

THE BASIC EPIDEMIOLOGY MODELS: MODELS, EXPRESSIONS FOR R_0 , PARAMETER ESTIMATION, AND APPLICATIONS

Herbert W. Hethcote
Department of Mathematics
University of Iowa
14 Maclean Hall
Iowa City, Iowa 52242, USA
E-mail: herbert-hethcote@uiowa.edu

Three basic models (SIS endemic, SIR epidemic, and SIR endemic) for the spread of infectious diseases in populations are analyzed mathematically and applied to specific diseases. Threshold theorems involving the basic reproduction number R_0 , the contact number σ , and the replacement number R are presented for these models and their extensions such as SEIR and MSEIRS. Values of R_0 and σ are estimated for various diseases in order to compare the percentages that must be vaccinated in order to achieve herd immunity.

Contents

1.1	Introduction	2
1.2	Why do epidemiology modeling?	6
1.3	Definitions, assumptions and model formulations	16
1.3.1	Formulating epidemiology models	19
1.3.2	Three threshold quantities: R_0 , σ , and R	21
1.4	The basic SIS endemic model	22
1.5	The basic SIR epidemic model	24
1.6	The basic SIR endemic model	28
1.7	Similar models with M and E epidemiological states	32
1.7.1	The SEIR epidemic model	33
1.7.2	The MSEIRS endemic model	35
1.8	Threshold estimates using the two SIR models	40
1.9	Comparisons of some directly transmitted diseases	41
1.10	The effectiveness of vaccination programs	43

1.10.1	Smallpox	43
1.10.2	Poliomyelitis	44
1.10.3	Measles	45
1.10.4	Rubella	46
1.10.5	Chickenpox (varicella)	47
1.10.6	Influenza	48
1.11	Other epidemiology models with thresholds	49
	References	52

1.1. Introduction

In this first chapter corresponding to my first two tutorial lectures,^a I introduce the terminology, notation, and standard results for epidemiology models by formulating and analyzing three basic epidemiology models. The three classic epidemiological models presented here provide an intuitive basis for understanding more complex epidemiology modeling results presented in other lectures in this course. This material is based on introductory material in previous publications [81, 76] with the addition of some new material.

The threshold for many epidemiology models is the basic reproduction number R_0 , which is defined as the average number of secondary infections produced when one infected individual is introduced into a host population where everyone is susceptible [45]. For many deterministic endemic models, an infection can get started in a fully susceptible population if and only if $R_0 > 1$. Thus the basic reproduction number R_0 is often considered as the threshold quantity that determines when an infection can invade and persist in a new host population. We demonstrate the role of R_0 in three basic models. Then thresholds are estimated from data on several diseases and the implications of the estimates are considered for diseases such as smallpox, polio, measles, rubella, chickenpox, and influenza.

Epidemiology models with variable population size are considered in Chapter 2. Infectious disease models that include both time t and age a as independent variables are more realistic than the three basic models. This is true because age groups mix heterogeneously, the recovered fraction usually increases with age, risks from an infection may be related to age, vaccination programs often focus on specific ages, and epidemiologic data is often age specific. Age-structured epidemiological models are formulated and analyzed in Chapter 3.

^ahttp://www.ims.nus.edu.sg/Programs/infectiousdiseases/files/hethcote_ln.pdf#1

Despite improved sanitation, antibiotics, and extensive vaccination programs, infectious diseases continue to be major causes of suffering and mortality. Moreover, infectious disease agents adapt and evolve, so that new infectious diseases have emerged and some existing diseases have re-emerged [112]. Newly identified diseases include Lyme disease (1975), Legionnaire's disease (1976), toxic-shock syndrome (1978), hepatitis C (1989), hepatitis E (1990), and hantavirus (1993). Popular books have given us exciting accounts of the emergence and detection of new diseases [64, 135, 137, 147]. The human immunodeficiency virus (HIV), which is the etiological agent for acquired immunodeficiency syndrome (AIDS), emerged in 1981 and has become an important sexually-transmitted disease throughout the world. Drug and antibiotic resistance have become serious issues for diseases such as tuberculosis, malaria, and gonorrhea. Malaria, dengue, and yellow fever have re-emerged and are spreading into new regions as climate changes occur. Diseases such as plague, cholera, and hemorrhagic fevers (Bolivian, Ebola, Lassa, Marburg, etc.) continue to erupt occasionally.

Surprisingly, new infectious agents called prions have recently joined the previously known agents: viruses, bacteria, protozoa, and helminths (worms). There is strong evidence that prions are the cause of spongiform encephalopathies, e.g. bovine spongiform encephalopathy (BSE, "mad cow disease"), Creutzfeldt–Jakob disease (CJD), kuru, and scrapie in sheep [135]. Biological terrorism with diseases such as smallpox or plague has become a new threat. In the 21st century, we have already encountered severe acute respiratory syndrome (SARS), a disease that emerged from a mutation of a wild animal coronavirus. An outbreak of foot and mouth disease in the United Kingdom in 2001 caused great economic hardships there [158, 60]. Avian influenza has devastated bird populations in SE Asia; moreover, there is the concern that a new human influenza strain will arise by recombination of an avian influenza strain with a human influenza strain. In the future we will undoubtedly face more new infectious disease challenges. It is clear that human or animal invasions of new ecosystems, global warming, environmental degradation, increased international travel, and changes in economic and social patterns will continue to provide opportunities for new and existing infectious diseases [122]. The emergence and re-emergence of novel and deadly forms of infectious diseases, global climate change, and population growth have increased the need for sound quantitative methods to guide disease interventions.

A model for smallpox was formulated and solved by Daniel Bernoulli in 1760 in order to evaluate the effectiveness of variolation of healthy people

with the smallpox virus [15] (the data are reproduced in Chapter 1 of [39]); however, deterministic epidemiology modeling seems to have started in the 20th century. Hamer formulated and analyzed a discrete time model in 1906 in his attempt to understand the recurrence of measles epidemics [70]. Hamer's model may have been the first to assume that the incidence (number of new cases per unit time) depends on the product of the densities of the susceptibles and infectives. Ross was interested in the incidence and control of malaria, so he developed differential equation models for malaria as a host-vector disease in 1911 [139]. Other deterministic epidemiology models were then developed in papers by Ross, Ross and Hudson, Martini, and Lotka [9, 44, 49]. Starting in 1926 Kermack and McKendrick published papers on epidemic models and obtained the epidemic threshold result that the density of susceptibles must exceed a critical value in order for an epidemic outbreak to occur [9, 106, 126].

Mathematical epidemiology seems to have grown exponentially starting in the middle of the 20th century (the first edition in 1957 of Bailey's book [9] is an important landmark), so that a tremendous variety of models have now been formulated, mathematically analyzed and applied to infectious diseases. Reviews of the literature [12, 24, 44, 48, 50, 78, 82, 83, 157] show the rapid growth of epidemiology modeling. Recent models have involved aspects such as passive immunity, gradual loss of vaccine and disease-acquired immunity, stages of infection, vertical transmission, disease vectors, macroparasitic loads, age structure, social and sexual mixing groups, spatial spread, vaccination, quarantine, and chemotherapy. Special models have been formulated for human diseases such as measles, rubella, chickenpox, whooping cough, diphtheria, smallpox, malaria, onchocerciasis, filariasis, rabies, gonorrhea, herpes, syphilis, and HIV/AIDS. The breadth of the subject is shown in the books on epidemiology modeling [4, 6, 7, 9, 10, 11, 13, 22, 23, 24, 38, 39, 43, 61, 62, 68, 85, 89, 99, 107, 111, 121, 132, 134, 139, 145, 154, 155].

Table 1 classifies diseases by agent (virus, bacteria, protozoa, helminths, prions) and method of transmission. This useful classification scheme is similar to one presented by K. Dietz [45]. The models considered in this chapter are suitable for viral and bacterial diseases which are transmitted directly from person to person. More complicated models with separate epidemiological compartments for nonhuman species must be used when there is transmission by insects called vectors or a reservoir of wild or domestic animal infectives.

Table 1. Classification of infectious diseases by agent and mode of transmission.

agent	person→person	person→environment environment→person	reservoir→vector vector→person	reservoir→person
virus (SIR type)	measles chickenpox mumps rubella smallpox influenza poliomyelitis herpes HIV (AIDS virus) SARS (coronavirus)		arboviruses: yellow fever dengue fever encephalitis tick fever sandfly fever West Nile virus	rabies hantavirus
bacteria (SIS or SIRS type)	gonorrhea tuberculosis pneumonia meningitis strep throat pertussis	typhoid fever cholera Legionnaire's disease	plague lyme disease	brucellosis tuleramia anthrax
protozoa	syphilis	amebiasis	malaria trypanosomiasis leishmaniasis	
helminths (worms)		dracunculiasis	schistosomiasis filariasis onchocerciasis	trichinosis
prions	kuru			BSE (mad cow disease) VCJD (in humans) scrapie

In recent years epidemiological modeling of infectious disease transmission has had an increasing influence on the theory and practice of disease management and control. Mathematical modeling of the spread of infectious diseases has become part of epidemiology policy decision making in many countries, including the United Kingdom, Netherlands, Canada, and the United States. Epidemiological modeling studies of diseases such as gonorrhea, HIV/AIDS, BSE, foot and mouth disease, measles, rubella, and pertussis have had an impact on public health policy in these countries. Thus modeling approaches have become very important for decision-making about infectious disease intervention programs. Recent approaches include deterministic models, computer simulations, Markov Chain Monte Carlo models, small world and other network models, stochastic simulation models, and microsimulations of individuals in a community. These techniques are often implemented computationally and use data on disease incidence

and population demographics. Sometimes the epidemiology, immunology, and evolution of a disease must all be considered. For example, some recent research has studied the rational design of influenza vaccines by considering the effects on the immunology of influenza immunity in individuals of the yearly epidemics of influenza A variants, the vaccine composition each year, and the yearly evolutionary drift of influenza A virus variants.

1.2. Why do epidemiology modeling?

Epidemiology is the study of the distribution and determinants of disease prevalence in humans. One function of epidemiology is to describe the distribution of the disease, i.e., find out who has how much of what, where and when. Another function is to identify the causes or risk factors for diseases in order to find out why everyone does not have the same thing uniformly. A third function of epidemiology is to build and test theories. A fourth function is to plan, implement and evaluate detection, control and prevention programs. Epidemiological modeling can play an important role in these last two functions. Here we focus on modeling infectious diseases in human populations and do not consider models for chronic diseases such as cancer and heart disease. In most of this section “epidemiological modeling” refers to dynamic, deterministic modeling where the population is divided into compartments based on their epidemiological status (e.g. susceptible, infectious, recovered), for which movements between compartments by becoming infected, progressing, recovering or migrating are specified by differential or difference equations.

Even though vaccines are available for many infectious diseases, these diseases still cause suffering and mortality in the world, especially in developing countries. In developed countries chronic diseases such as cancer and heart disease have received more attention than infectious diseases, but infectious diseases are still a more common cause of death in the world. Recently, emerging and re-emerging diseases have led to a revived interest in infectious diseases. The transmission mechanism from an infective to susceptibles is understood for nearly all infectious diseases and the spread of diseases through a chain of infections is known. However, the transmission interactions in a population are very complex, so that it is difficult to comprehend the large scale dynamics of disease spread without the formal structure of a mathematical model. *An epidemiological model uses a microscopic description (the role of an infectious individual) to predict the macroscopic behavior of disease spread through a population.*

Mathematical models have become important tools in analyzing the spread and control of infectious diseases. The model formulation process clarifies assumptions, variables, and parameters; moreover, models provide conceptual results such as thresholds, basic reproduction numbers, contact numbers, and replacement numbers. Mathematical models and computer simulations are useful experimental tools for building and testing theories, assessing quantitative conjectures, answering specific questions, determining sensitivities to changes in parameter values, and estimating key parameters from data. Understanding the transmission characteristics of infectious diseases in communities, regions and countries can lead to better approaches to decreasing the transmission of these diseases. Mathematical models are used in comparing, planning, implementing, evaluating, and optimizing various detection, prevention, therapy and control programs. Epidemiology modeling can contribute to the design and analysis of epidemiological surveys, suggest crucial data that should be collected, identify trends, make general forecasts, and estimate the uncertainty in forecasts [76, 85].

In many sciences it is possible to conduct experiments to obtain information and test hypotheses. Experiments with the spread of infectious diseases in human populations are often impossible, unethical or expensive. Data is sometimes available from naturally occurring epidemics or from the natural incidence of endemic diseases; however, the data is often incomplete due to underreporting. This lack of reliable data makes accurate parameter estimation difficult, so that it may only be possible to estimate a range of values for some parameters. Since repeatable experiments and accurate data are usually not available in epidemiology, mathematical models and computer simulations can be used to perform needed theoretical experiments. Calculations can easily be done for a variety of parameter values and data sets.

Comparisons can lead to a better understanding of the processes of disease spread. Modeling can often be used to compare different diseases in the same population, the same disease in different populations, or the same disease at different times. Epidemiological models are useful in comparing the effects of prevention or control procedures. Hethcote and Yorke [89] use models to compare gonorrhea control procedures such as screening, re-screening, tracing infectors, tracing infectees, post-treatment vaccination and general vaccination. Communicable disease models are often the only practical approach to answering questions about which prevention or control procedure is most effective. Quantitative predictions of epidemiological

models are always subject to some uncertainty since the models are idealized and the parameter values can only be estimated. However, predictions of the relative merits of several control methods are often robust in the sense that the same conclusions hold over a broad range of parameter values and a variety of models. Optimal strategies for vaccination can be found theoretically by using modeling.

An under-recognized value of epidemiological modeling is that it leads to a clear statement of the assumptions about the biological and sociological mechanisms which influence disease spread. The parameters used in an epidemiological model must have a clear interpretation such as a contact rate or a duration of infection. Models can be used to assess many quantitative conjectures. Epidemiological models can sometimes be used to predict the spread or incidence of a disease. For example, Hethcote [74] predicted that rubella and Congenital Rubella Syndrome will eventually disappear in the United States because the current vaccination levels using the combined measles–mumps–rubella vaccine are significantly above the threshold required for herd immunity for rubella. An epidemiological model can also be used to determine the sensitivity of predictions to changes in parameter values. After identifying the parameters which have the greatest influence on the predictions, it may be possible to design studies to obtain better estimates of these parameters.

Mathematical models have both limitations and capabilities that must be recognized. Sometimes questions cannot be answered by using epidemiological models, but sometimes the modeler is able to find the right combination of available data, an interesting question and a mathematical model which can lead to the answer.

Table 2 gives fifteen purposes of epidemiological modeling and then three limitations. One cannot infer from the numbers of purposes and limitations that the advantages overwhelm the limitations since the first limitation is very broad. More detailed explanations of these purposes and limitations are given in the paragraphs below.

A primary reason for infectious disease modeling is that it leads to clear statements of the assumptions about the biological and social mechanisms which influence disease spread. The model formulation process is valuable to epidemiologists and modelers because it forces them to be precise about the relevant aspects of transmission, infectivity, recovery and renewal of susceptibles. Epidemiological modelers need to formulate models clearly and precisely using parameters which have well-understood epidemiological interpretations such as a contact rate or an average duration of infection

Table 2. Purposes and limitations of epidemiological modeling.

Fifteen Purposes of Epidemiological Modeling	
1.	The model formulation process clarifies assumptions, variables and parameters.
2.	The behavior of precise mathematical models can be analyzed using mathematical methods and computer simulations.
3.	Modeling allows explorations of the effect of different assumptions and formulations.
4.	Modeling provides concepts such as a threshold, reproduction number, etc.
5.	Modeling is an experimental tool for testing theories and assessing quantitative conjectures.
6.	Models with appropriate complexity can be constructed to answer specific questions.
7.	Modeling can be used to estimate key parameters by fitting data.
8.	Models provide structures for organizing, coalescing and cross-checking diverse pieces of information.
9.	Models can be used in comparing diseases of different types or at different times or in different populations.
10.	Models can be used to theoretically evaluate, compare or optimize various detection, prevention, therapy and control programs.
11.	Models can be used to assess the sensitivity of results to changes in parameter values.
12.	Modeling can suggest crucial data which needs to be collected.
13.	Modeling can contribute to the design and analysis of epidemiological surveys.
14.	Models can be used to identify trends, make general forecasts, or estimate the uncertainty in forecasts.
15.	The validity and robustness of modeling results can be assessed by using ranges of parameter values in many different models.
Three Limitations of Epidemiological Modeling	
1.	An epidemiological model is not reality; it is an extreme simplification of reality.
2.	Deterministic models do not reflect the role of chance in disease spread and do not provide confidence intervals on results.
3.	Stochastic models incorporate chance, but are usually harder to analyze than the corresponding deterministic model.

[72, 78, 79]. Complete statements of assumptions are crucial so that the reasonableness of the model can be evaluated in the process of assessing the conclusions.

An advantage of mathematical modeling of infectious diseases is the economy, clarity and precision of a mathematical formulation. A model using difference, differential, integral or functional differential equations is not ambiguous or vague. Of course, the parameters must be defined precisely and each term in the equations must be explained in terms of mechanisms, but the resulting model is a definitive statement of the basic principles involved. Once the mathematical formulation is complete, there are many

mathematical techniques available for determining the threshold, equilibrium, periodic solutions, and their local and global stability. Thus the full power of mathematics is available for analysis of the equations. Moreover, information about the model can also be obtained by numerical simulation on digital computers of the equations in the model. The mathematical analyses and computer simulations can identify important combinations of parameters and essential aspects or variables in the model. In order to choose and use epidemiological modeling effectively on specific diseases, one must understand the behavior of the available formulations and the implications of choosing a particular formulation. Thus mathematical epidemiology provides a foundation for the applications [83, 82].

Modelers are able to explore and examine the effects of different assumptions and formulations. For example, they can compare an ordinary differential equations model, which corresponds to a negative exponential distribution for the infectious period, with a delay-differential equations model, which assumes that everyone has the same fixed, constant-length infectious period. They can examine the impact of assuming homogeneous mixing instead of heterogeneous mixing. They can study the influence of different mixing patterns between groups on the spread of a disease in a heterogeneous population [104, 24, 79]. They can decide if models with and without an exposed class for people in the latent period behave differently. The results of exploring different formulations may provide insights for epidemiologists and are certainly useful to modelers who are choosing models for specific diseases. For example, if the essential behaviors (thresholds and asymptotic behaviors) are the same for a formulation using ordinary differential equations and a formulation using delay differential equations, then the modeler would probably choose to use the ordinary differential equations model since it is simpler.

Epidemiologists need to understand the concept of a threshold which determines whether the disease persists or dies out. They need to know that the reproduction number is a combination of parameters which gives the number of secondary cases “reproduced” by a typical infective during the infectious period in a population where everyone is susceptible. They need to be aware that this reproductive number is the same as the contact number, which is the average number of adequate contacts of an infective person during the infectious period [84]. They need to realize that the average replacement number is one at an endemic equilibrium because, in an average sense, each infective replaces itself by one new infective in an equilibrium situation. Probably the most valuable contributions of mathematical model-

ing to epidemiologists are concepts like those above. Mathematical models serve as a framework to explain causes, relationships and ideas. Mathematical models are very useful in obtaining conclusions that have an easily understood interpretation. Helping epidemiologists internalize conceptual and intuitive understandings of disease processes, mechanisms and modeling results is an essential aspect of the work of mathematical modelers.

Epidemiological modeling is an important part of the epidemiologist's function to build and test theories. Mathematical and computer simulation models are the fundamental experimental tools in epidemiology. The only data usually available are from naturally occurring epidemics or from the natural incidence of endemic diseases; unfortunately, even these data are not complete since many cases are not reported. Since repeatable experiments and accurate data are usually not available in epidemiology, mathematical and computer simulation models must be used to perform necessary theoretical experiments with different parameter values and different data sets. It is easy in a computer simulation to find out what happens when one or several parameters are changed.

Another advantage of epidemiological modeling is the availability of models to assess quantitative conjectures. For example, models can be used to check the claim that a two-dose vaccination program for measles would lead to herd immunity whereas a one-dose program would not [74]. With a model one could see if the AIDS virus would eventually disappear in a population of homosexual men if half of them became celibate or practiced safer sex. Mathematical models are a formal, quantitative way of building and testing theories.

When formulating a model for a particular disease, it is necessary to decide which factors to include and which to omit. This choice often depends on the particular question that is to be answered. Simple models have the advantage that there are only a few parameters, but have the disadvantage of possibly being naive and unrealistic. Complex models may be more realistic, but they may contain many parameter values for which estimates cannot be obtained. The art of epidemiological modeling is to make suitable choices in the model formulation so that it is as simple as possible and yet is adequate for the question being considered. It is important to recognize both the capabilities and limitations of epidemiological modeling. Many important questions cannot be answered using a given class of models. The most difficult problem for a modeler is to find the right combination of available data, an interesting question and a mathematical model which can lead to the answer.

By fitting the incidence predicted by a model to actual incidence data (number of new cases per month or week or day), it is possible to estimate key parameters such as average durations, basic reproduction numbers, contact numbers, and replacement numbers. This is done for the three basic models covered in this chapter. Hethcote and Van Ark [84] present methods for estimating contact numbers for the groups in a multigroup model for disease transmission. A model can serve as the conceptual and quantitative unifier of all the available information. If parameters have already been estimated from the literature, then the modeling provides a check on these *a priori* estimates. Thus the model can be a way of coalescing parameter estimates and other pieces of data to see if they fit into a consistent framework.

Comparisons can lead to a better understanding of the process of disease spread. It may be enlightening to compare the same disease at different times, the same disease in different populations or different diseases in the same population. One method for comparing diseases is to estimate parameter values for the diseases and then compare the parameter values. This is done in Table 4 in Section 9 for various diseases such as measles, whooping cough, chicken pox, diphtheria, mumps, rubella, poliomyelitis and smallpox.

A very important reason for epidemiological modeling is the value of models for theoretical evaluations and comparisons of detection, prevention, therapy and control programs. Epidemiologists and politicians need to understand the effects of different policy decisions. Qualitative predictions of infectious disease models are always subject to some uncertainty since the models are idealized and the parameter values can only be estimated. However, quantifications of the relative merits of several control methods or of control versus no control are often robust in the sense that the same conclusions hold for a broad range of parameter values and a variety of models. Control strategies for gonorrhea such as screening, rescreening, infectee tracing, infector tracing and potential vaccines for both women and men were compared in Hethcote and Yorke [89]. Various vaccination strategies for rubella and measles have been compared using modeling [74, 77, 4].

Sometimes an optimization method can be used on some aspect of an epidemiological model. Hethcote and Waltman [87] used dynamic programming to find an optimal vaccination strategy to control an epidemic at least cost. Longini *et al.* [120] determined the best age or social groups to vaccinate for Hong Kong and Asian influenza if there were a limited amount of vaccine available. Hethcote [75] found theoretical optimal ages of vaccinations for measles in three geographic locations. A primary conclusion of

that paper was that better data is needed on vaccine efficacy as a function of age in order to better estimate the optimal age of vaccination. Thus epidemiological modeling can be used to identify crucial data that needs to be collected. See Wickwire [157] for a survey of modeling papers on the control of infectious diseases.

Predictions or quantitative conclusions of a model are said to be sensitive to a parameter if a small change in the parameter causes a large change in the outcome. A model is insensitive to a parameter if the outcomes are the same for a wide range of values of that parameter. An important reason for epidemiological modeling is the determination of the sensitive and insensitive parameters since this information provides insight into the epidemiological processes. Once the sensitive parameters are identified, it may be possible to make special efforts to gather data to obtain better estimates of these parameters. Thus the modeling process can help identify crucial data that should be collected. If the data and information are insufficient to estimate parameters in a model or to piece everything together, then the model formulation process can identify the missing information which must be gathered.

Modeling can enhance an epidemiological survey by providing a framework for the analysis of the survey results. In the design phase modeling can help identify important questions that need to be asked in order to have adequate data to obtain results. The statistical aspects of a survey design are now analyzed before the survey is conducted. Similarly, those design aspects relevant to parameter estimation in models should be analyzed before the survey is conducted to see if all significant data are being collected.

Another purpose of epidemiological modeling is to make forecasts about the future incidence of a disease. Although people often think of prediction as the first or only purpose of epidemiological modeling, the purposes given above are more important. Accurate forecasts are usually not possible because of the idealization in the models and the uncertainty in the parameter values. However, possible forecasts under various scenarios can sometimes be given or trends can be identified. One purpose of epidemiological modeling is to reduce the uncertainty in future incidence projections.

The validation of models and modeling results is difficult since there are rarely enough good data to adequately test or compare different models with data [7]. However, when the same concept or result is obtained for a variety of models and by several modelers, then the confidence in the validity of the result is increased. For example, three different papers

using different approaches to comparing two rubella vaccination strategies agree on the criteria determining situations when each strategy is best [77]. Similarly, when a quantitative result holds for a wide range of parameter values, then this result has some robustness. For example, if the same relative values of a variety of control procedures are obtained for wide ranges of parameter values, then the relative values may be robust even though the absolute values may vary [89].

After discussing the purposes and advantages of epidemiological modeling, it is also imperative to discuss the limitations. Epidemiologists and policy makers need to be aware of both the strengths and weaknesses of the epidemiological modeling approach.

The first and most obvious limitation is that all epidemiological models are simplifications of reality. For example, it is often assumed that the population is uniform and homogeneously mixing; this is always a simplification, but the deviation of reality from this simplification varies with the disease and the circumstances. This deviation from reality is rarely testable or measurable; however, it can sometimes be estimated intuitively from an understanding of the disease epidemiology. As described in the introduction, part of the reason for the differences between the model's behavior and the actual spread of the disease is that transmission is based on the interactions of human beings, and *homo sapiens* have highly variable behavior and interactions. People do not behave in reasonably predictable ways like molecules or cells or particles.

Because transmission models are simplifications with usually unknown relationships to actual disease spread, one can never be completely certain about the validity of findings obtained from modeling such as conceptual results, experimental results, answers to questions, comparisons, sensitivity results and forecasts. Even when models are made more complicated in order to better approximate actual disease transmission, they are still abstractions. For example, multigroup models based on risk groups or sexual activity levels or age structure still assume that the distinctions between groups are clear and sharp and that the mixing between groups follows some assumed pattern. The modeler must always exercise judgment in deciding which factors are relevant and which are not when analyzing a specific disease or question.

The modeler needs to know that an SIR model is usually adequate since the behavior is essentially the same for the SEIR model where E is the latent period class. The notation SEIR for a model means that susceptibles (S) move to an exposed class (E) (when they are in a latent period; i.e., they

are infected, but not yet infectious), then move to an infectious class (I), and finally to a removed class (R) when they recover and have permanent immunity. The modeler must realize that assuming that the infectivity is constant during the infectious period is reasonable for diseases like measles with a short infectious period (about 8 days), but is not reasonable for a disease like the human immunodeficiency virus (HIV) where the average time from HIV infection to AIDS is about 10 years. To incorporate age structure, the modeler must decide whether to use a model with time and age as continuous variables or to use discrete compartments corresponding to age brackets; the former requires the estimation of birth, death and aging rate coefficients as functions of time while the latter requires the estimation of birth, death and transfer rate constants. The modeler needs to decide whether the disease is epidemic in a short time period so that vital dynamics (births and deaths) can be excluded, or else is endemic over a long time period (many years) so that vital dynamics must be included. The modeler must be aware that as the complexity of a model is increased so that it better approximates reality, the number of parameters increases so that it becomes increasingly difficult to estimate values of all of these parameters. Thus the modeler must (consciously or unconsciously) make many decisions regarding the relevant aspects in choosing a model for a specific disease or question.

When describing the purposes for epidemiological modeling, we have been thinking primarily of deterministic models, but the purposes also apply to stochastic epidemiological models. Deterministic models are those which use difference, differential, integral or functional differential equations to describe the changes in time of the sizes of the epidemiological classes. Given the starting conditions for a well-posed deterministic model, the solutions as a function of time are unique. In stochastic models, there are probabilities at each time step of moving from one epidemiological class to another. When these models are simulated with the probabilities calculated using random number generators, the outcomes of different runs are different so that this approach is often called Monte Carlo simulation; conclusions are obtained by averaging the results of many computer simulations. Simple deterministic models for epidemics have a precise threshold which determines whether an epidemic will or will not occur. In contrast, stochastic models for epidemics yield quantities such as the probability that an epidemic will occur and the mean time to extinction of the disease. Thus the approach, concepts and appropriate questions can be quite different for stochastic models [9, 11, 39, 43, 49, 62, 100, 134].

Both deterministic and stochastic epidemiological models have other limitations besides being mere approximations of reality. Deterministic models do not reflect the role of chance in disease spread. Sometimes parameter values in deterministic models are set equal to the mean of observed values and the information on the variance is ignored. A set of initial conditions leads to exactly one solution in a deterministic model; thus no information is available on the reliability of or the confidence in the results. When quantities such as the basic reproduction number, contact number, replacement number, average infectious period, etc. are estimated by fitting output to data, confidence intervals on these estimates are not obtained. Confidence intervals also cannot be found for comparisons and forecasts. Some understanding of the dependence on parameter values is obtained through a sensitivity analysis where the effects of changes in parameter values on results are found. If the variance of the observed value of a parameter is high and the results are sensitive to that parameter, then the confidence in the results would be low. However, the lack of precise confidence intervals implies that the reliability of the results is uncertain.

Stochastic epidemiological models incorporate chance, but it is usually harder to get analytic results for these models. Moreover, computational results are also harder since Monte Carlo simulations require many computer runs (25 to 100 or more) in order to detect patterns and get quantitative results. Even for stochastic epidemiological models where parameters are estimated by fitting the mean of the simulations to data, it may not be possible to find confidence intervals on these parameter estimates.

1.3. Definitions, assumptions and model formulations

The study of disease occurrence is called epidemiology. An epidemic is an unusually large, short term outbreak of a disease. A disease is called endemic if it persists in a population. Thus epidemic models are used to describe rapid outbreaks that occur in less than one year, while endemic models are used for studying diseases over longer periods, during which there is a renewal of susceptibles by births or recovery from temporary immunity. The spread of an infectious disease involves not only disease-related factors such as the infectious agent, mode of transmission, latent period, infectious period, susceptibility and resistance, but also social, cultural, demographic, economic and geographic factors. The three basic models considered here are the simplest prototypes of three different types of epidemiological models. It is important to understand their behavior before considering general models incorporating more of the factors above.

The population under consideration is divided into disjoint classes whose sizes change with time t . In the basic epidemiological models presented here, it is assumed that the population has constant size N , which is sufficiently large so that the sizes of each class can be considered as continuous variables. Models with variable population size are considered in Chapter 2. If the model is to include vital dynamics, then it is assumed that births and natural deaths occur at equal rates and that all newborns are susceptible. The three basic models assume that the population is homogeneously mixing, but models for heterogeneously mixing populations have been formulated and analyzed [84, 79]. A basic concept in epidemiology is the existence of thresholds; these are critical values for quantities such as the basic reproduction number, contact number, population size, or vector density that must be exceeded in order for an epidemic to occur or for a disease to remain endemic. The formulations used here are somewhat different from the most classical formulations [9], since they involve the fractions of the populations in the classes instead of the numbers in the classes. These formulations have much more intuitive threshold conditions involving basic reproduction numbers or contact numbers instead of population sizes. See [72, 84] for further comparisons of formulations in terms of proportions and numbers in the classes.

Compartments with labels such as M, S, E, I, and R are often used for the epidemiological classes as shown in Figure 1. If a mother has been infected, then some IgG antibodies are transferred across the placenta, so that her newborn infant has temporary passive immunity to an infection. The class M contains these infants with passive immunity. After the maternal antibodies disappear from the body, the infant moves to the susceptible class S. Infants who do not have any passive immunity, because their mothers were never infected, also enter the class S of susceptible individuals; that is, those who can become infected. When there is an adequate contact of a susceptible with an infective so that transmission occurs, then the susceptible enters the exposed class E of those in the latent period; they are infected, but not yet infectious. After the latent period ends, the individual enters the class I of infectives, who are infectious in the sense that they are capable of transmitting the infection. When the infectious period ends, the individual enters the recovered class R consisting of those with permanent infection-acquired immunity.

The choice of which compartments to include in a model depends on the characteristics of the particular disease being modeled and the purpose of the model. The passively immune class M and the latent period

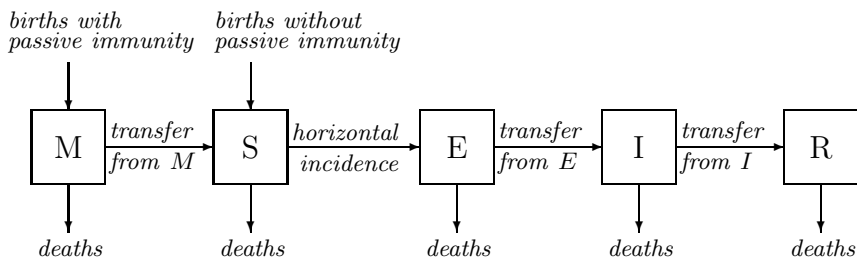


Fig. 1. The general transfer diagram for the MSEIR model with the passively-immune class M, the susceptible class S, the exposed class E, the infective class I, and the recovered class R.

class E are often omitted, because they are not crucial for the susceptible–infective interaction. Acronyms for epidemiology models are often based on the flow patterns between the compartments such as MSEIR, MSEIRS, SEIR, SEIRS, SIR, SIRS, SEI, SEIS, SI, and SIS. For example, in the MSEIR model shown in Figure 1, passively immune newborns first become susceptible, then exposed in the latent period, then infective, and then removed with permanent immunity. An MSEIRS model would be similar, but the immunity in the R class would be temporary, so that individuals would regain their susceptibility when the temporary immunity ends. An MSEIRS endemic model in a population with constant size is formulated and analyzed in Section 7.2 of this chapter.

The simplest model in which recovery does not give immunity is the SIS model, since individuals move from the susceptible class to the infective class and then back to the susceptible class upon recovery. If individuals recover with permanent immunity, then the simplest model is an SIR model. If individuals recover with temporary immunity so that they eventually become susceptible again, then the simplest model is an SIRS model [72, 83]. If individuals do not recover, then the simplest model is an SI model. In general, SIR models are appropriate for viral agent diseases such as measles, mumps, and smallpox, while SIS models are appropriate for some bacterial agent diseases such as meningitis, plague, and sexually transmitted diseases, and for protozoan agent diseases such as malaria and sleeping sickness (see Table 1).

The three most basic deterministic models for infectious diseases which are spread by direct person-to-person contact in a population are the SIS endemic, the SIR epidemic, and the SIR endemic models. We use these three models to illustrate the essential results and then discuss their extensions to the more general models listed in the previous paragraph. These

simplest models are formulated as initial value problems for systems of ordinary differential equations and are analyzed mathematically. The presentation here provides a sound intuitive understanding for the three most basic epidemiological models for microparasitic infections. Theorems are stated regarding the asymptotic stability regions for the equilibrium points, and phase plane portraits of solution paths are presented. Parameters are estimated for various diseases and are used to compare the vaccination levels necessary for herd immunity for these diseases. Although the three models presented are simple and their mathematical analyses are elementary, these models provide notation, concepts, intuition and foundation for considering more refined models. Other possible refinements are described briefly in the last section.

1.3.1. *Formulating epidemiology models*

The horizontal incidence shown in Figure 1 is the infection rate of susceptible individuals through their contacts with infectives, so it is the number of new cases per unit time. If $S(t)$ is the number of susceptibles at time t , $I(t)$ is the number of infectives, and N is the total population size, then $s(t) = S(t)/N$ and $i(t) = I(t)/N$ are the susceptible and infectious fractions, respectively. Let β be the average number of adequate contacts (i.e. contacts sufficient for transmission) of a person per unit time. Here we assume that the contact rate β is fixed and does not depend on the population size N or vary seasonally. The type of direct or indirect contact necessary to be adequate for transmission depends on the specific disease. From these definitions we see that $\beta I/N = \beta i$ is the average number of contacts with infectives per unit time of one susceptible, and $(\beta I/N)S = \beta N i s$ is the number of new cases per unit time amongst the $S = Ns$ susceptibles. We call this form of incidence function the standard incidence, because it is formulated from the basic principles above [72, 78]. Because this form of the incidence depends on the relative frequency I/N of infectives, $(\beta I/N)S$ is sometimes called frequency-dependent incidence [88].

The mass action law $\eta IS = \eta(Ni)(Ns)$, with η as a mass action coefficient, has sometimes been used for the horizontal incidence. The parameter η has no direct epidemiological interpretation, but comparing it with the standard formulation shows that $\beta = \eta N$, so that this form implicitly assumes that the contact rate β increases linearly with the population size. Naively, it might seem plausible that the population density and hence the contact rate would increase with population size, but the daily contact

patterns of people are often similar in large and small communities, cities, and regions. For human diseases the contact rate seems to be only very weakly dependent on the population size. Using an incidence of the form $\eta N^v SI/N$, data for five human diseases in communities with population sizes from 1,000 to 400,000 ([6], p. 157, [7], p. 306) implies that v lies between 0.03 and 0.07. This strongly suggests that the standard incidence corresponding to $v = 0$ is more realistic for human diseases than the simple mass action incidence corresponding to $v = 1$. This result is consistent with the concept that people are infected through their daily encounters and the patterns of daily encounters are largely independent of community size within a given country (e.g. students of the same age in a country usually have a similar number of daily contacts).

The standard incidence is also a better formulation than the simple mass action law for animal populations such as mice in a mouse-room or animals in a herd [40], because disease transmission primarily occurs locally from nearby animals. For more information about the differences in models using these two forms of the horizontal incidence, see [63, 66, 67, 72, 84, 128]. See [88] for a recent very detailed comparison of the standard (frequency-dependent) and mass action incidences. Vertical incidence, which is the infection rate of newborns by their mothers, is sometimes included in epidemiology models by assuming that a fixed fraction of the newborns are infected vertically [22]. Various forms of nonlinear incidences have been considered [86, 117, 118, 119]. See [82] for a survey of mechanisms including nonlinear incidences that can lead to periodicity in epidemiological models.

A common assumption is that the movements out of the M, E, and I compartments and into the next compartment are governed by terms like δM , ϵE , and γI in an ordinary differential equations model. It has been shown [83] that these terms correspond to exponentially distributed waiting times in the compartments. For example, the transfer rate γI corresponds to $P(t) = e^{-\gamma t}$ as the fraction that is still in the infective class t units after entering this class and $1/\gamma$ as the mean waiting time. For measles the mean period $1/\delta$ of passive immunity is about 6 to 9 months, while the mean latent period $1/\epsilon$ is 1–2 weeks and the mean infectious period $1/\gamma$ is about 1 week. Another possible assumption is that the fraction still in the compartment t units after entering is a non-increasing, piecewise continuous function $P(t)$ with $P(0) = 1$ and $P(\infty) = 0$. Then the rate of leaving the compartment at time t is $-P'(t)$ so the mean waiting time in the compartment is $\int_0^\infty t(-P'(t))dt = \int_0^\infty P(t)dt$. These distributed delays lead to epidemiology models with integral or integro-differential or functional

differential equations. If the waiting time distribution is a step function given by $P(t) = 1$ if $0 \leq t \leq \tau$, and $P(t) = 0$ if $\tau < t$, then the mean waiting time is τ and for $t \geq \tau$ the model reduces to a delay-differential equation [83]. Each waiting time in a model can have a different distribution, so there are many possible models [78].

Table 3. Summary of notation.

symbol	quantity (number, fraction, rate, or period)
M	passively-immune infants
S	susceptibles
E	exposed people in the latent period
I	infectives
R	recovered people with immunity
m, s, e, i, r	fractions of the population in the classes above
β	contact rate
$1/\delta$	average period of passive immunity
$1/\varepsilon$	average latent period
$1/\gamma$	average infectious period
R_0	basic reproduction number
σ	contact number
R	replacement number

1.3.2. Three threshold quantities: R_0 , σ , and R

Recall that the basic reproduction number R_0 is defined as the average number of secondary infections that occur when one infective is introduced into a completely susceptible host population [45]. Note that R_0 is also called the basic reproduction ratio [42] or basic reproductive rate [7]. It is implicitly assumed that the infected outsider is in the host population for the entire infectious period and mixes with the host population in exactly the same way that a population native would mix. The contact number σ is defined as the average number of adequate contacts of a typical infective during the infectious period [72, 84]. An adequate contact is one that is sufficient for transmission, if the individual contacted by the susceptible is an infective. The replacement number R is defined to be the average number of secondary infections produced by a typical infective during the entire period of infectiousness [72]. Note that the replacement number R changes as a function of time t as the disease evolves after the initial invasion. Some authors use the term reproduction number instead of replacement number, but it is better to avoid the name reproduction number since it is easily confused with the basic reproduction number. Note that these

three quantities R_0 , σ , and R are all equal at the beginning of the spread of an infectious disease when the entire population (except the infective invader) is susceptible. In recent epidemiological modeling literature, the basic reproduction number R_0 is often used as the threshold quantity which determines whether a disease can invade a population.

Although R_0 is only defined at the time of invasion, σ and R are defined at all times. For most models, the contact number σ remains constant as the infection spreads, so it is always equal to the basic reproduction number R_0 . In these models σ and R_0 can be used interchangeably and invasion theorems can be stated in terms of either quantity. But for some pertussis models [81], the contact number σ becomes less than the basic reproduction number R_0 after the invasion, because new classes of re-infected people with lower infectivity appear after the disease has entered the population. The replacement number R is the actual number of secondary cases from a typical infective, so that after the infection has invaded a population and everyone is no longer susceptible, R is always less than the basic reproduction number R_0 . Also after the invasion, the susceptible fraction is less than one, so that not all adequate contacts result in a new case. Thus the replacement number R is always less than the contact number σ after the invasion. Combining these results leads to

$$R_0 \geq \sigma \geq R,$$

with equality of the three quantities at the time of invasion. Note that $R_0 = \sigma$ for most models, and $\sigma > R$ after the invasion for all models.

1.4. The basic SIS endemic model

The basic SIS model is the simplest model for diseases in which an infection does not confer immunity. Thus susceptibles become infected and then become susceptible again upon recovery. It is an endemic model because the disease persists, in contrast to the epidemic model in the next section in which the disease eventually dies out. The SIS model with vital dynamics (births and deaths) is given by

$$\begin{aligned} dS/dt &= \mu N - \mu S - \beta IS/N + \gamma I, & S(0) &= S_o > 0, \\ dI/dt &= \beta IS/N - \gamma I - \mu I, & I(0) &= I_o > 0, \end{aligned} \tag{1.4.1}$$

with $S(t) + I(t) = N$. This SIS model has an inflow of newborns into the susceptible class at rate μN and deaths in the classes at rates μS and μI . The deaths balance the births, so that the population size N is constant.

The mean lifetime $1/\mu$ would be about 75 years in the United States. This model is appropriate for bacterial agent diseases such as gonorrhea, meningitis and streptococcal sore throat.

Dividing the equations in (1.4.1) by the constant total population size N and using $i(t) = I(t)/N$ and $s(t) = S(t)/N = 1 - i(t)$ yields

$$di/dt = \beta i(1 - i) - (\gamma + \mu)i, \quad i(0) = i_o > 0. \quad (1.4.2)$$

Note that this equation involves the infected fraction, but does not involve the population size N . Here only nonnegative solutions are considered since negative solutions have no epidemiological significance. Since (1.4.2) is a Bernoulli differential equation, the substitution $y = i^{-1}$ converts it into a linear differential equation from which the unique solution is found to be

$$i(t) = \begin{cases} \frac{e^{(\gamma+\mu)(\sigma-1)t}}{\sigma[e^{(\gamma+\mu)(\sigma-1)t} - 1]/(\sigma-1) + 1/i_o} & \text{for } \sigma \neq 1, \\ \frac{1}{\beta t + 1/i_o} & \text{for } \sigma = 1, \end{cases} \quad (1.4.3)$$

where $\sigma = \beta/(\gamma + \mu)$ is the contact number, which is the same as the basic reproduction number R_0 . The theorem below follows directly from the explicit solution.

Theorem 1: The solution $i(t)$ of (1.4.2) approaches 0 as $t \rightarrow \infty$ if $R_0 = \sigma \leq 1$ and approaches $1 - 1/\sigma$ as $t \rightarrow \infty$ if $R_0 = \sigma > 1$.

This theorem means that for a disease without immunity with any positive initial infective fraction, the infective fraction approaches a constant endemic value if the contact number exceeds 1; otherwise, the disease dies out. Here the contact number σ remains equal to the basic reproduction number R_0 for all time, because no new classes of susceptibles or infectives occur after the invasion. For this model the threshold quantity is given by $R_0 = \sigma = \beta/(\gamma + \mu)$, which is the contact rate β times the average death-adjusted infectious period $1/(\gamma + \mu)$, and the critical threshold value is 1. Although the model (1.4.1) reduces to a one-dimensional initial value problem (1.4.2), we describe the behavior in the (s, i) phase plane, so that it can be compared with the phase planes for the other two basic models. For this SIS model solution paths stay on the line $s + i = 1$ and all paths approach the point $(s, i) = (1/\sigma, 1 - 1/\sigma)$ when $R_0 = \sigma > 1$ and approach the point $(s, i) = (1, 0)$ when $R_0 = \sigma \leq 1$.

Note that the replacement number $R = \sigma s = R_0 s$ is 1 at the endemic equilibrium point, since $s = 1/\sigma = 1/R_0$ when $i = 1 - 1/\sigma$. This is plausible, since the prevalence would increase if the replacement number were greater than 1 and it would decrease if the replacement number were less than 1.

A threshold result for an SI model is obtained from the theorem above by taking the removal rate γ to be zero in the model. If both the removal rate γ and the birth and death rate μ are zero, then $\sigma = \beta/(\gamma + \mu) = \infty$, so that there is no threshold and eventually everyone is infected. This SI model with $\gamma + \mu = 0$ is called the “simple epidemic model” considered in Bailey ([9], p. 22).

The prevalence is defined as the number of cases of a disease at a given time so that it corresponds to I , while the fractional prevalence would correspond to i . Since the incidence is defined to be the number of new cases per unit time, it corresponds to the term $\beta IS/N$ in model (1.4.1). At an endemic equilibrium the prevalence is equal to the incidence times the average duration of infection $1/(\gamma + \mu)$ since the right side of the second equation in (1.4.1) is zero at an equilibrium.

The incidence and the prevalence of some diseases oscillate seasonally in a population. This oscillation seems to be caused by seasonal oscillation in the contact rate β . For example, the incidence of childhood diseases such as measles and rubella increases each year in the winter when children aggregate in schools [119, 159, 46]. If the contact rate β changes with time t , then β in the models above is replaced by $\beta(t)$. Hethcote [71] gives the asymptotic behavior of solutions $i(t)$ of (1.4.2) when $\beta(t)$ is periodic with period p . When the average contact number $\bar{\sigma} = \bar{\beta}/(\gamma + \mu)$ satisfies $\bar{\sigma} \leq 1$, $i(t)$ damps in an oscillatory manner to 0 for large t . However, if $\bar{\sigma} > 1$, then $i(t)$ approaches an explicit periodic solution for large t . See [71, 76] for figures illustrating these two cases. Gonorrhea is an example of a disease for which infection does not confer immunity, so it is of SIS type. See [89] for graphs of the seasonal oscillation of reported cases of gonorrhea from 1946 to 1984. Numerous models for gonorrhea transmission dynamics and control including a seasonal oscillation model are presented in Hethcote and Yorke [89].

1.5. The basic SIR epidemic model

The basic SIR epidemic model is given by the initial value problem

$$\begin{aligned} dS/dt &= -\beta IS/N, & S(0) &= S_o \geq 0, \\ dI/dt &= \beta IS/N - \gamma I, & I(0) &= I_o \geq 0, \\ dR/dt &= \gamma I, & R(0) &= R_o \geq 0, \end{aligned} \quad (1.5.1)$$

where $S(t)$, $I(t)$, and $R(t)$ are the numbers in these classes, so that $S(t) + I(t) + R(t) = N$. This SIR model is a special case of the MSEIR model in Figure 1, in which the passively immune class M and the exposed class E are omitted. This model uses the standard incidence and has recovery at

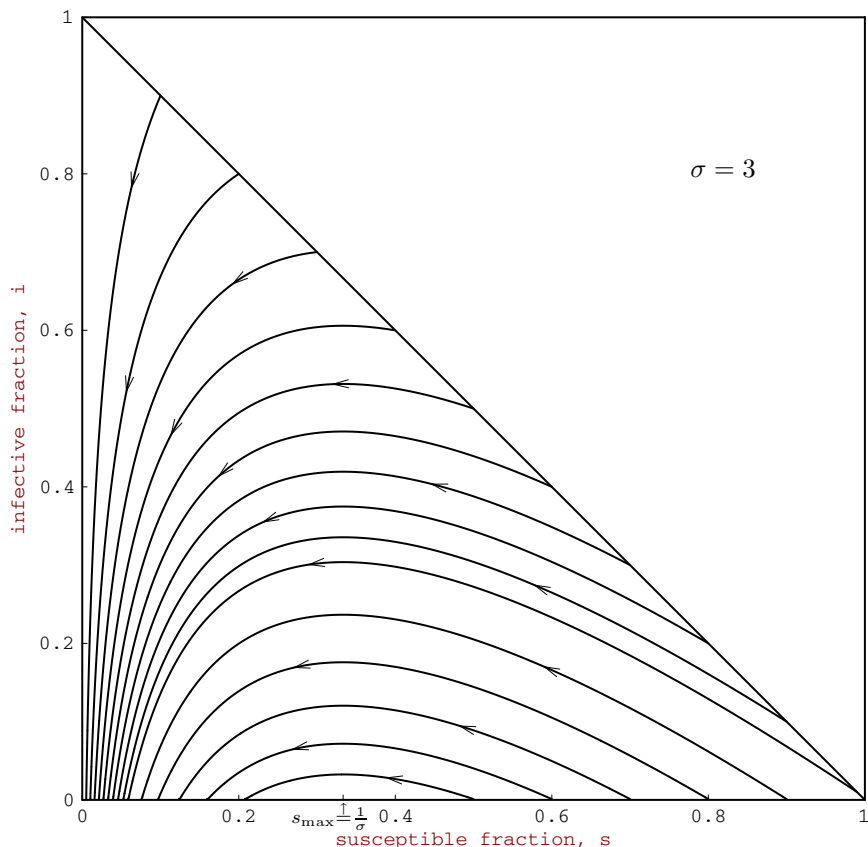


Fig. 2. Phase plane portrait for the classic SIR epidemic model with contact number $\sigma = 3$.

rate γI , corresponding to an exponential waiting time $e^{-\gamma t}$. Since the SIR epidemic model is used for outbreaks occurring in a short time period, this model has no vital dynamics (births and deaths). Dividing the equations in (1.5.1) by the constant total population size N yields

$$\begin{aligned} ds/dt &= -\beta is, & s(0) &= s_o \geq 0, \\ di/dt &= \beta is - \gamma i, & i(0) &= i_o \geq 0, \end{aligned} \quad (1.5.2)$$

with $r(t) = 1 - s(t) - i(t)$, where $s(t)$, $i(t)$, and $r(t)$ are the fractions in the classes. The triangle T in the (s, i) phase plane given by

$$T = \{(s, i) \mid s \geq 0, i \geq 0, s + i \leq 1\} \quad (1.5.3)$$

is positively invariant and unique solutions exist in T for all positive time, so that the model is mathematically and epidemiologically well posed [72]. Here the contact number $\sigma = \beta/\gamma$ is the contact rate β per unit time multiplied by the average infectious period $1/\gamma$, so it has the proper interpretation as the average number of adequate contacts of a typical infective during the infectious period. Here the replacement number at time zero is σs_o , which is the product of the contact number σ and the initial susceptible fraction s_o .

Theorem 2: Let $(s(t), i(t))$ be a solution of (1.5.2) in T . If $\sigma s_o \leq 1$, then $i(t)$ decreases to zero as $t \rightarrow \infty$. If $\sigma s_o > 1$, then $i(t)$ first increases up to a maximum value $i_{\max} = i_o + s_o - 1/\sigma - \lfloor \ln(\sigma s_o) \rfloor / \sigma$ and then decreases to zero as $t \rightarrow \infty$. The susceptible fraction $s(t)$ is a decreasing function and the limiting value s_∞ is the unique root in $(0, 1/\sigma)$ of the equation

$$i_o + s_o - s_\infty + \ln(s_\infty/s_o)/\sigma = 0. \quad (1.5.4)$$

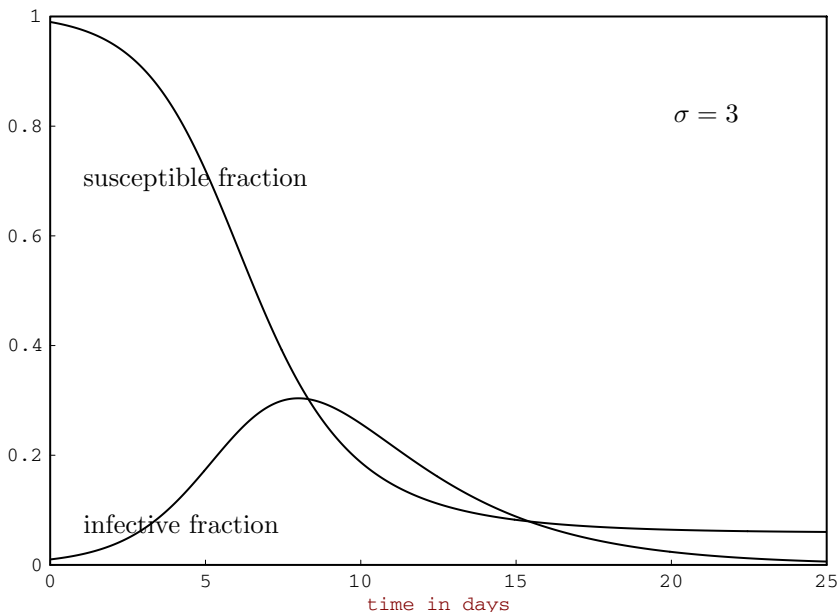


Fig. 3. Solutions of the classic SIR epidemic model with contact number $\sigma = 3$ and average infectious period $1/\gamma = 3$ days.

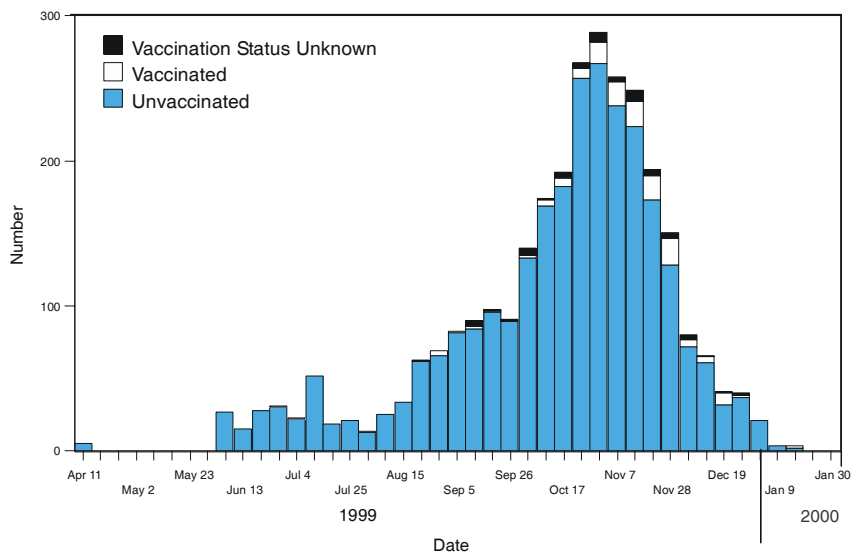


Fig. 4. Reported number of measles cases in the Netherlands by week of onset and vaccination status during April 1999–January 2000. Most of the unvaccinated cases were people belonging to a religious denomination that routinely does not accept vaccination. The 2961 measles cases included 3 measles-related deaths.

Typical paths in T are shown in Figure 2, and solutions as a function of time are shown in Figure 3. Note that the hallmark of a typical epidemic outbreak is an infective curve that first increases from an initial i_o near zero, reaches a peak, and then decreases towards zero as a function of time. For example, a recent measles epidemic in the Netherlands [35] is shown in Figure 4. The susceptible fraction $s(t)$ always decreases, but the final susceptible fraction s_∞ is positive. The epidemic dies out because, when the susceptible fraction $s(t)$ goes below $1/\sigma$, the replacement number $\sigma s(t)$ goes below one. The results in the theorem are epidemiologically reasonable, since the infectives decrease and there is no epidemic, if enough people are already immune so that a typical infective initially replaces itself with no more than one new infective ($\sigma s_o \leq 1$). But if a typical infective initially replaces itself with more than one new infective ($\sigma s_o > 1$), then infectives initially increase so that an epidemic occurs. The speed at which an epidemic progresses depends on the characteristics of the disease. See [76] for more examples of epidemic outbreak curves.

To prove the theorem, observe that the solution paths

$$i(t) + s(t) - [\ln s(t)]/\sigma = i_o + s_o - [\ln s_o]/\sigma$$

in Figure 2 are found from the quotient differential equation $di/ds = -1 + 1/(\sigma s)$. The equilibrium points along the s axis are neutrally unstable for $s > 1/\sigma$ and are neutrally stable for $s < 1/\sigma$. For a complete (easy) proof, see [72] or [76]. One classic approximation derived in [9] is that for small i_o and s_o slightly greater than $s_{\max} = 1/\sigma$, the difference $s_{\max} - s(\infty)$ is about equal to $s_o - s_{\max}$, so the final susceptible fraction is about as far below the susceptible fraction s_{\max} (the s value where the infective fraction is a maximum) as the initial susceptible fraction was above it (see Figure 2). Observe that the threshold result here involves the initial replacement number $R = \sigma s_o$ and does not involve the basic reproduction number R_0 .

1.6. The basic SIR endemic model

The classic endemic model is the SIR model with vital dynamics (births and deaths) given by

$$\begin{aligned} dS/dt &= \mu N - \mu S - \beta IS/N, & S(0) &= S_o \geq 0, \\ dI/dt &= \beta IS/N - \gamma I - \mu I, & I(0) &= I_o \geq 0, \\ dR/dt &= \gamma I - \mu R, & R(0) &= R_o \geq 0, \end{aligned} \quad (1.6.1)$$

with $S(t) + I(t) + R(t) = N$. This SIR model is almost the same as the SIR epidemic model (1.5.1) above, except that it has an inflow of newborns into the susceptible class at rate μN and deaths in the classes at rates μS , μI , and μR . The deaths balance the births, so that the population size N is constant. Dividing the equations in (1.6.1) by the constant total population size N yields

$$\begin{aligned} ds/dt &= -\beta is + \mu - \mu s, & s(0) &= s_o \geq 0, \\ di/dt &= \beta is - (\gamma + \mu)i, & i(0) &= i_o \geq 0, \end{aligned} \quad (1.6.2)$$

with $r(t) = 1 - s(t) - i(t)$. The triangle T in the (s, i) phase plane given by (1.5.3) is positively invariant, and the model is well posed [72]. Here the contact number σ remains equal to the basic reproduction number R_0 for all time, because no new classes of susceptibles or infectives occur after the invasion. For this model the threshold quantity is given by $R_0 = \sigma = \beta/(\gamma + \mu)$, which is the contact rate β times the average death-adjusted infectious period $1/(\gamma + \mu)$.

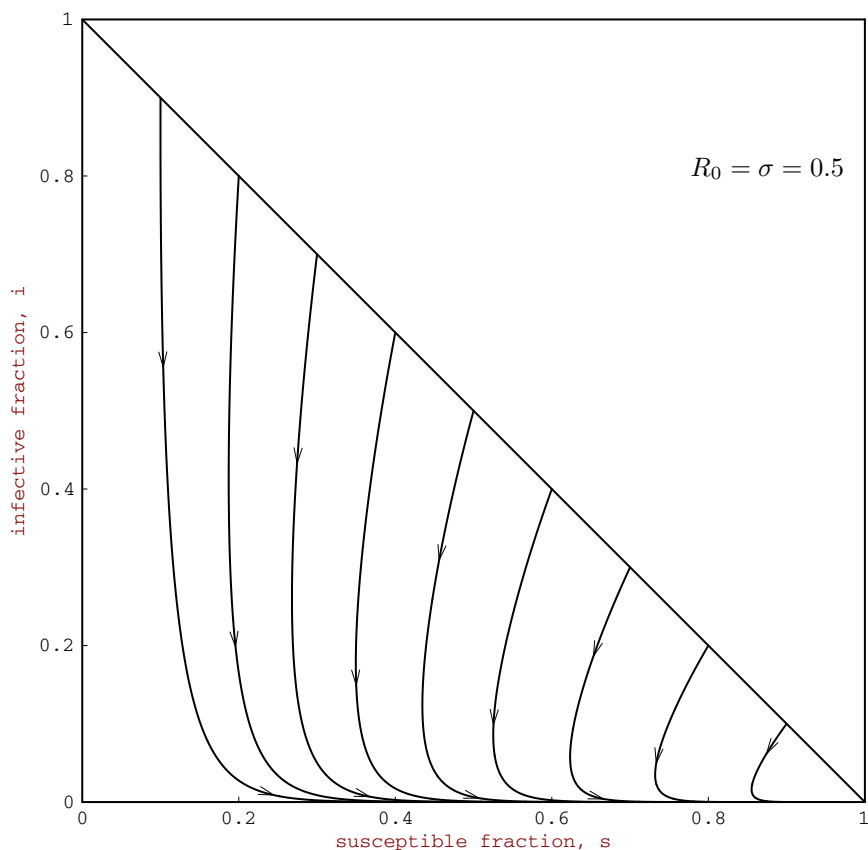


Fig. 5. Phase plane portrait for the classic SIR endemic model with contact number $\sigma = 0.5$.

Theorem 3: Let $(s(t), i(t))$ be a solution of (1.6.2) in T . If $R_0 = \sigma \leq 1$ or $i_o = 0$, then solution paths starting in T approach the disease-free equilibrium given by $(s, i) = (1, 0)$. If $R_0 = \sigma > 1$, then all solution paths with $i_o > 0$ approach the endemic equilibrium given by $s_e = 1/\sigma$ and $i_e = \mu(\sigma - 1)/\beta$.

Figures 5 and 6 illustrate the two possibilities given in the theorem. If $R_0 = \sigma \leq 1$, then the replacement number σs is less than one when $i_o > 0$, so that the infectives decrease to zero. Although the speeds of movement along the paths are not apparent from Figure 5, the infective

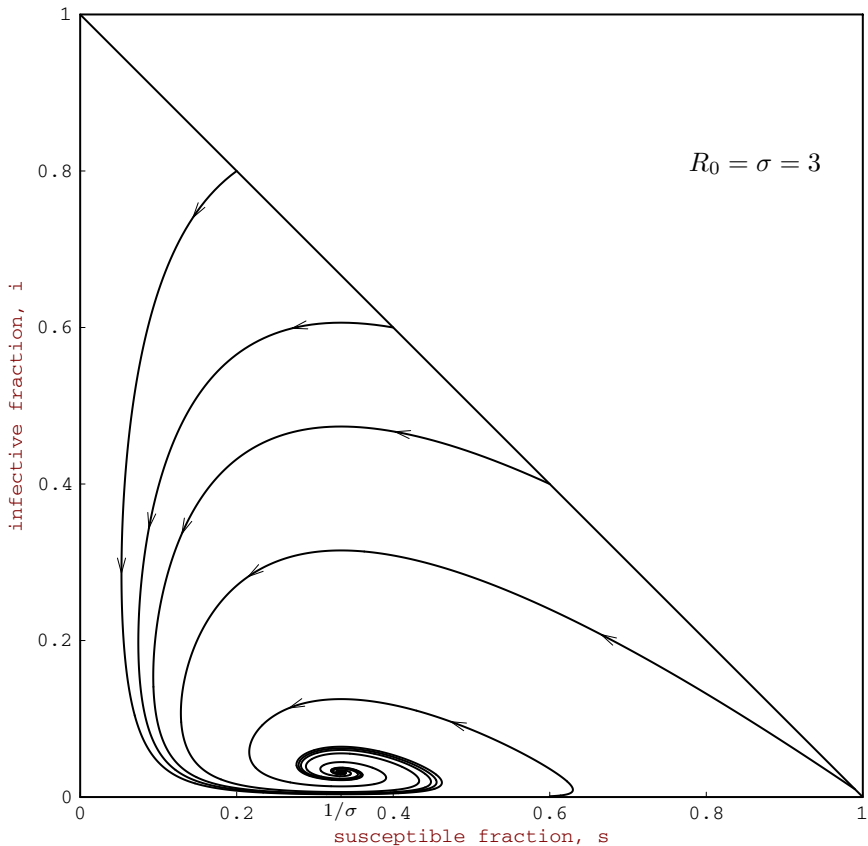


Fig. 6. Phase plane portrait for the classic SIR endemic model with contact number $\sigma = 3$, average infectious period $1/\gamma = 3$ days, and average lifetime $1/\mu = 60$ days. This unrealistically short average lifetime has been chosen so that the endemic equilibrium is clearly above the horizontal axis and the spiralling into the endemic equilibrium can be seen.

fraction decreases rapidly to very near zero, and then over 100 or more years, the recovered people slowly die off and the birth process slowly increases the susceptibles, until eventually everyone is susceptible at the disease-free equilibrium with $(s, i) = (1, 0)$. If $R_0 = \sigma > 1$, i_o is small, and s_o is large with $\sigma s_o > 1$, then $s(t)$ decreases and $i(t)$ increases up to a peak and then decreases, just as it would for an epidemic (compare Figure 6 with Figure 2). However, after the infective fraction has decreased to a low level,

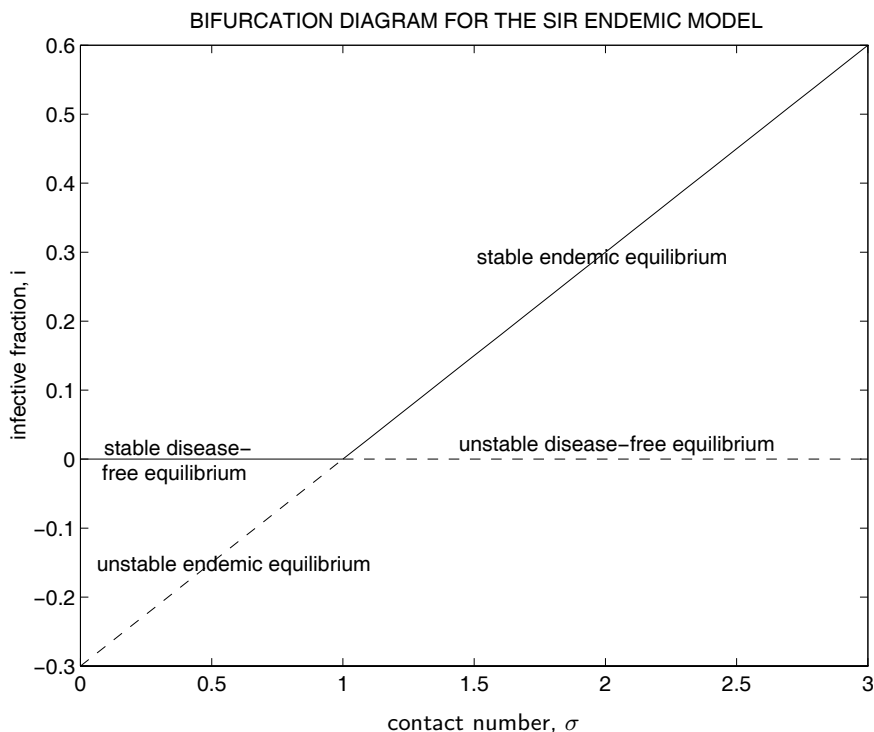


Fig. 7. The bifurcation diagram for the SIR endemic model, which shows that the disease free and endemic equilibria exchange stability when the contact number σ is 1.

the slow processes of the deaths of recovered people and the births of new susceptibles gradually (over about 10 or 20 years) increase the susceptible fraction until $\sigma s(t)$ is large enough that another smaller epidemic occurs. This process of alternating rapid epidemics and slow regeneration of susceptibles continues as the paths approach the endemic equilibrium given in the theorem. At this endemic equilibrium the replacement number σs_e is one, which is plausible since if the replacement number were greater than or less than one, the infective fraction $i(t)$ would be increasing or decreasing, respectively.

Theorem 3 is proved in [72] and in [76] using phase plane methods and Liapunov functions. For this SIR model there is a transcritical (stability exchange) bifurcation at $\sigma = 1$, as shown in Figure 7. Notice that the i_e coordinate of the endemic equilibrium is negative for $\sigma < 1$, coincides with

the disease-free equilibrium value of zero at $\sigma = 1$, and becomes positive for $\sigma > 1$. This equilibrium given by $(s_e, i_e) = (1/\sigma, \mu(\sigma - 1)/\beta)$ is unstable for $\sigma < 1$ and is locally asymptotically stable for $\sigma > 1$, while the disease-free equilibrium given by $(s, i) = (1, 0)$ is locally stable for $\sigma < 1$ and unstable for $\sigma > 1$. Thus these two equilibria exchange stabilities as the endemic equilibrium moves through the disease-free equilibrium when $\sigma = 1$ and becomes a distinct, epidemiologically feasible, locally asymptotically stable equilibrium when $\sigma > 1$.

The following interpretation of the results in the theorem and paragraph above is one reason why the basic reproduction number R_0 has become widely used in the epidemiology literature. If the basic reproduction number R_0 (which is equal to the replacement number R when the entire population is susceptible) is less than one, then the disease-free equilibrium is locally asymptotically stable and the disease cannot “invade” the population. But if $R_0 > 1$, then the disease-free equilibrium is unstable with a repulsive direction into the positive (s, i) quadrant, so the disease can “invade” in the sense that any path starting with a small positive i_0 moves into the positive (s, i) quadrant where the disease persists. Thus for this classic SIR endemic model and for many other more complex models [42], the behavior is almost completely dependent on the threshold quantity R_0 , which determines not only when the local stability of the disease-free equilibrium switches, but also when the endemic equilibrium enters the feasible region with a positive infective fraction. The latter condition is used to obtain expressions for R_0 in age-structured models in Chapter 3.

1.7. Similar models with M and E epidemiological states

An infected or vaccinated mother transfers some IgG antibodies across the placenta to her fetus, so that her newborn infant has temporary passive immunity to an infection. When these passive antibodies are gone (no new antibodies are produced by the infant), the infant moves from the passively immune state M to the susceptible state S. Infants who do not have any passive immunity, because their mothers were neither infected nor vaccinated, also enter the class S of susceptible individuals, who can become infected. When there is an adequate contact of a susceptible with an infective so that transmission occurs, then the susceptible enters the exposed class E of those in the latent period, who are infected, but not yet infectious. The incubation period is defined as the period from initial exposure to the appearance of symptoms. Since a person may become infectious before or

after symptoms appear, the incubation period is often different from the latent period. In infectious disease modeling, we are always interested in the latent period, since we want the period until the person becomes infectious. After the latent period ends, the individual enters the class I of infectives, who are infectious in the sense that they are capable of transmitting the infection. Models with the M and E epidemiological states have behaviors that are analogous to the models without these states.

Examples of models with different epidemiological states are MSEIR, MSEIRS, SEIR, SEIRS, SIR, SIRS, SEI, SEIS, SI, and SIS. Models such as MSEIRS, SEIRS, SIRS, SEIS, and SIS with a flow back into the susceptible class S are always endemic models and never epidemic models, since there is always a threshold above which the disease remains endemic. Models of the types MSEIR, SEIR, SIR, SEI, and SI without return flow into the susceptible class S are endemic if they have a birth and death process, and are epidemic if they are for a short time period without a vital dynamics process. Behaviors for these models are similar to those for the analogous basic models. We consider an SEIR epidemic model first and then an MSEIRS endemic model.

1.7.1. The SEIR epidemic model

Since an epidemic occurs in a short time period, we ignore loss of temporary immunity and the birth and death processes. Therefore we have no flow from the removed class back to the susceptible class and omit the passively immune class M into which newborns would move. Hence we consider the SEIR epidemic model, which has analogous behavior to that of the basic SIR epidemic model. For this model the initial value problem is

$$\begin{aligned} dS/dt &= -\beta SI/N, & S(0) &= S_o > 0, \\ dE/dt &= \beta SI/N - \varepsilon E, & E(0) &= E_o \geq 0, \\ dI/dt &= \varepsilon E - \gamma I, & I(0) &= I_o > 0, \\ dR/dt &= \gamma I, & R(0) &= R_o \geq 0, \end{aligned} \tag{1.7.1}$$

where $S(t)$, $E(t)$, $I(t)$, and $R(t)$ are the numbers in the susceptible, exposed, infectious, and removed classes, respectively, so that $S(t) + E(t) + I(t) + R(t) = N$. This model uses the standard incidence, has movement out of the exposed (latent) class E at rate εE (corresponding to an exponential waiting time $e^{-\varepsilon t}$), and has recovery at rate γI (corresponding to an exponential waiting time $e^{-\gamma t}$). As explained in the formulation section, the average latent period is $1/\varepsilon$ and the average infectious period is $1/\gamma$.

Dividing the equations in (1.7.1) by the constant total population size N yields

$$\begin{aligned} ds/dt &= -\beta si, & s(0) &= s_o > 0 \\ de/dt &= \beta si - \varepsilon e, & e(0) &= e_o \geq 0 \\ di/dt &= \varepsilon e - \gamma i, & i(0) &= i_o > 0 \end{aligned} \quad (1.7.2)$$

with $r(t) = 1 - s(t) - e(t) - i(t)$, where $s(t)$, $e(t)$, $i(t)$, and $r(t)$ are the fractions in the classes. The tetrahedron T in the (s, e, i) phase space given by

$$T = \{(s, e, i) \mid s \geq 0, e \geq 0, i \geq 0, s + e + i \leq 1\} \quad (1.7.3)$$

is positively invariant and unique solutions exist in T for all positive time, so that the model is mathematically and epidemiologically well posed. As in the basic SIR epidemic model, the contact number $\sigma = \beta/\gamma$ is the contact rate β per unit time multiplied by the average infectious period $1/\gamma$, so it has the proper interpretation as the average number of adequate contacts of a typical infective during the infectious period. Moreover, the replacement number at time zero is still σs_o , which is the product of the contact number σ and the initial susceptible fraction s_o .

Theorem 4: Let $(s(t), e(t), i(t))$ be a solution of (1.7.2) in T . If $\sigma s_o \leq 1$, then $e(t)$ and $i(t)$ decrease to zero as $t \rightarrow \infty$. If $\sigma s_o > 1$, then $e(t) + i(t)$ first increases up a maximum with value $e_{\max} + i_{\max} = e_o + i_o + s_o - [\ln(\sigma s_o)]/\sigma$ and then decreases to zero as $t \rightarrow \infty$. The susceptible fraction $s(t)$ is a decreasing function and the limiting value s_∞ is the unique root in the interval $(0, 1/\sigma)$ of the equation

$$e_o + i_o + s_o - s_\infty + [\ln(s_\infty/s_o)]/\sigma = 0.$$

Typical paths in the projection of the tetrahedron T on the (s, i) face are similar to those in Figure 2, and solutions as a function of time are similar to those in Figure 3. Thus we still have the typical epidemic outbreak with an infective curve that first increases from an initial i_o near zero, reaches a peak, and then decreases towards zero as a function of time.

As before, the susceptible fraction $s(t)$ always decreases, but the final susceptible fraction s_∞ is positive. The epidemic dies out because, when the susceptible fraction $s(t)$ goes below $1/\sigma$, the replacement number $\sigma s(t)$ goes below one. As before, if enough people are already immune so that a typical infective initially replaces itself with no more than one new infective ($\sigma s_o \leq 1$), then there is no epidemic outbreak. But if a typical infective

initially replaces itself with more than one new infective ($\sigma s_o > 1$), then infecteds initially increase so that an epidemic occurs. The speed at which an epidemic of a particular disease progresses depends on the contact rate β , the average latent period $1/\varepsilon$, and the average infectious period $1/\gamma$. Measles has a latent period of about 7 days and an infectious period of about 7 days, so the measles epidemic in Figure 4 evolved slowly and lasted for about 9 months. Because the latent period for influenza is only 1–3 days and the infectious period is only 2–3 days, an influenza epidemic can sweep through a city in less than 6 weeks.

The proof of this theorem is similar to that for the SIR epidemic model. The tetrahedron T is positively invariant, since no direction vectors at a boundary point are outward. The only equilibrium points, which are along the s axis where $(e, i) = (0, 0)$, are neutrally unstable for $s > 1/\sigma$ and are neutrally stable for $s < 1/\sigma$. The equation $ds/dt = -\beta si \leq 0$ implies that $s(t)$ is non-increasing and $s(t) \geq 0$, so that a unique limit s_∞ exists as $t \rightarrow \infty$. Since $dr/dt = \gamma i \geq 0$ and $r(t)$ is bounded above by 1, the limit $r(\infty)$ exists. Since $e(t) + i(t) = 1 - s(t) - r(t)$, the limit $e(\infty) + i(\infty)$ exists. From $d(e + i)/dt = (\beta s - \gamma)i$, we see that the limit $i(\infty)$ must exist and $i(\infty) = 0$ if $s(\infty) \neq 1/\sigma$. Alternatively, $i(\infty) = 0$, since otherwise $dr/dt = \gamma i/2$ for t sufficiently large, so that $r(\infty) = \infty$, which contradicts $r(\infty) \leq 1$. Also $e(\infty) = 0$, since $e(\infty) > 0$ would contradict $di/dt = \varepsilon e - \gamma i$. Solution paths

$$e(t) + i(t) + s(t) - [\ln s(t)]/\sigma = e_o + i_o + s_o - [\ln s_o]/\sigma \quad (1.7.4)$$

are found from the quotient differential equation $d(e + i)/ds = -1 + 1/(\sigma s)$. The theorem conclusions follow from (1.7.4) and the observation that $e(t) + i(t)$ has a maximum on the solution path when $\sigma s = 1$. Similar to the SIR epidemic model, the threshold result involves the initial replacement number σs_o and does not involve the basic reproduction number R_0 .

1.7.2. The MSEIRS endemic model

The transfer diagram for the MSEIRS model is shown in Figure 8. The simpler MSEIR with permanent immunity (so the R to S transition is missing) is suitable for a directly transmitted viral disease such as measles, rubella, or mumps. As in the SIR endemic model, let the birth and death rate constants be μ , so the population size $N(t)$ remains constant. An MSEIR model with exponentially changing population size is formulated and analyzed in [81]. The numbers of people in the epidemiological classes are denoted by $M(t)$, $S(t)$, $E(t)$, $I(t)$, and $R(t)$, where t is time, and the fractions of the population in these classes are $m(t)$, $s(t)$, $e(t)$, $i(t)$, and $r(t)$.

We are interested in finding conditions that determine whether the disease dies out (i.e. the fraction i goes to zero) or remains endemic (i.e. the fraction i remains positive).

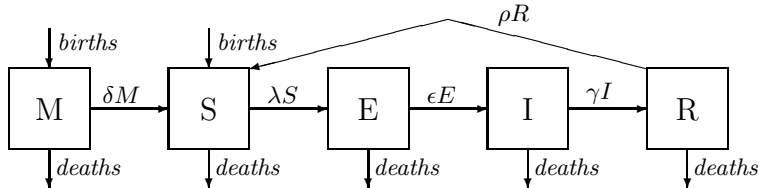


Fig. 8. Transfer diagram for the MSEIRS model with the passively-immune class M , the susceptible class S , the exposed class E , the infective class I , and the recovered class R .

The birth rate μS into the susceptible class of size S corresponds to newborns whose mothers are susceptible, and the other newborns $\mu(N - S)$ enter the passively immune class of size M , since their mothers were infected or had some type of immunity. Although all women would be out of the passively immune class long before their childbearing years, theoretically a passively immune mother would transfer some IgG antibodies to her newborn child, so the infant would have passive immunity. Deaths occur in the epidemiological classes at the rates μM , μS , μE , μI , and μR , respectively.

In this MSEIRS epidemiological model, the transfer out of the passively immune class is δM , the transfer out of the exposed class is εE , the recovery rate from the infectious class is γI , and the rate of loss of immunity is ρR . The linear transfer terms in the differential equations correspond to waiting times with negative exponential distributions, so that when births and deaths are ignored, the mean passively immune period is $1/\delta$, the mean latent period is $1/\varepsilon$, the mean infectious period is $1/\gamma$, and the mean period of infection-induced immunity is $1/\rho$ [83]. These periods would $1/\delta = 6$ months, $1/\varepsilon = 14$ days, and $1/\gamma = 7$ days in an MSEIR model for chickenpox [143]. For sexually transmitted diseases, it is useful to define both a sexual contact rate and the fraction of contacts that result in transmission, but for directly-transmitted diseases spread primarily by aerosol droplets, transmission may occur by entering a room, hallway, building, etc. that is currently or has been occupied by an infective. Since there is no clear definition of a contact or a transmission fraction, they are replaced by a definition that includes both. An adequate contact is a contact that is suf-

ficient for transmission of infection from an infective to a susceptible. Let the contact rate β be the average number of adequate contacts per person per unit time, so that the force of infection $\lambda = \beta i$ is the average number of contacts with infectives per unit time. Then the incidence (the number of new cases per unit time) is $\lambda S = \beta i S = \beta SI/N$, since it is the number of contacts with infectives per unit time of the S susceptibles. As described earlier, this standard form $\beta SI/N$ for the incidence is consistent with numerous studies which show that the contact rate β is nearly independent of the population size.

The system of differential equations for the numbers in the epidemiological classes is

$$\begin{aligned} dM/dt &= \mu(N - S) - (\delta + \mu)M, \\ dS/dt &= \mu S + \delta M - \beta SI/N - \mu S + \rho R, \\ dE/dt &= \beta SI/N - (\varepsilon + \mu)E, \\ dI/dt &= \varepsilon E - (\gamma + \mu)I, \\ dR/dt &= \gamma I - (\rho + \mu)R. \end{aligned} \tag{1.7.5}$$

It is convenient to convert to differential equations for the fractions in the epidemiological classes with simplifications by dividing by the constant population size N and eliminating the differential equation for s by using $s = 1 - m - e - i - r$. Then the ordinary differential equations for the MSEIRS model are

$$\begin{aligned} dm/dt &= \mu(e + i + r) - \delta m \\ de/dt &= \beta i(1 - m - e - i - r) - (\varepsilon + \mu)e \\ di/dt &= \varepsilon e - (\gamma + \mu)i \\ dr/dt &= \gamma i - (\rho + \mu)r. \end{aligned} \tag{1.7.6}$$

A suitable domain is

$$\mathfrak{D} = \{(m, e, i, r) : m \geq 0, e \geq 0, i \geq 0, r \geq 0, m + e + i + r \leq 1\}.$$

The domain \mathfrak{D} is positively invariant, because no solution paths leave through any boundary. The right sides of (1.7.6) are smooth, so that initial value problems have unique solutions that exist on maximal intervals [69]. Since paths cannot leave \mathfrak{D} , solutions exist for all positive time. Thus the model is mathematically and epidemiologically well posed.

The basic reproduction number R_0 for this MSEIRS model is the same as the contact number σ given by

$$R_0 = \sigma = \frac{\beta\varepsilon}{(\gamma + \mu)(\varepsilon + \mu)}. \quad (1.7.7)$$

This R_0 is the product of the contact rate β per unit time, the average infectious period adjusted for population growth of $1/(\gamma + \mu)$, and the fraction $\varepsilon/(\varepsilon + \mu)$ of exposed people surviving the latent class E . Thus R_0 has the correct interpretation that it is the average number of secondary infections due to an infective during the infectious period, when everyone in the population is susceptible. The equations (1.7.6) always have a disease-free equilibrium given by $m = e = i = r = 0$, so that $s = 1$. If $R_0 > 1$, there is also a unique endemic equilibrium in \mathfrak{D} given by

$$\begin{aligned} m_e &= \frac{\mu}{\delta + \mu} \left(1 - \frac{1}{R_0} \right), \\ e_e &= \frac{\delta(\gamma + \mu)(\rho + \mu)}{(\delta + \mu)[(\rho + \mu)(\gamma + \varepsilon + \mu) + \gamma\varepsilon]} \left(1 - \frac{1}{R_0} \right), \\ i_e &= \frac{\delta\varepsilon(\rho + \mu)}{(\delta + \mu)[(\rho + \mu)(\gamma + \varepsilon + \mu) + \gamma\varepsilon]} \left(1 - \frac{1}{R_0} \right), \\ r_e &= \frac{\delta\varepsilon\gamma}{(\delta + \mu)[(\rho + \mu)(\gamma + \varepsilon + \mu) + \gamma\varepsilon]} \left(1 - \frac{1}{R_0} \right), \end{aligned} \quad (1.7.8)$$

where $s_e = 1/R_0 = 1/\sigma$. Note that the replacement number σs_e is 1 at the endemic equilibrium.

By linearization, the disease-free equilibrium is locally asymptotically stable if $R_0 < 1$ and is an unstable hyperbolic equilibrium with a stable manifold outside \mathfrak{D} and an unstable manifold tangent to a vector into \mathfrak{D} when $R_0 > 1$. The disease-free equilibrium can be shown to be globally asymptotically stable in \mathfrak{D} if $R_0 \leq 1$ by using the Liapunov function $V = \varepsilon e + (\varepsilon + \mu)i$ as follows. The Liapunov derivative is $\dot{V} = [\beta\varepsilon s - (\varepsilon + \mu)(\gamma + \mu)]i \leq 0$, since $\beta\varepsilon \leq (\varepsilon + \mu)(\gamma + \mu)$. The set where $\dot{V} = 0$ is the face of \mathfrak{D} with $i = 0$, but $di/dt = \varepsilon e$ on this face, so that i moves off the face unless $e = 0$. When $e = i = 0$, $dr/dt = -\mu r$, so that $r \rightarrow 0$. When $e = i = r = 0$, then $dm/dt = -\delta m$, so $m \rightarrow 0$. Because the origin is the only positively invariant subset of the set with $\dot{V} = 0$, all paths in \mathfrak{D} approach the origin by the Liapunov–Lasalle theorem ([69], p. 296). Thus if $R_0 \leq 1$, then the disease-free equilibrium is globally asymptotically stable in \mathfrak{D} .

The characteristic equation corresponding to the Jacobian at the endemic equilibrium is a fourth degree polynomial. Using a symbolic algebra

program, it can be shown that the Routh–Hurwitz criteria are satisfied if $R_0 > 1$, so that the endemic equilibrium (1.7.8) is locally asymptotically stable when it is in \mathfrak{D} . Thus if $R_0 > 1$, then the disease-free equilibrium is unstable and the endemic equilibrium is locally asymptotically stable. The system (1.7.6) can be defined to be uniformly persistent if $\liminf_{t \rightarrow \infty} i(t) \geq c$ for some $c > 0$ for all initial points such that $e(0) + i(0) > 0$. The properties of the disease-free equilibrium and Theorem 4.5 in [152] imply that the system (1.7.6) is uniformly persistent if $R_0 > 1$. Based on results for the SIR and SEIR models, we expect (but have not proved rigorously) that all paths in \mathfrak{D} with some initial latents or infectives go to the endemic equilibrium if $R_0 > 1$. Then we have the usual behavior for an endemic model, in the sense that the disease dies out below the threshold, and the disease goes to a unique endemic equilibrium above the threshold.

Similar models also have the endemic threshold property above. The MSEIR model is similar to the MSEIRS model, but $\rho = 0$ so that the immunity after an infection is permanent. This MSEIR model has a simpler endemic equilibrium, but it has the same basic reproduction number R_0 given by (1.7.7). If $\delta \rightarrow \infty$, then heuristically the M class disappears (think of people moving through the M class with infinite speed), so that the MSEIR and MSEIRS models become SEIR and SEIRS models [117] with the same basic reproduction number R_0 given by (1.7.7). If $\varepsilon \rightarrow \infty$, then the E class disappears, leading to an MSIR and MSIRS models with $R_0 = \beta/(\gamma + \mu)$. If an SEIRS model has an ρR transfer term from the removed class R to the susceptible class S and $\rho \rightarrow \infty$, then the R class disappears, leading to an SEIS model with R_0 given by (1.7.7). If $\varepsilon \rightarrow \infty$, then the E class disappears and the SEIS model becomes the basic SIS endemic model with $R_0 = \beta/(\gamma + \mu)$. If $\delta \rightarrow \infty$ and $\varepsilon \rightarrow \infty$, then both the M and E classes disappear in the MSEIRS model leading to an SIRS model with $R_0 = \beta/(\gamma + \mu)$.

Global stability has been proved for some of these models. For the SEIR model with constant population size, the global stability below the threshold was proved in [117] and the global stability above the threshold was proved using clever new methods [113]. Global stability of the endemic equilibrium has been proved for the SEIRS model with short or long period of immunity [115], and for the SEIR model with exponentially changing population size under a mild restriction [114]. The global stabilities of the SIR, SIRS, SEIS models with constant population sizes are proved by standard phase plane methods [72, 76]. Epidemiology models with variable population size are considered in Chapter 2.

1.8. Threshold estimates using the two SIR models

The two basic SIR models in previous sections are very important as conceptual models (similar to predator-prey and competing species models in ecology). The SIR epidemic modeling yields the useful concept that the replacement number $R = \sigma s_o$ is the threshold quantity which determines when an epidemic occurs. It also yields formulas for the peak infective fraction i_m and the final susceptible fraction s_∞ . The SIR endemic model yields $R_0 = \sigma$ as the threshold quantity which determines when the disease remains endemic, the concept that the infective replacement number σs_e is 1 at the endemic equilibrium, and the explicit dependence of the infective fraction i_e on the parameters. However, these simple SIR models have obvious limitations. Although the assumption that the population is uniform and homogeneously mixing is plausible since many children have similar patterns of preschool and school attendance, mixing patterns can depend on many factors including age (children usually have more adequate contacts per day than adults). Moreover, different geographic and social-economic groups have different contact rates. Despite their limitations, the classic SIR models can be used to obtain some useful estimates and comparisons.

From equation (1.5.4) for s_∞ in the classic SIR epidemic model, the approximation

$$\sigma \approx \frac{\ln(s_o/s_\infty)}{s_o - s_\infty} \quad (1.8.1)$$

follows because i_o is negligibly small. By using data on the susceptible fractions s_o and s_∞ at the beginning and end of epidemics, this formula can be used to estimate contact numbers for specific diseases [76]. Using blood samples from freshmen at Yale University [56], the fractions susceptible to rubella at the beginning and end of the freshman year were found to be 0.25 and 0.090, so the epidemic formula (1.8.1) gives $\sigma \approx 6.4$. The fractions $s_o = 0.49$ and $s_\infty = 0.425$ for Epstein-Barr virus (related to mononucleosis) lead to $\sigma \approx 2.2$, and the fractions $s_o = 0.911$ and $s_\infty = 0.514$ for influenza (H3N2 type A “Hong Kong”) lead to $\sigma \approx 1.44$. For the 1957 “Asian Flu” (H2N2 type A strain of influenza) in Melbourne, Australia, the fractions $s_o = 1$ and $s_\infty = 0.55$ from [20] (p. 129) yield the contact number estimate $\sigma \approx 1.33$. Thus the easy theory for the basic SIR epidemic model yields the formula above that can be used to estimate contact numbers from epidemic data.

The basic SIR endemic model can also be used to estimate contact numbers. If blood samples in a serosurvey are tested for antibodies to a virus and it is assumed that the SIR model above holds in the population with the disease at an endemic equilibrium, then the contact number can be estimated from $\sigma = 1/s_e$, where s_e is the fraction of the samples that are not seropositive, since $s_e = 1 - i_e - r_e$. This approach is somewhat naive, because the average seropositivity in a population decreases to zero as the initial passive immunity declines and then increases as people age and are exposed to infectives. Thus the ages of those sampled are critical in using the estimate $\sigma = 1/s_e$.

For an age-structured SIR model with negative exponential survival considered in Chapter 3, one estimation formula for the basic reproduction number is $R_0 = 1 + L/A$, where L is the average lifetime $1/\mu$ and A is the average age of infection. This estimation formula can also be derived heuristically from the basic SIR endemic model. The incidence rate at the endemic equilibrium is $\beta i_e s_e$, so that βi_e is the incidence rate constant, which with exponential waiting time implies that the average age of infection (the mean waiting time in S) is $A = 1/\beta i_e = 1/[\mu(\sigma - 1)]$. Using $\mu = 1/L$, this leads to $R_0 = \sigma = 1 + L/A$, since $R_0 = \sigma$ for this model.

1.9. Comparisons of some directly transmitted diseases

Data on average ages of infection and average lifetimes in developed countries have been used to estimate basic reproduction numbers R_0 for some viral diseases ([7], p. 70 and [76]). From Table 4, we see that these estimates of R_0 are about 16 for measles, 11 for varicella (chickenpox), 8 for mumps, 7 for rubella, and 5 for poliomyelitis and smallpox. Because disease-acquired immunity is only temporary for bacterial diseases such as pertussis (whooping cough), diphtheria, and scarlet fever, the formula $R_0 = \sigma = 1 + L/A$ cannot be used to estimate R_0 for these diseases (see [81] for estimates of R_0 and σ for pertussis).

Herd immunity occurs for a disease if enough people have disease-acquired or vaccination-acquired immunity, so that the introduction of one infective into the population does not cause an invasion of the disease. Intuitively, if the contact number is σ , so that the typical infective has adequate contacts with σ people during the infectious period, then the replacement number σs must be less than one so that the disease does not spread. This means that s must be less than $1/\sigma$, so the immune fraction r must be satisfy $r > 1 - 1/\sigma = 1 - 1/R_0$. For example, if $R_0 = \sigma = 10$, then the

Table 4. Estimates of basic reproduction numbers and fractions for herd immunity.

Disease	Location	A	L	$R_0 = \sigma = 1 + L/A$	Min fraction r for herd immunity	Vacc. efficacy VE	Min fraction vaccinated for herd immunity
Measles	England and Wales, 1956–59	4.8	70	15.6	0.94	0.95	0.99
	USA, 1912–1928	5.3	60	12.3	0.92	0.95	0.97
	Nigeria, 1960–68	2.5	40	17.0	0.94	0.95	0.99
Chickenpox (varicella)	Maryland, USA, 1943	6.8	70	11.3	0.91	0.90	1.01
Mumps	Maryland, USA, 1943	9.9	70	8.1	0.88	0.95	0.93
Rubella	England and Wales, 1979	11.6	70	7.0	0.86	0.95	0.91
	West Germany, 1972	10.5	70	7.7	0.87	0.95	0.92
Poliomyelitis	USA, 1955	17.9	70	4.9	0.80	?	
	Netherlands, 1960	11.2	70	4.3	0.86	?	
Smallpox	India	12.0	50	5.2	0.81	0.95	0.85

immune fraction must satisfy $r > 1 - 1/10 = 0.9$, so that the replacement number is less than one and the disease does not invade the population. Table 4 contains data [4] on the average age of infection for various directly transmitted diseases and average lifetimes. See [7] for more data. Table 4 also contains estimates of basic reproduction numbers R_0 and the minimum fractions that must be immune in order to achieve herd immunity. Using the estimates above for R_0 , the minimum immune fractions for herd immunity in these populations are about 0.94 for measles, 0.88 for mumps, 0.86 for rubella, and about 0.80 for poliomyelitis and smallpox.

Primary vaccination failure occurs when a vaccinated individuals does not become immune. Secondary vaccination failure occurs when the immunity from an initially-successful vaccination wanes, so that the person becomes susceptible to infection again at a later time. The fraction of those vaccinated who initially become immune is called the vaccine efficacy VE. The vaccine efficacy for many vaccines is about 0.95, so that primary vaccination failure occurs in about 5% of those vaccinated. Secondary vaccination failure is rare for vaccines for most viral diseases in Table 2. In

order to calculate the minimum fraction that must be vaccinated to achieve herd immunity, one divides the minimum fraction that must be immune by the vaccine efficacy. These values are shown in the last column in Table 4.

The minimum fractions that must be vaccinated to achieve herd immunity in these populations are about 0.99 for measles, 0.93 for mumps, 0.91 for rubella, and about 0.85 for smallpox. Although these values give only crude, ballpark estimates for the vaccination level in a community required for herd immunity, they are useful for comparing diseases. For example, these numbers suggest that it should be easier to achieve herd immunity for mumps, rubella, poliomyelitis, and smallpox than for measles and chickenpox. Indeed, the theoretical vaccination levels of 0.99 for measles and 1.01 for varicella (chickenpox) suggests that herd immunity would be very hard to achieve for these diseases with a one dose program, because the percentages not vaccinated would have to be below 1% for measles, below 7% for mumps, and below 9% for rubella. Because vaccinating all but 1% against measles would be difficult to achieve, a two dose program for measles is used in many countries [34, 74, 75]. With a varicella (chickenpox) vaccination program whose goal is to give each child one dose, it is impossible to achieve vaccination of 101% of the population. Moreover, immunity from the varicella (chickenpox) vaccine wanes, so that children vaccinated may become susceptible again about 20 years after vaccination [144].

1.10. The effectiveness of vaccination programs

The conclusions from Table 4 are justified by the actual effectiveness of vaccination programs in reducing, locally eliminating, and eradicating these diseases (eradication means elimination throughout the world). The information in this section verifies that smallpox has been eradicated worldwide and polio should be eradicated worldwide within a few years, while the diseases of rubella and measles still persist at low levels in the United States and at higher levels in many other countries. This section provides historical context and verifies the disease comparisons obtained in Table 4 from the basic SIR endemic model.

1.10.1. *Smallpox*

Smallpox is believed to have appeared in the first agricultural settlements around 6,000 BC. For centuries the process of variolation with material from smallpox pustules was used in Africa, China, and India before arriving in Europe and the Americas in the 18th century. Edward Jenner, an English

country doctor, observed over 25 years that milkmaids who had been infected with cowpox did not get smallpox. In 1796 he started vaccinating people with cowpox to protect them against smallpox [135]. This was the world's first vaccine (*vacca* is the Latin word for cow). Two years later, the findings of the first vaccine trials were published, and by the early 1800s, the smallpox vaccine was widely available. Smallpox vaccination was used in many countries in the 19th century, but smallpox remained endemic. When the World Health Organization (WHO) started a global smallpox eradication program in 1967, there were about 15 million cases per year, of which 2 million died and millions more were disfigured or blinded by the disease [58]. The WHO strategy involved extensive vaccination programs, surveillance for smallpox outbreaks, and containment of these outbreaks by local vaccination programs. There are some interesting stories about the WHO campaign including the persuasion of African chiefs to allow their tribes to be vaccinated and monetary bounty systems for finding hidden smallpox cases in India. Smallpox was slowly eliminated from many countries, with the last case in the Americas in 1971. The last case worldwide was in Somalia in 1977, so smallpox has been eradicated throughout the world [14, 58, 135]. The WHO estimates that the elimination of worldwide smallpox vaccination saves over two billion dollars per year. The smallpox virus has been kept in USA and Russian government laboratories; the USA keeps it so that vaccine could be produced if smallpox were ever used in biological terrorism [146]. See [41] for a paper on smallpox attacks that compares the effects of quarantine, isolation, vaccination, and behavior change.

1.10.2. *Poliomyelitis*

Most cases of poliomyelitis are asymptomatic, but a small fraction of cases result in paralysis. In the 1950s in the United States, there were about 60,000 paralytic polio cases per year. In 1955 Jonas Salk developed an injectable polio vaccine from an inactivated polio virus. This vaccine provides protection for the person, but the person can still harbor live viruses in their intestines and can pass them to others. In 1961 Albert Sabin developed an oral polio vaccine from weakened strains of the polio virus. This vaccine provokes a powerful immune response, so the person cannot harbor the “wild-type” polio viruses, but a very small fraction (about one in 2 million) of those receiving the oral vaccine develop paralytic polio [14, 135]. The Salk vaccine interrupted polio transmission and the Sabin vaccine eliminated polio epidemics in the United States, so there have been no indigenous cases

of naturally-occurring polio since 1979. In order to eliminate the few cases of vaccine-related paralytic polio each year, the United States now recommends the Salk injectable vaccine for the first four polio vaccinations, even though it is more expensive [34]. In the Americas, the last case of paralytic polio caused by the wild virus was in Peru in 1991. In 1988 WHO set a goal of global polio eradication by the year 2000 [142]. Most countries are using the live-attenuated Sabin oral vaccine, because it is inexpensive (8 cents per dose) and can be easily administered into a mouth by an untrained volunteer. The WHO strategy includes routine vaccination, National Immunization Days (during which many people in a country or region are vaccinated in order to interrupt transmission), mopping-up vaccinations, and surveillance for acute flaccid paralysis [91]. Polio has disappeared from many countries in the past ten years, so that by 1999 it is concentrated in the Eastern Mediterranean region, South Asia, West Africa and Central Africa. It is likely that polio will be eradicated worldwide soon. WHO estimates that eradicating polio will save about \$1.5 billion each year in immunization, treatment, and rehabilitation around the globe [29].

1.10.3. *Measles*

Measles is a serious disease of childhood that can lead to complications and death. For example, measles caused about 7500 deaths in the United States in 1920 and still causes about 1 million deaths worldwide each year [31, 32]. Measles vaccinations are given to children between 6 and 18 months of age, but the optimal age of vaccination for measles seems to vary geographically [75].

Consider the history of measles in the United States. In the prevaccine era, every child had measles, so the incidences were approximately equal to the sizes of the birth cohorts. After the measles vaccine was licensed in 1963 in the United States, the reported measles incidence dropped in a few years to around 50,000 cases per year. In 1978 the US adopted a goal of eliminating measles and vaccination coverage increased, so that there were fewer than 5000 reported cases per year between 1981 and 1988. Pediatric epidemiologists at meetings at Centers for Disease Control in Atlanta in November 1985 and February 1988 decided to continue the one-dose program for measles vaccinations instead of changing to a more expensive two-dose program. But there were about 16,000, 28,000, and 17,000 reported measles cases in the United States in 1989, 1990, and 1991, respectively; there were also measles outbreaks in Mexico and Canada during these years

[92]. Because of this major measles epidemic, epidemiologists decided in 1989 that the one-dose vaccination program for measles, which had been used for 26 years, should be replaced with a two-dose program with the first measles vaccination at age 12–15 months and the second vaccination at 4–6 years, just before children start school [34]. Reported measles cases declined after 1991 until there were only 137, 100, and 86 reported cases in 1997, 1998, and 1999, respectively. Each year some of the reported cases are imported cases and these imported cases can trigger small outbreaks. The proportion of cases not associated with importation has declined from 85% in 1995, 72% in 1996, 41% in 1997, to 29% in 1998. Analysis of the epidemiologic data for 1998 suggests that measles is no longer an indigenous disease in the United States [31]. Measles vaccination coverage in 19–35 month old children was only 92% in 1998, but over 99% of children had at least one dose of measles-containing vaccine by age 6 years. Because measles is so easily transmitted and the worldwide measles vaccination coverage was only 72% in 1998 [32, 135], this author does not believe that it is feasible to eradicate measles worldwide using the currently available measles vaccines.

1.10.4. *Rubella*

Rubella (also called 3-day measles or German measles) is a mild disease with few complications, but a rubella infection during the first trimester of pregnancy can result in miscarriage, stillbirth, or infants with a pattern of birth defects called Congenital Rubella Syndrome (CRS) [14]. Because the goal of a rubella vaccination program is to prevent rubella infections in pregnant women, special vaccination strategies such as vaccination of 12–14 year old girls are sometimes used [74, 77]. In the prevaccine era, rubella epidemics and subsequent congenital rubella syndrome (CRS) cases occurred about every 4 to 7 years in the United States. During a major rubella epidemic in 1964, it is estimated that there were over 20,000 CRS cases in the United States with a total lifetime cost of over \$2 billion [74, 77]. Since the rubella vaccine was licensed in 1969, the incidences of rubella and CRS in the United States have decreased substantially. Since many rubella cases are subclinical and unreported, we consider only the incidence of Congenital Rubella Syndrome (CRS). The yearly incidences of CRS in the United States were between 22 and 67 in the 1970s, between 0 and 50 in the 1980s, 11 in 1990, 47 in 1991, 11 in 1992, and then between 4 and 8 between 1993 and 1999 [28]. Although there have been some increases in CRS cases associated with occasional rubella outbreaks, CRS has been at

a relatively low level in the United States in recent years. In recent rubella outbreaks in the United States, most cases occurred among unvaccinated persons aged at least 20 years and among persons who were foreign born, primarily Hispanics (63% of reported cases in 1997) [30]. Although it does not solve the problem of unvaccinated immigrants, the rubella vaccination program for children has reduced the incidence of rubella and CRS in the United States to very low levels. Worldwide eradication of rubella is not feasible, because over two-thirds of the population in the world is not yet routinely vaccinated for rubella. Indeed, the policies in China and India of not vaccinating against rubella may be the best policies for those countries, because most women of childbearing age in these countries already have disease-acquired immunity [65].

1.10.5. *Chickenpox (varicella)*

The varicella zoster virus (VZV) is the agent for varicella, commonly known as chickenpox. Chickenpox is usually a mild disease in children that lasts about 4–7 days with a body rash of several hundred lesions. After a case of chickenpox, the VZV becomes latent in the dorsal root ganglia, but VZV can reactivate in the form of zoster, commonly known as shingles. Shingles is a painful vesicular rash along one or more sensory root nerves that usually occurs when the immune system is less effective due to illness or aging [14]. People with shingles are less infectious than those with chickenpox, but they can transmit the VZV. Indeed, it was found that some isolated Amazon tribes had no antibodies to diseases such as measles, mumps and rubella, but they did have antibodies to VZV [16]. Thus it appears that the persistence of VZV in these small isolated populations has occurred because VZV can be dormant in people for many years and then be spread in the population by a case of shingles. Because of the transmission by those with both chickenpox and shingles, the expression for R_0 is more complicated than for the MSEIR model [144]. A varicella vaccine was licensed in the United States in 1995 and is now recommended for all young children. But the vaccine-immunity wanes, so that vaccinated children can get chickenpox as adults. Two possible dangers of this new varicella vaccination program are more chickenpox cases in adults, when the complication rates are higher, and an increase in cases of shingles. An age-structured epidemiologic-demographic model has been used with parameters estimated from epidemiological data to evaluate the effects of varicella vaccination programs [143]. Although the age distribution of varicella cases does shift

in the computer simulations, this shift does not seem to be a problem since many of the adult cases occur after vaccine-induced immunity wanes, so they are mild varicella cases with fewer complications. In the computer simulations, shingles incidence increases in the first 30 years after initiation of a varicella vaccination program, because people are more likely to get shingles as adults when their immunity is not boosted by frequent exposures, but after 30 years the shingles incidence starts to decrease as the population includes more previously vaccinated people, who are less likely to get shingles. Thus the simulations validate the second danger that the new vaccination program could lead to more cases of shingles in the first several decades [143].

1.10.6. *Influenza*

Type A influenza has three subtypes in humans (H1N1, H2N2, and H3N2) that are associated with widespread epidemics and pandemics (i.e. worldwide epidemics). Types B and C influenza tend to be associated with local or regional epidemics. Influenza subtypes are classified by antigenic properties of the H and N surface glycoproteins, whose mutations lead to new variants every few years [14]. An infection or vaccination for one variant may give only partial immunity to another variant of the same subtype, so that flu vaccines must be reformulated almost every year. Sometimes the experts do not choose the correct variant for the vaccine. For example, the A/California/7/2004(H3N2) variant was the dominant type A influenza in the 2004–05 flu season in the United States [36]. However, the 2004–05 influenza vaccine still contained an A/Wyoming/3/2003(H3N2)-like virus, which was the dominant type A influenza in the 2003–04 flu season in the United States. An A/California/7/2004(H3N2)-like virus has been recommended as the H3 component for the 2005–06 Northern Hemisphere flu vaccine. If the influenza virus variant did not change, then it should be easy to eradicate, because the contact number for flu has been estimated above to be only about 1.4. But the frequent drift of the A subtypes to new variants implies that flu vaccination programs cannot eradicate them because the target is constantly moving.

Completely new A subtypes (antigenic shift) emerge occasionally from unpredictable recombinations of human with swine or avian influenza antigens. These new subtypes can lead to major pandemics. A new H1N1 subtype led to the 1918–19 pandemic that killed over one-half million people in the United States and over 20 million people worldwide. Pandemics also

occurred in 1957 from the Asian Flu (an H2N2 subtype) and in 1968 from the Hong Kong flu (an H3N2 subtype) [105]. When 18 confirmed human cases with 6 deaths from an H5N1 chicken flu occurred in Hong Kong in 1997, there was great concern that this might lead to another antigenic shift and pandemic. At the end of 1997, veterinary authorities slaughtered all (1.6 million) chickens present in Hong Kong, and importation of chickens from neighboring areas was stopped. Fortunately, the H5N1 virus has so far not evolved into a form that is readily transmitted from person to person [149, 156]. But outbreaks of avian influenza continue to occur in many countries, particularly in Southeast Asia, so that the danger of the occurrence of a new type A influenza that could spread in humans continues.

1.11. Other epidemiology models with thresholds

Mathematical epidemiology has now evolved into a separate area of population dynamics that is parallel to mathematical ecology. Epidemiology models are now used to combine complex data from various sources in order to study equally complex outcomes. In this chapter we have focused on the role of the basic reproduction number R_0 , which is defined as the average number of people infected when a typical infective enters an entirely susceptible population. We have illustrated the significance of R_0 by obtaining explicit expressions for R_0 and proving threshold results which imply that a disease can invade a completely susceptible population if and only if $R_0 > 1$. Using the SIR endemic model, the estimates of R_0 for various diseases show that some diseases are more easily spread than others, so that they are more difficult to control or eradicate. We considered the effects of these differences for six diseases (smallpox, polio, measles, rubella, chickenpox, and influenza).

The results presented in this chapter provide a theoretical background for reviewing some previous results. In this section we do not attempt to cite all papers on infectious disease models with heterogeneity, and spatial structure, but primarily cite sources that consider thresholds and the basic reproduction number R_0 . The cited papers reflect the author's interests, but additional references are given in these papers and in the books and survey papers listed in the Introduction. We refer the reader to other sources for information on stochastic epidemiology models [9, 11, 39, 43, 49, 62, 100, 134], discrete time models [2, 3], models involving macroparasites [7, 43, 68], genetic heterogeneity [7, 68], plant disease models [107, 154], and wildlife disease models [68]. Models with age structure are surveyed in Chapter 3.

Irregular and biennial oscillations of measles incidences have led to various mathematical analyses including the following seven modeling explanations, some of which involve age structure. Yorke and London [159] proposed SEIR models with seasonal forcing in delay differential equations. Dietz [46] proposed subharmonic resonance in a seasonally forced SEIR model using ordinary differential equations. Schenzle [141] used computer simulations to show that the measles outbreak patterns in England and Germany could be explained by the primary school yearly calendars and entry ages. Olson and Schaffer [136] proposed chaotic behavior in simple deterministic SEIR models. Bolker and Grenfell [18] proposed realistic age structured models with seasonal forcing and stochastic terms. Ferguson, Nokes, and Anderson [59] proposed finely age-stratified models with stochastic fluctuations that can shift the dynamics between biennial and triennial cycle attractors. Earn, Rohani, Bolker, and Grenfell [53] proposed a simple, time-forced SEIR model with slow variation in the average rate of recruitment of new susceptibles.

In recent years the human immunodeficiency virus (HIV), which leads to acquired immunodeficiency syndrome (AIDS), has emerged as an important new infectious disease. Many models have been developed for HIV/AIDS. May and Anderson [123] found R_0 for some simple HIV transmission models. Bongaarts [19] and May *et al.* [125] used models with age structure to examine the demographic effects of AIDS in African countries. The book [24] by Castillo contains a review of HIV/AIDS modeling papers including single-group models, multiple-group models, and epidemiologic-demographic models. It also contains papers on AIDS models with HIV class age, variable infectivity, distributions for the AIDS incubation period, heterogeneity, and structured mixing. Busenberg and Castillo [21] found an R_0 expression for an HIV model with variable infectivity and continuous chronological and HIV class age structure and proportionate mixing. Hyman, Li and Stanley [95] generalized these results on R_0 to HIV models with non-proportionate mixing and discrete or continuous risk.

For many infectious diseases the transmission occurs in a diverse population, so the epidemiological model must divide the heterogeneous population into subpopulations or groups, in which the members have similar characteristics. This division into groups can be based not only on mode of transmission, contact patterns, latent period, infectious period, genetic susceptibility or resistance, and amount of vaccination or chemotherapy, but also on social, cultural, economic, demographic or geographic factors. For these models it is useful to find R_0 from the threshold conditions for invasion

and endemicity, and to prove stability of the equilibria. For the SIS model with n groups, the threshold was first found in terms of whether $s(A) \leq 0$ or $s(A) > 0$, where $s(A)$ is the largest real part of the eigenvalues of the Jacobian matrix A at the disease-free equilibrium. The seminal paper [110] of Lajmanovich and Yorke found this threshold condition and proved the global stability of the disease-free and endemic equilibria using Liapunov functions. This approach has been extended to SIR, SEIR, and SEIRS models with n groups [73, 150, 151]. For these models R_0 can be shown to be the spectral radius of a next generation matrix that is related to the Jacobian matrix A [79, 84]. This next generation operator approach has also been used for epidemiology models with a variety of features such as proportionate mixing, preferred mixing, heterosexual transmission, host-vector groups, multiple mixing groups, vaccination, and age structure [42, 43]. For proportionate mixing models with multiple interacting groups, the basic reproduction number R_0 is the contact number σ , which is the weighted average of the contact numbers in the groups [79, 84, 89]. The sexual transmission of diseases often occurs in a very heterogeneous population, because people with more sexual partners have more opportunities to be infected and to infect others. The basic reproduction number R_0 has been determined for many different models with heterogeneous mixing involving core, social, and sexual mixing groups [89, 101, 103, 108, 109, 148]. It has been shown that estimates of R_0 , under the false assumption that a heterogeneously mixing population is homogeneously mixing, are not greater than the actual R_0 for the heterogeneous population [1, 79]. Many models with heterogeneity in the form of competing strains of infectious agents have been considered for diseases such as influenza, dengue, and myxomatosis [8, 25, 26, 27, 47, 52, 54, 55, 57, 124, 129].

HIV/AIDS is spread in a very heterogeneous population by heterosexual intercourse, homosexual intercourse, and sharing of needles by injecting drug users. Because of the great diversity and heterogeneity among those at risk to HIV/AIDS, modeling this disease is a challenging task [5, 7, 24, 80, 90, 93, 94, 98, 102]. For HIV/AIDS models with a continuous distribution of sexual activity levels and with various preference mixing functions, the proportionate mixing has been shown to be the only separable solution, and expressions for the basic reproduction number R_0 in the proportionate mixing case has been found [17, 21]. Expressions for R_0 have also been found for HIV/AIDS models using groups of people based on their sexual behavior, e.g. homosexual men, bisexual men, heterosexual women, and heterosexual men, with further subdivisions based on their

numbers of sexual or needle-sharing partners. For staged progression models for HIV/AIDS with many infectious classes with different infectivities, the basic reproduction number R_0 is often the weighted average of the basic reproduction numbers in the infectious classes, where the weights involve the fraction of contacts (or partners) that result in an infection and the probability of reaching that infectious stage [85, 96, 97, 104, 116].

There is clear evidence that infectious diseases spread geographically and maps with isodate spread contours have been produced [7, 38, 127, 133]. Some estimated speeds of propagation are 30–60 kilometers per year for fox rabies in Europe starting in 1939 [133], 18–24 miles per year for raccoon rabies in the Eastern United States starting in 1977 [33], about 140 miles per year for the plague in Europe in 1347–1350 [133], and worldwide in one year for influenza in the 20th century [140]. Epidemiology models with spatial structures have been used to describe spatial heterogeneity [7, 72, 84] and the spatial spread of infectious diseases [23, 37, 43, 68, 133, 153]. There seem to be two types of spatial epidemiology models [131, 153]. Diffusion epidemiology models are formulated from non-spatial models by adding diffusion terms corresponding to the random movements each day of susceptibles and infectives. Dispersal-kernel models are formulated by using integral equations with kernels describing daily contacts of infectives with their neighbors. For both types of spatial epidemiology models in infinite domains, one often determines the thresholds (sometimes in terms of R_0) above which a traveling wave exists, finds the minimum speed of propagation and the asymptotic speed of propagation (which is usually shown to be equal to the minimum speed), and determines the stability of the traveling wave to perturbations [130, 138]. For spatial models in finite domains, stationary states and their stability have been investigated [23]. For stochastic spatial models there is also a threshold condition, so that the disease dies out below the threshold and approaches an endemic stationary distribution above the threshold [51].

Acknowledgements: The author thanks Daryl Daley for his editorial suggestions in Chapters 1–3, and thanks Brian Treadway for his manuscript preparation assistance.

References

1. F. R. Adler, The effects of averaging on the basic reproduction ratio, *Math. Biosci.*, 111 (1992), pp. 89–98.
2. L. J. S. Allen, Some discrete-time SI, SIR, and SIS epidemic models, *Math. Biosci.*, 124 (1994), pp. 83–105.

3. L. J. S. Allen and A. M. Burgin, Comparison of deterministic and stochastic SIS and SIR models in discrete-time, *Math. Biosci.*, 163 (2000), pp. 1-33.
4. R. M. Anderson, ed., *Population Dynamics of Infectious Diseases*, Chapman and Hall, London, 1982.
5. ———, The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS, *J. AIDS*, 1 (1988), pp. 241-256.
6. R. M. Anderson and R. M. May, eds., *Population Biology of Infectious Diseases*, Springer-Verlag, Berlin, Heidelberg, New York, 1982.
7. ———, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.
8. V. Andreasen, J. Lin, and S. A. Levin, The dynamics of cocirculating influenza strains conferring partial cross-immunity, *J. Math. Biol.*, 35 (1997), pp. 825-842.
9. N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases*, Hafner, New York, second ed., 1975.
10. ———, *The Biomathematics of Malaria*, Charles Griffin, London, 1982.
11. M. Bartlett, *Stochastic Population Models in Ecology and Epidemiology*, Methuen, London, 1960.
12. N. Becker, The use of epidemic models, *Biometrics*, 35 (1978), pp. 295-305.
13. ———, *Analysis of Infectious Disease Data*, Chapman and Hall, New York, 1989.
14. A. S. Benenson, *Control of Communicable Diseases in Man*, American Public Health Association, Washington, D.C., sixteenth ed., 1995.
15. D. Bernoulli, Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir, in *Mémoires de Mathématiques et de Physique*, Académie Royale des Sciences, Paris, 1760, pp. 1-45.
16. F. L. Black, W. J. Hierholzer, F. d. P. Pinheiro, et al., Evidence for the persistence of infectious disease agents in isolated human populations, *Am. J. Epidemiol.*, 100 (1974), pp. 230-250.
17. S. P. Blythe and C. Castillo-Chavez, Like with like preference and sexual mixing models, *Math. Biosci.*, 96 (1989), pp. 221-238.
18. B. M. Bolker and B. T. Grenfell, Chaos and biological complexity in measles dynamics, *Proc. Roy. Soc. London Ser. B*, 251 (1993), pp. 75-81.
19. J. Bongaarts, A model of the spread of HIV infection and the demographic impact of AIDS, *Stat. Med.*, 8 (1989), pp. 103-120.
20. M. Burnett and D. O. White, *Natural History of Infectious Disease*, Cambridge University Press, Cambridge, fourth ed., 1974.
21. S. Busenberg and C. Castillo-Chavez, A general solution of the problem of mixing of subpopulations and its application to risk- and age-structured models for the spread of AIDS, *IMA J. Math. Appl. Med. Biol.*, 8 (1991), pp. 1-29.
22. S. Busenberg and K. L. Cooke, *Vertically Transmitted Diseases*, vol. 23 of *Biomathematics*, Springer-Verlag, Berlin, 1993.
23. V. Capasso, *Mathematical Structures of Epidemic Systems*, vol. 97 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1993.

24. C. Castillo-Chavez, ed., *Mathematical and Statistical Approaches to AIDS Epidemiology*, vol. 83 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1989.
25. C. Castillo-Chavez, H. W. Hethcote, V. Andreasen, S. A. Levin, and W. M. Liu, Epidemiological models with age structure, proportionate mixing, and cross-immunity, *J. Math. Biol.*, 27 (1989), pp. 233–258.
26. C. Castillo-Chavez, W. Huang, and J. Li, Competitive exclusion in gonorrhea models and other sexually-transmitted diseases, *SIAM J. Appl. Math.*, 56 (1996), pp. 494–508.
27. ———, Competitive exclusion and multiple strains in an SIS STD model, *SIAM J. Appl. Math.*, 59 (1999), pp. 1790–1811.
28. Centers for Disease Control and Prevention, Summary of notifiable diseases, United States, 1998, *MMWR*, 47 (1999), pp. 78–83.
29. ———, Progress toward global poliomyelitis eradication, 1997–1998, *MMWR*, 48 (1999), pp. 416–421.
30. ———, Rubella outbreak—Westchester County, New York, 1997–1998, *MMWR*, 48 (1999), pp. 560–563.
31. ———, Epidemiology of measles—United States, 1998, *MMWR*, 48 (1999), pp. 749–753.
32. ———, Global measles control and regional elimination, 1998–99, *MMWR*, 48 (1999), pp. 1124–1130.
33. ———, Update: Raccoon rabies epizootic—United States and Canada, 1999, *MMWR*, 49 (2000), pp. 31–35.
34. ———, Recommended childhood immunization schedule—United States, 2000, *MMWR*, 49 (2000), pp. 35–47.
35. ———, Measles outbreak—Netherlands, April 1999–January 2000, *MMWR*, 49 (2000), pp. 299–303.
36. ———, Update: Influenza Activity—United States and Worldwide, 2004–05 season, *MMWR*, 54 (2005), pp. 631–634.
37. A. D. Cliff, Incorporating spatial components into models of epidemic spread, in *Epidemic Models: Their Structure and Relation to Data*, D. Mollison, ed., Cambridge University Press, Cambridge, 1996, pp. 119–149.
38. A. D. Cliff and P. Haggett, *Atlas of Disease Distributions: Analytic Approaches to Epidemiological Data*, Blackwell, London, 1988.
39. D. J. Daley and J. Gani, *Epidemic Modelling: an Introduction*, Cambridge University Press, Cambridge, 1999.
40. M. C. M. De Jong, O. Diekmann, and J. A. P. Heesterbeek, How does transmission depend on population size?, in *Human Infectious Diseases, Epidemic Models*, D. Mollison, ed., Cambridge University Press, Cambridge, 1995, pp. 84–94.
41. S. Del Valle, H. W. Hethcote, J. M. Hyman, and C. Castillo-Chavez, Effects of Behavioral Changes in a Smallpox Attack Model, *Mathematical Biosciences* 195 (2005) 228–251.
42. O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious

- diseases in heterogeneous populations, *J. Math. Biol.*, 28 (1990), pp. 365–382.
43. O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases*, Wiley, New York, 2000.
44. K. Dietz, Epidemics and rumours: A survey, *J. Roy. Statist. Soc. Ser. A*, 130 (1967), pp. 505–528.
45. ———, Transmission and control of arbovirus diseases, in *Epidemiology*, K. L. Cooke, ed., SIAM, Philadelphia, 1975, pp. 104–121.
46. ———, The incidence of infectious diseases under the influence of seasonal fluctuations, in *Mathematical Models in Medicine*, J. Berger, W. Buhler, R. Repges, and P. Tautu, eds., vol. 11 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1976, pp. 1–15.
47. ———, Epidemiologic inference of virus populations, *J. Math. Biol.*, 8 (1979), pp. 291–300.
48. ———, Density dependence in parasite transmission dynamics, *Parasit. Today*, 4 (1988), pp. 91–97.
49. ———, The first epidemic model: A historical note on P. D. En'ko, *Austral. J. Statist.*, 30A (1988), pp. 56–65.
50. K. Dietz and D. Schenzle, *Mathematical models for infectious disease statistics*, in *A Celebration of Statistics*, A. C. Atkinson and S. E. Feinberg, eds., Springer-Verlag, New York, 1985, pp. 167–204.
51. R. Durrett, Stochastic spatial models, *SIAM Rev.*, 41 (1999), pp. 677–718.
52. G. Dwyer, S. A. Levin, and L. Buttel, A simulation model of the population dynamics and evolution of myxomatosis, *Ecological Monographs*, 60 (1990), pp. 423–447.
53. D. J. D. Earn, P. Rohani, B. M. Bolker, and B. T. Grenfell, A simple model for complex dynamical transitions in epidemics, *Science*, 287 (2000), pp. 667–670.
54. L. Esteva and C. Vargas, Analysis of a dengue disease transmission model, *Math. Biosci.*, 150 (1998), pp. 131–151.
55. ———, A model for dengue disease with variable human population, *J. Math. Biol.*, 38 (1999), pp. 220–240.
56. A. S. Evans, *Viral Infections of Humans*, Plenum Medical Book Company, New York, second ed., 1982.
57. Z. Feng and J. X. Velasco-Hernandez, Competitive exclusion in a vector-host model for the dengue fever, *J. Math. Biol.*, 35 (1997), pp. 523–544.
58. F. Fenner, D. A. Henderson, I. Arita, Z. Jezek, and I. D. Ladnyi, *Smallpox and its Eradication*, World Health Organization, Geneva, 1988.
59. N. M. Ferguson, D. J. Nokes, and R. M. Anderson, Dynamical complexity in age-structured models of the transmission of measles virus: Epidemiological implications of high levels of vaccine uptake, *Math. Biosci.*, 138 (1996), pp. 101–130.
60. N. M. Ferguson, C. A. Donnelly, R. M. Anderson, The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions, *Science*, 292 (2001) 1155–1160.

61. J. C. Frauenthal, *Mathematical Modeling in Epidemiology*, Springer-Verlag Universitext, Berlin, 1980.
62. J. P. Gabriel, C. Lefever, and P. Picard, eds., *Stochastic Processes in Epidemic Theory*, Springer-Verlag, Berlin, 1990.
63. L. Q. Gao and H. W. Hethcote, Disease transmission models with density-dependent demographics, *J. Math. Biol.*, 30 (1992), pp. 717–731.
64. L. Garrett, *The Coming Plague*, Penguin, New York, 1995.
65. L. Q. Gao and H. W. Hethcote, Simulations of rubella vaccination strategies in China, *Math. Biosciences*, to appear.
66. L. Q. Gao, J. Mena-Lorca, and H. W. Hethcote, Four SEI endemic models with periodicity and separatrices, *Math. Biosci.*, 128 (1995), pp. 157–184.
67. ———, Variations on a theme of SEI endemic models, in *Differential Equations and Applications to Biology and to Industry*, M. Martelli et al., eds., World Scientific Publishing, Singapore, 1996, pp. 191–207.
68. B. T. Grenfell and A. P. Dobson, eds., *Ecology of Infectious Diseases in Natural Populations*, Cambridge University Press, Cambridge, 1995.
69. J. K. Hale, *Ordinary Differential Equations*, Wiley-Interscience, New York, 1969.
70. W. H. Hamer, Epidemic disease in England, *Lancet*, 1 (1906), pp. 733–739.
71. H. W. Hethcote, Asymptotic behavior in a deterministic epidemic model, *Bull. Math. Biology* 35 (1973) 607–614.
72. H. W. Hethcote, Qualitative analyses of communicable disease models, *Math. Biosci.*, 28 (1976), pp. 335–356.
73. ———, An immunization model for a heterogeneous population, *Theoret. Population Biol.*, 14 (1978), pp. 338–349.
74. ———, Measles and rubella in the United States, *Am. J. Epidemiol.*, 117 (1983), pp. 2–13.
75. ———, Optimal ages of vaccination for measles, *Math. Biosci.*, 89 (1988), pp. 29–52.
76. ———, Three basic epidemiological models, in *Applied Mathematical Ecology*, L. Gross, T. G. Hallam, and S. A. Levin, eds., Springer-Verlag, Berlin, 1989, pp. 119–144.
77. ———, Rubella, in *Applied Mathematical Ecology*, L. Gross, T. G. Hallam, and S. A. Levin, eds., Springer-Verlag, Berlin, 1989, pp. 212–234.
78. ———, A thousand and one epidemic models, in *Frontiers in Theoretical Biology*, S. A. Levin, ed., vol. 100 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1994, pp. 504–515.
79. ———, Modeling heterogeneous mixing in infectious disease dynamics, in *Models for Infectious Human Diseases*, V. Isham and G. F. H. Medley, eds., Cambridge University Press, Cambridge, 1996, pp. 215–238.
80. ———, Modeling AIDS prevention programs in a population of homosexual men, in *Modeling the AIDS Epidemic: Planning, Policy and Prediction*, E. H. Kaplan and M. L. Brandeau, eds., Raven Press, New York, 1994, pp. 91–107.
81. ———, The Mathematics of Infectious Diseases, *SIAM Review* 42 (2000) 599–653.

82. H. W. Hethcote and S. A. Levin, Periodicity in epidemiological models, in *Applied Mathematical Ecology*, L. Gross, T. G. Hallam, and S. A. Levin, eds., Springer-Verlag, Berlin, 1989, pp. 193–211.
83. H. W. Hethcote, H. W. Stech, and P. van den Driessche, Periodicity and stability in epidemic models: A survey, in *Differential Equations and Applications in Ecology, Epidemics and Population Problems*, S. N. Busenberg and K. L. Cooke, eds., Academic Press, New York, 1981, pp. 65–82.
84. H. W. Hethcote and J. W. Van Ark, Epidemiological models with heterogeneous populations: Proportionate mixing, parameter estimation and immunization programs, *Math. Biosci.*, 84 (1987), pp. 85–118.
85. ———, *Modeling HIV Transmission and AIDS in the United States*, vol. 95 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1992.
86. H. W. Hethcote and P. van den Driessche, Some epidemiological models with nonlinear incidence, *J. Math. Biol.*, 29 (1991), pp. 271–287.
87. H. W. Hethcote and P. Waltman, Optimal vaccination schedules in a deterministic epidemic model, *Math. Biosci.* 18 (1973) 365–382.
88. H. W. Hethcote, W. Wang, and Y. Li, Species coexistence and periodicity in host-host pathogen models, *J. Math. Biol.* 51 (2005) 629–660.
89. H. W. Hethcote and J. A. Yorke, *Gonorrhea Transmission Dynamics and Control*, vol. 56 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1984.
90. W. Huang, K. L. Cooke, and C. Castillo-Chavez, Stability and bifurcation for a multiple-group model for the dynamics of HIV/AIDS transmission, *SIAM J. Appl. Math.*, 52 (1992), pp. 835–854.
91. H. F. Hull, N. A. Ward, B. F. Hull, J. B. Milstein, and C. de Quatros, Paralytic poliomyelitis: Seasoned strategies, disappearing disease, *Lancet*, 343 (1994), pp. 1331–1337.
92. S. S. Hutchins, L. E. Markowitz, W. L. Atkinson, E. Swint, and S. Hadler, Measles outbreaks in the United States, 1987 through 1990, *Pediatr. Infect. Dis. J.*, 15 (1996), pp. 31–38.
93. J. M. Hyman and E. A. Stanley, Using mathematical models to understand the AIDS epidemic, *Math. Biosci.*, 90 (1988), pp. 415–473.
94. ———, The effect of social mixing patterns on the spread of AIDS, in *Mathematical Approaches to Problems in Resource Management and Epidemiology*, C. Castillo-Chavez, S. A. Levin, and C. Shoemaker, eds., vol. 81 of *Lecture Notes in Biomathematics*, Springer-Verlag, New York, 1989.
95. J. M. Hyman, J. Li, and E. A. Stanley, Threshold conditions for the spread of HIV infection in age-structured populations of homosexual men, *J. Theor. Biol.*, 166 (1994), pp. 9–31.
96. ———, The differential infectivity and staged progression models for the transmission of HIV, *Math. Biosci.*, 155 (1999), pp. 77–109.
97. J. M. Hyman and J. Li, An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations. *Math. Biosci.*, to appear.

98. V. Isham, Mathematical modeling of the transmission dynamics of HIV infection and AIDS: A review, *J. Roy. Statist. Soc. Ser. A*, 151 (1988), pp. 5–31.
99. V. Isham and G. Medley, eds., *Models for Infectious Human Diseases*, Cambridge University Press, Cambridge, 1996.
100. J. A. Jacquez and C. P. Simon, The stochastic SI model with recruitment and deaths I: Comparison with the closed SIS model, *Math. Biosci.*, 117 (1993), pp. 77–125.
101. J. A. Jacquez, C. P. Simon, and J. S. Koopman, The reproduction number in deterministic models of contagious diseases, *Curr. Topics in Theor. Biol.*, 2 (1991), pp. 159–209.
102. J. A. Jacquez, J. S. Koopman, C. P. Simon, and I. M. Longini, Role of the primary infection in epidemics of HIV infection in gay cohorts, *J. AIDS*, 7 (1994), pp. 1169–1184.
103. J. A. Jacquez, C. P. Simon, and J. S. Koopman, Core groups and the R_0 's for subgroups in heterogeneous SIS models, in *Epidemic Models: Their Structure and Relation to Data*, D. Mollison, ed., Cambridge University Press, Cambridge, 1996, pp. 279–301.
104. J. A. Jacquez, C. P. Simon, J. S. Koopman, L. Sattenspiel, and T. Perry, Modeling and analyzing HIV transmission: The effect of contact patterns, *Math. Biosci.*, 92 (1988), pp. 119–199.
105. M. M. Kaplan and R. G. Webster, The epidemiology of influenza, *Scientific American*, 237 (1977), pp. 88–105.
106. W. O. Kermack and A. G. McKendrick, Contributions to the mathematical theory of epidemics, part 1, *Proc. Roy. Soc. London Ser. A*, 115 (1927), pp. 700–721.
107. J. Kranz, ed., *Epidemics of Plant Diseases: Mathematical Analysis and Modelling*, Springer-Verlag, Berlin, 1990.
108. C. M. Kribs-Zaleta, Structured models for heterosexual disease transmission, *Math. Biosci.*, 160 (1999), pp. 83–108.
109. ———, Core recruitment effects in SIS models with constant total populations, *Math. Biosci.*, 160 (1999), pp. 109–158.
110. A. Lajmanovich and J. A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, *Math. Biosci.*, 28 (1976), pp. 221–236.
111. H. A. Lauwerier, *Mathematical Models of Epidemics*, Mathematisch Centrum, Amsterdam, 1981.
112. R. Levins, T. Awerbuch, U. Brinkman, et al., The emergence of new diseases, *American Scientist*, 82 (1994), pp. 52–60.
113. M. Y. Li and J. S. Muldowney, Global stability for the SEIR model in epidemiology, *Math. Biosci.*, 125 (1995), pp. 155–164.
114. M. Y. Li, J. R. Graef, L. Wang, and J. Karsai, Global dynamics of a SEIR model with varying population size, *Math. Biosci.*, 160 (1999), pp. 191–213.
115. M. Y. Li, J. S. Muldowney, and P. van den Driessche, Global stability for the SEIRS models in epidemiology, *Canadian Quart. Appl. Math.*, submitted.

116. X. Lin, H. W. Hethcote, and P. van den Driessche, An epidemiological model for HIV/AIDS with proportional recruitment, *Math. Biosci.*, 118 (1993), pp. 181–195.
117. W. M. Liu, H. W. Hethcote, and S. A. Levin, Dynamical behavior of epidemiological models with nonlinear incidence rates, *J. Math. Biol.*, 25 (1987), pp. 359–380.
118. W. M. Liu, S. A. Levin, and Y. Iwasa, Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models, *J. Math. Biol.*, 23 (1986), pp. 187–204.
119. W. P. London and J. A. Yorke, Recurrent outbreaks of measles, chickenpox and mumps I: Seasonal variation in contact rates, *Am. J. Epidemiol.*, 98 (1973), pp. 453–468.
120. I. P. Longini, E. Ackerman, L. R. Elveback, An optimization model for influenza A epidemics. *Math. Biosci.* 38 (1978) 141–157.
121. D. Ludwig and K. L. Cooke, eds., *Epidemiology*, SIMS Utah Conference Proceedings, SIAM, Philadelphia, 1975.
122. P. Martens, How will climate change affect human health?, *American Scientist*, 87 (1999), pp. 534–541.
123. R. M. May and R. M. Anderson, Transmission dynamics of HIV infection, *Nature*, 326 (1987), pp. 137–142.
124. ———, Parasite-host coevolution, *Parasitology*, 100 (1990), pp. S89–S101.
125. R. M. May, R. M. Anderson, and A. R. McLean, Possible demographic consequences of HIV/AIDS, *Math. Biosci.*, 90 (1988), pp. 475–506.
126. A. G. McKendrick, Applications of mathematics to medical problems, *Proc. Edinburgh Math. Soc.*, 44 (1926), pp. 98–130.
127. W. H. McNeill, *Plagues and People*, Blackwell, Oxford, 1976.
128. J. Mena-Lorca and H. W. Hethcote, Dynamic models of infectious diseases as regulators of population sizes, *J. Math. Biol.*, 30 (1992), pp. 693–716.
129. J. Mena-Lorca, J. X. Velasco-Hernandez, and C. Castillo-Chavez, Density-dependent dynamics and superinfection in an epidemic model, *IMA J. Math. Appl. Med. Biol.*, 16 (1999), pp. 307–317.
130. J. A. J. Metz and F. van den Bosch, Velocities of epidemic spread, in *Epidemic Models: Their Structure and Relation to Data*, D. Mollison, ed., Cambridge University Press, Cambridge, 1996, pp. 150–186.
131. D. Mollison, Dependence of epidemic and population velocities on basic parameters, *Math. Biosci.*, 107 (1991), pp. 255–287.
132. ———, ed., *Epidemic Models: Their Structure and Relation to Data*, Cambridge University Press, Cambridge, 1996.
133. J. D. Murray, *Mathematical Biology*, Springer-Verlag, Berlin, 1989.
134. I. Nasell, *Hybrid Models of Tropical Infections*, Springer-Verlag, Berlin, 1985.
135. M. B. A. Oldstone, *Viruses, Plagues, and History*, Oxford University Press, New York, 1998.
136. L. F. Olson and W. M. Schaffer, Chaos versus noisy periodicity: Alternative hypotheses for childhood epidemics, *Science*, 249 (1990), pp. 499–504.
137. R. Preston, *The Hot Zone*, Random House, New York, 1994.

138. J. Radcliffe and L. Rass, Spatial Deterministic Epidemics. Mathematical Surveys and Monographs, American Mathematical Society, Providence, to appear, 2000.
139. R. Ross, The Prevention of Malaria, Murray, London, second ed., 1911.
140. L. A. Rvachev and I. M. Longini, A mathematical model for the global spread of influenza, *Math. Biosci.*, 75 (1985), pp. 3–22.
141. D. Schenzle, An age-structured model of pre- and post-vaccination measles transmission, *IMA J. Math. Appl. Med. Biol.*, 1 (1984), pp. 169–191.
142. L. Schlien, Hunting down the last of the poliovirus, *Science*, 279 (1998), p. 168.
143. M. C. Schuette and H. W. Hethcote, Modeling the effects of varicella vaccination programs on the incidence of chickenpox and shingles, *Bull. Math. Biol.*, 61 (1999), pp. 1031–1064.
144. M. Schuette, A quantitative analysis of the transmission of the varicella-zoster virus. *Math. Biosci.* 182 (2003) 113–126.
145. M. E. Scott and G. Smith, eds., *Parasitic and Infectious Diseases*, Academic Press, San Diego, 1994.
146. D. Shalala, Bioterrorism: How prepared are we?, *Emerg. Infect. Dis.*, 5 (1999). Available from URL: <http://www.cdc.gov/ncidod/eid/vol5no4/shalala.htm>.
147. R. Shilts, *And The Band Played On*, St. Martin's Press, New York, 1987.
148. C. P. Simon and J. A. Jacquez, Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations, *SIAM J. Appl. Math.*, 52 (1992), pp. 541–576.
149. R. Snacken, A. P. Kendal, L. R. Haaheim, and J. M. Wood, The next influenza pandemic: Lessons from Hong Kong, 1997, *Emerg. Infect. Dis.*, 5 (1999). Available from URL: <http://www.cdc.gov/ncidod/eid/vol5no2/snacken.htm>.
150. H. R. Thieme, Global asymptotic stability in epidemic models, in *Equadiff 82 Proceedings*, H. W. Knobloch and K. Schmitt, eds., vol. 1017 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1983, pp. 609–615.
151. ———, Local stability in epidemic models for heterogeneous populations, in *Mathematics in Biology and Medicine*, V. Capasso, E. Grosso, and S. L. Paveri-Fontana, eds., vol. 97 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1985, pp. 185–189.
152. ———, Persistence under relaxed point-dissipativity (with an application to an endemic model), *SIAM J. Math. Anal.*, 24 (1993), pp. 407–435.
153. F. van den Bosch, J. A. J. Metz, and O. Diekmann, The velocity of spatial population expansion, *J. Math. Biol.*, 28 (1990), pp. 529–565.
154. J. E. Vanderplank, *Plant Diseases: Epidemics and Control*, Academic Press, New York, 1963.
155. P. E. Waltman, Deterministic Threshold Models in the Theory of Epidemics, vol. 1 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, 1974.

156. R. G. Webster, Influenza: An emerging disease, *Emerg. Infect. Dis.*, 4 (1998). Available from URL: <http://www.cdc.gov/ncidod/eid/vol4no3/webster.htm>.
157. K. Wickwire, Mathematical models for the control of pests and infectious diseases: A survey, *Theoret. Population Biol.*, 11 (1977), pp. 182–238.
158. M. Woolhouse, M. Chase-Topping, D. Haydon, J. Friar, L. Matthews, G. Hughes, D. Shaw, J. Wilesmith, A. Donaldson, S. Cornel, M. Keeling, and Bryan Grenfell, Foot-and-mouth disease under control in the UK. *Nature* 411(2001) 258–259.
159. J. A. Yorke and W. P. London, Recurrent outbreaks of measles, chickenpox, and mumps II, *Am. J. Epidemiol.*, 98 (1973), pp. 469–482.