MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **3**, Number **1**, January **2006**

http://www.mbejournal.org/

pp. **1–15**

SOME SIMPLE EPIDEMIC MODELS

Fred Brauer

Department of Mathematics
University of British Columbia
Vancouver, BC V6T 1Z2, Canada

Abstract. The SARS epidemic of 2002–3 led to the study of epidemic models including management measures and other generalizations of the original 1927 epidemic model of Kermack and McKendrick. We consider some natural extensions of the Kermack-McKendrick model and show that they share the main properties of the original model.

In honor of Professor Zhien Ma's 70th birthday

1. **Introduction.** Almost since the beginning of recorded history there have been epidemics. An epidemic may be described as a sudden outbreak of a disease that infects a substantial portion of the population in a region before it disappears. In the nineteenth century, recurrent waves of cholera killed millions in India. The influenza epidemic of 1918–1919 killed at least 20 million people overall, more than half a million in the United States. One of the questions that first attracted the attention of scientists interested in the study of the spread of communicable diseases was why diseases would suddenly develop in a community and then disappear just as suddenly without infecting the entire community.

One of the early triumphs of mathematical epidemiology was the formulation of a simple model that predicted just such behavior. Kermack and McKendrick [18] formulated a model with the population under study being divided into compartments, namely a susceptible class S, an infective class I, and a removed class R. A special case of their model has become known as the Kermack-McKendrick model. In this model R is determined once S and I are known, and thus we can drop the R equation from the model, leaving the system of two equations:

$$\begin{aligned} S' &= -\beta SI \\ I' &= (\beta S - \alpha)I. \end{aligned} \tag{1}$$

The system (1) has been used successfully to fit data from many epidemics, and it is very easily analyzed qualitatively. This analysis has been described in many references, including [6].

The model is based on the following assumptions:

i. An average member of the population makes contact sufficient to transmit infection with βN others per unit time, where N represents total population size.
ii. A fraction α of infectives leave the infective class per unit time.

2000 *Mathematics Subject Classification.* 92D30.
Key words and phrases. epidemic models, Kermack-McKendrick, vaccination.

iii. There is no entry into or departure from the population, except possibly through death from the disease.

Since the fraction of contacts made by an infective that are with a susceptible and may therefore produce a new infective is S/N, the rate of new infections is

$$\beta N \frac{S}{N} I = \beta SI.$$

If there are disease deaths, N is a function of t. Let us think of a population of size K into which a small number of infectives is introduced, so that $S(0) \approx K$, $I(0) \approx 0$, $N(0) = K$.

The model (1) has the following two basic properties shared by more complicated epidemic models.

I. There is a *basic reproduction number* $R_0 = \beta K/\alpha$ such that if $R_0 < 1$ there is no epidemic, while if $R_0 > 1$ there is an epidemic.

II. An epidemic eventually dies out, leaving part of the population untouched.

By an epidemic we mean a situation in which I first increases to a maximum and then decreases to zero. The other possibility is that I decreases monotonically to zero (no epidemic). The definition of the basic reproduction number R_0 is that the basic reproduction number is the number of secondary infections caused by a single infective introduced into a wholly susceptible population of size K over the course of the infection of this single infective. In this situation, an infective makes βK contacts in unit time, all with susceptibles and producing new infections, and the mean infective period is $1/\alpha$; thus, the basic reproduction number is $\beta K/\alpha$.

To establish the first property, we find orbits of the system (1) in the (S, I) plane by integrating

$$\frac{dI}{dS} = \frac{I'}{S'} = \frac{(\beta S - \alpha)I}{-\beta SI} = -1 + \frac{\alpha}{\beta S}.$$

Every orbit is a curve in the (S, I) plane, terminating at a point $(S_\infty, 0)$, with $S_\infty > 0$ and satisfying

$$K - \frac{\alpha}{\beta} \log S_0 = S_\infty - \frac{\alpha}{\beta} \log S_\infty, \tag{2}$$

the *final size equation*. From this we deduce that $\lim_{t\to\infty} S(t) = S_\infty > 0$. To show that $I(t) \to 0$, we observe that $S + I$ is a non-negative monotone nonincreasing function, which has a limit as $t \to \infty$. Since $(S + I)$ is a smooth function, its derivative must approach zero, and this shows that $\lim_{t\to\infty} I(t) = 0$.

Initially, the number of infectives grows exponentially if and only if $R_0 > 1$, because the equation for I may be approximated for $t \to 0$ by

$$I' = (\beta K - \alpha)I,$$

and the initial exponential growth rate is

$$r = \beta K - \alpha = \alpha(R_0 - 1).$$

This initial growth rate r may be determined experimentally when an epidemic begins. Then, since K and α may be measured, β may be calculated as

$$\beta = \frac{r + \alpha}{K}$$

However, because of incomplete data and underreporting of cases, this estimate may not be very accurate. This inaccuracy is even more pronounced for an outbreak of a previously unknown disease, where early cases are likely to be misdiagnosed.

Until the SARS epidemic of 2002–3, most of the work in mathematical epidemiology concentrated on studies of specific diseases or on the interplay between epidemiological and demographic effects. There were studies of epidemic models, including [4, 5, 6, 13, 14, 15, 16, 22, 24], which concentrated on analysis of the course of an epidemic but did not examine the effects of control measures. Studies of the SARS epidemic, such as the dynamic models in [2, 9], began with the original so-called Kermack-McKendrick model (1) but included control measures to make the models more realistic. The purpose of this work is to show that a large variety of general epidemic models, including those which have come out of the SARS epidemic, have the basic properties I and II.

2. **General contact rates.** The assumption in the model (1) of a rate of contacts per infective proportional to population size N, called *mass action incidence*, was used in all the early epidemic models. It is more realistic to assume a contact rate that is a nonincreasing function of total population size. For example, a situation in which the number of contacts per infective in unit time is constant, called *standard incidence*, is probably a more accurate description for sexually transmitted diseases.

We generalize the model (1) by replacing assumption (i) with the assumption that an average infective makes $C(N)$ contacts in unit time with $C'(N) \geq 0$, [1, 7], and we define

$$\beta(N) = \frac{C(N)}{N}.$$

It is reasonable to assume $\beta'(N) \leq 0$ to express the idea of saturation in the number of contacts. Then bilinear incidence corresponds to the choice $C(N) = \beta N, \beta(N) = \beta$, and standard incidence corresponds to the choice $C(N) = \lambda, \beta(N) = \lambda/N$. The assumptions $C(N) = N\beta(N), C'(N) \geq 0$ imply that

$$\beta(N) + N\beta'(N) \geq 0. \tag{3}$$

Some disease transmission models have used a Michaelis-Menten type of interaction of the form

$$C(N) = \frac{aN}{1 + bN}$$

[7]. Another form based on a mechanistic derivation for pair formation [12] leads to an expression of the form

$$C(N) = \frac{aN}{1 + bN + \sqrt{1 + 2bN}}.$$

Data for diseases transmitted by contact in cities of moderate size [21] suggests a good fit is obtained using the form

$$C(N) = \lambda N^a,$$

with $a = 0.05$. All of these forms satisfy the conditions $C'(N) \geq 0, \beta'(N) \leq 0$. Additional remarks on contact modeling may be found in [19]. If there are no disease deaths, so that the total population size is constant, all incidence assumptions involving total population size are equivalent, but if there are disease deaths the behavior of a model might depend on the form of the incidence.

An epidemic model in which the incidence is assumed to depend on total population size must include an equation for total population size. This forces us to distinguish between members of the population who die of the disease and members of the population who recover with immunity against reinfection. Such a distinction was not needed in the original Kermack-McKendrick model, even in its general

form. We assume that a fraction f of the members leaving the infective class at time t recover and the remaining fraction $(1-f)$ die of disease.

To generalize (1) to include a general contact rate, we use S, I, and N as variables, with $R = N-S-I$. It will be convenient to use N as a model variable in place of R because the rate of new infections is now $\beta(N)SI$. We obtain a three-dimensional model:

$$\begin{aligned} S' &= -\beta(N)SI \\ I' &= \beta(N)SI - \alpha I \\ N' &= -(1-f)\alpha I. \end{aligned} \tag{4}$$

We wish to show that model (4) has the same qualitative behavior as model (1), namely that it has the properties I and II.

For model (4), the basic reproduction number is given by

$$R_0 = \frac{K\beta(K)}{\alpha}$$

because a single infective introduced into a wholly susceptible population makes $C(K) = K\beta(K)$ contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\alpha$. In addition to the basic reproduction number R_0, there is also a time-dependent running reproduction number, that we call R^*, representing the number of secondary infections caused by a single individual in the population who becomes infective at time t. In this situation, an infective makes $C(N) = N\beta(N)$ contacts in unit time, of which a fraction S/N are with susceptibles and thus produce new infections. Then it is easy to see that for model (4) the running reproduction number is given by

$$R^* = \frac{S\beta(N)}{\alpha}.$$

If $R^* < 1$ for all large t, the epidemic will die out. We may calculate the rate of change of the running reproduction number with respect to time, using (3) and (4) to find that

$$\begin{aligned} \frac{d}{dt}R^* &= \frac{S'(t)\beta(N) + S(t)\beta'(N)N'(t)}{\alpha} = \frac{(-\beta(N))^2 SI - S\alpha(1-f)\beta'(N)}{\alpha} \\ &\leq \frac{\beta(N)SI}{\alpha} \cdot \left[\beta(N) - \frac{(1-f)\alpha}{N}.\right] \end{aligned}$$

Thus, $\frac{d}{dt}R^* < 0$ if $N\beta(N) > \alpha(1-f)$, or $R^* > (1-f)S/N$. This means that R^* decreases whenever $R^* > 1$. Thus, if $R^* < 1$ for $t = T$, then $R^* < 1$ for $t > T$. If $R_0 > 1$, then $I'(0) = \alpha(R_0 - 1)I(0) > 0$, and an epidemic begins. However, R^* decreases until it is less than 1 and then remains less than 1. Thus, the epidemic will die out. If $R_0 < 1$, then $I'(0) = \alpha(R_0 - 1)I(0) < 0$, $R^* < 1$ for all t, and there is no epidemic.

If we use the same approach as was used for (1) to show that $S_\infty > 0$, we obtain

$$\frac{dI}{dS} = -1 + \frac{\alpha}{S\beta(N)},$$

and we are unable to proceed because of the dependence on N. However, since

$$\beta(N) \leq \beta(0),$$

we have

$$\frac{dI}{dS} \geq -1 + \frac{\alpha}{S\beta(0)},$$

We may now use the following comparison theorem established by R. Conti [3].

Theorem 1. Let $y(x), z(x)$ be the respective solutions of the initial value problems

$$\begin{aligned} y' &= f(x,y), \quad y(x_0) = y_0 \\ z' &= g(x,y), \quad z(x_0) = y_0 \end{aligned}$$

and suppose that $g(x,y) \leq f(x,y)$ for all x, y. Then

$$\begin{aligned} z(x) \leq y(x), \quad x \geq x_0 \\ z(x) \geq y(x), \quad x \leq x_0. \end{aligned}$$

If we apply this comparison theorem to the solutions $I(S)$ of

$$\frac{dI}{dS} = -1 + \frac{\alpha}{S\beta(N)}$$

and $I_0(S)$ of

$$\frac{dI}{dS} = -1 + \frac{\alpha}{S\beta(0)},$$

we see that for $S \leq K$ we have $I(S) \leq I_0(S)$. Since we have seen in our study of (1) that the the graph of $I_0(S)$ does not reach $S = 0$, it follows that the same is true for the graph of $I(S)$, and this shows that $S_\infty > 0$. We should note that we are assuming here that $\beta(0)$ is finite.

If $\beta(N) \to \infty$ as $N \to 0$, we must use a different approach to analyze the limiting behavior. If $f = 1$, the total population size remains equal to the constant K, and model (4) reduces to the simpler model (1) with β replaced by the constant $\beta(K)$. If $f < 1$, then N is a decreasing function of t, which may be inverted to give t as a decreasing function of N, and then we may consider S as a function of N. This function satisfies the separable differential equation

$$\frac{dS}{dN} = \frac{S'}{N'} = \frac{\beta(N)S}{\alpha(1-f)}. \tag{5}$$

We may solve (5) with initial condition $S = N = K$ corresponding to $t = 0$ by separation of variables, obtaining

$$S(t) = K \exp\left[\frac{-\int_{N(t)}^{K} \beta(N)dN}{\alpha(1-f)}\right] \tag{6}$$

We deduce from (6) that $S_\infty = 0$ is possible only if $N \to 0$ and $\int_0^K \beta(N)dN$ diverges. If $f > 0$, it is clear that it is not possible to have $N \to 0$, since $N(t) \geq R(t)$ and $R(t)$ takes positive values. Thus $S_\infty = 0$ is possible only if $f = 0$.

Our conclusion is that for model (4), some susceptibles escape the epidemic, as with the simpler model (1), unless all infectives die of disease and $\int_0^K \beta(N)dN$ diverges. This divergence condition means that $\beta(N)$ must be unbounded as $N \to 0$, a condition that is biologically unreasonable. In particular, standard incidence is not realistic for small population sizes. A more realistic assumption would be that the number of contacts per infective in unit time is linear for small population size and saturates for larger population sizes, which rules out the possibility that the epidemic could sweep through the entire population [19].

In many infectious diseases, there is an exposed period after the transmission of infection from susceptibles to potentially infective members but before these potential infectives can transmit infection. A generalization of the epidemic model (4) that incorporates an exponentially distributed exposed period with compartments S, E, I, R, and total population size $N = S + E + I + R$ is

$$\begin{aligned} S' &= -\beta(N)SI \\ E' &= \beta(N)SI - \kappa E \\ I' &= \kappa E - \alpha I \\ N' &= -(1-f)\alpha I. \end{aligned} \tag{7}$$

We obtain, much as in the analysis of (1),

$$\frac{d(E+I)}{dS} = -1 + \frac{\alpha}{S\beta(N)},$$

and we may deduce from this that $S_\infty > 0$, as with (4). It follows that (7) and (4) have the same asymptotic behavior.

Some diseases have an asymptomatic stage in which there is some infectivity rather than an exposed period. This may be modeled by assuming infectivity reduced by a factor ϵ_A during an exposed stage. The analogue of model (7) with infectivity during the exposed stage and a density-dependent contact rate is

$$\begin{aligned} S' &= -\beta(N)S(I + \epsilon_A A) \\ A' &= \beta(N)S(I + \epsilon_A A) - \kappa A \\ I' &= \kappa A - \alpha I \\ N' &= -(1-f)\alpha I. \end{aligned} \tag{8}$$

Here, we have used A, for asymptomatic, in place of E as one of the compartments. For this model

$$R_0 = \frac{K\beta(K)}{\alpha} + \epsilon_A \frac{K\beta(K)}{\kappa}.$$

If we attempt to use the approach used when there is no infectivity during the exposed stage, we obtain

$$\frac{d}{dI}(A+I) = -1 + \frac{\alpha I}{\beta(N)(I + \epsilon_A A)S},$$

which we are unable to integrate. However, we can use an alternate approach which actually comes from the original Kermack-McKendrick analysis.

We begin by writing

$$I' = -S' - A' - \alpha I, \quad I(0) = 0$$

and solving this linear first order differential equation to give

$$I(t) = -\int_0^t [S'(s) + A'(s)]e^{-\alpha(t-s)}ds.$$

A similar integration of the initial value problem

$$A' = -\kappa A - S', \quad E(0) = 0$$

gives

$$A(t) = -\int_0^t S'(s)e^{-\kappa(t-s)}ds.$$

Then we have

$$-\frac{S'(t)}{S(t)} = -\beta(N(t))\int_0^t [S'(s)+A'(s)]e^{-\alpha(t-s)}ds + \beta\epsilon_A\int_0^t [-S'(s)]e^{-\kappa(t-s)}ds,$$

which we integrate with respect to t from 0 to ∞, obtaining

$$\begin{aligned}
\log\frac{K}{S_\infty} &= \int_0^\infty \beta(N(t))\int_0^t [-S'(s)][e^{-\alpha(t-s)}+\epsilon_A e^{-\kappa(t-s)}]dsdt \\
&\quad -\beta(N(t))\int_0^\infty\int_0^t A'(s)e^{-\alpha(t-s)}dsdt \\
&\le \beta(0)\int_0^\infty\int_t^\infty [-S'(s)][e^{-\alpha(t-s)}+\epsilon_A e^{-\kappa(t-s)}]dtds \\
&\quad -\beta(K)\int_0^\infty\int_t^\infty A'(s)e^{-\alpha(t-s)}dtds \\
&= \beta(0)\int_0^\infty\int_0^\infty [-S'(s)][e^{-\alpha(u)}+\epsilon_A e^{-\kappa u}]duds \\
&\quad -\beta(K)\int_0^\infty\int_0^\infty A'(s)e^{-\alpha u}duds.
\end{aligned}$$

Since $\int_0^\infty A'(s)ds = 0$, this reduces to

$$\log\frac{K}{S_\infty} \le \beta(0)(K-S_\infty)\frac{\kappa+\epsilon_A\alpha}{\kappa\alpha}. \tag{9}$$

This is the analogue of the limit equation, (2), for the simple model, (1), with mass action incidence, and shows that $S_\infty > 0$.

3. **Models incorporating public-health responses.** An actual epidemic differs considerably from the idealized models, (1) or (4), as was shown by the SARS epidemic of 2002–3. Some notable differences include the following:

1. Diagnosed infectives may be hospitalized, both for treatment and to isolate them from the rest of the population.
2. Contact tracing of diagnosed infectives may identify people at risk of becoming infective, who may be quarantined (instructed to remain at home and avoid contacts) and monitored so that they may be isolated immediately if and when they become infective.
3. Isolation may be imperfect; in-hospital transmission of infection was a major problem in the SARS epidemic.

In the SARS epidemic of 2002–3 in-hospital transmission of disease from patients to healthcare workers or visitors because of imperfect isolation accounted for many of the cases. This points to an essential heterogeneity in disease transmission that must be included whenever there is any risk of such transmission.

All these generalizations have been considered in studies of the SARS epidemic of 2002–3. While the ideas were suggested in SARS modelling, they are in fact relevant to any epidemic. One beneficial effect of the SARS epidemic has been to draw attention to epidemic modeling which may be of great value in coping with future epidemics.

When no vaccine is available, isolation and quarantine are the main measures available for attempting to manage an outbreak of a new disease. We assume that an epidemic has started, but that the number of infectives is small and almost

all members of the population are still susceptible, and we formulate a model to describe the course of an epidemic when management measures are begun under the following assumptions:

1. Asymptomatic members may be infective with infectivity reduced by a factor $\epsilon_A, 0 \leq \epsilon_A < 1$.
2. Asymptomatic members who are not isolated become infective at rate κ_1.
3. We introduce a class Q of quarantined members and a class J of isolated members.
4. Asymptomatic members are quarantined at a rate $\gamma_1 A$ (in practice, a quarantine will also be applied to many susceptibles, but we ignore this in the model). The effect of this assumption is that some susceptibles make fewer contacts than the model assumes. Quarantine is not perfect, but reduces the contact rate by a factor ϵ_Q.
5. There may be transmission of disease by isolated members, with an infectivity factor of ϵ_J.
6. Infectives are diagnosed and isolated at a rate $\gamma_2 I$. In addition, quarantined members are monitored and isolated immediately when they develop symptoms, at a rate $\kappa_2 Q$.
7. Infectives leave the infective class at rate $\alpha_1 I$, and a fraction f_1 of these recover, and isolated members leave the isolated class at rate $\alpha_2 J$, with a fraction f_2 recovering.

These assumptions lead to the $SAQIJR$ model

$$\begin{aligned} S' &= -\beta(N)S[\epsilon_A A + \epsilon_A\epsilon_Q Q + I + \epsilon_J J] \\ A' &= \beta(N)S[\epsilon_A A + \epsilon_A\epsilon_Q Q + I + \epsilon_J J] - (\kappa_1 + \gamma_1)A \\ Q' &= \gamma_1 A - \kappa_2 Q \\ I' &= \kappa_1 A - (\alpha_1 + \gamma_2)I \\ J' &= \kappa_2 Q + \gamma_2 I - \alpha_2 J \\ N' &= -(1 - f_1)\alpha_1 I - (1 - f_2)\alpha_2 J. \end{aligned} \tag{10}$$

The model before management measures are begun is (8), the special case

$$\gamma_1 = \gamma_2 = \kappa_2 = \alpha_2 = f_2 = 0, \quad \kappa_1 = \kappa, \quad \alpha_1 = \alpha$$

of (10). The model (10) is equivalent to the SARS model of [9] except for the extension to a general contact rate in place of standard incidence and the omission of immigration and natural death rate terms. The model of [2] is closely related to model (10), having isolation but not quarantine, but also distinguishing two classes of susceptibles with different susceptibilities.

We define the *control reproduction number* R_c to be the number of secondary infections caused by a single infective in a population consisting essentially only of susceptibles with the control measures in place. It is analogous to the basic reproduction number, but instead of describing the very beginning of the disease outbreak, it describes the beginning of the recognition of the epidemic. We assume that this occurs soon enough so that the total population size is still approximately K. The basic reproduction number is the value of the control reproduction number before management measures are implemented.

In addition, there is a time-dependent *effective reproduction number* R^* that continues to track the number of secondary infections caused by a single infective as the epidemic continues with management measures (quarantine of asymptomatics and

isolation of symptomatics) in place. It is not difficult to show that if the inflow into the population from travellers and new births is small (i.e., if the epidemiological time scale is much faster than the demographic time scale), our model implies that R^* will become and remain less than unity, so that the epidemic will always die out. Even if $R_c > 1$, the epidemic will abate eventually when the effective reproduction number becomes less than unity. However, it should be remembered that if the epidemic takes so long to die out that there are enough new births and immigrants to keep $R^* > 1$, there will be an endemic equilibrium, meaning that the disease will establish itself and remain in the population.

We may calculate R_c in the same way as we calculate R_0 but using the full model with quarantined and isolated classes. We obtain

$$\begin{aligned} R_0 &= \frac{\epsilon_A K\beta(K)}{\kappa_1} + \frac{K\beta(K)}{\alpha_1} \\ R_c &= \frac{\epsilon_Q\epsilon_A K\beta(K)}{D_1} + \frac{K\beta(K)\kappa_1}{D_1D_2} + \frac{\epsilon_Q K\beta(K)\gamma_1}{D_1\kappa_2} + \frac{\epsilon_J K\beta(K)\kappa_1\gamma_2}{\alpha_2 D_1 D_2} + \frac{\epsilon_J K\beta(K)\gamma_1}{\alpha_2 D_1}, \end{aligned}$$

where $D_1 = \gamma_1 + \kappa_1, \quad D_2 = \gamma_2 + \alpha_1$.

The running reproduction number R^* is the control reproduction number with K replaced by N to reflect the change in total population size and multiplied by S/N to reflect the fact that the fraction of contacts by an infected member which are with a susceptible and thus can produce a new infection is S/N. Thus

$$\begin{aligned} R^* &= \left[\frac{\epsilon_Q\epsilon_A N\beta(N)}{D_1} + \frac{N\beta(N)\kappa_1}{D_1D_2} + \frac{\epsilon_Q N\beta(N)\gamma_1}{D_1\kappa_2} + \frac{\epsilon_J N\beta(N)}{\alpha_2 D_1 D_2}(\kappa_1\gamma_2 + \gamma_1 D_2)\right]\frac{S}{N} \\ &= S\beta(N)[\frac{\epsilon_Q\epsilon_A}{D_1} + \frac{\kappa_1}{D_1D_2} + \frac{\epsilon_Q\gamma_1}{D_1\kappa_2} + \frac{\epsilon_J\kappa_1\gamma_2}{\alpha_2 D_1 D_2} + \frac{\epsilon_J\gamma_1}{\alpha_2 D_1}]. \end{aligned}$$

Each term of R_c has an epidemiological interpretation. The mean duration in A is $1/D_1$ with contact rate $\epsilon_A\beta$, giving a contribution to R_c of $\epsilon_E K\beta(K)/D_1$. A fraction κ_1/D_1 goes from A to I, with contact rate β and mean duration $1/D_2$, giving a contribution of $K\beta(K)\kappa_1/D_1D_2$. A fraction γ_1/D_1 goes from A to Q, with contact rate $\epsilon_Q\epsilon_Q\beta$ and mean duration $1/\kappa_2$, giving a contribution of $\epsilon_Q K\beta(K)\gamma_1/D_1\kappa_2$. A fraction $\kappa_1\gamma_2/D_1D_2$ goes from A to I to J, with a contact rate of $\epsilon_J\beta$ and a mean duration of $1/\alpha_2$, giving a contribution of $\epsilon_J K\beta(K)\kappa_1\gamma_2/\alpha_2 D_1 D_2$. Finally, a fraction γ_1/D_1 goes from E to Q to J with a contact rate of $\epsilon_J\beta$ and a mean duration of $1/\alpha_2$, giving a contribution of $\epsilon_J K\beta(K)\gamma_1/D_1\alpha_2$. The sum of these individual contributions gives R_c.

The linearization of (10) at the disease-free equilibrium $(K, 0, 0, 0, 0, K)$ has the 4×4 matrix

$$\begin{bmatrix} \epsilon_A K\beta(K) - (\kappa_1 + \gamma_1) & \epsilon_A\epsilon_Q K\beta(K) & K\beta(K) & \epsilon_J K\beta(K) \\ \gamma_1 & -\kappa_2 & 0 & 0 \\ \kappa_1 & 0 & -(\alpha_1 + \gamma_2) & 0 \\ 0 & \kappa_2 & \gamma_2 & -\alpha_2 \end{bmatrix},$$

obtained by dropping the zero rows and columns corresponding to S and N. The corresponding characteristic equation is a fourth-degree polynomial equation whose constant term is positive if $R_c < 1$ and negative if $R_c > 1$. Thus, if $R_c > 1$, there is a positive eigenvalue, corresponding to an initial exponential growth rate of solutions of (10). If $R_c < 1$, it is possible to show that all eigenvalues of the coefficient matrix have negative real part, and thus solutions of (10) die out exponentially [23].

An argument similar to the one used for (4) but technically more complicated may be used to show that $S_\infty > 0$ for the treatment model (10). If $\epsilon_A = 0$—that is, if there is no infectivity in the exposed or asymptomatic stage,—we can give a simpler argument, as follows.

Let $U = A + Q + I + J$. Then we may consider U as a function of S, and

$$\frac{dU}{dS} = -1 + \frac{\alpha_1 I + \alpha_2 J}{\beta(N)S(I + \epsilon_J J)}.$$

If we take $c_1 = min(\alpha_1, \alpha_2/\epsilon_J)$ we obtain

$$\frac{dU}{dS} \geq -1 + \frac{\alpha}{\beta(N)S},$$

and we obtain $S_\infty > 0$ by the same analysis as was used for (4), provided $\beta(0)$ is finite. If $\beta(N) \to \infty$ as $N \to 0$, we may also use the analysis of (4) since

$$\frac{dS}{dN} \leq \frac{S\beta(N)}{c_2}$$

if $c_2 \leq min(\alpha_1(1 - f_1), \alpha_2(1 - f_2)/\epsilon_J)$.

Our conclusion is that the asymptotic behavior of the treatment model (10) is the same as that of the simpler model (4). If the control reproduction number R_c is less than 1, the disease dies out and if $R_c > 1$, there is an epidemic that will die out leaving some members of the population untouched.

To relate S_∞ and N_∞, we begin by integrating the equations for $S + A, Q, I, J$, and N of (10) with respect to t from $t = 0$ to $t = \infty$, using the initial conditions

$$S(0) + A(0) = N(0) = K, \qquad Q(0) = I(0) = J(0) = 0.$$

We obtain

$$\begin{aligned}
K - S_\infty &= (\kappa_1 + \gamma_1)\int_0^\infty A(s)ds \\
\gamma_1 \int_0^\infty A(s)ds &= \kappa_2 \int_0^\infty Q(s)ds \\
\kappa_1 \int_0^\infty A(s)ds &= (\alpha_1 + \gamma_2)\int_0^\infty I(s)ds \\
\kappa_2 \int_0^\infty Q(s)ds &= \alpha_2 \int_0^\infty J(s)ds - \gamma_2 \int_0^\infty I(s)ds \\
K - N_\infty &= (1 - f_1)\alpha_1 \int_0^\infty I(s)ds + (1 - f_2)\alpha_2 \int_0^\infty J(s)ds.
\end{aligned}$$

Now we need to express $\int_0^\infty I(s)ds$ and $\int_0^\infty J(s)ds$ in terms of $\int_0^\infty A(s)ds$. From the above relations for integrals we obtain

$$\begin{aligned}
(\alpha_1 + \gamma_2)\int_0^\infty I(s)ds &= \kappa_1 \int_0^\infty A(s)ds \\
\alpha_2 \int_0^\infty J(s)ds &= \frac{\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2}{\alpha_1 + \gamma_2}\int_0^\infty A(s)ds.
\end{aligned}$$

Thus we have

$$\begin{aligned}
K - N_\infty &= \frac{(1 - f_1)\alpha_1\kappa_1 + (1 - f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{\alpha_1 + \gamma_2}\int_0^\infty A(s)ds \\
&= \frac{(1 - f_1)\alpha_1\kappa_1 + (1 - f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{(\kappa_1 + \gamma_1)(\alpha_1 + \gamma_2)}[K - S_\infty]. \qquad (11)
\end{aligned}$$

This has the form

$$K - N_\infty = c[K - S_\infty] \tag{12}$$

with c, the disease death rate, given by

$$c = \frac{(1-f_1)\alpha_1\kappa_1 + (1-f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{(\kappa_1+\gamma_1)(\alpha_1+\gamma_2)}.$$

The mean disease death rate may be measured, and this expression gives information about some of the parameters in the model that cannot be measured directly. It is easy to see that $0 \le c \le 1$ with $c = 0$ if and only if $f_1 = f_2 = 1$, that is, if and only if there are no disease deaths, and $c = 1$ if and only if $f_1 = f_2 = 0$, that is, if and only if the disease is universally fatal. The argument used to establish (11) becomes considerably simpler when applied to model (4) and gives (12) with $c = 1 - f$

The computations in this section show that the asymptotic behavior of the treatment model (10) is the same as that of the simpler model (4). If the control reproduction number R_c is less than 1, the disease dies out, and if $R_c > 1$, there is an epidemic that will die out, leaving some members of the population untouched.

4. **A vaccination model.** If a vaccine is available for a disease that threatens an epidemic outbreak, a vaccinated class that is protected at least partially against infection should be included in a model. While this is probably not relevant for an outbreak of a new disease, it would be an important aspect for modeling influenza outbreaks or a bioterrorist outbreak of smallpox.

If there is an outbreak of a disease for which a vaccine is available, then it is natural to include the effect of vaccination in an epidemic model. To model (4), we add the assumption that in unit time a fraction ϕ of the susceptible class is vaccinated. The vaccination may reduce but not completely eliminate susceptibility to infection. We model this by including a factor σ, $0 \le \sigma \le 1$, in the infection rate of vaccinated members, with $\sigma = 0$ meaning that the vaccine is perfectly effective and $\sigma = 1$ meaning that the vaccine has no effect. We assume also that the vaccination loses effect at a proportional rate θ. We describe the new model by including a vaccinated class V, with

$$\begin{aligned} S' &= -\beta(N)SI - \phi S + \theta V \\ V' &= \phi S - \sigma\beta(N)VI - \theta V \\ I' &= \beta(N)(S + \sigma V)I - \alpha I \\ N' &= -(1-f)\alpha I. \end{aligned} \tag{13}$$

We may view model (13) as a management model for an epidemic modeled initially by (4) that is different from model (10). We think of introducing a single infective into a population of total size K consisting of S susceptibles and V vaccinated members. This infective makes $\beta(K)(S + \sigma V)$ contacts in unit time, and the mean infective period is $1/\alpha$. Thus the corresponding reproduction number is

$$\frac{\beta}{\alpha} \cdot (S(0) + \sigma V(0)). \tag{14}$$

There are two essentially different scenarios. The first is an outbreak of a new disease for which a vaccine (presumably developed originally for some other disease) is available. Then the population would not have been vaccinated before the

beginning of the disease outbreak and we would take $S(0) \approx K, V(0) = 0$. Then the reproduction number given by (14) would be

$$R_0 = \frac{K\beta(K)}{\alpha}.$$

The second scenario is a disease against which the population has been prevaccinated; in this scenario we would assume the population size has reached the disease-free equilibrium of (13); namely,

$$S = \frac{\theta}{\theta + \phi}K, \quad V = \frac{\phi}{\theta + \phi}K.$$

Then the control reproduction number given by (13) is

$$R_\phi = \frac{K\beta(K)}{\alpha} \cdot \frac{\theta + \sigma\phi}{\theta + \phi}.$$

From the equation for I in (13), we see that initially I grows exponentially, and there is an epidemic if and only if

$$\beta(K)(S(0) + \sigma V(0)) > \alpha,$$

that is, if and only if the reproduction number is greater than 1.

The next step in the analysis of model (13) is to integrate the equations

$$(S + V + I)' = -\alpha I$$

and

$$N' = -(1 - f)\alpha I$$

to obtain

$$K - N_\infty = (1 - f)[K - S_\infty - V_\infty].$$

Next we introduce a new variable $T = S + V$ and then, using $S + \sigma V \leq T$, we have

$$\begin{aligned} \frac{dI}{dT} &= -1 + \frac{\alpha}{\beta(N)(S + \sigma V)} \\ &\geq -1 + \frac{\alpha}{\beta(0)T}. \end{aligned}$$

The comparison theorem used previously shows that T is less than the solution of

$$\frac{dI}{dT} = -1 + \frac{\alpha}{\beta(0)T},$$

and thus we have $T_\infty > 0$ if $\beta(0)$ is finite.

If $\beta(N) \to \infty$ as $N \to 0$, we may follow the analysis of (7) with T in place of S to conclude that the quantity $T = S + V$, representing the number of members of the population not touched by the disease, remains positive unless $f = 0$ and $\int_0^K \beta(N)dN$ diverges. Thus the vaccination model (13) retains the basic properties (i) and (ii) of the simple epidemic model (1). In particular, in an epidemic model with vaccination, it is not possible to have a backward bifurcation. This contrasts with the situation in models with births and deaths in which vaccination may lead to endemic equilibria when $R_0 < 1$ and backward bifurcations (multiple endemic equilibria or endemic equilibria) when $R_0 < 1$ [8, 10, 11, 20].

If a vaccine is available for a particular disease, it would be straightforward to formulate a model incorporating both a quarantine and isolation program and vaccination. The same approach may be used to establish the basic properties (i) and (ii). A natural question, which could be answered using such a model, would be

whether vaccination, quarantine and isolation, or a combination of both measures would be the most effective management strategy. Presumably, the answer would depend on the parameters for the specific disease.

5. **Models incorporating responses by individuals to an epidemic.** An aspect of epidemics not addressed in any of the models in earlier sections is the possibility that some members may reduce the number of contacts they make with other members of the population. One way this may occur is if infectives are sick enough to be unable to continue their normal activities, even if they do not require hospitalization. Another possibility is that if it is known that there is a disease outbreak, some susceptibles will voluntarily reduce their contacts and take hygienic measures to decrease the danger of becoming infected. In fact, there is reason to believe that educational initiatives to encourage safer behavior may be a useful public health measure for coping with a disease outbreak [17].

We will formulate a model that generalizes (4) and (7) by incorporating both a decrease in the number of contacts made by infectives because of a reduced level of activity and a decrease in the number of contacts made by susceptibles in response to their awareness of infection. We begin with the reduction of contacts by infectives. We replace the assumption that all members of the population make $C(N)$ contacts in unit time by the assumption that infectives make $qC(N)$ contacts in unit time while other members of the population make $C(N)$ contacts in unit time, where $0 \le q \le 1$. Then the total number of contacts in unit time is $(N - I)C(N) + qIC(N)$, of which a fraction

$$\frac{qI}{N-(1-q)I}$$

is with an infective. This implies that the number of new infectives in unit time is

$$\frac{qSIC(N)}{N-(1-q)I}.$$

To model the reduction of contacts by susceptibles we replace the contact rate $C(N)$ by a function $C(N, I)$, satisfying the condition

$$\frac{\partial}{\partial I}C(N,I) \le 0$$

in addition to the analogues of the conditions imposed earlier on $C(N)$, namely,

$$\frac{\partial}{\partial N}C(N,I) \ge 0, \quad \frac{\partial}{\partial N}\frac{C(N,I)}{N} \le 0, \quad \lim_{N\to 0}\frac{C(N,0)}{N} = \beta(0) < \infty.$$

The resulting generalization of model (7) is

$$\begin{aligned} S' &= -\theta(N,I)SI \\ E' &= \theta(N,I)SI - \kappa E \\ I' &= \kappa E - \alpha I \\ N' &= -(1-f)\alpha I \end{aligned} \tag{15}$$

with

$$\theta(N,I) = \frac{qC(N,I)}{N-(1-q)I}.$$

System (15) is identical to the system (7), with $\beta(N)$ replaced by $\theta(N, I)$, and has

$$R_0 = \frac{qC(K,0)}{\alpha}.$$

To show that $S_\infty > 0$ for model (15), we note that $\theta(N, I)$ is a nondecreasing function of q for all N, I and thus takes its maximum value $C(N, I)/N$ when $q = 1$. Since $C(N, I) \leq C(N, 0)$ we have

$$\theta(N, I) \leq \frac{C(N, I)}{N} \leq \frac{C(N, 0)}{N} \leq \beta(0) < \infty,$$

and now we may apply the comparison theorem as with (7) to show that $S_\infty > 0$. Thus the qualitative behavior of model (15) is the same as that of (7), and model (15) has the basic properties I and II.

6. **Discussion.** We have established that general epidemic models behave in the same way asymptotically in the sense that there is a basic reproduction number which determines whether there will be an epidemic and that an epidemic will pass through a population leaving some members untouched. We conjecture that this would remain true for more complicated models with more compartments and more stages, including models with heterogeneity of mixing. Of course, our models assume that the course of the epidemic is rapid enough that demographic effects may be ignored. If this is not true, then it would be possible for a disease to become endemic.

The underlying assumptions in all the models we have described are that the sizes of the compartments are large enough that deterministic models are appropriate and that the mixing of members is homogeneous. While these assumptions are probably reasonable once an epidemic is well underway, events at the beginning of an epidemic may be quite different. To model events with a small number of infectives in a population of susceptibles, we should use a branching process stochastic model.

We have been considering an epidemic in a single location, ignoring travel of individuals who may be infective between locations. Modern transportation has permitted the rapid transfer of infectious diseases over great distances, and an aspect of epidemic management that has become important is the screening of travellers who may be infective. Epidemic models which include some movement into and out of populations are a natural extension of the models considered here.

Acknowledgements. This work was supported by Mathematics of Information Technology and Complex Systems (MITACS) and a National Science and Engineering Research Council (NSERC) research grant, and enhanced by discussions with Carlos Castillo-Chavez, Christopher Kribs-Zaleta, and James Watmough.

REFERENCES

[1] C. Castillo-Chavez, K. Cooke, W. Huang, and S. A. Levin (1989). The role of long incubation periods in the dynamics of HIV/AIDS. Part 1: Single Populations Models, *J. Math. Biol.,* **27:** 373–98.

[2] G. Chowell, P. W. Fenimore, M. A. Castillo-Garsow, and C. Castillo-Chavez (2003). SARS outbreaks in Ontario, Hong Kong and Singapore: The role of diagnosis and isolation as a control mechanism, *J. Theor. Biol.*, **224**: 1–8.

[3] R. Conti (1956). Sulla prolungabilità delle soluzioni di un sistema di equazioni differenziali ordinarie, *Boll. Unione Ital.*, **11**: 510–514.

[4] K. L. Cooke (1967). Functional differential equations: Some models and perturbation problems. In *Differential Equations and Dynamical Systems*, J. K. Hale and J. P. LaSalle, eds., Academic Press, New York, pp. 167–183.

[5] O. Diekmann (1977). Limiting behavior in an epidemic model, *Nonlin. Anal.*, **1**: 459–470.

[6] O. Diekmann and J. A. P. Heesterbeek (2000). *Mathematical Epidemiology of Infectious Diseases*, Wiley, Chichester.
[7] K. Dietz (1982). Overall patterns in the transmission cycle of infectious disease agents. In *Population Biology of Infectious Diseases*, R.M. Anderson and R.M. May, eds., Life Sciences Research Report **25**, Springer-Verlag, Berlin, pp. 87–102.
[8] J. Dushoff, W. Huang, and C. Castillo-Chavez (1998). Backwards bifurcations and catastrophe in simple models of total diseases, *J. Math. Biol.*, **36**: 227–248.
[9] A. Gumel, S. Ruan, T. Day, J. Watmough, F. Brauer, P. van den Driessche, D. Gabrielson, C. Bowman, M. E. Alexander, S. Ardal, J. Wu, and B. M. Sahai (2004). Modeling strategies for controlling SARS outbreaks in Toronto, Hong Kong, Singapore and Beijing, *Proc. Roy. Soc. London, Ser. B*, **271**: 2223–2233.
[10] K. P. Hadeler and C. Castillo-Chavez (1995). A core group model for disease transmission,*Math Biosc.*, **128**: 41–55.
[11] K. P. Hadeler and P. van den Driessche (1997). Backward bifurcation in epidemic control, *Math. Biosc.*, **146**: 15–35.
[12] J. A. P. Heesterbeek and J. A. J. Metz (1993). The saturating contact rate in marriage and epidemic models, *J. Math. Biology*, **31**: 529–539.
[13] H. W. Hethcote (1970) Note on determining the limiting susceptible population in an epidemic model, *Math. Biosc.*, **9**: 161–163.
[14] F. C. Hoppensteadt (1975) *Mathematical Theories of Populations: Demographics, Genetics and Epidemics*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
[15] F. C. Hoppensteadt and P. Waltman (1970). A problem in the theory of epidemics, *Math. Biosc.*, **9**: 71–91.
[16] F. C. Hoppensteadt and P. Waltman (1970) A problem in the theory of epidemics II, *Math. Biosc.*, **12**: 133–145.
[17] Y.-H. Hsieh, C.-C. King, C. W. S. Chu, M-S. Ho, S-B. Hsu, and Y-C. Wu (2005). Impact of intervention measures and public response for SARS outbreak, to appear.
[18] W. O. Kermack and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics, *Proc. Royal Soc. London*, **115**: 700–721.
[19] C. M. Kribs-Zaleta (2004) To switch or taper off: The dynamics of saturation, *Math. Biosc.*, **192**: 137–152.
[20] C. M. Kribs-Zaleta and J. X. Velasco-Hernandez (2000). A simple vaccination model with multiple endemic states, *Math Biosc.*, **164**: 183–201.
[21] J. Mena-Lorca and H. W. Hethcote (1992). Dynamic models of infectious diseases as regulators of population size, *J. Math. Biology*, **30**: 693–716.
[22] H. R. Thieme (1977). A model for the spatial spread of an epidemic, *J. Math. Biol.*, **4**: 337–351.
[23] P. van den Driessche and J. Watmough (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosc.*, **180**: 29–48.
[24] L. O. Wilson (1972) An epidemic model involving a threshold, *Math. Biosc.*, **15**: 109–121.

Received on 20 December, 2004. Revised on April 2, 2005.

E-mail address: brauer@math.ubc.ca