*_i$ on $\overline{\Delta}$ and $H > G^*$ on $\partial\Delta$. By Theorem 3.4, it follows that $H > \frac{1}{3} - \epsilon$ on Δ, and since ϵ is arbitrary, that $H \geqslant \frac{1}{3}$. We can improve on this bound by selecting $G(s + i) = 1/(s + i + 1 + \epsilon)^2$ which satisfies $G_s = G_i$ on $\overline{\Delta}$ and $H > G$ on $\partial\Delta$. Hence $H > G$ on Δ, and letting $\epsilon \to 0$, we get $H \geqslant 1/(s + i + 1)^2$. Finally, observe that $H(s, i) = 1/(s^2 + 2i + 1)$ satisfies all of these hypotheses and conclusions. Thus, the known boundary conditions of H can sometimes be used to design an appropriate function G.

4 An SIRI model

In this section we present a brief analysis for the special SIRI case of the general model in system (3.2)–(3.4) when $\varrho = \delta = 0$, $\Phi(s, i) = \phi s$, and $\Psi(r, i) = \psi r$. Observe that in Sect. 3 we showed that $\ell_s \Phi = \ell_r \Psi = 0$, so that no periodic solutions are possible. Since $s + i + r = 1$, the system is essentially two-dimensional, so we can eliminate one of the variables. Eliminating the r variable, we obtain the system

$$s' = b(1 - s) - (\phi - \epsilon)si, \tag{4.1}$$

$$i' = i[(\phi - \psi)s + (\epsilon - \psi)i - (\epsilon + b + \gamma - \psi)]. \tag{4.2}$$

Note that $\overline{\Delta} = \{(s, i) \in \mathbb{R}^2_+ : 0 \leqslant s + i \leqslant 1\}$ is invariant: if $s = 0$ then $s' = b > 0$, if $i = 0$ then $i' = 0$, and on $s + i = 1$ we have $(s + i)' = -\gamma i \leqslant 0$, so that trajecto-

ries beginning in $\overline{\Delta}$ remain in $\overline{\Delta}$. System (4.1)–(4.2) has an equilibrium point at $(1, 0)$ corresponding to the disease free equilibrium (DFE). The eigenvalues of the Jacobian at the DFE are $-b$ and $\phi-(\epsilon+b+\gamma)$, so that the DFE is a locally stable node whenever the threshold parameter $R_0=\phi/(\epsilon+b+\gamma)<1$, and an unstable saddle whenever $R_0>1$.

To determine whether there are any endemic equilibria (s_*, i_*) with $s_*, i_*>0$, observe that the $s'=0$ isocline is a hyperbola through the DFE with asymptotes $s=0$ and $i=b/(\epsilon-\phi)$. This hyperbola intersects Δ if and only if $\phi>b+\epsilon$. The $i'=0$ isoclines are the s-axis and the line

$$(\phi-\psi)s+(\epsilon-\psi)i=(\epsilon+b+\gamma-\psi). \tag{4.3}$$

The analysis depends on four quantities: $\epsilon+b+\gamma$, ϕ, ϵ, ψ, and the first two must exceed the third, so there are eight possible alternatives. There is no equilibrium besides the globally asymptotically stable DFE if $\epsilon+b+\gamma$ is the largest of the four quantities, because the line (4.3) either does not intersect the hyperbola, or intersects it outside of Δ. When ϕ is the largest quantity (so that the DFE is a saddle), there is at most one equilibrium point in Δ. This is easy to see if $\psi>\epsilon$, since (4.3) has positive slope and intersects Δ; the equilibrium point so obtained is stable since the trace of the Jacobian is negative and the determinant is positive. When $\epsilon>\psi$ so that the line (4.3) has negative slope, the line may intersect the hyperbola at two points, the lower of which is clearly in Δ. To see that the other intersection is outside of Δ, consider the intersections of the isoclines with the line $s+i=1$. Substituting $1-s=i$ into the hyperbola, we get $s=b/(\phi-\epsilon)$, while the intersection of (4.3) with $s+i=1$ yields $s=(b+\gamma)/(\phi-\epsilon)$. An isocline analysis of this case shows that near the origin in $\mathbb{R}^2_+$ we have $s'>0$ and $i'<0$, indicating that the direction field revolves counter-clockwise around the equilibrium point at (s_*, i_*) in Δ. The s-axis is a separatrix going to the saddle at the DFE and the separatrix leaving the DFE has initial slope $-(\phi-\epsilon-\gamma)/(\phi-\epsilon)$. This separatrix must approach the equilibrium point (s_*, i_*) as otherwise there would have to be a periodic orbit, which is impossible as shown above (see Fig. 2(a)). Thus, the equilibrium point (s_*, i_*) is a stable focus.

Finally, when ψ is the largest quantity, the line (4.3) intersects the hyperbola in one point interior to Δ when $R_0>1$, and in (possibly) two points inside Δ when $R_0<1$. In both cases the direction field in $\mathbb{R}^2_+$ near the origin satisfies $s'>0, i'>0$. If $\phi>\epsilon+b+\gamma$, then $R_0>1$ so that the DFE is a saddle point. An isocline analysis (see Fig. 2(b)) shows that the equilibrium point (s_*, i_*) in Δ must be a stable node. If $\phi<\epsilon+b+\gamma$, then $R_0<1$ and the DFE is a stable node. An isocline analysis reveals that when the line (4.3) intersects the hyperbola in two points inside Δ, the one with largest i value must be a stable node, while the other must be a saddle point (see Fig. 2(c)). The separatrices going to this saddle point divide Δ into two basins of attraction. Figure 2(c) shows that one of these separatrices is very close to the s-axis: this indicates that if the initial number of infectives is very low the disease dies out as the trajectory lies in the basin of attraction of the DFE, but a slightly larger number of initial infectives may initiate a trajectory in the basin of attraction of (s_*, i_*), leading to an endemic situation. Thus, this separatrix serves as a threshold curve determining the basin of attraction the trajectory will lie in. Hence, this SIRI model has a behavior more complicated than the corresponding SIRS model in Busenberg and van den Driessche (1990) and the constant population SIRI model in Tudor (1990), although none can have periodic solutions.

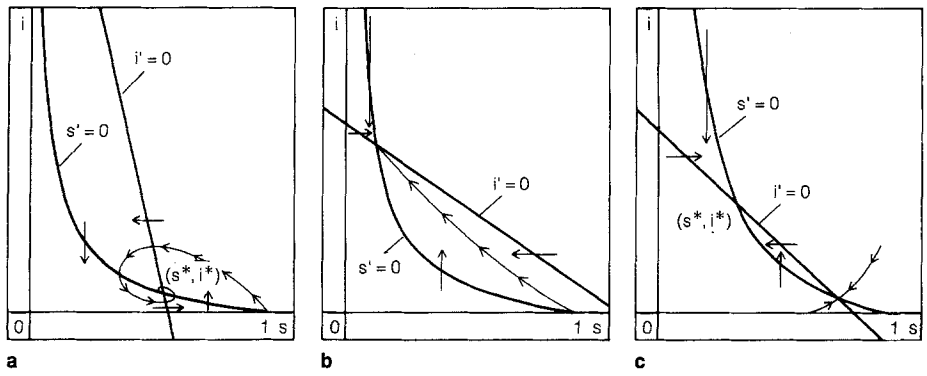

Fig. 2 Phase planes showing isoclines and separatrices for the $\varrho = \delta = 0$, $\Phi(s, i) = \phi s$, $\Psi(r, i) = \psi r$ case. (a) Parameters: $\phi = 0.3$, $\epsilon = 0.15$, $\psi = 0.1$, $\gamma = 0.05$, $b = 0.01$. Separatrix emanates from saddle at $(1, 0)$ and goes to stable focus at $(0.5355, 0.0578)$. (b) Parameters: $\phi = 0.25$, $\epsilon = 0.15$, $\psi = 0.4$, $\gamma = 0.07$, $b = 0.01$. Separatrix emanates from saddle at $(1, 0)$ and goes to stable node at $(0.1442, 0.5935)$. (c) Parameters: $\phi = 0.2$, $\epsilon = 0.15$, $\psi = 0.4$, $\gamma = 0.07$, $b = 0.01$. Separatrices to saddle at $(0.7791, 0.0567)$ split $\bar{\Delta}$ into two basins of attraction: one to the DFE, the other to the stable node at $(0.3209, 0.4233)$. Note separatrix gets arbitrarily close to s-axis

5 Periodic solutions

Before studying a situation that leads to periodic solutions, it is convenient to analyze the general case by eliminating the s variable from Eqs. (3.3) and (3.4). One equilibrium point is immediate: the disease free equilibrium at the origin. Any others depend on the locus of the functions $\Phi(1 - i - r, i)$ and $\Psi(r, i)$ in the triangular region $\bar{\Delta}_1 = \{(i, r) \in \mathbb{R}^2_+ : 0 \leqslant i + r \leqslant 1\}$. We denote any other such equilibrium point (if it exists) by the notation (i_*, r_*).

The eigenvalues of the Jacobian for the DFE are

$$-(\varrho + b + \delta) \quad \text{and} \quad (\epsilon + b + \gamma)(R_0 - 1),$$

where $R_0 = (\Phi(1, 0) + \Psi(0, 0))/(\epsilon + b + \gamma)$, so that the DFE is locally stable whenever $R_0 < 1$. For $R_0 > 1$, the DFE is a saddle point.

At any interior equilibrium point (i_*, r_*), the Jacobian is given by

$$J = \begin{bmatrix} i(\epsilon + [\Phi(1 - i - r, i) + \Psi(r, i)]_i) & i(\delta + [\Phi(1 - i - r, i) + \Psi(r, i)]_r) \\ \gamma + \epsilon r - \Psi(r, i) - i\Psi_i(r, i) & \delta r + i((\Psi(r, i) - \gamma)/r - \Psi_r(r, i)) \end{bmatrix}. \quad (5.1)$$

In the rest of this section we assume that $\Phi(s, i) = \phi si$ and $\Psi(r, i) \equiv 0$. Then $\ell_s \Phi = \phi/r$ and $\ell_r \Psi = 0$, so that by (3.10)

$$\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1} = \frac{1}{sir}\left(\phi si - \frac{\gamma i}{r} - \frac{bi}{s} - \frac{(b + \varrho)r}{s}\right). \quad (5.2)$$

Observe that if $b > \phi$ or $\gamma > \phi$, then (5.2) is negative and no periodic solutions are possible. However, for small values of b and $\gamma < \phi$, the term in parenthesis may be either positive or negative depending on the sizes of s, i, r, and ϱ. Thus, periodic solutions are possible. For this case the Jacobian becomes

$$J = \begin{bmatrix} i(\epsilon + \phi(1 - 2i - r)) & i(\delta - \phi i) \\ \gamma + \epsilon r & \delta r - i\gamma/r \end{bmatrix}. \quad (5.3)$$

The interior equilibrium points (i_*, r_*) are located at the intersections of the hyperbolas

$$\phi i^2 + \phi ir - (\phi + \epsilon)i - \delta r + (b + \epsilon + \gamma) = 0, \tag{5.4}$$

$$\epsilon ir + \delta r^2 + \gamma i - (b + \delta + \varrho)r = 0. \tag{5.5}$$

Using these expressions, we can write the trace and determinant of (5.3) in the following ways:

$$\operatorname{tr} J = (\epsilon i + \delta r) - i[\gamma/r + \phi(2i + r - 1)], \tag{5.6}$$

$$\operatorname{tr} J = (b + \epsilon + \gamma) - i[\gamma/r + \phi i], \tag{5.7}$$

$$\frac{\det J}{i} = \phi(2i + 2r - 1)(\gamma i/r - \delta r) + (\epsilon i + \delta r)(\phi r - \gamma/r), \tag{5.8}$$

where $(i, r) = (i_*, r_*)$. In what follows, we require the following inequalities:

$$3\epsilon - 2\delta < \phi < \min\{4\epsilon, 16(\epsilon - \delta), 16\delta\}, \tag{5.9}$$

$$0 < 2\sigma < 2\epsilon < \delta + \sqrt{\delta\phi}. \tag{5.10}$$

The parameters $\phi = 0.7$, $\epsilon = 0.3$, and $\delta = 0.25$ satisfy (5.9)–(5.10), and will be used in the numerical simulations in Sect. 6. The condition that $\delta < \epsilon$ implies that the disease related deaths are likely to be greater in class I than in class R, a reasonable assumption.

We begin by considering the situation when $b = \gamma = 0$, and extend the results we obtain by homotopy to small positive values of b and then γ. Observe that the region $\overline{\Delta}_1$ is invariant: setting $i = 0$ in (3.3) implies that $i' = 0$, setting $r = 0$ in (3.4) yields $r' = 0$, while on $i + r = 1$ we have $(i + r)' = -\varrho r \leqslant 0$ since $s = 0$. Thus, any orbit beginning in $\overline{\Delta}_1$ must remain in $\overline{\Delta}_1$. The $i' = 0$ isoclines are the axis $i = 0$ and the hyperbola $r = (1 - i)(\phi i - \epsilon)/(\phi i - \delta)$, while the $r' = 0$ isocline in Eq. (5.5) degenerates into two lines: the axis $r = 0$ and $r = (\delta + \varrho - \epsilon i)/\delta$. Solving the hyperbola and this last line simultaneously for i, we get the quadratic

$$(\epsilon - \delta)\phi i^2 - \varrho\phi i + \delta(\varrho - (\epsilon - \delta)) = 0. \tag{5.11}$$

Let $\varrho = k(\epsilon - \delta)$, then (5.11) becomes $i^2 - ki - (1 - k)\delta/\phi = 0$, which has the root

$$i = (k/2)[1 + \sqrt{1 + 4\delta(1 - k)/(\phi k^2)}, \tag{5.12}$$

so there is at most one equilibrium point in $\mathbb{R}^+_2$ when $0 < k < 1$. To see that there is an equilibrium point (i_*, r_*) in $\Delta_1 \cap \mathbb{R}^+_2$, observe that

$$r = (1 - k) + (k\epsilon/2\delta)[1 - \sqrt{1 + 4\delta(1 - k)/(\phi k^2)}] > 0,$$

$$s = 1 - i - r = (i - k)(\epsilon - \delta)/\delta > 0, \tag{5.13}$$

since $\delta\phi > \epsilon^2$ by (5.10). Thus, there are four equilibrium points in $\overline{\Delta}_1$: the DFE $(0, 0)$, $(\epsilon/\phi, 0)$, $(1, 0)$, and (i_*, r_*).

It is easy to check that the DFE is always a stable node ($R_0 = 0$), $(\epsilon/\phi, 0)$ is always a saddle point, and $(1, 0)$ is a saddle for $k < 1$. In Sect. 6 we will need the sum of the traces of the Jacobian at $(\epsilon/\delta, 0)$ and $(1, 0)$:

$$\operatorname{tr} J(\epsilon/\delta, 0) + \operatorname{tr} J(1, 0) = 3\epsilon - 2\delta - \phi - 2\varrho < 0, \tag{5.14}$$

for all ϱ by (5.9). We now consider the nature of (i_*, r_*) at $k = 0$ and $k = 1/2$.

If $k = 0$, then $i^2 = \delta/\phi$, so that $i = \delta/\sqrt{\delta\phi}$, $r = 1 - \epsilon i/\delta = 1 - \epsilon/\sqrt{\delta\phi}$, and $s = (\epsilon - \delta)/\sqrt{\delta\phi} > 0$, implying that $s < 1$ by (5.10). By (5.7) we have that $\operatorname{tr} J = \epsilon - \delta > 0$, and, by (5.8), $\det J = \phi i r[\epsilon i - \delta r - 2\delta i + \delta] = 2ir(\epsilon - \sigma)\sqrt{\delta\phi} > 0$. But the discriminant equals $(\operatorname{tr} J)^2 - 4 \det J = (\epsilon - \delta)[(\epsilon - \delta) - 8\delta(1 - \epsilon/\sqrt{\delta\phi})] < 0$, since by (5.10)

$$\frac{\epsilon}{\delta} + \frac{8\epsilon}{\sqrt{\delta\phi}} - 9 < \tfrac{1}{2}(8\sqrt{\delta/\phi} - 1)(1 - \sqrt{\phi/\delta}),$$

and the first term in parenthesis is positive by (5.9), while the latter is negative. Thus, (i_*, r_*) is an unstable focus.

If $k = 1/2$, then $i = [1 + \sqrt{1 + 8\delta/\phi}]/4$, $r = (1/2) + (\epsilon/\delta)[1 - \sqrt{1 + 8\delta/\phi}]/4$, and $\operatorname{tr} J = (2\epsilon - \delta - \phi i)/2 < 0$, by (5.10), because $i > \sqrt{\delta/\phi}$ from (5.11). Furthermore, $\det J = \phi i r(\epsilon - \delta)(\sqrt{1 + 8\delta/\phi})/2 > 0$, and the discriminant equals

$$\left(\frac{\phi i}{2} - \left(\epsilon - \frac{\delta}{2}\right)\right)^2 - 2(\epsilon - \delta)(\sqrt{1 + 8\delta/\phi})\left(\frac{\phi i}{2} - \frac{\epsilon}{2}\right),$$

which is negative since

$$0 < \left(\frac{\phi i}{2} - \left(\epsilon - \frac{\delta}{2}\right)\right) < \left(\frac{\phi i}{2} - \frac{\epsilon}{2}\right), 2(\epsilon - \delta)(\sqrt{1 + 8\delta/\phi}),$$

by (5.9) and (5.10). Hence, (i_*, r_*) is a stable focus. Thus, a Hopf bifurcation occurs between $k = 0$ and $k = 1/2$. It is not difficult to find the value of k at which the Hopf bifurcation occurs: substitute (5.12) into $\operatorname{tr} J = 0$ using (5.7) and solve the resulting quadratic equation getting

$$k = \frac{\epsilon - \delta}{\sqrt{\epsilon\phi} - \delta}. \tag{5.15}$$

We use (5.8), (5.12), (5.13), and (5.15) to compute the eigenvalues $\pm i\omega$ at the Hopf bifurcation. A messy, but routine, calculation yields

$$\omega^2 = (\epsilon - \delta)\left[\epsilon + \delta - \frac{2\epsilon\delta}{\sqrt{\epsilon\phi}}\right]\left(\frac{\sqrt{\epsilon\phi} - \epsilon}{\sqrt{\epsilon\phi} - \delta}\right)^2, \tag{5.16}$$

so that the period of the periodic solution is $2\pi/\omega$.

Now assume that $b > 0 = \gamma$. The $i' = 0$ isoclines again consist of the axis $i = 0$ and a hyperbola which lies below the previous hyperbola for $i > \delta/\phi$. For $(\phi - \epsilon)^2/4\phi > b$, the new hyperbola will intersect the i-axis at two points i_1 and i_2 between ϵ/ϕ and 1. The $r' = 0$ isoclines again consist of the axis $r = 0$ and a straight line $\epsilon i + \delta r = b + \delta + \varrho$. For $i_i < (b + \delta + \varrho)/\epsilon < i_2$, this line intersects the hyperbola at a point (i_*, r_*) with i_* between the roots of the hyperbola. We again find four equilibria: the DFE at the origin, $(i_1, 0)$, $(i_2, 0)$, and (i_*, r_*). The DFE is again a stable node with eigenvalues $-(b + \epsilon)$, $-(b + \delta + \varrho)$, and since

$$J(i_j, 0) = \begin{bmatrix} 2\phi i_j\left(\dfrac{1 + \epsilon/\phi}{2} - i_j\right) & i_j(\delta - \phi i_j) \\ 0 & \epsilon\left(i_j - \dfrac{b + \delta + \varrho}{\epsilon}\right) \end{bmatrix}, \quad j = 1, 2,$$

it follows that $(i_1, 0)$ and $(i_2, 0)$ are saddles. If we calculate the sum of the traces of the Jacobian at these points we get

$$\operatorname{tr} J(i_1, 0) + \operatorname{tr} J(i_2, 0) = 3\epsilon - 2\delta - \phi - 2\varrho + 2b,$$

which, like (5.14), will be negative for all ϱ provided b is sufficiently small.

The analysis at (i_*, r_*) is very similar to that done above and shows that for $\varrho = k(\epsilon - \delta) = 0$ we again have an unstable focus, while for $k = 1/2$ and sufficiently small b, we again have a stable focus. Thus a Hopf bifurcation occurs between these values.

If we continue this homotopy process to sufficiently small positive values of both b and γ, the $i' = 0$ isocline again consists of the axis $i = 0$ and a hyperbola lying below the previous ones for $i > \delta/\phi$, but the $r' = 0$ isocline becomes the hyperbola (5.5). For sufficiently small γ, this hyperbola is closely approximated by the two straight line isoclines of the $b > 0 = \gamma$ case. This time we have only three equilibria: the DFE, the point (i_*, r_*), and a point P close to $(i_1, 0)$. The equilibrium corresponding to $(i_2, 0)$ now lies in the fourth quadrant, so has no biological meaning. The trace at P is given by (5.6) or (5.7), and an isocline analysis reveals that P is a saddle point. The point (i_*, r_*) is an unstable focus for k near zero and a stable focus for k near 1/2, so that a Hopf bifurcation again occurs in between.

To investigate the stability of the periodic oscillations associated with the Hopf bifurcations occurring above, we could invoke the criterion for super- or sub-critical Hopf bifurcations given by Guckenheimer and Holmes (1983) or its equivalent formulation given by Eq. (A.16) in Liu et al. (1986). It is difficult to test this condition analytically at these Hopf bifurcations, because it contains dozens of terms. In the next section, we use two other numerical methods to determine the stability of the periodic oscillations.

6 Numerical simulations

In this final section we will provide numerical evidence to show that stable periodic solutions exist for the model in Sect. 5. We use the values $\phi = 0.7$, $\epsilon = 0.3$, $\delta = 0.25$, and make computations for the cases $b = \gamma = 0$, $b = 0.001 > \gamma = 0$, and $b = \gamma = 0.001$.

The analysis above indicates the existence of a single Hopf bifurcation. We show that as ϱ increases stable periodic solutions arise as a consequence of a saddle connection, for the cases $b = \gamma = 0$ and $b = 0.001 > \gamma = 0$, and from a homoclinic orbit, for the case $b = \gamma = 0.001$, and these stable periodic solutions vanish as a consequence of the Hopf bifurcation.

We have used PhasePlane (Ermentrout 1990) to obtain the first three graphs in Fig. 3, showing the separatrices emanating from each saddle point for the case $b = \gamma = 0$. These graphs were produced using a Runge-Kutta integrator and a time step of 0.05. In Fig. 3(a) we have set $\varrho = 0.005$ ($k = 0.1$); a separatrix emanates from the saddle at (1, 0), bounds the unstable focus at (0.6191, 0.2770) and goes to the origin. The saddle at (3/7, 0) has separatrices going to the origin and to (1, 0) and a separatrix coming from the unstable focus. Figure 3(b) shows the saddle connection which occurs at approximately $\varrho = 0.007278$ according to PhasePlane, and confirmed by AUTO (Doedel 1986) (see Fig. 3(d)). This structurally unstable connection, from the saddle at (1, 0) to the saddle at $(\epsilon/\phi, 0)$, encloses an unstable focus at (0.6300, 0.2732) and gives birth to stable periodic orbits of arbitrarily large periods by (5.14) (see p. 295 of Guckenheimer and Holmes 1983). Using (5.15), we calculate that the Hopf bifurcation occurs at approximately $\varrho = 0.012$, so that periodic orbits exist in the range (0.007278, 0.012). Figure 3(c) shows that when $\varrho = 0.009$, the separatrix emanating from (1, 0) is attracted to a region containing the unstable focus at

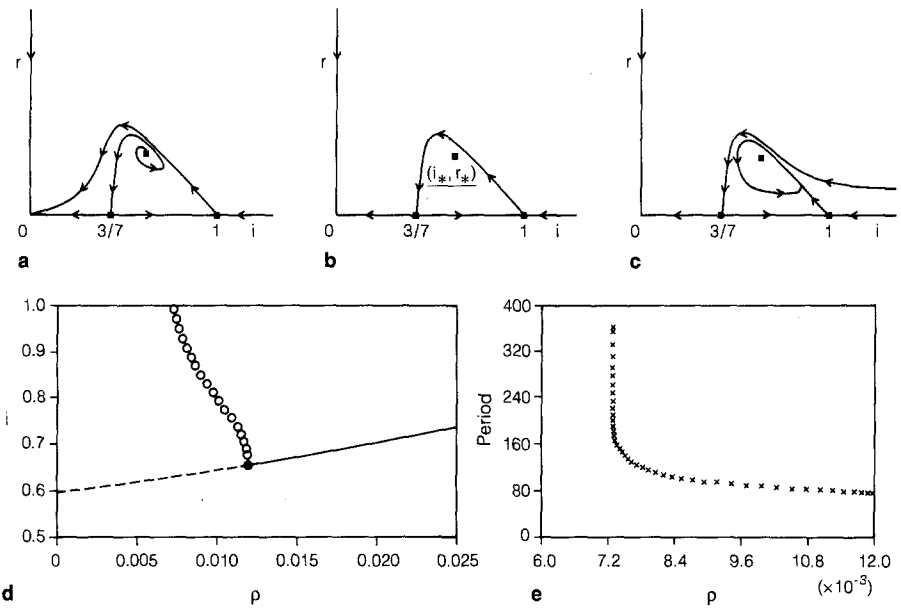

Fig. 3 Separatrix analysis and bifurcation diagram for the $b = \gamma = 0$, $\Phi(s, i) = \phi si$, $\Psi(r, i) = 0$ case. The following parameters are held constant: $\phi = 0.7$, $\epsilon = 0.3$, $\delta = 0.25$. The parameter ϱ varies. (a) $\varrho = 0.005$. Separatrix emanates from saddle at $(1, 0)$ and goes to DFE. Separatrix emanates from unstable focus at $(0.6191, 0.2770)$ and goes to saddle at $(3/7, 0)$. (b) Saddle connection occurs at $\varrho \approx 0.007278$, creating stable periodic orbits of arbitrarily large period by (5.14). (c) Here $\varrho = 0.009$. Separatrix to $(3/7, 0)$ separates $\bar{\Delta}$ into two basins of attraction: to the DFE, and to the stable periodic orbit. (d) Bifurcation diagram with ϱ on abscissa and i on ordinate. Dashed lines correspond to unstable focus, solid line stable focus at (i_*, r_*). Open circles denote maximum i value of stable periodic orbit. Observe max i tends to 1 as ϱ decreases to 0.0073, where the periodic orbits originate, and decreases to i_* as ϱ tends to the Hopf bifurcation at 0.012. (e) AUTO graph plotting periods of stable periodic orbits for values in the range $0.007278 < \varrho < 0.012$. Periods decrease from 361.0856 at $\varrho = 0.007278$ to 78.3611 at $\varrho = 0.012$. Data points marked by an x

$(0.6386, 0.2697)$, while the separatrix going to $(3/7, 0)$ emanates from the line $i + r = 1$. Here there are two basins of attraction: one to the DFE above this separatrix, the other (below the separatrix) going to the stable periodic orbit. Figure 3(d), produced by AUTO, tracks the equilibrium point (i_*, r_*) for a range of values ϱ; the abscissa shows the values of ϱ, the ordinate the values of i. AUTO uses a broken line for unstable and a solid line for stable equilibria, and marks the Hopf bifurcation by a solid box. The circles in Fig. 3(d) indicate the maximum value of i for the stable periodic orbits cast off by the Hopf bifurcation. Figure 3(e) also produced by AUTO graphs the periods of the stable periodic orbits. Observe that the period increases as ϱ decreases to 0.007278. For $\varrho = 0.012$, PhasePlane computes eigenvalues of $\pm 0.08001i$ at $(0.6546, 0.2624)$, which are very close to the exact values $\pm 0.08018i$ obtained using (5.16). The corresponding exact period of the orbit is 78.3605, in very close agreement with the value 78.3611 obtained by AUTO.

We repeated these calculations for the case $b = 0.001 > \gamma = 0$, with the same values for ϕ, ϵ, and δ. Figure 4(a) with $\varrho = 0$ shows the isoclines and separatrices through the saddles at $(i_1, 0)$ and $(i_2, 0)$. There is an unstable focus at

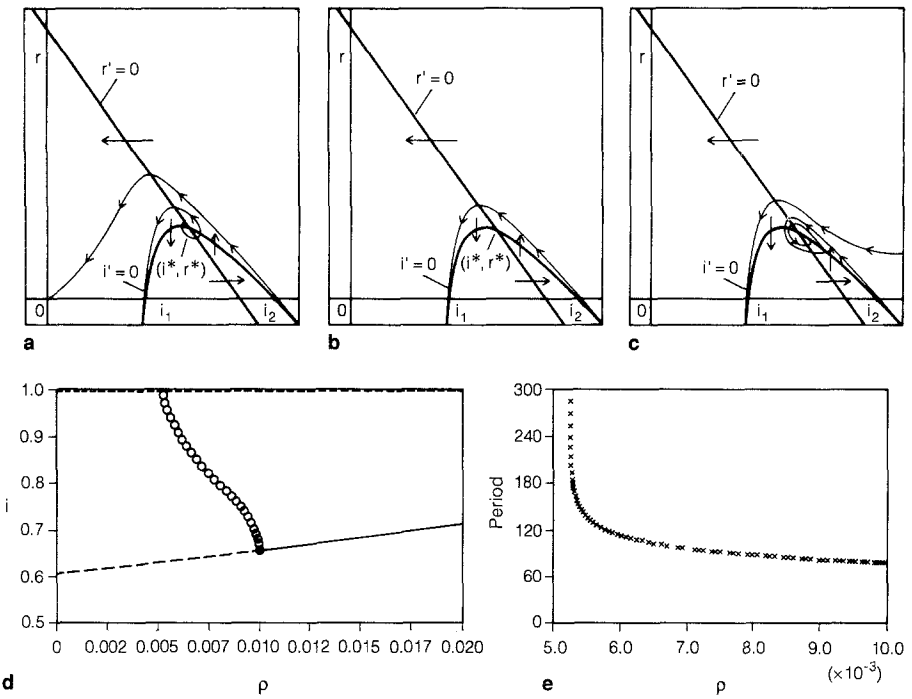

Fig. 4 Isoclines, separatrices, and bifurcation diagram for the $b = 0.001 > \gamma = 0$, $\Phi(s, i) = \phi si$, $\Psi(r, i) = 0$ case. Constant parameters used for $\phi = 0.7$, $\epsilon = 0.3$, $\delta = 0.25$. Again ϱ varies. (a) Here $\varrho = 0$, $i_1 = 0.4311$ and $i_2 = 0.9975$. Separatrix from $(i_2, 0)$ goes to DFE. Single basin of attraction. (b) Saddle connection occurs at $\varrho \approx 0.005249$. Here $i_1 = 0.4311$, $i_2 = 0.9975$ and the unstable focus is at $(0.6313, 0.2674)$. Two basins: one to DFE, the other to the saddle connection. (c) $\varrho = 0.01$. Two basins of attraction: one to DFE, the other to the stable focus at $(0.6557, 0.2571)$. (d) Bifurcation diagram with ϱ on abscissa and i on ordinate. Dashed line represents an unstable focus, solid line a stable focus. Open circle give maximum i value for stable periodic orbits. Note they originate near 0.0052 and vanish at the Hopf bifurcation, indicated by black square. (e) AUTO graph showing period of stable periodic orbits for ϱ in the range $0.005 < \varrho < 0.01$. Period reaches 285.6755 at $\varrho = 0.005249$ and decreases to a period of 78.2918 at $\varrho = 0.01$. Data points marked by an x

$(0.6077, 0.2748)$ which sends a trajectory to $(i_1, 0)$. In Fig. 4(b) the saddle connection between $(i_1, 0)$ and $(i_2, 0)$ is formed at about $\varrho = 0.005249$. Here $(0.6313, 0.2674)$ is an unstable focus. The saddle connection produces stable periodic orbits of arbitrarily large period. Figure 4(c), with $\varrho = 0.01$, shows two basins of attraction: one to the DFE, the other to a stable focus at $(0.6557, 0.2571)$, so the Hopf bifurcation occurs in the interval $0.005249 < \varrho < 0.01$. PhasePlane finds (by repeated trials) that the Hopf bifurcation occurs when $\varrho \approx 0.009944$ at $(0.6670, 0.2516)$, with eigenvalues of $\pm 0.07906i$, implying a period of 79.4696. Using AUTO (see Figs. 4(d) and 4(e)), we find the Hopf bifurcation occurring at $\varrho = 0.01$ with a period of 78.2918. AUTO tracks stable periodic orbits to $\varrho = 0.005249$ with a period of 285.6755, in very close agreement with PhasePlane and the theory we have done in Sect. 5. Hence, stable periodic solutions exist in the range $0.005249 < \varrho < 0.01$.

When $b = \gamma = 0.001$, the $r' = 0$ isocline no longer degenerates into two lines (see (5.5)), so that we lose the saddle at $(1, 0)$. We now have a stable node at the

origin, a saddle at the point P near $(\epsilon/\phi, 0)$ where the two hyperbolic isoclines intersect, and an equilibrium point at (i_*, r_*). The first assertion is easily shown from the Jacobian matrix and the second follows from the direction field. By homotopy, for k near zero we should have an unstable focus and for k near 1/2 a stable focus at (i_*, r_*). In Fig. 5(a) we set $\varrho = 0$ and note that the separatrix emanating to the right from $P = (0.4341, 0.0036)$ encircles the unstable focus at (0.6044, 0.2698) and proceeds to the origin. A separatrix emanates from the unstable focus and goes to P. We set $\varrho = 0.006393$ to get a homoclinic orbit as shown in Fig. 5(b). Essentially, the homotopy has transformed the saddle connection we had when $b = 0.001 > \gamma = 0$ into this homoclinic orbit. The same theory (see Theorem 6.1.1 of Guckenheimer and Holmes 1983) applies to homoclinic orbits as to saddle connections: they cast off periodic orbits with arbitrarily large periods. Here the trace at P is positive, so unstable periodic orbits are cast off by the homoclinic orbit. However, since $(i_*, r_*) = (0.6315, 0.2622)$ is an unstable focus, a stable periodic orbit must lie between it and the homoclinic orbit. We should expect AUTO to find at least one limit point between the periodic orbits to indicate changes from stable periodic orbits to unstable periodic orbits (see Fig. 5(d)). At least two basins of attraction must exist. Figure 5(c) for $\varrho = 0.0111$ shows that the separatrix emanating from $P = (0.4341, 0.0033)$ approaches the stable focus at (0.6542, 0.2530), and that a separatrix originates on the line $i + r = 1$ and goes to P. Thus there are now exactly two basins of attraction: the DFE and the stable focus. PhasePlane locates the Hopf bifurcation (by repeated trials) at approximately $\varrho = 0.011$, with eigenvalues $\pm 0.08185i$, implying a period of 76.7618. Figure 5(d), produced by AUTO with ϱ on the abscissa and i on the ordinate, shows that a Hopf bifurcation occurs at $\varrho = 0.01106$ at the point (0.6540, 0.2531). AUTO finds several limit points near $\varrho = 0.011$ and $\varrho = 0.006393$, indicating the existence of multiple stable and unstable periodic orbits (see Figs. 5(e) and 5(f)). The periods of the orbits near $\varrho = 0.006393$ become progressively larger; those near $\varrho = 0.011$ are approximately 76.349, in very close agreement with PhasePlane. Thus, stable periodic solutions exist for the range of values $0.006393 < \varrho < 0.01106$. Their maximum i values are plotted in Fig. 5(d) by the open circles.

The results above can be summarized briefly as follows. Setting $\varrho = 0$ in our model (3.2)–(3.4), with $\Phi(s, i) = \phi si$, $\Psi(r, i) \equiv 0$, and the other parameters used, yields an SIR model whose endemic equilibrium (i_*, r_*) is an unstable focus. Thus, the infected and recovered proportions tend to zero as $t \to \infty$. By contrast, for a corresponding SIRS model with $\varrho = (\epsilon - \delta)/2$, the endemic equilibrium (i_*, r_*) is a stable focus. Then, depending on initial conditions, the infective and recovered proportions can tend to endemic values. For intermediate values of ϱ, stable periodic solutions (as shown in Figs. (3)–(5)) are possible.

As suggested by a referee, we have also found periodic solutions for our model with a different incidence function $\Phi(s, i) = \phi\sqrt{s}$, with $\Psi(r, i) \equiv 0$ and the parameter values used above. Observe that (3.10) yields

$$\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1} = \frac{1}{sir}\left(\frac{\phi i}{2\sqrt{s}} - \frac{\gamma i}{r} - \frac{bi}{s} - \frac{(b+\varrho)r}{s}\right)$$

in $\mathscr{D}^0$, so that for small values of b, γ, and ϱ, $\operatorname{curl} \boldsymbol{g} \cdot \mathbf{1}$ may be positive. This incidence function corresponds to that used for constant populations by Liu et al. (1986, 1987) with $p = 1$, $q = \frac{1}{2}$, where no periodic solutions occur. Numerical

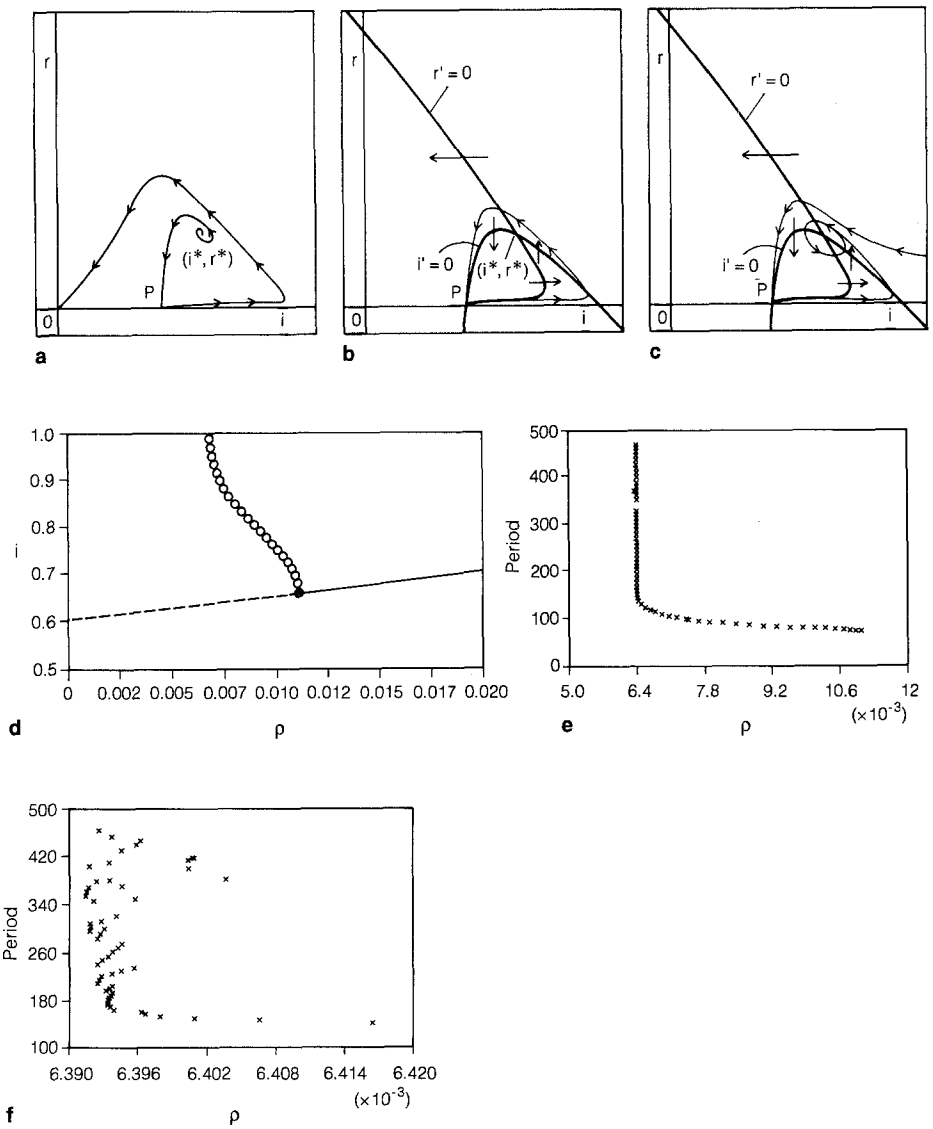

Fig. 5 Isoclines, separatrices, and bifurcation diagram for the $b = \gamma = 0.001$, $\Phi(s, i) = \phi si$, $\Psi(r, i) = 0$ case. Constant parameters $\phi = 0.7$, $\epsilon = 0.3$, $\delta = 0.25$. The parameter ϱ varies. (a) Using $\varrho = 0$ we get a saddle at $P = (0.4341, 0.0036)$ with separatrix going around the unstable focus at $(0.6044, 0.2698)$ and proceeding to the DFE. Note that the track of this trajectory is close to the separatrices from $(i_1, 0)$ to $(i_2, 0)$ and from $(i_2, 0)$ to the DFE. (b) Homoclinic orbit occurring at $\varrho = 0.006393$ for $P = (0.4341, 0.0034)$ and surrounding the unstable focus at $(0.6315, 0.2622)$. (c) Using $\varrho = 0.0111$, separatrix from $P = (0.4341, 0.0033)$ proceeds to stable focus at $(0.6542, 0.2530)$. Separatrix from $i + r = 1$ goes to P splitting domain into two basins of attraction: one to the DFE, the other to the stable focus. (d) AUTO bifurcation diagram indicating creation of stable periodic orbits shown as open circles near $\varrho = 0.0064$, which then vanish at the Hopf bifurcation point marked by a black box near $\varrho = 0.0111$. Unstable foci denoted by dashed line, stable foci by solid line. (e) Bifurcation diagram showing periods of periodic orbits for different values of ϱ in the range $0.006393 < \varrho < 0.0111$. Each data point marked by an x. (f) Magnification of the data in Fig. 5(e) in a range about $0.00639 < \varrho < 0.00642$

calculations with $\varrho = 0$ yield a stable limit cycle in the ir-plane in the region $(0.0, 0.1] \times [0.9, 1.0)$ about the unstable focus at $(0.0255, 0.9694)$ for $b = \gamma = 0$, or $(0.0457, 0.9489)$ for $b = \gamma = 0.001$.

References

Busenberg, S., van den Driessche, P.: Analysis of a disease transmission model in a population with varying size. J. Math. Biol. **28**, 257–270 (1990)

Busenberg, S., van den Driessche, P. Nonexistence of periodic solutions for a class of epidemiological models. In: Busenberg, S., Martelli, M. (eds.) Biology, Epidemiology, and Ecology. (Lect. Notes Biomath., vol. 92, pp. 70–79) Berlin Heidelberg New York: Springer 1991

Castillo-Chavez, C., Hethcote, H. W., Andreasen, V., Levin, S. A., Liu, W. M.: Epidemiological models with age structure, proportionate mixing, and cross-immunity. J. Math. Biol. **27**, 233–258 (1989)

Diekmann, O., Kretzschmar, M.: Patterns in the effects of infectious diseases on population growth. J. Math. Biol. **29**, 539–570 (1991)

Doedel, E.: AUTO: Software for Continuation and Bifurcation Problems in Ordinary Differential Equations. California Institute of Technology (1986)

Ermentrout, B.: PhasePlane: The Dynamical System's Tool, Version 3.0. Pacific Grove, CA: Brooks/Cole 1990

Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Berlin Heidelberg New York: Springer 1983

Hahn, W.: Stability of Motion. Berlin Heidelberg New York: Springer 1967

Hethcote, H. W., Levin S. A.: Periodicity in epidemiological models. In: Levin, S. A., Hallam, T. G., Gross, L. J. (eds.) Applied Mathematical Ecology. (Biomath., vol. 18). Berlin Heidelberg New York: Springer 1989

Jacquez, J. A., Simon, C. P., Koopman, J., Sattenspiel, L., Perry, T.: Modelling and analyzing HIV transmission: the effect of contact patterns. Math. Biosci. **92**, 119–199 (1988)

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

Liu, W.-M., Hethcote, H. W., Levin, S. A.: Dynamical behavior of epidemiological models with nonlinear incidence rates. J. Math. Biol. **25**, 359–380 (1987)

Mena-Lorca, J., Hethcote, H. W.: Dynamic models of infectious diseases as regulators of population size. J. Math. Biol. **30**, 693–716 (1992)

Nold, A.: Heterogeneity in disease-transmission modeling. Math. Biosci. **52**, 227–240 (1980)

Revelle, C., Lynn, W. R., Feldmann, F.: Mathematical models for the economic allocation of tuberculosis control activities in developing nations. Am. Rev. Respir. Dis. **96**, 893–909 (1967)

Tudor, D.: A deterministic model for herpes infections in human and animal populations. SIAM Rev. **32**, 136–139 (1990)

Westphal, H.: Zur Abschätzung der Lösungen nichlinearer parabolischer Differentialgleichungen. Math. Z. **51**, 690–695 (1947/49)