J. Math. Biol. (1992) 30: 717–731

Journal of
Mathematical
Biology

Disease transmission models with density-dependent demographics

Linda Q. Gao and Herbert W. Hethcote*
Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA

Received May 14, 1991; received in revised form July 22, 1991

Abstract. The models considered for the spread of an infectious disease in a population are of SIRS or SIS type with a standard incidence expression. The varying population size is described by a modification of the logistic differential equation which includes a term for disease-related deaths. The models have density-dependent restricted growth due to a decreasing birth rate and an increasing death rate as the population size increases towards its carrying capacity. Thresholds, equilibria and stability are determined for the systems of ordinary differential equations for each model. The persistence of the infectious disease and disease-related deaths can lead to a new equilibrium population size below the carrying capacity and can even cause the population to become extinct.

Key words: Epidemiological model – Density-dependent logistic growth – Thresholds – Stability

1 Introduction

The asymptotic behavior of solutions of an infectious disease transmission model depends not only on the epidemiological formulation, but also on the demographic process incorporated into the model. The simplest epidemiological models often assume that the total population size is constant. For epidemics (rapid short-term outbreaks of a disease) the population is often assumed to be fixed and closed. For modeling an endemic situation (long-term persistence of a disease), there are often births and balancing deaths so the total population size remains constant. See Hethcote (1976, 1989) for results on basic models with fixed population size; surveys on epidemiological modeling are given in Hethcote et al. (1981) and Hethcote and Levin (1989). Sometimes there are a significant number of deaths caused by infection with the disease, which affect the population size. Sometimes the population size is growing or decreasing significantly due to other factors. In these cases it is not reasonable to assume that the population size is constant so that the model must incorporate demographic features which allow the population size to vary.

* Research supported in part by Centers for Disease Control contract 200-87-0515

Some epidemiological models with varying population size assume constant immigration and deaths proportional to the population size so that the population approaches an equilibrium size. Other models assume a more natural demographic process where the birth and death rates are proportional to the population size. Anderson and May (1978, 1979) and May and Anderson (1978, 1979) proposed a variety of models for infectious diseases with varying population sizes and applied some of these to data on diseases in laboratory populations of mice. Several epidemiological models with varying population size are analyzed mathematically in Busenberg and van den Driessche (1990), Busenberg and Hadeler (1990) and Mena-Lorca and Hethcote (1992). Many models for AIDS have varying population size (Hyman and Stanley 1988, Jacquez et al. 1988, May et al. 1988, Anderson et al. 1988, Castillo-Chavez et al. 1989).

A disadvantage of the models with birth and death rates proportional to the population size are that the population size decreases or grows exponentially except in the special case when births exactly balance deaths. Extinction of the population by exponential decay is demographically unlikely; also exponential growth to infinity is unrealistic in human and animal populations since finite resources always eventually limit the growth. Models with restricted growth due to density dependence have been considered by Anderson et al. (1981), Brauer (1989, 1990), Bremermann and Thieme (1989) and Pugliese (1990). New epidemiological models with density-dependent growth are analyzed here.

A demographic structure with density-dependent restricted population growth is given by the logistic equation

$$dN/dt = r(1 - N/K)N \tag{1.1}$$

where $N(t)$ is the total population size as a function of time t, r is the positive growth rate constant and K is the carrying capacity of the environment (see Edelstein 1988 for a discussion of the logistic equation). All of the epidemiological models in this paper are built from the logistic equation (1.1). The substitution $u = 1/N$ converts (1.1) to a linear differential equation. From the explicit solution it is found that all solutions with positive initial population size N_0 approach the carrying capacity K; of course, the population size remains at zero if N_0 is zero. Although it is less common, it is possible to consider (1.1) with $r = 0$ or $r < 0$. For $r = 0$, the Eq. (1.1) is trivial and the population size $N(t)$ remains at N_0. For $r < 0$, solutions $N(t)$ with $N_0 > K$ grow to infinity and solutions with $N_0 < K$ decrease to zero. When $r < 0$, only solutions with $N_0 < K$ are considered in this paper since the others are inconsistent with the concept of density-dependent restricted growth.

An epidemiological model is of SIRS type if susceptibles become infectious, then removed with immunity after recovery from infection and then susceptible again when the temporary immunity fades away. The numbers of individuals who are susceptible, infectious and removed at time t are denoted by $X(t)$, $Y(t)$ and $Z(t)$, respectively. All individuals are in one of these classes so that they add up to the total population size $N(t)$.

The incidence in an epidemiological model is the rate at which susceptibles become infectious. Thus the incidence is the number of new infections per day or per other time unit. The daily contact rate λ is the average number of adequate contacts of an infective per day. An adequate contact is a contact

with another individual which is sufficient for transmission of infection if the other individual is susceptible. Since $X(t)/N(t)$ is the susceptible fraction of the population, $\lambda X(t)/N(t)$ is the average number of infection transmissions per infective per day. The incidence is $\lambda X(t)Y(t)/N(t)$, which is the average number of infection transmissions per day by all infectives $Y(t)$. This same incidence also results if λ is the average number of adequate contacts of a susceptible, since $Y(t)/N(t)$ is the infectious fraction of the population.

The incidence above assumes that the mixing patterns and contact rate are independent of population size. It also assumes that the population is uniform, homogeneous and randomly mixing. This random mixing corresponds approximately to the collision of "susceptible and infectious" molecules in an excited gas. Although the incidence expression above seems to be the standard one, other bilinear and nonlinear expressions have been used (Anderson and May 1979, Liu et al. 1987, Hethcote and van den Driessche 1991, Mena-Lorca and Hethcote 1992).

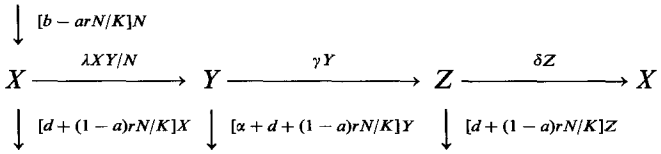
In the epidemiological models here the recovery rate is γY, the rate of loss of immunity is δZ and the disease-related death rate of infectives is αY. Linear terms in a differential equation correspond to residence times in a state with negative exponential distributions (Hethcote et al. 1981). Thus if the natural death rates were dY and dZ, then the mean (death-adjusted) infectious period would be $1/(\gamma + \alpha + d)$ and the mean (death-adjusted) period of temporary immunity would be $1/(\delta + d)$.

An SIRS model for a population satisfying the logistic equation with disease-related deaths is formulated and analyzed in Sects. 2–5. Three threshold quantities determine the appearance and stability of four equilibria. The results for this model are summarized in Table 1. Section 6 states results for an extension of the SIRS model which includes vertical transmission. The analogous SIS model is analyzed in Sect. 7. The discussion in Sect. 8 emphasizes how the epidemiological and demographic processes affect each other.

2 The SIRS model with density-dependent demographics

The epidemiological model formulated here has population dynamics corresponding to the logistic equation where the restricted growth is due to density-dependence in both the birth and death rates. The birth rate decreases and the death rate increases as the population size increases towards its carrying capacity. Vertical transmission from infected females to thier offspring before, during or just after birth is not included here, but is incorporated into the model in Sect. 6. Thus all newborns are susceptible. It is assumed that infection does not affect fertility so that the birth rates are the same for women in all three epidemiological states.

The transfer diagram for the numbers X, Y and Z of susceptibles, infectives and removed individuals, respectively, is



The parameters in the model are:

a = convex combination constant
b = natural birth rate constant
d = natural death rate constant
$r = b - d$ = growth rate constant
K = carrying capacity of the environment
λ = daily contact rate
α = disease-related death rate constant
γ = recovery rate constant
δ = loss of immunity rate constant.

We are following the convention of using Roman letters for demographic parameters and Greek letters for epidemiological parameters. We assume that b, d, K, λ, and γ are positive, a is in the interval [0, 1], and α and δ are nonnegative.

The autonomous differential equations corresponding to the transfer diagram are:

$$\begin{aligned} X'(t) &= [b - arN/K]N - \lambda XY/N - [d + (1-a)rN/K]X + \delta Z \\ Y'(t) &= \lambda XY/N - [\gamma + \alpha + d + (1-a)rN/K]Y \\ Z'(t) &= \gamma Y - [\delta + d + (1-a)rN/K]Z \\ N'(t) &= r[1 - N/K]N - \alpha Y \end{aligned} \tag{2.1}$$

where one of the equations is redundant since $N = X + Y + Z$. In the absence of disease the differential equation for N is the logistic equation (1.1); the second term in this differential equation for N corresponds to disease-related deaths. For $0 < a < 1$ the birth rate decreases and the death rate increases as N increases to its carrying capacity K; these are consistent with the limited resources associated with density-dependence. The birth rate is density independent when $a = 0$ and the death rate is density independent when $a = 1$.

First consider the case where $r > 0$ so the logistic equation (1.1) really does describe restricted growth. The birth rate in the model does not make sense if it is negative so we consider the positively invariant subset of the first octant in XYZ space where $N < bK/ar$. Since $N'(t) < 0$ for $N > K$, all solution paths in the subset above approach, enter or stay in the subset where $N = X + Y + Z \leqslant K$. Solution paths with $N_0 > K$ which do not enter the region $N \leqslant K$ in finite time must have their omega limit sets on the $N = K$ plane. Thus for $r > 0$ it suffices to analyze solution paths and omega limits sets in the subset of the first octant where $N \leqslant K$. If $r = 0$ and there is no disease, then the population size remains constant. Thus if $r = 0$, we consider (2.1) in the entire positively-invariant first octant in XYZ space. If $r < 0$ and there is no disease, then the population size decreases to zero if $N_0 < K$ and increases to infinity if $N_0 > K$. Here we consider only the subspace of the first octant where $N < K$ when $r < 0$.

It is convenient to reformulate the model (2.1) in terms of the fractions $I = Y/N$ and $R = Z/N$ of the population, which are infectious and removed, respectively. The susceptible fraction $S = X/N$ satisfies $S = 1 - I - R$. The last

three differential equations in (2.1) become

$$\begin{aligned} I'(t) &= [\lambda - (\gamma + \alpha + b) - (\lambda - \alpha)I - \lambda R + arN/K]I, \\ R'(t) &= \gamma I - (\delta + b)R + arNR/K + \alpha IR, \\ N'(t) &= [r(1 - N/K) - \alpha I]N. \end{aligned} \tag{2.2}$$

In IRN space the positively invariant subset corresponding to the subset $X + Y + Z \leqslant K$ in the first octant of XYZ space is

$$D = \{(I, R, N) \mid I \geqslant 0, R \geqslant 0, I + R \leqslant 1, 0 \leqslant N \leqslant K\}. \tag{2.3}$$

The continuity of the right side of (2.2) and its derivatives implies that unique solutions exist on maximal time intervals. Since for $r > 0$ solution paths approach, enter or stay in D, these paths are always bounded and continuable so they exist for all positive time (Hale 1969, pp. 18–27). Thus the initial value problem for system (2.2) is mathematically well-posed and is epidemiologically reasonable since the fractions I and R remain between 0 and 1. It is also well-posed in the first octant when $r = 0$ and in the subset where $N < K$ if $r < 0$.

3 Thresholds and the four equilibria

In the absence of disease the population size approaches its carrying capacity K if $r > 0$ and $N_0 > 0$. If disease is initially present, then the population size can go to zero, approach an equilibrium size below the carrying capacity or approach the carrying capacity, and the disease can die out or persist (remain endemic) depending on the values of several threshold quantities. The contact number is the average number of adequate contacts of an infective during the infectious period. In a totally susceptible population the contact number is the average number of new infections (secondary cases) produced per infective, so it is sometimes called the reproduction number (Anderson and May 1979). Here the contact rate is λ, and the average infectious period when the population is at its carrying capacity is $1/(\gamma + \alpha + d + (1 - a)r)$ where $r = b - d$ so that the contact number is

$$\sigma = \lambda/(\gamma + \alpha + b - ar). \tag{3.1}$$

A closely-related threshold quantity is the modified contact number θ given by

$$\theta = \lambda/(\gamma + \alpha + b). \tag{3.2}$$

Note that σ and θ coincide when $a = 0$. When $r > 0$ and the disease persists, the net growth threshold

$$\phi = \frac{r\lambda}{\alpha[\lambda - (\gamma + \alpha + d)]}\left(1 + \frac{\gamma}{\delta + d}\right) \tag{3.3}$$

determines whether the population size decreases to zero or approaches a constant size. The net growth threshold ϕ primarily reflects the relative effects of the disease-related death rate constant α and the growth rate constant r.

The system (2.2) can have up to four equilibria in the subregion D of IRN space defined by (2.3). They are found by setting the right sides of (2.2) equal to zero. System (2.2) always has the equilibria $P_1 = (0, 0, 0)$ and $P_2 = (0, 0, K)$ corresponding to fadeout of the disease with the population size at zero or at the carrying capacity K.

For $N = 0$ the right sides of (2.2) are equal to zero if

$$\begin{aligned} I &= [\lambda - (\gamma + \alpha + b) - \lambda R]/(\lambda - \alpha), \\ I &= (\delta + b)R/(\gamma + \alpha R). \end{aligned} \tag{3.4}$$

Analysis of the graphs of these equations shows that they intersect and yield a distinct equilibrium $P_3 = (I_3, R_3, 0)$ in D iff the modified contact number θ defined by (3.1) satisfies $\theta > 1$. Note that $P_3 \to P_1$ as $\theta \to 1^+$.

For nonzero I and N values at equilibrium, use $rN/K = r - \alpha I$ in the equilibrium equations for I and R to obtain

$$\begin{aligned} I &= [\lambda - (\gamma + \alpha + b - ar) - \lambda R]/(\lambda - (1 - a)\alpha), \\ I &= (\delta + b - ar)R/[\gamma + (1 - a)\alpha R]. \end{aligned} \tag{3.5}$$

The graphs of the equations in (3.5) intersect at a nonzero point in D iff the contact number σ satisfies $\sigma > 1$. The Eqs. (3.5) imply that R_4 is the positive root of the quadratic

$$\begin{aligned} &\lambda(1 - a)\alpha R^2 + [(1 - a)\alpha(\gamma + \alpha - \delta - \lambda) + \lambda(\gamma + \delta + b - ar)]R \\ &\quad + \gamma(\gamma + \alpha + b - ar - \lambda) = 0. \end{aligned} \tag{3.6}$$

To obtain an equation for N_4, we find equations for X_4, Y_4 and Z_4 in terms of N_4 and set their sum to N_4. In (2.1) the X_4 equation comes from the second equation, the Y_4 equation comes from the fourth equation and Z_4 in terms of Y_4 comes from the third equation. The resulting quadratic equation for $\chi = rN_4/K$ is

$$\begin{aligned} &(1 - a)[\lambda - (1 - a)\alpha]\chi^2 + [\lambda(\gamma + \delta + d) - (1 - a)\alpha(\gamma + \alpha + \delta + 2d) \\ &\quad + \lambda(1 - a)(\alpha - r)]\chi + \alpha(\lambda - \gamma - \alpha - d)(\delta + d) - \lambda r(\gamma + \delta + d) = 0. \end{aligned} \tag{3.7}$$

The Eq. (3.7) has a positive solution if the left side is negative at $\chi = 0$ which is equivalent to $\phi > 1$ where the net growth threshold ϕ is defined by (3.3). Thus $P_4 = (I_4, R_4, N_4)$ is a distinct equilibrium in D iff $\sigma > 1$ and $\phi > 1$. Note from Eq. (3.5) that $P_4 \to P_2$ as $\sigma \to 1^+$. Note from (3.7) and (3.6) that $P_4 \to P_3$ as $\phi \to 1^+$ for $\sigma > 1$. Note also that the equations in (3.4) are the same as those in (3.5) with $a = 0$. Thus R_3 in the equilibrium P_3 is the larger root of the quadratic equation (3.6) with $a = 0$.

In the special case when $a = 1$ so that the birth rate is density dependent and the death rate is density independent, the coordinates of the equilibrium P_4 are found explicitly as

$$\begin{aligned} I_4 &= (1 - 1/\sigma)(\delta + d)/(\gamma + \delta + d), \\ R_4 &= (1 - 1/\sigma)\gamma/(\gamma + \delta + d), \\ N_4 &= K(1 - 1/\phi), \\ \sigma &= \lambda/(\gamma + \alpha + d). \end{aligned} \tag{3.8}$$

4 Asymptotic behavior for the SIRS model

In proving the stability results summarized in Table 1, recall from Sect. 2 that it suffices to analyze the stability in region D. For $r > 0$ the equilibrium $P_1 = (0, 0, 0)$ is always a saddle whose unstable manifold includes the N axis. The

Table 1. Stability results* in IRN space for model (2.2) with $r > 0$

		$P_1 = (0, 0, 0)$	$P_2 = (0, 0, K)$	$P_3 = (I_3, R_3, 0)$	$P_4 = (I_4, R_4, N_4)$
$\theta \leqslant 1$	$\sigma \leqslant 1$	saddle	GAS[1]	NID[2]	NID[2]
	$\sigma > 1$	saddle	saddle	NID[2]	LAS[3]
$\theta > 1$	$\phi < 1$	saddle	saddle	LAS[3]	NID[2]
	$\phi = 1$	saddle	saddle	NAS[4]	NID[2]
	$\phi > 1$	saddle	saddle	saddle	LAS[3]

* All stability results hold globally if $a = 0$ or $\alpha = 0$

[1] GAS means globally asymptotically stable

[2] NID means not in D or not a distinct equilibrium in D

[3] LAS means locally asymptotically stable

[4] NAS means numerical calculations suggest asymptotic stability

equilibrium $P_2 = (0, 0, K)$ is locally asymptotically stable (LAS) if $\sigma < 1$ and is a saddle if $\sigma > 1$ with the unstable manifold in the I direction and the $I = 0$ plane as the stable manifold. The theorem below completes the results for the equilibria P_1 and P_2 in Table 1 by proving that P_2 is globally asymptotically stable (GAS) if $\sigma \leqslant 1$. Note that $\sigma \leqslant 1$ implies that $\theta \leqslant 1$.

Theorem 4.1 *If $\sigma \leqslant 1$, then the set D except the $N = 0$ face is an asymptotic stability region for $P_2 = (0, 0, K)$.*

Proof. The cases $\gamma + \alpha + b - ar \geqslant \lambda > \alpha$ and $\lambda \leqslant \alpha$ are treated separately. In the first case consider the Liapunov function $V = I$ with Liapunov derivative

$$V' = [\lambda - (\gamma + \alpha + b - ar)]I - ar(1 - N/K)I - (\lambda - \alpha)I^2 - \lambda IR \leqslant 0.$$

The Liapunov–Lasalle theorem (Hale 1969, p. 296–7) implies that all paths in D approach the largest positively invariant subset of the set E where $V' = 0$. Here the set E is the $I = 0$ face of D which is positively invariant. Although the R axis is in the stable manifold for the saddle P_1, all paths in E with $N_0 > 0$ approach the equilibrium P_2. Uniqueness and continuous dependence on the initial data imply that all paths in the set D except the $N = 0$ face must approach P_2. In the second case when $\lambda \leqslant \alpha$, consider the Liapunov function $V = I + R$ with Liapunov derivative

$$V' = -(I + R)[b(1 - aN/K) + adN/K] - (\alpha - \lambda)I(1 - I - R) - \delta R \leqslant 0.$$

Here the set E is the N axis where $I = 0$ and $R = 0$. All paths on the N axis with $N_0 > 0$ approach P_2 so that all paths in the set D except the $N = 0$ face must also approach P_2. □

Now consider the equilibrium $P_3 = (I_3, R_3, 0)$ where I_3 and R_3 satisfy (3.4). Recall that P_3 is a distinct equilibrium in D if $\theta > 1$. The local stability of P_3 is given in the theorem below; numerical calculations suggest that P_3 is always asymptotically stable when $\phi = 1$.

Theorem 4.2 *For $\theta > 1$ the equilibrium $P_3 = (I_3, R_3, 0)$ in D is locally asymptotically stable* (LAS) *if $\phi < 1$, and for $\phi > 1$ it is a saddle with the $N = 0$ face as the two-dimensional stable manifold and a one-dimensional unstable manifold whose tangent at P_3 is in the N direction.*

Proof. The Jacobian of system (2.2) at P_3 is

$$J(P_3) = \begin{bmatrix} -(\lambda-\alpha)I_3 & -\lambda I_3 & arI_3/K \\ \gamma+\alpha R_3 & -\delta-b+\alpha I_3 & arR_3/K \\ 0 & 0 & r-\alpha I_3 \end{bmatrix}. \tag{4.1}$$

Two eigenvalues of $J(P_3)$ are roots of the quadratic equation

$$\begin{aligned} x^2+px+q&=0, \\ p&=(\lambda-\alpha)I_3+\delta+b-\alpha I_3, \\ q&=(\lambda-\alpha)I_3[\delta+b-\alpha I_3]+\lambda I_3(\gamma+\alpha R_3). \end{aligned} \tag{4.2}$$

Now $\theta>1$ implies that $\lambda>\alpha$ and (3.4) implies that $(\delta+b-\alpha I_3)R_3=\gamma I_3$ so that $\delta+b-\alpha I_3>0$. Thus p and q are both positive so that both roots of (4.2) have negative real parts.

Thus the stability of P_3 depends on the sign of the third eigenvalue $r-\alpha I_3$. We now prove that this eigenvalue has the same sign as $\phi-1$. The first equation in (3.4) implies that

$$r-\alpha I_3=(\lambda r+\alpha(\gamma+\alpha+d-\lambda)+\lambda\alpha R_3)/(\lambda-\alpha), \tag{4.3}$$

where $\lambda>\alpha$ since $\theta>1$. Suppose $r-\alpha I_3>0$, then $r>\alpha I_3$ and $\delta+b-\alpha I_3>\delta+d$ so that

$$1>\frac{\alpha I_3(\delta+d)}{r(\delta+b-\alpha I_3)}=\frac{\alpha R_3(\delta+d)}{\gamma r}. \tag{4.4}$$

Using $\phi<1$ and the estimate (4.4) in (4.3) implies that $r-\alpha I_3<0$ which is a contradiction. Thus $\phi<1$ implies that $r-\alpha I_3<0$ so that all eigenvalues have negative real parts and P_3 is LAS. Similarly, $\phi>1$ implies that $r-\alpha I_3>0$ and P_3 is a saddle whose unstable manifold has a tangent at P_3 in the N direction. □

Recall that $P_4=(I_4, R_4, N_4)$ is a distinct equilibrium in D iff $\sigma>1$ and $\phi>1$. Since $r=b-d>0$, then $\sigma>\theta$ so P_4 exists (see Table 1) for $\theta>1$ and $\phi>1$. Also $\theta\leqslant 1$ and $\sigma>1$ imply that $\phi>1$ because ϕ can be changed to

$$\phi=\frac{\lambda}{\alpha}\frac{b-d}{\lambda-(\gamma+\alpha+d)}\frac{\gamma+\delta+d}{\delta+d}$$

where each factor has one as a lower bound. Thus P_4 also exists (see Table 1) for $\theta\leqslant 1$ and $\sigma>1$.

Theorem 4.3 *The equilibrium $P_4=(I_4, R_4, N_4)$ is locally asymptotically stable* (LAS) *if $\sigma>1$ and $\phi>1$.*

Proof. The Jacobian of system (2.2) at P_4 is

$$J(P_4) = \begin{bmatrix} -(\lambda-\alpha)I_4 & -\lambda I_4 & arI_4/K \\ \gamma+\alpha R_4 & -\gamma I_4/R_4 & arR_4/K \\ -\alpha N_4 & 0 & -rN_4/K \end{bmatrix}.$$

The cubic characteristic equation is

$$x^3 + c_2x^2 + c_1x + c_0 = 0,$$

$$c_2 = (\lambda - \alpha)I_4 + \gamma I_4/R_4 + rN_4/K > 0,$$

$$c_1 = \gamma(\lambda - \alpha)I_4^2/R_4 + (\lambda - \alpha)I_4 rN_4/K + \gamma r I_4 N_4/KR_4 + \alpha a r I_4 N_4/K + \lambda I_4(\gamma + \alpha R_4) > 0,$$

$$c_0 = (\lambda - \alpha)r\gamma I_4^2 N_4/KR_4 + \lambda I_4(\gamma + \alpha R_4)rN_4/K + a\alpha r\gamma I_4^2 N_4/KR_4 - \lambda\alpha r a I_4 R_4 N_4/K > 0.$$

Thus all roots have negative real parts iff $c_2c_1 - c_0 > 0$ by the Routh–Hurwitz criteria (Miller and Michel 1982). Here all terms of c_2 and c_1 are positive and the three negative terms in $-c_0$ are cancelled out by term 1 in c_2 times term 3 in c_1, term 3 in c_2 times term 5 in c_1 and term 2 in c_2 times term 4 in c_1. Thus $c_2c_1 - c_0$ is positive. □

If $\alpha = 0$ and $N_0 > 0$, then $N \to K$ and the system (2.2) reduces in the K plane to a basic system in IR space (Hethcote 1976, 1989) with $(I, R) \to (0, 0)$ if $\sigma \leqslant 1$ and $(I, R) \to (I_e, R_e)$ if $I_0 > 0$ and $\sigma > 1$. Note that the Liapunov–Lasalle theorem with $V = (K - N)^2/2$ implies that these stability results in the $N = K$ plane are global in D. Thus if the disease-related death rate constant α is zero, then $\phi = \infty$, $N_4 = K$ and the LAS results for P_4 in Table 1 are GAS results for P_4 in D minus the $N = 0$ and $I = 0$ faces.

The special cases when $r < 0$ and $r = 0$

For $r < 0$ we only consider solution paths with $N_0 < K$. In this case solution paths of model (2.2) satisfy $N(t) \to 0$ as $t \to \infty$. Moreover all paths approach $P_1 = (0, 0, 0)$ if $\theta \leqslant 1$ and approach $P_3 = (I_3, R_3, 0)$ if $\theta > 1$. To prove this use the Liapunov function $V = N$ so the Liapunov derivative satisfies $V' \leqslant 0$ with equality only if $N = 0$. Then all solutions with $N_0 < K$ approach the largest positively invariant subset in the $N = 0$ plane by the Liapunov–Lasalle theory (Hale 1969, p. 296–7). The result above follows since if $N = 0$, then the system (2.2) for I and R reduces to a system where all solutions approach $(0, 0)$ if $\theta \leqslant 1$ and (I_3, R_3) if $\theta > 1$ and $I_0 > 0$ (Mena-Lorca and Hethcote 1992, Sect. 4).

If the growth rate constant $r = b - d$ is zero, then the nonlinear terms involving K disappear so the model (2.2) reduces to a model with balanced exponential births and deaths (Mena-Lorca and Hethcote 1992, Sect. 4). If the disease-related death rate constant α is zero, the population size remains at N_0. For $\alpha = 0$ and $\theta = \sigma \leqslant 1$, the line $(0, 0, N_0)$ of equilibria in IRN space is neutrally stable. For $\alpha = 0$ and $\theta = \sigma > 1$, the line $(0, 0, N_0)$ is neutrally unstable and the line (I_e, R_e, N_e) is neutrally stable. When the disease-related death rate constant α is positive and $\theta = \sigma < 1$, the line $(0, 0, N_e)$ of equilibria is neutrally stable. For $\alpha > 0$ and $\theta = \sigma = 1$, numerical calculations suggest that the equilibrium $P_1 = (0, 0, 0)$ is globally asymptotically stable (GAS). For $\alpha > 0$ and $\theta = \sigma > 1$, the equilibrium $P_3 = (I_3, R_3, 0)$ is GAS. In these last two cases the population size would remain constant if there were no disease, but the presence of the disease and disease-related deaths cause the population size to go to zero.

5 The SIRS model with density-dependent deaths

In the special case of the SIRS model in Sect. 2 where $a = 0$, the birth rate is density independent and the death rate is density dependent. In this case the local stability results for P_3 and P_4 in Table 1 can be proved globally. This is possible because the differential equations for I and R in system (2.2) do not involve N so they are completely uncoupled from the population size N. The contact number σ is the same as the modified contact number θ if $a = 0$ so the case $\theta \leqslant 1$ and $\sigma > 1$ in Table 1 does not occur. Thus if $a = 0$, the modified contact number above determines whether the disease dies out ($\theta \leqslant 1$) or remains endemic ($\theta > 1$).

The differential equations for I and R in (2.2) with $a = 0$ also occurred in models analyzed by Busenberg and van den Driessche (1990) and by Mena-Lorca and Hethcote (1992, Sect. 4). Let T be the triangle in the first quadrant of IR space where $I + R \leqslant 1$. If the modified contact number θ satisfies $\theta \leqslant 1$, then $(0, 0)$ is the only equilibrium in T and is globally asymptotically stable (GAS). If $\theta > 1$, then $(0, 0)$ is a saddle and the equilibrium (I_e, R_e) is GAS in T except along the $I = 0$ edge where paths go to $(0, 0)$. Since the Eqs. (3.5) for I_4 and R_4 coincide with those in (3.4) for I_3 and R_3 when $a = 0$, we use I_e and R_e here for these common equilibrium values. In order to analyze the asymptotic behavior of $N(t)$ in D, we need the following lemma on the perturbed logistic equation.

Lemma 5.1 *Consider* $N'(t) = g(t)N - cN^2$ *where* c *is a positive constant,* $g(t) \in C[0, \infty)$ *and the limit* g_∞ *of* $g(t)$ *exists as* t *approaches infinity. If* $g_\infty < 0$, *then all solutions with* $N_0 \geqslant 0$ *approach* 0 *as* $t \to \infty$. *If* $g_\infty > 0$, *then all solutions with* $N_0 > 0$ *approach* g_∞/c *as* $t \to \infty$.

Proof. The change $u = 1/N$ leads to a linear differential equation $u' = -g(t)u + c$ whose solution is

$$u(t) = \left[u_0 + c \int_0^t \exp\left(\int_0^\tau g(v)\, dv \right) d\tau \right] \Big/ \exp\left(\int_0^t g(v)\, dv \right). \tag{5.1}$$

If $g_\infty < 0$, then there is a t_0 such that $g(t) \leqslant g_\infty/2$ for $t \geqslant t_0$ and

$$0 \leqslant \exp\left(\int_0^t g(v)\, dv \right) \leqslant \exp\left(\int_0^{t_0} g(v)\, dv + g_\infty (t - t_0)/2 \right) \to 0$$

as $t \to \infty$. Since the numerator in (5.1) has a finite positive limit and the denominator goes to zero, $u(t) \to \infty$ and $N(t) \to 0$ as $t \to \infty$. If $g_\infty > 0$, then there is a $t_0 > 0$ such that $g(t) \geqslant g_\infty/2$ for $t \geqslant t_0$ and

$$\exp\left(\int_0^t g(v)\, dv \right) \geqslant \exp\left(\int_0^{t_0} g(v)\, dv + g_\infty (t - t_0)/2 \right) \to \infty$$

as $t \to \infty$. Since the numerator and denominator in (5.1) both approach infinity, we apply L'Hôpital's rule to obtain $u(t) \to c/g_\infty$ and $N(t) \to g_\infty/c$ as $t \to \infty$. □

Lemma 5.1 is now used to prove the global asymptotic stability (GAS) for the equilibria P_2, P_3 and P_4. Lemma 5.1 does not apply when $\theta > 1$ and $\phi = 1$, but numerical calculations in this case suggest that D except the $I = 0$ face is an asymptotic stability region for P_3.

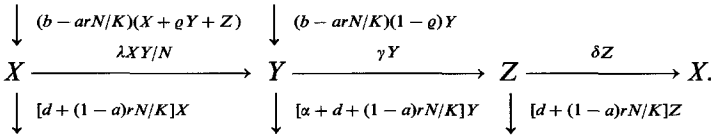
Theorem 5.2 *Consider the model* (2.2) *with* $a = 0$. *If* $\theta \leqslant 1$, *then* D *except the* $N = 0$ *face is an asymptotic stability region for* P_2. *If* $\theta > 1$ *and* $\phi < 1$, *then* D *except the* $I = 0$ *face is an asymptotic stability region for* P_3. *If* $\theta > 1$ *and* $\phi > 1$, *then* D *except the* $I = 0$ *and* $N = 0$ *faces is an asymptotic stability region for* P_4.

Proof. When $\theta \leqslant 1$ so that $I(t) \to 0$, Lemma 5.1 applies with $g_\infty = r > 0$ and $c = r/K$ so that $N(t) \to K$. Thus $P_2 = (0, 0, K)$ is GAS in D minus the $N = 0$ face. For $\theta > 1$ and $\phi < 1$ so that $I(t) \to I_e$ if $I_0 > 0$, we use $g_\infty = r - \alpha I_e < 0$ (see Sect. 4) and $c = r/K$ so that $N(t) \to 0$. Hence $P_3 = (I_e, R_e, 0)$ is GAS in D minus the $I = 0$ face. For $\theta > 1$ and $\phi > 1$ so that $I(t) \to I_e$ if $I_0 > 0$, we use $g_\infty = r - \alpha I_e > 0$ and $c = r/K$ so that $N(t) \to K(1 - \alpha I_e/r) = N_4$ as $t \to \infty$. Thus $P_4 = (I_e, R_e, N_4)$ is GAS in D minus the $I = 0$ and $N = 0$ faces. □

6 Modification of the SIRS model to include vertical transmission

For some diseases there may be vertical transmission of infection from some infected mothers to their offspring before, during or just after birth. However, vertical transmission is usually less important than horizontal transmission since infectiousness and parturition usually occur at different times. In Sects. 2–5 we focussed on the interactive effects of horizontal disease transmission and density-dependent demographics without complications such as vertical transmission. However, a model with vertical transmission can be easily formulated and analyzed.

Let ϱ be the fraction of newborns who are not infected by an infectious mother so $1 - \varrho$ is the fraction who are infected by vertical transmission. The transfer diagram in Sect. 2 is now modified to



The differential equations corresponding to this transfer diagram are similar to those in Sect. 2. The contact number σ, modified contact number θ and growth threshold ϕ are

$$\sigma = [\lambda + (1 - \varrho)(b - ar)]/(\gamma + \alpha + b - ar), \tag{6.1}$$

$$\theta = [\lambda + (1 - \varrho)b]/(\gamma + \alpha + b), \tag{6.2}$$

$$\phi = \frac{r\lambda}{\alpha[\lambda + (1 - \varrho)b - (\gamma + \alpha + d)]}\left(1 + \frac{\gamma}{\delta + d}\right). \tag{6.3}$$

As in Sect. 2, these expressions have epidemiological intepretations. For example, the mean infectious period at the carrying capacity K is $1/(\gamma + \alpha + b - ar)$, the horizontal transmission contact rate is λ and the vertical transmission rate constant when $N = K$ is $(1 - \varrho)(b - ar)$, so that the contact number σ is the average number of horizontal and vertical adequate contacts of an infective during the infectious period.

The details of the analysis of this SIRS model are not given since the essential ideas and proofs can be seen for the simpler model with only horizontal transmission. Indeed, every threshold, equilibrium, equation and theorem in Sects. 2–5 can be proved with obvious modifications for this SIRS model with

both horizontal and vertical transmission, but the algebraic expressions are more complicated. Thus all of the results in Table 1 hold for this new model with σ, θ and ϕ given above.

7 The SIS model with density-dependent demographics

Infection with some diseases does not confer any immunity so that infectives become susceptible again upon recovery. Using the notation and definitions in previous sections, the transmission diagram for an SIS disease with both horizontal and vertical transmission is

$$\begin{array}{ccccc} \downarrow{\scriptstyle [b-arN/K](X+\varrho Y)} & & \downarrow{\scriptstyle (b-arN/K)(1-\varrho)Y} & & \\ X & \xrightarrow{\lambda XY/N} & Y & \xrightarrow{\gamma Y} & X. \\ \downarrow{\scriptstyle [d+(1-a)rN/K]X} & & \downarrow{\scriptstyle [\alpha+d+(1-a)rN/K]Y} & & \end{array}$$

The differential equations analogous to (2.7) for the fraction $I = Y/N$ and the population size N are

$$\begin{aligned} I'(t) &= [\lambda - (\gamma + \alpha) - (\lambda - \alpha)I - \varrho(b - arN/K)]I, \\ N'(t) &= [r(1 - N/K) - \alpha I]N. \end{aligned} \tag{7.1}$$

As in Sect. 2 it can be shown that this model is well-posed and that it suffices to consider solutions in

$$D = \{(I, N) \,|\, 0 \leqslant I \leqslant 1, 0 \leqslant N \leqslant K\}. \tag{7.2}$$

The contact number σ and modified contact number θ are still given by (6.1) and (6.2), but the net growth threshold ϕ is now given by

$$\phi = \frac{r\lambda}{\alpha[\lambda + (1 - \varrho)b - (\gamma + \alpha + d)]}. \tag{7.3}$$

Note that this is the limit of ϕ in (6.3) as $\delta \to \infty$, corresponding roughly to instantaneous movement through the removed class. This SIS model as the intuitive limit of an SIRS model is consistent with the model reduction in Table 1 in Liu et al. (1987).

The analysis of model (7.1) is quite similar to that of the SIRS model except that there is no removed class. The four equilibria in IN space are $P_1 = (0, 0)$, $P_2 = (0, K)$, $P_3 = (I_3, 0)$ where

$$I_3 = (\gamma + \alpha + b)(\theta - 1)/(\lambda - \alpha), \tag{7.4}$$

and $P_4 = (I_4, N_4)$ where

$$\begin{aligned} I_4 &= (\gamma + \alpha + b - ar)(\sigma - 1)/[\lambda - \alpha(1 - \varrho a)], \\ N_4 &= K(1 - \alpha I_4/r). \end{aligned} \tag{7.5}$$

As in Sect. 3, P_3 is a distinct equilibrium in D iff $\theta > 1$ and P_4 is a distinct equilibrium in D iff $\sigma > 1$ and $\phi > 1$.

For the two dimensional system (7.1) with $r > 0$, it is possible to prove all stability results globally.

Theorem 7.1 *If $\sigma \leqslant 1$, then the set D except the $N = 0$ side is an asymptotic stability region for $P_2 = (0, K)$.*

Proof. This proof is similar to that of Theorem 4.1. Using the Liapunov function $V = I$, the Liapunov derivative is

$$\begin{aligned} V' &= [\lambda - (\gamma + \alpha + \varrho(b - ar))]I - \varrho ar(1 - N/K)I - (\lambda - \alpha)I^2 \\ &= -(\alpha - \lambda)I(1 - I) - \varrho I[b(1 - aN/K) + adN/K]. \end{aligned} \tag{7.6}$$

The first expression for V' is nonpositive if $\lambda > \alpha$ and the second is nonpositive if $\lambda \leqslant \alpha$. The largest positively invariant subset where $V' = 0$ is the $I = 0$ face and paths in this face with $N_0 > 0$ approach P_2. Thus paths in D with $N_0 > 0$ also approach the equilibrium P_2. □

Theorem 7.2 *For $\theta > 1$ and $\phi < 1$, D except the $I = 0$ face is an asymptotic stability region for $P_3 = (I_3, 0)$. For $\theta > 1$ and $\phi > 1$, D except the $I = 0$ and $N = 0$ faces is an asymptotic stability region for $P_4 = (I_4, N_4)$.*

Proof. Using $\lambda > \alpha$ which follows from $\theta > 1$, it can be shown that there are no limit cycles or cycle graphs in D by Dulac's test (Jordan and Smith 1987) with multiplying factor $(IN)^{-1}$. For $\theta > 1$, $\phi \leqslant 1$ and $I_0 > 0$, paths cannot approach the repeller P_1 or saddle P_2, so that they must approach the locally asymptotically stable equilibrium P_3 by the Poincaré–Bendixson theory. For $\theta > 1$, $\phi > 1$, $I_0 > 0$ and $N_0 > 0$, paths in D except the $I = 0$ and $N = 0$ faces cannot approach the repeller P_1, or the saddles P_2 and P_3 so they must approach the locally asymptotically stable equilibrium P_4. □

Note that $\theta \leqslant 1$ and $\sigma > 1$ imply that $\phi > 1$. Thus all of the stability results in Table 1 hold globally for the SIS model.

8 Discussion

The epidemiological and demographic processes in a dynamic model affect each other. Mena-Lorca and Hethcote (1992) studied epidemiological models with two demographic processes: immigration with the death rate proportional to the population size, and both the birth and death rates proportional to the population size. Using the same notation and terminology, this paper considers infectious disease models with density-dependent restricted growth corresponding to the logistic equation. The models here are really an entire range of models with $0 \leqslant a \leqslant 1$ where, roughly speaking, the fraction a of the density-dependence is allocated to reducing the birth rate and the fraction $1 - a$ is allocated to increasing the death rate. We find that the demographic and epidemiological aspects of the models affect each other by altering the expected asymptotic behaviors.

The population size in a logistic demographic model usually approaches the carrying capacity K, but in all except the first case in Table 1, the persistence of the disease and the disease-related deaths either lower the asymptotic population size or cause the population to approach extinction. The competing effects of the positive growth rate constant r and the disease-related death rate constant α are measured by the net growth threshold ϕ given by (3.3), (6.3) or (7.3). If $\theta > 1$ and $\phi \leqslant 1$, then the disease-related deaths overcome the intrinsic growth rate corresponding to positive r and cause the population size to decrease to zero. In

this case the equilibrium $P_3 = (I_3, R_3, 0)$ corresponds to population extinction due to disease-related deaths and the persistence of the disease. If $\phi > 1$, the intrinsic growth rate corresponding to r is large enough so that the population persists, but the disease-related deaths cause the asymptotic population size N_4 to be lower than the carrying capacity K. Hence the equilibrium $P_4 = (I_4, R_4, 0)$ corresponds to a decreased equilibrium population size due to disease-related deaths and the persistence of the disease. In the two cases above, the infectious disease dynamics clearly affect the population size dynamics.

The population size dynamics also affect the infectious disease dynamics. In the model of Mena-Lorca and Hethcote (1992, Sect. 4) with exponential growth demographics, the modified contact number θ always determines whether the disease dies out ($\theta \leqslant 1$) or remains endemic ($\theta > 1$). This is generally true for the models here with logistic growth demographics, but for $\sigma > 1$ in the models in Sects. 2, 6 and 7, the disease remains endemic even though the modified contact number θ satisfies $\theta \leqslant 1$. In this case the logistic demographic structure causes the disease to persist even though it would normally have died out with the exponential growth demographic structure. Thus the demographic aspects affect the epidemiological aspects. Note that the epidemiological thresholds are also affected by the logistic demographics since the contact number σ, the modified contact number θ and the net growth threshold ϕ given by (3.1) to (3.3) or by (6.1) to (6.3) involve the demographic parameters a, b, d and r.

The net growth threshold ϕ plays a similar role here and in the SIRS model with exponential growth demographics (Mena-Lorca and Hethcote 1992, Sect. 4). In these models with $r > 0$ and $\theta > 1$, the population size would normally approach the carrying capacity or grow exponentially, but in both models the disease-related deaths cause the population to become extinct if $\phi < 1$. For $\phi > 1$ the disease-related deaths cause the population size to grow at a slower rate in the exponential demographic model and to approach a size below the carrying capacity in the logistic demographic models here.

Based on numerical calculations with a variety of parameter sets and our global stability results in the special cases when $a = 0$ or $\alpha = 0$, we conjecture that for the SIRS model the local stability results in Table 1 are actually global. Proofs of these global stability results would be of some interest, but our results here have already revealed the main concepts about density-dependent demographics in infectious disease models. The SIRS model reduces to an SIR model when the immunity loss rate constant δ is zero and both the SIRS and SIS models reduce to an SI model when the recovery rate constant γ is zero. All of the results in Table 1 hold for these special cases provided b is positive so that there is some inflow into the susceptible class.

Brauer (1989, 1990) has used density-dependent demographics described by $N'(t) = [B(N) - D(N)]N$ in disease models where $B(N) - D(N)$ has properties which make the solutions behave like solutions of the logistic equation. The results in this paper could have been obtained using this more general formulation, but we have chosen to consider the logistic equation since our results for it are more complete, concise and understandable. The analysis of SEIRS models with logistic demographics has not been done here because the details are very complicated. Analyzing these SEIRS models is probably not a good open problem since the results are predictable from those for the SIRS model and the details would be too complicated to be interesting.

Acknowledgements. We thank Jinshi Zhou for suggestions and an improved proof of Lemma 5.1.

References

Anderson, R. M., Jackson, H. C., May, R. M., Smith, A. D. M.: Population dynamics of fox rabies in Europe. Nature **289**, 765–777 (1981)

Anderson, R. M., May, R. M.: Regulation and stability of host-parasite interactions. J. Anim. Ecol. **47**, 219–247 (1978)

Anderson, R. M., May, R. M.: Population biology of infectious diseases I. Nature **180**, 361–367 (1979)

Anderson, R. M., May, R. M., McLean, A. R.: Possible demographic consequences of AIDS in developing countries. Nature **332**, 228–234 (1988)

Brauer, F.: Epidemic models in populations of varying size. In: Castillo-Chavez, C., Levin, S. A., Shoemaker, C. (eds.) Mathematical approaches to ecological and environmental problem solving, pp. 109–123. Berlin Heidelberg New York: Springer 1989

Brauer, F.: Models for the spread of universally fatal diseases. J. Math. Biol. **28**, 451–462 (1990)

Bremermann, H. J., Thieme, H. R.: A competitive exclusion principle for pathogen virulence. J. Math. Biol. **27**, 179–190 (1989)

Busenberg, S. N., 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. N., Hadeler, K. P.: Demography and epidemics. Math. Biosci. **101**, 41–62 (1990)

Castillo-Chavez, C. C., Cooke, K. L., Huang, L., Levin, S. A.: On the role of long periods of infectiousness in the dynamics of AIDS, Part 1, Single population models. J. Math. Biol. **27**, 373–398 (1989)

Edelstein-Keshet, L.: Mathematical models in biology. New York: Random House 1988

Hale, J. K.: Ordinary differential equations. New York: Wiley-Interscience 1969

Hethcote, H. W.: Qualitative analysis for communicable disease models. Math. Biosci. **28**, 335–356 (1976)

Hethcote, H. W.: Three basic epidemiological models. In: Gross, L., Hallam, T. G., Levin, S. A. (eds.) Applied mathematical ecology, pp. 119–144. Berlin Heidelberg New York: Springer 1989

Hethcote, H. W., Levin, S. A.: Periodicity in epidemiological models. In: Gross, L., Hallam, T. G. Levin, S. A. (eds.) Applied mathematical ecology, pp. 193–211. Berlin Heidelberg New York: Springer 1989

Hethcote, H. W., Stech, H. W., van den Driessche, P.: Periodicity and stability in epidemic models: a survey. In: Busenberg, S. N., Cooke, K. L. (eds.) Differential equations and applications in ecology, epidemic and populations problems, pp. 65–82. New York: Academic Press 1981

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

Hyman, J. M., Stanley, E. A.: Using mathematical models to understand the AIDS epidemic. Math. Biosci. **90**, 415–473 (1988)

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

Jordan, D. W., Smith, P.: Nonlinear ordinary differential equations. Oxford: Clarendon Press 1987

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)

May, R. M., Anderson, R. M.: Regulation and stability of host-parasite population interactions. II. Destabilizing processes. J. Anim. Ecol. **47**, 248–267 (1978)

May, R. M., Anderson, R. M.: Population biology of infectious diseases II. Nature **280**, 455–461 (1979)

May, R. M., Anderson, R. M., McLean, A. R.: Possible demographic consequences of HIV/AIDS epidemics. Math. Biosci. **90**, 475–505 (1988)

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

Miller, R. K., Michel, A. N.: Ordinary differential equations. New York: Academic Press 1982

Pugliese, A.: Population models for diseases with no recovery. J. Math. Biol. **28**, 65–82 (1990)