Epidemiological Models with Non-Exponentially Distributed Disease Stages and Applications to Disease Control

Zhilan Feng[a,*], Dashun Xu[b], Haiyun Zhao[a]

[a] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA*
[b] *Department of Mathematics, Southern Illinois University, Carbondale, IL 62901, USA*

Abstract *SEIR* epidemiological models with the inclusion of quarantine and isolation are used to study the control and intervention of infectious diseases. A simple ordinary differential equation (ODE) model that assumes exponential distribution for the latent and infectious stages is shown to be inadequate for assessing disease control strategies. By assuming arbitrarily distributed disease stages, a general integral equation model is developed, of which the simple ODE model is a special case. Analysis of the general model shows that the qualitative disease dynamics are determined by the reproductive number $\mathcal{R}_c$, which is a function of control measures. The integral equation model is shown to reduce to an ODE model when the disease stages are assumed to have a gamma distribution, which is more realistic than the exponential distribution. Outcomes of these models are compared regarding the effectiveness of various intervention policies. Numerical simulations suggest that models that assume exponential and non-exponential stage distribution can produce inconsistent predictions.

1. Introduction

The mathematical theory of infectious diseases pioneered by Ross, MacDonald, Kermack, McKendrick and others has played a major role in the study of the control and prevention of infectious diseases (see, for example, Ross, 1911; Kermack and McKendrick, 1927). More recently, mathematical models have been used to investigate how to more effectively control SARS via various disease control measures including vaccination, quarantine, and isolation (see, for

---

*Corresponding author.
E-mail address: zfeng@math.purdue.edu (Z. Feng).

example, Chowell et al., 2003; Lipsitch et al., 2003; Riley et al., 2003; McLean et al., 2005).

Many of these studies have taken the approach of using simple ODE models to draw conclusions regarding the effectiveness of various disease control programs. The simplicity of these models is often achieved by making the assumption that the disease stages are exponentially distributed. When the models do not include quarantine and/or isolation, or when the isolation is assumed to be perfect (i.e., isolated individuals do not transmit the disease), the exponential distribution assumption (EDA) and the models that use this assumption have been shown to provide valuable information and important insights into the disease dynamics. However, as demonstrated in this article, the EDA may not be appropriate in models for diseases with relatively long latent and/or infectious periods when isolation is not completely effective.

Here, we discuss the problem by considering a simple ODE model that is a commonly used *SEIR*-type model. In a standard *SEIR* model, the whole population N is divided into four sub-classes: susceptible (S), exposed but not yet infectious (E), infectious (I), and immune or recovered (R) individuals. Susceptibles become exposed (latent) at the rate $\lambda_1(t)S(t) = \beta S(t)I(t)/N$ where β is the disease transmission coefficient in the absence of interventions and N is the total population size. Latent individuals progress to the infectious stage at a constant rate α_1 and infectious individuals recover at a constant rate δ_1. A transmission diagram for this case is shown in Fig. 1a. All variables and parameters are listed in Table 1.

To incorporate control measures such as quarantine and isolation, additional sub-classes can be included. For example, quarantine and isolation can be modeled as follows. Let $\lambda(t)$ denote the force of infection (a specific form is given below). Assume that a fraction b of contacts (susceptible individuals who have had contacts with an infectious person) are actually infected, and that the other fraction $(1 - b)$ of contacts remain susceptible who will be quarantined (S_Q) and will return to the S class at a rate r (see, e.g., Lipsitch et al., 2003). Among the infected individuals ($b\lambda(t)S$) a fraction γ will be quarantined (Q) at the time of infection (i.e., there is a rate, $\gamma b\lambda(t)S$, from S to Q directly). The fraction $(1 - \gamma)$ of the exposed (and infected) individuals (E) who are not quarantined at the time of infection will be quarantined at a constant rate χ throughout the latent period. (We remark that although for most diseases quarantine is not considered as the exposed individuals show no disease symptoms, the situation for SARS is different in which quarantine was implemented in several places including Hong Kong, Taiwan, and China.) The non-quarantined and quarantined (exposed) individuals will progress to the
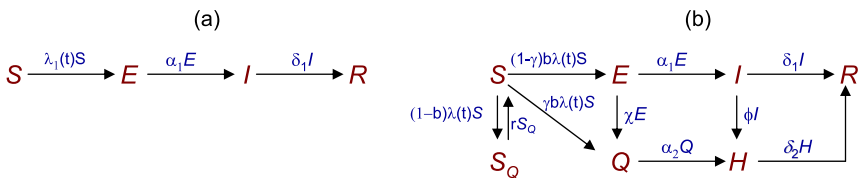
Fig. 1 Disease transmission diagrams (birth and death omitted). (a) No disease control. (b) Quarantine and isolation are included.

Table 1 Definitions of frequently used symbols

Symbol	Definition
$S(t)$	Number of susceptible individuals at time t
$S_Q(t)$	Number of susceptible individuals quarantined at time t
$E(t)$	Number of exposed (not yet infectious) individuals at time t
$Q(t)$	Number of quarantined (exposed) individuals at time t
$I(t)$	Number of susceptible individuals at time t
$H(t)$	Number of isolated (infectious) individuals at time t
$R(t)$	Number of recovered individuals at time t
N	Total population size (constant)
$C(t)$	Number of cumulative new infections at time t
$\lambda(t)$	Force of infection at time t
β	Transmission coefficient
$\alpha_1,\ \alpha_2$	Rate at which non-quarantined, quarantined individuals become infectious
α	Same as α_1
$\delta_1,\ \delta_2$	Rate at which non-isolated, isolated individuals become recovered
δ	Same as δ_1
μ	Natural death rate
χ, ϕ	Rate of quarantine, isolation
ρ	Isolation efficiency ($0 \le \rho \le 1$)
b	Fraction of contacts infected ($b = 1$ in this paper)
γ	Fraction of infecteds quarantined at time of exposure ($\gamma = 0$ in this paper)
$p_i(s), P_i(s)$	Probability that disease stage i lasts longer than s time units ($i = E, I$)
$k(s), l(s)$	Probability of not being quarantined, isolated at stage age s
$T_E,\ T_I$	Mean of $p_E(s) = \mathrm{e}^{-\alpha s}$, $p_I(s) = \mathrm{e}^{-\delta s}$ ($T_E = 1/\alpha$, $T_I = 1/\delta$)
$\mathcal{M}_i(s), M$	Expected remaining sojourn at age s: $\displaystyle\int_0^\infty P_i(t\vert s)\,\mathrm{d}t$ ($i = E, I$), $M = \mathcal{M}(0)$
$\mathcal{T}_E$	Probability of surviving and becoming infectious: $\displaystyle\int_0^\infty [-\dot{P}_E(s)]\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{T}_{E_k}$	"Quarantine-adjusted" probability (similar to $\mathcal{T}_E$): $\displaystyle\int_0^\infty [-\dot{P}_E(s)k(s)]\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{T}_I$	Probability an infectious person survives and recovers: $\displaystyle\int_0^\infty [-\dot{P}_I(s)]\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{T}_{I_l}$	"Isolation-adjusted" probability (similar to $\mathcal{T}_I$): $\displaystyle\int_0^\infty [-\dot{P}_I(s)l(s)]\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{D}_E$	Mean time in exposed stage (adjusted by death): $\displaystyle\int_0^\infty P_E(s)\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{D}_{E_k}$	"Quarantine-adjusted" mean time in exposed stage: $\displaystyle\int_0^\infty P_E(s)k(s)\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{D}_I$	Mean time in infectious stage (adjusted by death): $\displaystyle\int_0^\infty P_I(s)\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{D}_{I_l}$	"Isolation-adjusted" mean time in infectious stage: $\displaystyle\int_0^\infty P_I(s)l(s)\,\mathrm{e}^{-\mu s}\,\mathrm{d}t$
$\mathcal{R}_0$	The basic reproductive number
$\mathcal{R}_c$	The reproductive number under control measures
EDA, GDA	Exponential distribution assumption, Gamma distribution assumption
EDM, GDM	Exponential distribution model, Gamma distribution model

infectious stage at constant rates α_1 and α_2 respectively (the relationship between α_1 and α_2 will be discussed later). Infectious individuals will be isolated (H) at a rate ϕ and individuals in the H class will recover at a rate δ_2 (the relationship between δ_1 and δ_2 will be discussed later). Since we are considering the case of imperfect isolation, the new infections are now produced at the rate $\lambda(t)S(t)$ with

$$\lambda(t) = \beta \left[\frac{I(t) + (1 - \rho)H(t)}{N} \right], \tag{1}$$

where $\rho \in [0, 1]$ is the fraction of reduction in the transmission rate of isolated individuals with $\rho = 1$, $\rho = 0$, and $0 < \rho < 1$ representing a completely effective, completely ineffective, and partially effective isolation, respectively. The corresponding transmission diagram is shown in Fig. 1b.

In this paper, we consider only the case that $\gamma = 0$ by assuming that the fraction of infected contacts that can be traced and quarantined at the time of infection is very small, and hence most infected people are quarantined during the latency period. In addition, we assume a large population size N (in comparison with the size of the infected population), in which case the quarantine of susceptibles is unlikely to have a significant impact on the disease transmission dynamics and hence will be ignored. This is equivalent to assuming that $b = 1$. (It may not be appropriate to ignore the S_Q class if one is concerned with the cost associated with quarantine, which is not the case in this paper.) For simplicity, the disease-induced death is ignored and the per-capita birth rate and the natural death rate are assumed to be equal. Hence, the total population size N remains constant. Then the corresponding ODE model is given by the following system

$$
\begin{aligned}
S' &= \mu N - \beta S \tfrac{I + (1 - \rho)H}{N} - \mu S, \\
E' &= \beta S \tfrac{I + (1 - \rho)H}{N} - (\chi + \alpha_1 + \mu)E, \\
Q' &= \chi E - (\alpha_2 + \mu)Q, \\
I' &= \alpha_1 E - (\phi + \delta_1 + \mu)I, \\
H' &= \alpha_2 Q + \phi I - (\delta_2 + \mu)H, \\
R' &= \delta_1 I + \delta_2 H - \mu R.
\end{aligned}
\tag{2}
$$

" $'$ " denotes the derivative with respect to time t. All involved parameters are nonnegative constants, and all variables and parameters are listed in Table 1.

In the next section, we discuss some of the drawbacks of the simple model (2) when used to evaluate intervention policies. We argue that the main reason for these problems is due to the simplifying assumption of exponential distributions for the disease stages, which is used in the model. This provides a motivation for using more realistic stage distributions. Non-exponential distributions have been considered in epidemiological models (see, for example, Hethcote and Tudor, 1980; Hethcote et al., 1981; Plant and Wilson, 1986; Taylor and Karlin, 1998; Feng and Thieme, 2000a,b; Feng et al., 2001; Lloyd, 2001a,b). However, none of these studies focuses on the evaluation of intervention policies.

In this paper, we develop a general model with arbitrarily distributed disease stages. The general setting allows us to identify new models that are improvements to the simple model (2) while keeping the improved models as simple as possible. We show that in the case of exponential distributions, the general model reduces to the simple model (2) with appropriate constraints on model parameters. We also consider a particular non-exponential stage distribution, the gamma distribution, in which case the general model reduces to another ODE model. Analysis

for both the general model and the model with the gamma distribution assumption (GDA) are provided. We demonstrate that the model under GDA is indeed an improvement on the model under EDA.

2. Drawbacks of EDA and model (2)

One of the main roles of model (2) and its variants is to evaluate various disease control measures. Relevant parameters that represent disease intervention are χ, ϕ, ρ, α_2 and δ_2. It is very important that these parameters have well-defined meanings in order to connect them with epidemiological data and to determine their appropriate values, and to ensure that the model predictions are reasonable regarding the effect of various control strategies. Otherwise, the results obtained from the model might be misleading, as demonstrated below.

One of the quantities that can be used to assess the impact of various control measures is the cumulative number, $C(t)$, of infections determined by the equation

$$C'(t) = \lambda(t)S(t).$$

Let $C(0) = 0$ so that $C(t)$ is the cumulative number of new infections at the end of an epidemic (in the case that the disease is driven into extinction). One would expect that the C value will be reduced if we increase the value of any of the control parameters. However, Fig. 2 (see (a) and (b)) shows that C increases with increasing rates of quarantine (χ) and isolation (ϕ). The parameter values used in Fig. 2 are the following: $\alpha_1 = 0.2$ and $\delta_1 = 0.15$, which correspond to a latency period of $1/\alpha_1 = 5$ days and an infectious period of $1/\delta_1 \approx 7$ days, respectively. These values are in the realistic range of many infectious diseases. The transmission coefficient is chosen to be $\beta = 0.13$ which corresponds to a reproductive number (calculated using the formula (43)) that is equal to approximately 0.9 (so that the disease will die out). The isolation efficiency is $\rho = 0.3$ and other parameters have different values in Figs. 2a–d depending on the assumptions. In Figs. 2a and 2b, no additional constraints are imposed on α_2 and δ_2 which have values 0.17 and 0.1, respectively.

The lack of constraints on α_2 and δ_2 may be responsible for the problem shown in Figs. 2a and 2b. Our simulation results show that the problem can be avoided if the following constraints are imposed

$$\alpha_2 = \alpha_1, \quad \delta_2 = \delta_1. \tag{3}$$

A more rigorous argument for constraint (3) will be provided in Section 3.2. Here, we give only a heuristic argument. The ordinary differential equation model (2) implicitly assumes the exponential distribution for the latent and infectious stages. More precisely, the exponential functions

$$p_E(s) = \mathrm{e}^{-\alpha_1 s} \quad \text{and} \quad p_I(s) = \mathrm{e}^{-\delta_1 s}$$
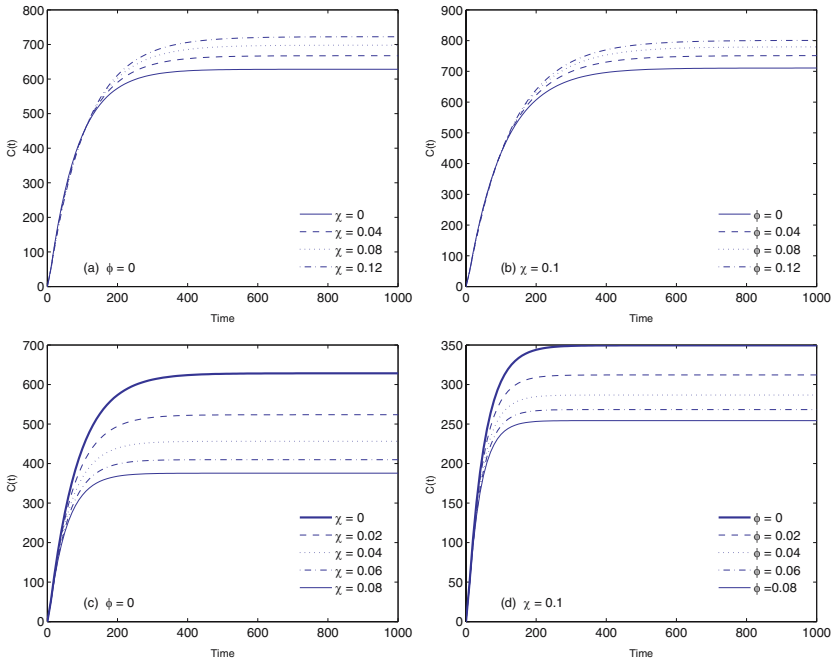
Fig. 2 Numerical simulations of the model (2). The number of cumulative new infections $C(t)$ is plotted for various values of the control parameters χ and ϕ. (a) and (b) are for the case of no constraints on the parameter values. (c) and (d) are for the case when constraint (3) is used.

have been used to describe the probability of remaining in the latent stage and the infectious stage, respectively, and the mean durations of latent and infectious stages are

$$T_E = \int_0^\infty p_E(s)\,\mathrm{d}s = \frac{1}{\alpha_1} \quad \text{and} \quad T_I = \int_0^\infty p_I(s)\,\mathrm{d}s = \frac{1}{\delta_1}. \tag{4}$$

Similarly, the mean sojourn times in the Q and H classes are respectively

$$T_Q = \frac{1}{\alpha_2} \quad \text{and} \quad T_H = \frac{1}{\delta_2}. \tag{5}$$

A fundamental property of the exponential distribution is the memory-less property, which requires that the remaining expected sojourn in the H (or Q) class is independent of the time already elapsed before entering it. This property implies that

$$T_H = T_I \quad \text{and} \quad T_Q = T_E,$$

which is equivalent to the condition given in (3) (see (4) and (5)).

Another argument for the use of (3) is the following. The average time individuals (both isolated and non-isolated individuals) stay in the I class is equal to

$1/(\delta_1 + \phi)$, and the average time an isolated individual stays in the H class is $1/\delta_2$. Notice that $\phi/(\delta_1 + \phi)$ and $\delta_1/(\delta_1 + \phi)$ are fractions of isolated and non-isolated individuals, respectively. Then the weighted average time an individual stays in the infectious stage is

$$\frac{\phi}{\delta_1 + \phi} \left(\frac{1}{\delta_1 + \phi} + \frac{1}{\delta_2} \right) + \frac{\delta_1}{\delta_1 + \phi} \left(\frac{1}{\delta_1 + \phi} \right), \tag{6}$$

which is equal to $\frac{1}{\delta_1}$ (i.e., the infectious period) only if we set $\delta_2 = \delta_1$ in accordance with constraint (3). It follows from a similar argument that the weighted average time an individual stays in the exposed stage (when quarantine is present) is equal to the latency period only if $\alpha_2 = \alpha_1$. Figs. 2c and 2d illustrate that the value of $C(t)$ reduces as the values of control parameters increase, showing the improvement compared to Figs. 2a and 2b. For Figs. 2c and 2d, all parameter values are the same as in Figs. 2a and 2b except that the constraint (3) holds.

Constraint (3) seems to provide a partial solution to the problem exhibited in Figs. 2a and 2b. However, it creates a different problem. Since the isolated or quarantined individuals have already spent some time in the infectious or latent stage before entering the H or the Q class, the condition (3) amounts to allowing for a prolonged period of infectiousness for isolated individuals and a prolonged period of latency for quarantined individuals. In reality, if an infectious person already spent some time in the I class before being isolated, then the expected remaining sojourn in the H class should be shorter than the infectious period. Therefore, the model assumption (EDA) conflicts with biological constraints. A similar argument applies to quarantined individuals. In Section 5.2, we demonstrate how the predictions of the model (2) constrained by (3) may be in disagreement with that of models using more realistic stage distributions.

It should be pointed out that the purpose of this article is not to argue which assumptions/constraints are more appropriate than others or whether they are correct or not. Our goal is to point out the weakness of models that assume exponential distributions for one or both disease stages and to demonstrate possible problems with either constraints of type (3) or no constraint. Therefore, models with more realistic stage distributions may need to be considered.

The above examples demonstrate some of the drawbacks of the EDA and the simple model (2). Epidemiological models with non-exponential distributions such as the gamma distribution have been previously studied (see, for example, Hethcote and Tudor, 1980; Plant and Wilson, 1986; Taylor and Karlin, 1998; Lloyd, 2001a,b). In these studies, the authors discussed other objections to the EDA. For example, it is pointed out that constant recovery is a poor description of real-world infections, and they show that in models with more realistic distributions of disease stages, less stable behavior may be expected and disease persistence may be diminished (see Hethcote et al., 1981; Lloyd, 2001a,b). However, these studies do not focus on the impact of imperfect isolation. In the rest of the paper, we consider models with more realistic disease stage distributions and study the properties of these models.

for the evaluation of control measures. The detailed description of $\mathcal{R}_c$ obtained from the general model (12) also provides important information about the role of the model parameters in the disease transmission dynamics. The comparison of reproductive numbers for the two models under different intervention policies indicate that the EDM again may generate outcomes which conflicts with that of the GDM (see Fig. 5). disease control policies (described by χ and ϕ) using the number of cumulative infections C. Our simulation results from the models (14) and (39) suggest that for many sets of parameter values, the two models predict contradictory outcomes (see Figs. 5 and 6).

To summarize, the results of our study in this article suggest that standard *SEIR*-type ODE models (such as models (2) and (14)), while capable of capturing many essential features of the disease transmission dynamics in the absence of quarantine and isolation, may produce results that are inconsistent with those from models with non-exponentially distributed disease stages. We considered one such model (39) by using the gamma distribution with $n = 3$. Obviously, other types of more realistic stage distributions (P_E and P_I) can be used to derive different improved models. In addition to the choice of P_E and P_I, modifications on the model (14) can be obtained by making different assumptions on $k(s)$ and $l(s)$ as well. These functions obviously play a very important role for the study of disease intervention as these are the parameters (or parameter functions) that reflect the control measures.

Finally, it should be pointed out that the biology of the infection ultimately determines the most appropriate forms for the latent and infectious periods. For some infections it might well be that an n ($n > 1$) stage gamma distribution is a better choice, for others it might even be that an exponential distribution provides a reasonable approximation. In order to determine which model is more appropriate for the problem under investigation, we need to have a deep understanding of the advantages and limitations of these models, as well as the biology of the disease.

Acknowledgements

References

Anderson, R.M., May, R.M., 1979. Prevalence of schistosome infections within molluscan populations: Observed patterns and theoretical predictions. Parasitology 79, 63–94.

Chowell, G., Fenimorea, P., Castillo-Garsowc, M., Castillo-Chavez, C., 2003. SARS outbreaks in Ontario, Hong Kong and Singapore: The role of diagnosis and isolation as a control mechanism. J. Theo. Bio. 224, 1–8.

Feng, Z., Huang, W., Castillo-Chavez, C., 2001. On the role of variable latent periods in mathematical models for tuberculosis. J. Dyn. Diff. Eqs. 13, 435–452.

Feng, Z., Thieme, H.R., 2000a. Endemic models for the spread of infectious diseases with arbitrarily distributed disease stages I: General theory. SIAM J. Appl. Math. 61(3), 803–833.

Feng, Z., Thieme, H.R., 2000b. Endemic models for the spread of infectious diseases with arbitrarily distributed disease stages II: Fast disease dynamics and permanent recovery. SIAM J. Appl. Math. 61(3), 983–1012.

Hethcote, H., Tudor, D., 1980. Integral equation models for endemic infectious diseases. J. Math. Biol. 9, 37–47.

Hethcote, H., Stech, H.W., van den Driessche, P., 1981. Nonlinear oscillations in epidemic models. SIAM J. Appl. Math. 40(1), 1–9.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. A 115, 700–721.

Kermack, W.O., McKendrick, A.G., 1932. Contributions to the mathematical theory of epidemics. II. The problem of endemicity. Proc. Roy. Soc. A 138, 55–83.

Kermack, W.O., McKendrick, A.G., 1933. Contributions to the mathematical theory of epidemics. III. Further studies of the problem of endemicity. Proc. Roy. Soc. A 141, 94–122.

Lipsitch, M., Cohen, T., Copper, B., et al., 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science 300, 1966–1970.

Lloyd, A., 2001a. Realistic distributions of infectious periods in epidemic models. Theor. Pop. Biol. 60, 59–71.

Lloyd, A., 2001b. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proc. R. Soc. Lond. B. 268, 985–993.

MacDonald, N., 1978. Time Lags in Biological Models. Springer-Verlag, New York.

McLean, A.R., Robert, M.M., Pattison, J., Weiss, R.A., 2005. A Case Study in Emerging Infections. Springer-Verlag, Oxford University Press, New York.

Plant, R.E., Wilson, L.T., 1986. Models for age-structured populations with distributed maturation rates. J. Math. Biol. 23, 247–262.

Riley, S., Fraser, C., Donnelly, C.A., Ghani, A.C., et al., 2003. Transmission dynamics of the etiological agent of SARS in Hong Kong—impact of public health interventions. Science 300, 1961–1966.

Ross, R., 1911. The Prevention of Malaria. John Murray, London.

Taylor, H.M., Karlin, S., 1998. An Introduction to Stochastic Modeling, 3rd edition. Academic Press, San Diego.

Thieme, H., 2003. Mathematics in Population Biology. Princeton University Press, Princeton.