10. Dynamics of Infectious Diseases: Epidemic Models and AIDS

10.1 Historical Aside on Epidemics

The history of epidemics is an ever fascinating area; the 14th century Black Death is just the most famous epidemic historically (see Chapter 13, Volume II, which deals with the spatial spread of epidemics, for a brief history of it). In Europe, which had a population of around 85 million at the time, about a third of the population died.

One epidemic which has exercised classical scholars for a very long time is the Plague of Athens (430–428 BC) described in great detail by Thucydides including the symptoms and disease progression. He also gave some exact figures such as that 1050 of 4000 soldiers on an expedition died of the disease. The disease described so minutely by Thucydides, even to the fact that dogs who ate the dead bodies also suffered, has been the source of numerous articles over some hundreds of years with cases being made (with great conviction and defended vehemently) for an incredible range of diseases such as bubonic plague, measles, Malta fever, smallpox, scarlet fever, typhus, typhoid fever and many others. The symptoms described by Thucydides are (i) heat in the head, (ii) inflammation of the eyes, (iii) suffusion with blood of the tongue and throat, (iv) foetid breath, (v) hoarseness with violent coughing, (vi) vomiting of bile, (vii) retching and convulsions, (viii) pustular and ulcerating eruptions of the skin, (ix) total body hyperaesthesia and restlessness, (x) irresistible desire for water to assuage thirst and immersion therein to alleviate body heat, (xi) terminal exhaustion apparently produced by diarrhoea, (xii) loss of toes, fingers and genitalia, (xiii) destruction of eyes and, (xiv) if recovery occurs, amnesia, the latter no doubt a blessing. Based on the symptoms none of the above suggestions seems to fit the Athens disease. Whatever it was it was certainly very nasty. An interesting review article on the Athens plague is given by Poole and Holladay (1979). They conclude that it has either become extinct or has been modified over the millennia. Since then other articles have appeared with yet other possiblities.

One of the interesting aspects of Thucydides' account is that there is no mention of person-to-person contagion which we now accept so freely with diseases. It was only in the 19th century that it was beginning to be discussed. Evil exhalations from the earth, aerial miasmata and so on were generally accepted. The latter explanation for some diseases, or rather illnesses, is not as ridiculous as it might at first appear when you think of the number of people, with the same epidemiclike medical problems, who live on contaminated ground or in regions where the water is iodine-deficient resulting in goitres to mention just two examples. Many South-East Asians can be forgiven for believing

that the smog and smoke belching from the forest fires in Indonesia are responsible for the large upsurge of dengue fever, carried by the mosquito, *Aedes aegypti*. This is a man-made mosquito in effect since it breeds in urban areas in water gathering in plastic, rubber and metallic containers that litter many poor urban areas.

The study of epidemics with its long history has come up with an astonishing number and variety of models and explanations for the spread and cause of epidemic outbreaks. Even today they are often attributed to evil spirits or displeased gods. For example, AIDS (autoimmune deficiency syndrome), the dominant epidemic of the past 20 years and the major one since the 1918[1] influenza pandemic have been ascribed by many as a punishment sent by God. Hippocrates (459–377 BC), in his essay on 'Airs, Waters and Localities' wrote that one's temperament, personal habits and environment were important factors—not unreasonable even today, particularly so in view of the comments in the last paragraph. Somewhat less relevant, but not without its moments of humour, is Alexander Howe's (1865) book in which he sets out his 'Laws of Pestilence' in 31 propositions of which the following, proposition 2, is typical: 'The length of the interval between successive periodic visitations corresponds with the period of a single revolution of the lunar node, and a double revolution of the lunar apse time.'

The first major epidemic in the U.S.A. was the Yellow Fever epidemic in Philadelphia in 1793 in which about 5000 people died out of a population of around 50,000, although estimates suggest that about 20,000 fled the city; see the interesting *Scientific American* article by Foster et al. (1998) and the book by Powell (1993). The epidemic story here is a saga of wild, as well as sensible, theories as to cause and treatment, petty jealousies with disastrous consequencies, genuine humanity and fomented controversies. A leading physician was the strongest advocate of bleeding as the appropriate treatment while others recommended cleanliness, rest, Peruvian bark and wine. This epidemic had a major impact on the subsequent life and politics of the country.

The landmark book by McNeill (1989) is a fascinating story of the relation between disease and people. More recently there have been several books which try to explain various aspects of diseases from the triumphs of medicine (Oldstone 1998) to the socioeconomic (Watts 1998). The latter is written from a very anti-European, western-imperial-colonialists-are-responsible-for-it-all, viewpoint. Europeans are blamed for most of the world's problems with infectious diseases. Leaving aside some of his wilder assertions,[2] the polemics and the emotional outbursts, he has diligently researched historical data and unearthed some dreadful examples of how diseases have been spread by the stupidity of certain colonial western nations with horrifying consequences. Watts'

[1] The influenza epidemic in 1918–1919 is the most deadly pandemic (that is, a world epidemic) per unit time in recorded history and somewhat surprisingly has been to a large extent ignored in historical studies until relatively recently. The Black Death palls in comparison with its severity. The original estimate of the number that died is continually being upgraded. A meeting on the epidemic in 1998 concluded that as many as 100 million people died. Coming towards the end of World War I some people at the time thought it was perhaps germ warfare. If a similar virulent influenza struck in the U.S.A. now, on the order of 1.5 million would die, although current medical treatments could possibly reduce that figure if vaccine could be produced quickly enough. It is about 20 years since the last flu epidemic and many epidemioloigists feel the next is overdue in the cycle of such outbreaks.

[2] For example, Watts asserts that syphilis in the 17th to 19th centuries was a consequence of the Christians' opposition to masturbation.

book is an important contribution to the history and current global relevance of infectious diseases.

Since the end of World War II, public health strategy has focused on the elimination and control of organisms which cause disease. The advent of new antibiotics changed the whole ethos of disease control. Just over 20 years ago, in 1978, the United Nations signed the 'Health for All, 2000' accord which set the ambitious goal of the eradication of disease by the year 2000. AIDS at the time had not yet been discovered, or perhaps recognised is a better word, and in the year before, the last known case of smallpox had been treated. There was certainly cause for optimism albeit short lived. Scientists thought that microbes were biologically stationary targets and hence would not mutate in resistance to drugs and other biological influences.

This comforting image of unchanging microbes started to change shortly after this time with the emergence of microbes that could swim in a pool of bleach, grow on a bar of soap, and ignore doses of penicillin logarithmically larger than those effective in the 1950's (Garrett 1996). The practical reality of bacterial mutation is dramatically seen in New York City with tuberculosis. Control of the W-strain of the disease, which first appeared in the city in 1992, is resistant to every available drug and kills over half its victims, has already cost more than $1 billion. It was only 20 years ago that it was predicted that tuberculosis would be eradicated in the world by 2000.

Another aspect in the current spread of disease is with the modern era of transportation allowing more than a million people a day to cross international borders, the threat of a major outbreak of exotic diseases is very real. The population explosion, especially in underdeveloped countries, is another factor in the microbes' favour. These played key roles in the proliferation of HIV (human immunodeficiency virus) in the 1980's. Recently the World Health Organization (WHO) estimated that over 30 million people worldwide are currently infected with HIV. Information on global and country-specific disease statistics can be found on the Web pages of places such as the WHO (www.who.org) and the Centers for Disease Control (CDC: www.cdc.gov) in Atlanta.

Diseases (including such as heart disease and cancer) cause orders of magnitude more deaths in the world than anything else, even wars and famines. The appearance of new diseases, and resurgence of old ones, makes the case for interdisciplinary involvement ever more pressing. Modelling can play an increasingly significant role. Historians can also play a role. Like the plague of Athens much has been written about the 'sweating sickness' of the late 15th and first half of the 16th centuries in England.[3] The symptoms of the progression of the disease are, among others, high temperatures, body filling with fluid, particularly the lungs, the apparently well-being of a person in the morning and death the same day or within a day or two. The symptoms are so similar to those of the hanta virus in the 1993 outbreak in the Southwest U.S.A. that there is a plausible case they are the same disease but which has been dormant for several hundred years. There is some justification in believing that some of the new diseases are in fact reappearances of old ones.

[3]Henry VIII of England succeeded to the throne because his older brother died of the sweating sickness, and changed the course of history. Henry, for example, dissolved the monasteries, helped usher in the Reformation and developed the British Navy as a professional service which was the basis for the later development of the British Empire.

There are four main disease-causing microrganisms: viruses, bacteria, parasites and fungi. In this chapter, we describe some models for the population dynamics of disease agents and later (in Chapter 13, Volume II) the spatiotemporal spread of infections. Such models have been commonly used to model the spread of viral, bacterial and parasitic infections but considerably less so with fungal infections. We shall discuss several models and then try to exploit the models in the control, or ideally the eradication, of the disease or infection we are considering. The practical use of such models must rely heavily on the realism put into the models. As usual, this does not mean the inclusion of all possible effects, but rather the incorporation in the model mechanisms, in as simple a way as possible, of what appear to be the major components. Like most models they generally go through several versions before qualitative phenomena can be explained or predicted with any degree of confidence. Great care must be exercised before practical use is made of any epidemic models. However, even simple models should, and frequently do, pose important questions with regard to the underlying process and possible means of control of the disease or epidemic. One such case study, which went through various hypothetical scenarios, is the model proposed by Capasso and Paveri-Fontana (1979) for the 1973 cholera epidemic in the port city of Bari in southern Italy.[4]

An interesting early mathematical model, involving a nonlinear ordinary differential equation, by Bernoulli (1760), considered the effect of cow-pox inoculation on the spread of smallpox. The article has some interesting data on child mortality at the time. It is probably the first time that a mathematical model was used to assess the practical advantages of a vaccination control programme. Thucydides mentions immunity in connection with the Athens plague and there is evidence of an even more ancient Chinese custom where children were made to inhale powders made from the crusts of skin lesions of people recovering from smallpox.

Models can also be extremely useful in giving reasoned estimates for the level of vaccination for the control of directly transmitted infectious diseases. We discuss one case study later in the chapter when modelling bovine tuberculosis; see, for example, Anderson and May (1982, 1985, 1991), and Herbert et al. (1994). The recent paper by Schuette and Hethcote (1999) discusses vaccination protocols in connection with chickenpox and shingles and highlights certain dangers of extensive vaccination. Among other things, they evaluate with their models the effects of different vaccination programmes. The classical theoretical papers on epidemic models by Kermack and McKendrick (1927, 1932, 1933) have had a major influence in the development of mathematical models and are still relevant in a suprising number of epidemic situations; we

[4]In the epidemic, cases of cholera were most common in the poorer areas near the port. At the time raw sewage from the hospital that treated the cholera patients went directly into the sea. One suggestion was that the bacteria infected local people bathing in the area. On investigation this did not seem to be borne out. Another thought was that the water in the stand pipes, commonly in use in these districts, was contaminated. Again this was found not to be the case. Yet another thought was that the cholera entered the mussel population which was caught in the shore areas near the port and which was sold and eaten at the local stalls and shops by the local inhabitants as a delicacy, thus passing it on to humans. However, after a few hours away from direct bacterial contact mussels actually kill the cholera bacteria so this was also discarded since several hours elapsed between catching and selling. The solution was finally found to be indeed in the infected sea water. The stall holders kept a bucket of (contaminated) sea water with which they regularly doused the displayed mussels to make them look fresh and succulent. It was the bacteria in the 'fresh' sea water sprayed on the shells which caused the cholera infection.

describe some of these in this chapter. The modelling literature is now extensive and growing very quickly. Although now quite old, a good introduction to the variety of problems and models for the spread and control of infectious diseases is the book by Bailey (1975). The article by Hethcote (1994) reviews three basic epidemiological models. The book by Diekmann and Heesterbeek (2000) is a good introduction to the field. For example, they discuss how to use biological assumptions in constructing models and present applications; they cover both deterministic and stochastic modelling. Other sources are to be found in the above references and in the papers referred to in the rest of this chapter. Particularly useful sources for the latest information on specific diseases, either globally or for a specific country, are the WHO (http://www.who.org/) and the CDC (http://www.cdc.gov/); their search and information features are very efficient.

In this chapter we discuss several quite different models for very different diseases which incorporate some general aspects of epidemiological modelling of disease transmission, time evolution of epidemics, acquired resistance to infection, vaccination strategies and so on. The use of mathematical modelling in immunology and virology is also growing very quickly. We discuss in some detail models for the dynamics of HIV infections and relate them to patient data. We also discuss a bacterial infection and one involving parasites. In Chapter 13, Volume II we consider the geographic spread of infectious diseases and describe in detail a practical model for the spatial spread of rabies, a possible means of its control and the effect of including immunity. The modelling of infectious diseases involves the concepts of population dynamics which we have dicussed in earlier chapters. Although the detailed forms of the equations are different the essential elements and analysis are very similar.

At the basic level we consider two types of models. In one the total population is taken to be approximately constant with, for example, the population divided into susceptible, infected and immune groups: other groupings are also possible, depending on the disease. We first discuss models in this category. In the other, the population size is affected by the disease via the birth rate, mortality and so on. Host–parasite interacting populations often come into this category. We only discuss deterministic models which are deficient in certain situations—eradication of a disease is one, since here the probability that the last few infected individuals will infect another susceptible is not deterministic. Nevertheless it is perhaps surprising how useful, and quantitatively predictive, deterministic models can often be; the examples below are only a very few examples where this has proven to be the case.

10.2 Simple Epidemic Models and Practical Applications

In the classical (but still highly relevant) models we consider here the total population is taken to be constant. If a small group of infected individuals is introduced into a large population, a basic problem is to describe the spread of the infection within the population as a function of time. Of course this depends on a variety of circumstances, including the actual disease involved, but as a first attempt at modelling directly transmitted diseases we make some not unreasonable general assumptions.

Consider a disease which, after recovery, confers immunity which, if lethal, includes deaths: dead individuals are still counted. Suppose the disease is such that the

population can be divided into three distinct classes: the susceptibles, S, who can catch the disease; the infectives, I, who have the disease and can transmit it; and the removed class, R, namely, those who have either had the disease, or are recovered, immune or isolated until recovered. The progress of individuals is schematically represented by

$$S \longrightarrow I \longrightarrow R.$$

Such models are often called *SIR* models. The number of classes depends on the disease. *SI* models, for example, have only susceptible and infected classes while *SEIR* models have a suceptible class, S, a class in which the disease is latent, E, an infectious class, I, and a recovered or dead class, R.

The assumptions made about the transmission of the infection and incubation period are crucial in any model; these are reflected in the terms in the equations and the parameters. With $S(t)$, $I(t)$ and $R(t)$ as the number of individuals in each class we assume here that: (i) The gain in the infective class is at a rate proportional to the number of infectives and susceptibles, that is, rSI, where $r > 0$ is a constant parameter. The susceptibles are lost at the same rate. (ii) The rate of removal of infectives to the removed class is proportional to the number of infectives, that is, aI where $a > 0$ is a constant; $1/a$ is a measure of the time spent in the infectious state. (iii) The incubation period is short enough to be negligible; that is, a susceptible who contracts the disease is infective right away.

We now consider the various classes as uniformly mixed; that is, every pair of individuals has equal probability of coming into contact with one another. This is a major assumption and in many situations does not hold as in most sexually transmitted diseases (STD's). The model mechanism based on the above assumptions is then

$$\frac{dS}{dt} = -rSI, \tag{10.1}$$

$$\frac{dI}{dt} = rSI - aI, \tag{10.2}$$

$$\frac{dR}{dt} = aI, \tag{10.3}$$

where $r > 0$ is the infection rate and $a > 0$ the removal rate of infectives. This is the classic Kermack–McKendrick (1927) model. We are, of course, only interested in non-negative solutions for S, I and R. This is a basic model but, even so, we can make some highly relevant general comments about epidemics and, in fact, adequately describe some specific epidemics with such a model.

The constant population size is built into the system (10.1)–(10.3) since, on adding the equations,

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \quad \Rightarrow \quad S(t) + I(t) + R(t) = N, \tag{10.4}$$

where N is the total size of the population. Thus, S, I and R are all bounded above by N. The mathematical formulation of the epidemic problem is completed given initial

conditions such as

$$S(0) = S_0 > 0, \quad I(0) = I_0 > 0, \quad R(0) = 0. \tag{10.5}$$

A key question in any epidemic situation is, given r, a, S_0 and the initial number of infectives I_0, whether the infection will spread or not, and if it does how it develops with time, and crucially when it will start to decline. From (10.2),

$$\left[\frac{dI}{dt}\right]_{t=0} = I_0(rS_0 - a) \quad \begin{cases} > 0 \\ < 0 \end{cases} \quad \text{if} \quad S_0 \quad \begin{cases} > \rho \\ < \rho \end{cases}, \quad \rho = \frac{a}{r}. \tag{10.6}$$

Since, from (10.1), $dS/dt \leq 0$, $S \leq S_0$ we have, if $S_0 < a/r$,

$$\frac{dI}{dt} = I(rS - a) \leq 0 \quad \text{for all} \quad t \geq 0, \tag{10.7}$$

in which case $I_0 > I(t) \to 0$ as $t \to \infty$ and so the infection dies out; that is, no epidemic can occur. On the other hand if $S_0 > a/r$ then $I(t)$ initially increases and we have an epidemic. The term 'epidemic' means that $I(t) > I_0$ for some $t > 0$; see Figure 10.1. We thus have a *threshold phenomenon*. If $S_0 > S_c = a/r$ there is an epidemic while if $S_0 < S_c$ there is not. The critical parameter $\rho = a/r$ is sometimes called the *relative removal rate* and its reciprocal $\sigma(= r/a)$ the infection's *contact rate*.

We write

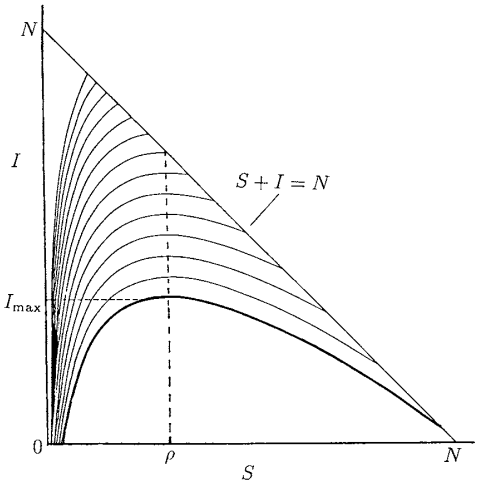


Figure 10.1. Phase trajectories in the susceptibles (S)-infectives (I) phase plane for the *SIR* model epidemic system (10.1)–(10.3). The curves are determined by the initial conditions $I(0) = I_0$ and $S(0) = S_0$. With $R(0) = 0$, all trajectories start on the line $S + I = N$ and remain within the triangle since $0 < S + I < N$ for all time. An epidemic situation formally exists if $I(t) > I_0$ for any time $t > 0$; this always occurs if $S_0 > \rho(= a/r)$ and $I_0 > 0$.

$$R_0 = \frac{rS_0}{a},$$

where R_0 is the basic *reproduction rate* of the infection, that is, the number of secondary infections produced by one primary infection in a wholly susceptible population. Here $1/a$ is the average infectious period. If more than one secondary infection is produced from one primary infection, that is, $R_0 > 1$, clearly an epidemic ensues. The whole question of thresholds in epidemics is obviously important. The definition and derivation or computation of the basic reproduction rate is crucial and can be quite complicated. One such example is if the population is heterogeneous (Diekman et al. 1990).

The basic reproduction rate is a crucial parameter grouping for dealing with an epidemic or simply a disease which is currently under control with vaccination, for example. Although the following arguments are based on R_0 they are quite general. Clearly one way to reduce the reproduction rate is to reduce the number of sucepti-bles, S_0. Vaccination is the common method of doing this and it has been successful in eradicating smallpox. In the U.S.A. it reduced the incidence of measles from 894,134 reported cases in 1941 to 135 in 1997 and for polio from 21,269 in 1952 to the last indigenously acquired case of wild-virus polio reported in 1979 (the Western hempisphere was officially certified polio-free in 1994) with similar reductions in other childhood diseases. Mass vaccination is the cheapest and most effective means of disease control. However, although vaccines are generally extremely safe, no medicine is totally risk-free, however small the risk may be. (There have, however, been a few cases of instant death from diptheria and tetanus vaccines and there is currently much controversy about the vaccine for Anthrax for the military.) As people in the West forget the ravages of polio, measles, diptheria, rubella and so on, many will become less keen to have their children vaccinated because of the risk even if very small. Vaccination not only provides protection for the individual it also provides it for the community at large since it keeps the effective reproduction rate below the level which would allow an epidemic to start. This is the so-called 'herd immunity.' The point is that once the threshold herd immunity level of R_0 has been reached and memory of former diseases fades there is the possibility that people will not have their children vaccinated but have a free ride instead; the unvaccinated have effectively the same immunity. In this situation the best, but unethical, strategy for parents is to urge all other parents to have their children vaccinated but free ride with their own. The important point to keep in mind, however, is that an epidemic can start and rise very quickly if the reproduction rate increases beyond the critical value for an epidemic so in the end free-riding is not without its own risks. (This happened with the Conquistadors in Mexico.)

We can derive some other useful analytical results from this simple model. From (10.1) and (10.2)

$$\frac{dI}{dS} = -\frac{(rS - a)I}{rSI} = -1 + \frac{\rho}{S}, \quad \rho = \frac{a}{r}, \quad (I \neq 0).$$

The singularities all lie on the $I = 0$ axis. Integrating the last equation gives the (I, S) phase plane trajectories as

$$I + S - \rho \ln S = \text{constant} = I_0 + S_0 - \rho \ln S_0, \tag{10.8}$$

where we have used the initial conditions (10.5). The phase trajectories are sketched in Figure 10.1. Note that with (10.5), all initial values S_0 and I_0 satisfy $I_0 + S_0 = N$ since $R(0) = 0$ and so for $t > 0$, $0 \le S + I < N$.

If an epidemic exists we would like to know how severe it will be. From (10.7) the maximum I, $I_{\max}$, occurs at $S = \rho$ where $dI/dt = 0$. From (10.8), with $S = \rho$,

$$\begin{aligned} I_{\max} &= \rho \ln \rho - \rho + I_0 + S_0 - \rho \ln S_0 \\ &= I_0 + (S_0 - \rho) + \rho \ln\left(\frac{\rho}{S_0}\right) \\ &= N - \rho + \rho \ln\left(\frac{\rho}{S_0}\right). \end{aligned} \tag{10.9}$$

For any initial values I_0 and $S_0 > \rho$, the phase trajectory starts with $S > \rho$ and we see that I increases from I_0 and hence an epidemic ensues. It may not necessarily be a severe epidemic as is the case if I_0 is close to $I_{\max}$. It is also clear from Figure 10.1 that if $S_0 < \rho$ then I decreases from I_0 and no epidemic occurs.

Since the axis $I = 0$ is a line of singularities, on all trajectories $I \to 0$ as $t \to \infty$. From (10.1), S decreases since $dS/dt < 0$ for $S \neq 0$, $I \neq 0$. From (10.1) and (10.3),

$$\begin{aligned} &\frac{dS}{dR} = -\frac{S}{\rho} \\ &\Rightarrow \quad S = S_0 e^{-R/\rho} \ge S_0 e^{-N/\rho} > 0 \\ &\Rightarrow \quad 0 < S(\infty) \le N. \end{aligned} \tag{10.10}$$

In fact from Figure 10.1, $0 < S(\infty) < \rho$. Since $I(\infty) = 0$, (10.4) implies that $R(\infty) = N - S(\infty)$. Thus, from (10.10),

$$S(\infty) = S_0 \exp\left[-\frac{R(\infty)}{\rho}\right] = S_0 \exp\left[-\frac{N - S(\infty)}{\rho}\right]$$

and so $S(\infty)$ is the positive root $0 < z < \rho$ of the transcendental equation

$$S_0 \exp\left[-\frac{N - z}{\rho}\right] = z. \tag{10.11}$$

We then get the total number of susceptibles who catch the disease in the course of the epidemic as

$$I_{\text{total}} = I_0 + S_0 - S(\infty), \tag{10.12}$$

where $S(\infty)$ is the positive solution z of (10.11). An important implication of this analysis, namely, that $I(t) \to 0$ and $S(t) \to S(\infty) > 0$, is that the disease dies out from a lack of infectives and *not* from a lack of susceptibles.

The threshold result for an epidemic is directly related to the relative removal rate, ρ: if $S_0 > \rho$ an epidemic ensues whereas it does not if $S_0 < \rho$. For a given disease, the relative removal rate varies with the community and hence determines whether an epidemic may occur in one community and not in another. The number of susceptibles

S_0 also plays a major role, of course. For example, if the density of susceptibles is high and the removal rate, a, of infectives is low (through ignorance, lack of medical care, inadequate isolation and so on) then an epidemic is likely to occur. Expression (10.9) gives the maximum number of infectives while (10.12) gives the total number who get the infection in terms of $\rho(= a/r)$, I_0, S_0 and N.

In most epidemics it is difficult to determine how many new infectives there are each day since only those that are removed, for medical aid or whatever, can be counted. Public Health records generally give the number of infectives per day, week or month. So, to apply the model to actual epidemic situations, in general we need to know the number removed per unit time, namely, dR/dt, as a function of time.

From (10.10), (10.4) and (10.3) we get an equation for R alone; namely,

$$\frac{dR}{dt} = aI = a(N - R - S) = a\left(N - R - S_0 e^{-R/\rho}\right), \quad R(0) = 0, \tag{10.13}$$

which can only be solved analytically in a parametric way: the solution in this form however is not particularly convenient. Of course, if we know a, r, S_0 and N it is a simple matter to compute the solution numerically. Usually we do not know all the parameters and so we have to carry out a best fit procedure assuming, of course, the epidemic is reasonably described by such a model. In practice, however, it is often the case that if the epidemic is not large, R/ρ is small—at least $R/\rho < 1$. Following Kermack and McKendrick (1927) we can then approximate (10.13) by

$$\frac{dR}{dt} = a\left[N - S_0 + \left(\frac{S_0}{\rho} - 1\right)R - \frac{S_0 R^2}{2\rho^2}\right].$$

Factoring the right-hand side quadratic in R, we can integrate this equation to get, after some elementary but tedious algebra, the solution

$$\begin{aligned} R(t) &= \frac{r^2}{S_0}\left[\left(\frac{S_0}{\rho} - 1\right) + \alpha \tanh\left(\frac{\alpha a t}{2} - \phi\right)\right] \\ \alpha &= \left[\left(\frac{S_0}{\rho} - 1\right)^2 + \frac{2S_0(N - S_0)}{\rho^2}\right]^{1/2}, \quad \phi = \frac{\tanh^{-1}\left(\frac{S_0}{\rho} - 1\right)}{\alpha}. \end{aligned} \tag{10.14}$$

The removal rate is then given by

$$\frac{dR}{dt} = \frac{a\alpha^2\rho^2}{2S_0} \operatorname{sech}^2\left(\frac{\alpha a t}{2} - \phi\right), \tag{10.15}$$

which involves only 3 parameters, namely, $a\alpha^2\rho^2/(2S_0)$, αa and ϕ. With epidemics which are not large, it is this function of time which we should fit to the public health records. On the other hand, if the disease is such that we know the actual number of the removed class then it is $R(t)$ in (10.14) we should use. If R/ρ is not small, however, we must use the differential equation (10.13) to determine $R(t)$.

We now apply the model to two very different epidemic situations.

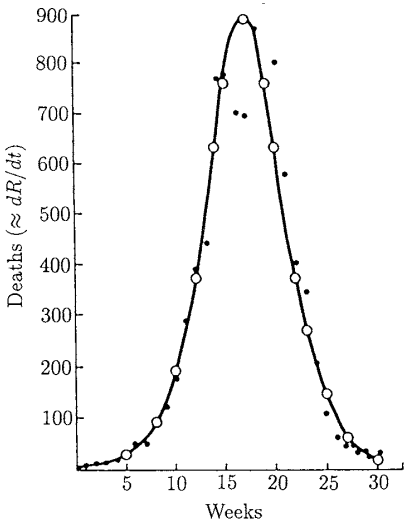


Figure 10.2. Bombay plague epidemic of 1905–1906. Comparison between the data (•) and theory (○) from the (small) epidemic model and where the number of deaths is approximately dR/dt given by (10.16). (After Kermack and McKendrick 1927)

Bombay Plague Epidemic 1905–1906

This plague epidemic lasted for almost a year. Since most of the victims who got the disease died, the number removed per week, that is, dR/dt, was approximately equal to the number of deaths per week. On the basis that the epidemic was not severe (relative to the population size), Kermack and McKendrick (1927) compared the actual data with (10.15) and determined the best fit for the three parameters which resulted in

$$\frac{dR}{dt} = 890\,\mathrm{sech}^2\,(0.2t - 3.4). \tag{10.16}$$

This is illustrated in Figure 10.2 together with the actual epidemic data.

Influenza Epidemic in an English Boarding School 1978

In 1978 in the British medical journal, *The Lancet*, there was a report with detailed statistics of a flu epidemic in a boys' boarding school with a total of 763 boys. Of these 512 were confined to bed during the epidemic, which lasted from 22nd January to 4th February 1978. It seems that one infected boy initiated the epidemic. This situation has many of the requirements assumed in the above model derivation. Here, however, the epidemic was severe and the full system has to be used. Also, when a boy was infected he was put to bed and so we have $I(t)$ directly from the data. Since in this case we have no analytical solution for comparison with the data, a best fit numerical technique was used directly on the equations (10.1)–(10.3) for comparison of the data. Figure 10.3 illustrates the resulting time evolution for the infectives, $I(t)$, together with the epidemic statistics. The R-equation (10.3) is uncoupled; the solution for $R(t)$ is simply proportional to the area under the $I(t)$ curve.

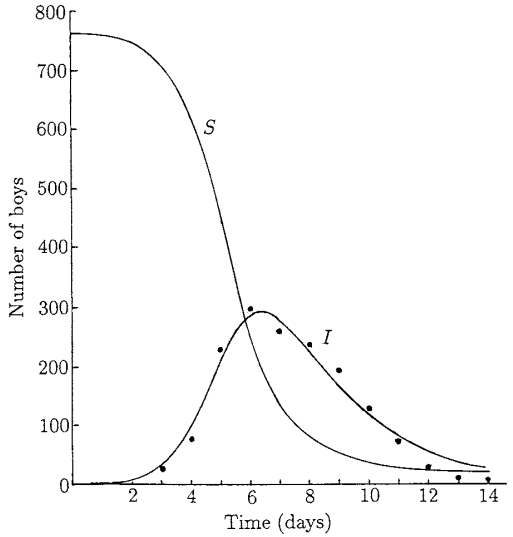


Figure 10.3. Influenza epidemic data (•) for a boys' boarding school as reported in the British medical journal, *The Lancet*, 4th March 1978. The continuous curves for the infectives (I) and susceptibles (S) were obtained from a best fit numerical solution of the *SIR* system (10.1)–(10.3): parameter values $N = 763$, $S_0 = 762$, $I = 1$, $\rho = 202$, $r = 2.18 \times 10^{-3}$/day. The conditions for an epidemic to occur, namely, $S_0 > \rho$, are clearly satisfied and the epidemic is severe since R/ρ is not small.

Plague in Eyam, England 1665–1666

There was an outbreak of plague in the village of Eyam in England from 1665 to 1666. In this remarkable altruistic incident, the village sealed itself off when plague was discovered, so as to prevent it spreading to the neighbouring villages, and it was successful. By the end of the epidemic only 83 of the original population of 350 survived. Raggett (1982) applied the *SIR* model (10.1)–(10.3) to this outbreak. Here, $S(\infty) = 83$ out of an initial $S_0 = 350$. This is another example, like the school flu epidemic, where the epidemic was severe. Raggett (1982) shows how to determine the parameters from the available data and knowledge of the etiology of the disease. He reiterates the view that although the initial form was probably bubonic plague, the pneumonic form most likely became prevalent; the latter form can be transmitted from the cough of a victim (see Chapter 13, Volume II for a brief description of the plague and its history). The comparison between the solutions from the deterministic model and the Eyam data is very good. The comparison is much better than that obtained from the corresponding stochastic model, which Raggett (1982) also considered. We discuss a model for the spatial spread of plague in Chapter 13, Volume II.

If a disease is *not* of short duration then (10.1), the equation for the susceptibles, should include birth and death terms. Mortality due to natural causes should also be included in equation (10.2) for the infectives and in (10.3) for the removed class. The

resulting models can be analysed in a similar way to that used here and in Chapter 3 on interacting populations: they are still systems of ordinary differential equations. It is not surprising, therefore, that oscillatory behaviour in disease epidemics is common; these are often referred to as epidemic waves. Here they are *temporal* waves. *Spatial* epidemic waves appear as an epidemic spreads geographically. The latter are also common and we consider them in detail in Chapter 13, Volume II.

Many diseases have a latent or incubation period when a susceptible has become infected but is not yet infectious. Measles, for example, has an 8- to 13-day latent period. The incubation time for AIDS, on the other hand, is anything from a few months to years after the patient has been shown to have antibodies to the human immunodeficiency virus (HIV). We can, for example, incorporate this as a delay effect, or by introducing a new class, $E(t)$ say, in which the susceptible remains for a given length of time before moving into the infective class. Such models give rise to integral equation formulations and they can exhibit oscillatory behaviour as might be expected from the inclusion of delays. Some of these are described by Hoppensteadt (1975, see also 1982). Nonlinear oscillations in such models have been studied by Hethcote et al. (1981); see also Hethcote (1994). Alternative approaches recently used in modelling AIDS are discussed below. Finally age, a, is often a crucial factor in disease susceptibility and infectiousness. The models then become partial differential equations with independent variables (t, a); we consider one such model later in this chapter.

There are many modifications and extensions which can and often must be incorporated in epidemic models; these depend critically on the disease and location. In the following sections we discuss a few more general models to illustrate different but important points. The books and references already cited describe numerous models and go into them in considerable detail.

10.3 Modelling Venereal Diseases

The incidence of sexually transmitted diseases (STDs), such as gonorrhea (*Neisseria gonorrhoeae*), chlamydia, syphilis and, of course, AIDS, is a major health problem in both developed and developing countries. In the U.S.A., for example, as reported by the Centers for Disease Control (www.cdc.gov), in 1996 there were over 300,000 cases of gonorrhea reported and over 11,000 cases of syphilis and nearly 500,000 cases of chlamydia. Whereas the rate has been decreasing for gonorrhea and syphilis it is growing for chlamydia. We give some of the numbers for HIV incidence in the AIDS sections below.

STDs have certain characteristics which are different from other infections, such as measles or rubella (German measles). One difference is that they are mainly restricted to the sexually active community so the assumption of uniform mixing in the whole population is not really justified. Another is that often the carrier is asymptomatic (that is, the carrier shows no overt symptoms) until quite late on in the development of the infection. A third crucial difference is that STDs induce little or no acquired immunity following an infection. Equally important in virus infections is the lack of present knowledge of some of the parameters which characterise the transmission dynamics.

Although gonorrhea, syphilis and AIDS are well known, with the latter growing alarmingly, one of the STDs which has far outstripped gonorrhea is the less well-known *Chlamydia trachomatis*, which in 1996 struck more than gonorrhea and syphilis put together and is on the increase. It can produce sterility in women without their ever showing any overt symptoms. Diagnostic techniques are now sufficiently refined to make diagnosis more accurate and less expensive and could account in part for the increase in reported cases.[5] The asymptomatic character of this disease among women is serious. Untreated, it causes pelvic inflammatory disorders (PID) which are often accompanied by chronic pain, fever and sterility. With pregnancy, PID, among other complications, can often cause premature delivery and ectopic pregnancies (that is, the fertilised egg is implanted outside the womb) which are life threatening. Untreated gonorrhea, for example, can also cause blindness, PID, heart failure and ultimately death. STDs are a major cause of sterility in women. The consequences of untreated STDs in general are very unpleasant. The vertical transmission of STDs from mother to newborn children is another of the threats and tragedies of many STDs. Another problem is the appearance of new strains: in connection with AIDS, HIV-1 is the common virus but a relatively new one, HIV-2 has now been found. With gonorrhea the relatively new strain, *Neisseria gonorrhoeae*, which was discovered in the 1970's proved resistant to penicillin.

In this section we present a simple classical epidemic model which incorporates some of the basic elements in the heterosexual spread of venereal diseases. We have in mind such diseases as gonorrhea; AIDS we discuss separately later in the chapter. The monograph by Hethcote and Yorke (1984) is still a good survey of models used for the spread and control of gonorrhea. They show how models and data can be used to advantage; the conclusions they arrived at are specifically aimed at public health workers.

For the model here we assume there is uniformly promiscuous behaviour in the population we are considering. As a simplification we consider only heterosexual encounters. The population consists of two interacting classes, males and females, and infection is passed from a member of one class to the other. It is a criss-cross type of disease in which each class is the disease host for the other. In all of the models we have assumed homogeneous mixing between certain population subgroups. Dietz and Hadeler (1988), for example, considered epidemic models for STDs in which there is heterogeneous mixing. More complex models can include the pairing of two susceptibles, which confers temporary immunity, several subgroups and so on. We discuss a multi-group example later in this section.

Criss-cross infection is similar in many ways to what goes on in malaria[6] and bilharzia, for example, where two criss-cross infections occur. In bilharzia it is between

[5] One U.S. Public Health official when asked some years ago about the high incidence of chlamydia and what doctors were doing about it, is said to have remarked 'Doing about it? Most of them can't even spell it.'

[6] A very interesting, exciting and potentially important new and cheap treatment for malaria, which kills around 2.7 million people a year, has been discovered by Dr. Henry Lai, and his colleagues in Bioengineering in the University of Washington. They found that the malarial parasite *Plasmodium falciparum* (the deadliest of the four malarial parasites) can lose vigour and die when subjected to small oscillating magnetic fields (of the order of the earth's field). They suggest it may be due to the movement caused in the very small iron particles inside the parasite which damages the parasites by disrupting their feeding process which involves the haemoglobin in the red blood cells of the host. They found that exposed samples of the parasite ended up with 33–70% fewer parasites as compared to unexposed samples.

humans and a particular type of snail. Bilharzia, or schistosomiasis, has been endemic in Africa for a very long time. (See footnote 1 in Chapter 3.) We discuss in detail later in this chapter a more complex practical example of a criss-cross infection between badgers and cattle, namely, bovine tuberculosis.

Since the incubation period for venereal diseases is usually quite short—in gonorrhea, for example, it is three to seven days—when compared to the infectious period, we use an extension of the simple epidemic model in Section 10.2. We divide the promiscuous male population into susceptibles, S, infectives, I, and a removed class, R; the similar female groups we denote by S^*, I^* and R^*. If we do not include any transition from the removed class to the susceptible group, the infection dynamics is schematically

$$\begin{array}{ccccc} S & \longrightarrow & I & \longrightarrow & R \\ & \searrow\!\swarrow & & & \\ S^* & \longrightarrow & I^* & \longrightarrow & R^*. \end{array} \tag{10.17}$$

Here I^* infects S and I infects S^*.

As we noted above, the contraction of gonorrhea does not confer immunity and so an individual removed for treatment becomes susceptible again after recovery. In this case a better dynamics flow diagram for gonorrhea is

$$\begin{array}{ccccc} S & \longrightarrow & I & \longrightarrow & R \\ & \searrow\!\swarrow & & & \\ S^* & \longrightarrow & I^* & \longrightarrow & R^*. \end{array} \tag{10.18}$$

An even simpler version involving only susceptibles and infectives is

$$\begin{array}{ccc} S & \longrightarrow & I \\ & \searrow\!\swarrow & \\ S^* & \longrightarrow & I^*. \end{array} \tag{10.19}$$

which, by way of illustration, we now analyse. It is a criss-cross SI model.

We take the total number of males and females to be constant and equal to N and N^* respectively. Then, for (10.19),

$$S(t) + I(t) = N, \quad S^*(t) + I^*(t) = N^*. \tag{10.20}$$

As before we now take the rate of decrease of male susceptibles to be proportional to the male susceptibles times the infectious female population with a similar form for the female rate. We assume that once infectives have recovered they rejoin the susceptible class. A model for (10.19) is then (10.20) together with

$$\begin{aligned} \frac{dS}{dt} &= -rSI^* + aI, \quad \frac{dS^*}{dt} = -r^*S^*I + a^*I^* \\ \frac{dI}{dt} &= rSI^* - aI, \quad \frac{dI^*}{dt} = r^*S^*I - a^*I^*, \end{aligned} \tag{10.21}$$

where r, a, r^* and a^* are positive parameters. We are interested in the progress of the

disease given initial conditions

$$S(0) = S_0, \quad I(0) = I_0, \quad S^*(0) = S_0^*, \quad I^*(0) = I_0^*. \tag{10.22}$$

Although (10.21) is a 4th-order system, with (10.20) it reduces to a 2nd-order system in either S and S^* or I and I^*. In the latter case we get

$$\frac{dI}{dt} = rI^*(N - I) - aI, \quad \frac{dI^*}{dt} = r^*I(N^* - I^*) - a^*I^*, \tag{10.23}$$

which can be analysed in the (I, I^*) phase plane in the standard way (cf. Chapter 3). The equilibrium points, that is, the steady states of (10.23), are $I = 0 = I^*$ and

$$I_s = \frac{NN^* - \rho\rho^*}{\rho + N^*}, \quad I_s^* = \frac{NN^* - \rho\rho^*}{\rho^* + N}, \quad \rho = \frac{a}{r}, \quad \rho^* = \frac{a^*}{r^*}. \tag{10.24}$$

Thus nonzero positive steady state levels of the infective populations exist only if $NN^*/\rho\rho^* > 1$: this is the *threshold condition* somewhat analogous to that found in Section 10.2.

With the experience gained from Chapter 3, we now expect that, if the positive steady state exists then the zero steady state is unstable. This is indeed the case. The eigenvalues λ for the linearisation of (10.23) about $I = 0 = I^*$ are given by

$$\begin{vmatrix} -a - \lambda & rN \\ r^*N^* & -a^* - \lambda \end{vmatrix} = 0$$

$$\Rightarrow \quad 2\lambda = -(a + a^*) \pm \left[(a + a^*)^2 + 4aa^* \left(\frac{NN^*}{\rho\rho^*} - 1 \right) \right]^{1/2}.$$

So, if the threshold condition $NN^*/\rho\rho^* > 1$ holds, $\lambda_1 < 0 < \lambda_2$ and the origin is a saddle point in the (I, I^*) phase plane. If the threshold condition is not satisfied, that is, $(0 <) NN^*/\rho\rho^* < 1$, then the origin is stable since both $\lambda < 0$. In this case positive I_s and I_s^* do not exist.

If I_s and I_s^* exist, meaning in the context here that they are positive, then linearising (10.23) about it, the eigenvalues λ satisfy

$$\begin{vmatrix} -a - rI_s^* - \lambda & rN - rI \\ r^*N^* - r^*I^* & -a^* - r^*I_s - \lambda \end{vmatrix} = 0;$$

that is,

$$\lambda^2 + \lambda[a + a^* + rI_s^* + r^*I_s] + [a^*rI_s^* + ar^*I_s + rr^*(I^*N + IN^*) + aa^* - rr^*NN^*] = 0,$$

the solutions of which have $\mathrm{Re}\,\lambda < 0$ and so the positive steady state (I_s, I_s^*) in (10.24) is stable.

The threshold condition for a nonzero steady state infected population is $NN^*/\rho\rho^* = (rN/a)(r^*N^*/a^*) > 1$. We can interpret each term as follows. If every male is susceptible then rN/a is the average number of males contacted by a female infective during her infectious period; a reciprocal interpretation holds for r^*N^*/a^*. These quan-

tities, rN/a and r^*N^*/a^*, are the maximal male and female *contact rates* respectively.

Although parameter values for contacts during an infectious stage are notoriously unreliable from individual questionnaires, what is abundantly clear from the statistics since 1950 is that an epidemic has occurred in a large number of countries and so $NN^*/\rho\rho^* > 1$. From data given by a male and a female infective, in the U.S.A. in 1973, regarding the number of contacts during a period of their infectious state, figures of maximal contact rates of $N/\rho \approx 0.98$ and $N^*/\rho^* \approx 1.15$ were calculated for the male and female respectively which give $NN^*/\rho\rho^* \approx 1.127$.

10.4 Multi-Group Model for Gonorrhea and Its Control

Although the SI model in the last section is a particularly simple one, it is not too unrealistic. In the case of gonorrheal infections, however, it neglects many relevant factors. For example, as already mentioned a large proportion of females, although infected and infectious, show no obvious symptoms; that is, they form an asymptomatic group. There are, in fact, various population subgroups. For example, we could reasonably have susceptible, symptomatic, treated infective and untreated infective groups. Lajmanovich and Yorke (1976) proposed and analysed an 8-group model for gonococcal infections consisting of sexually (i) very active and (ii) active females (males) who are asymptomatic when infectious and (iii) very active and (iv) active females (males) who are symptomatic when infectious.

If the total populations of active male and female are N and N^*, assumed constant, we can normalise the various group populations as fractions of N and N^*. Denote the groups of women with indices 1, 3, 5, 7 and the men with indices 2, 4, 6, 8. Then if N_i, $i = 1, 2, \ldots, 8$ denote the *normalised* populations

$$N_1 + N_3 + N_5 + N_7 = 1, \quad N_2 + N_4 + N_6 + N_8 = 1. \tag{10.25}$$

Since neither immunity nor resistance is acquired in gonococcal infections we consider only two classes, susceptibles and infectives. If $I_i(t)$, $i = 1, 2, \ldots, 8$ denote the fractions infectious at any time t, the fractional numbers of susceptibles at that time are then $1 - I_i(t)$, $i = 1, 2, \ldots, 8$.

We again assume homogeneous mixing. For each group let D_i be the mean length of time (in months) of the infection in group i. Then, there is a $1/D_i$ chance of an infective recovering each month. This implies that the removal rate per month is I_i/D_i.

Let L_{ij} be the number of effective contacts per month of an infective in group j with an individual in group i. Since the model here considers only heterosexual (as opposed to homosexual) contacts we have

$$L_{ij} = 0 \quad \text{if } i + j \text{ even.}$$

The matrix $[L_{ij}]$ is called the *contact matrix*. Although there are seasonable variations in the L_{ij} we take them to be constant here. Then the average number of susceptibles infected per unit time (month) in group i by group j is $L_{ij}(1 - I_i)$. Thus the model

differential equation system is

$$\underbrace{\frac{d(N_i I_i)}{dt}}_{\substack{\text{rate of new}\\ \text{infectives}}} = \underbrace{\sum_{j=1}^{8} L_{ij}(1 - I_i)N_j I_j}_{\substack{\text{rate of new}\\ \text{infectives (incidence)}}} - \underbrace{\frac{N_i I_i}{D_i}}_{\substack{\text{recovery rate}\\ \text{of infectives}}} \tag{10.26}$$

with given initial conditions $I_i(0) = I_{i0}$.

By considering the linearisation about the nonzero steady state the effect of varying the parameters can be assessed and hence the effects of various control strategies. This model is analysed in detail by Lajmanovich and Yorke (1976).

Major aims in control include of course the reduction in incidence and an increase in detection, each of which affects the long term progress of the spread of the disease. So, screening, detection and treatment of infectives is the major first step in control. The paper by Hethcote et al. (1982) compares various control methods for gonorrhea; it also has references to other models which have been proposed.

As an example, suppose C is a parameter proportional to the number of women screened and CR_i is the rate at which infected women are detected in group i. Let EP_i be the general supplementary detection rate where E is a measure of the effort put in and P_i is the population of a group i: E depends on the control strategy. Then, in place of (10.26) we have the control model

$$\frac{d(N_i I_i)}{dt} = \sum_{j=1}^{8} L_{ij}(1 - I_i)N_j I_j - \frac{N_i I_i}{D_i} - CR_i - EP_i. \tag{10.27}$$

Different control methods imply different R_i and P_i.

Suppose there is general screening of women (the major control procedure in the U.S.A.). On the basis that the number of infected women detected is directly proportional to the number infected and the supplementary programme is general screening of the women population, we have

$$P_i = R_i = I_i N_i, \quad i = 1, 3, 5, 7; \qquad P_i = R_i = 0, \quad i = 2, 4, 6, 8.$$

If the programme is for men, the odd and even number range is interchanged.

These and other control procedures are discussed in the paper by Hethcote et al. (1982); see also Hethcote and Yorke (1984). They also discuss the important problem of parameter estimation and finally carry out a comparison of various control strategies. The cost and social range of screening are not negligible factors in the practical implementation of such programmes. The political and sociological considerations can also be rather sensitive.

It should be emphasised again, that venereal disease models, which are to be used in control programmes, must have a realistic validation, which can only come from a comparison of their solutions and predictions with actual data. This should, of course, apply to all disease control models.

10.5 AIDS: Modelling the Transmission Dynamics of the Human Immunodeficiency Virus (HIV)

Some Background, Myths, Statistics and Polemics

One aspect of the AIDS (autoimmune deficiency syndrome) epidemic is the myth of denial, a not uncommon phenomenon with certain diseases where, for example, there is a perceived social stigma or a strong economic element; the brief highly pertinent article by Weiss (1996) discusses some recent examples of this regarding AIDS and suggests some of the modern reasons for it. He also quotes some astonishing statements such as one by a British Government Home Office minister who said that HIV could not possibly be transmitted in prisons because drugs and sex were not permitted. Another is by a medical journalist writing in the respected British newspaper *The Independent* who said 'The government has wasted £150 million of our taxpayers' money anathematizing the innocent pleasures of casual heterosexual intercourse.' The belief that HIV does not cause AIDS is subscribed to in the book by the scientist Duesberg (1996) of the University of California at Berkeley. Among other things, he says that AIDS is not only not contagious but that it is caused by drugs taken for the express purpose of blocking HIV. Lauritsen (1993) attributes AIDS to the medical-industrial complex whose aim is profiteering and genocide: he claims that the medicine AZT, used in the treatment of HIV, actually causes AIDS! The problem of how so much AZT could have got into sub-Saharan Africa, the major problem area, is not discussed. A recent book on the origin of the disease by Hooper (1999) makes the controversial case for AIDS being caused by an experimental oral vaccine for polio which was given to around a million people in Rwanda, Burundi and the Congo from 1957 to 1960. This area is the epicentre of the African epidemic. He argues that the vaccine might have been made with chimpanzee tissue which was contaminated with an ancestor of the virus. This has subsequently been denied by doctors involved in the programme. His argument is carefully researched but the evidence is still circumstantial.

The major horror of the AIDS epidemic is in Africa where around 70% of the total AIDS deaths in the world have occurred and, as recently stated (July, 1999) by Dr. Peter Piot, head of the United Nations AIDS (UNAIDS) programmes, half of all newborn babies in Africa are HIV positive. The regular early ludicrous denials in the 1980's of its existence by some African leaders ('There is no AIDS in my country.') and others in positions of responsibility, however, began to change in the mid-1990's. In sub-Saharan Africa up to 1998 HIV has infected 34 million people and killed 11.5 million since 1981 and approximately 1.8 million in 1998 alone. In 1999 an estimated 5.6 million adults and children became infected with the virus with a worldwide total estimated at 50 million infected since 1981 of which 16 million have died; around 2.6 million died in 1999 alone, the highest number of any year. In Zimbabwe, Malawi and Botswana perhaps the countries worst afflicted with HIV infection, it is a human and economic catastrophe: in Zimbabwe at least 20% are HIV positive while in Botswana it is more than 35%. Its extremely rapid growth in South Africa (where as many as 20% of the population is HIV positive) is alarming.[7] Life expectancy which increased dramatically

[7] In the case of South Africa it was certainly not helped by the Health Minister saying (in the National Assembly in the week of 16th November 1999) that AZT may be too dangerous to use. (Perhaps it has to be accepted that a drug to treat a fatal disease is more toxic than drinking herbal tea.) If that was not enough, in April 2000 President Thabo Mbeki astonishingly and depressingly said that he wished to discuss HIV and AIDS specifically with those scientists who say there is no connection; he simply refused to accept the connection between HIV and AIDS.

in South Africa in the 1990's is now plummeting. Malthus (see Chapter 1) may well be right about his disease prognosis and population control. A broad picture of the world scene and several important aspects of the disease is given in the special report on AIDS in *Scientific American* (1998) by various authors dealing with such issues as prevention, ethical dilemmas, children with HIV, drug resistance, vaccines and others.

AIDS, unlike its early image as a homosexual disease, is now very much a heterosexual disease. In a UNAIDS report for World AIDS Day, 1st December 1999, it says that of the 22.3 million adults in sub-Saharan Africa with HIV, 55% of them are women. In South and Southeast Asia it is estimated that 30% are women and in North America 20%. In Africa it is mainly transmitted by heterosex whereas in the U.S. it is mainly homosexual transmission.

The most important aspect of defense against infectious diseases is unquestionably surveillance which characterises the pattern of each disease. Although there are social problems associated with gathering data on the number of people who have the HIV, it is unlikely that the epidemic will be contained if this information is not made available.[8] Widespread surveillance of human tuberculosis (*Mycobacterium tuberculosis*) in the 1950's essentially eradicated the disease in many developed countries. However, new human strains are now appearing including in the developed world: it is already a significant problem in New York. Tuberculosis is still a major killer in the world; between 1990 and 1999 approximately 30 million people have died (Cosivi et al. 1998) from the disease.

The lack of knowledge about HIV creates enormous difficulties in designing effective control programmes, not to mention those for health care facilities. Education programmes as to how it can spread are the minimum requirement. Those that have been pursued have had some success but even their continuing use and new ones have often been blocked by the religious establishments (and not just the loony right). Without a knowledge of the reservoir of the disease, it is extremely difficult to evaluate effective prevention and control strategies. According to a depressing UNAIDS Report (Global HIV/AIDS Epidemic December 1997) there are an estimated 16,000 new cases a day and that around 27 million people are HIV positive but do not know it. AIDS is just one disease where surveillance has been disastrously inadequate. Another in which the lack of surveillance is going to cause serious problems in the very near future is the misuse of antibiotics which is giving rise to resistant strains of bacteria.

[8] In a class on modelling epidemics I once had the students construct a model for the spread of an hypothetical disease which was based, in fact, on the spread of HIV but which I took pains to hide. After they produced a reasonable first model I then asked the class to discuss strategies for its control. Without exception everybody in the class agreed that the only way was to have universal surveillance with everyone being tested for the virus. I then took their model and at each step I related it directly to the present AIDS epidemic. The reaction in the class was what I had expected (but not the intensity of feeling): the class immediately launched into a very heated discussion among the civil libertarians, the politicians, the humanists, the religious group taking the moral high ground and the pragmatists. (I kept out of the discussion and was only the moderator.) In the end the students were unified only in believing that I had deliberately conned them into saying that clearly everybody should be tested for HIV positivity—they were absolutely right, of course; it was my intention.

Other than the new strains of HIV there is an increasing number of new or newly identified diseases or old agents in new locations, such as Vibrio cholera 0139 (new agent 1992) which is a variant of cholera, Hepatitis E virus (new 1990), Hemorrhagic fever (1991) in Venezuela, Hantavirus (1993) in Southwest U.S.A., Anthrax (1993) in the Carribean, Lasa fever (1992) in West Africa and numerous others. The book by Garrett (1994) specifically deals with newly emerging diseases which he refers to as 'the coming plague.' In spite of the appearance of other new diseases, recurrence of old ones and the other major killing diseases, AIDS is arguably the major epidemic of the 20th century and perhaps of all time. Its progression has exceeded the gloomy view expressed in the first edition of this book in 1989 and now in the year 2000 can only give pessimists cause for optimism.

Human Immunodeficiency Virus (HIV)—Background

The human immunodeficiency virus, HIV, leads to acquired immune deficiency syndrome, AIDS. HIV is a retrovirus and like most of the viruses in this family of viruses, the *Retroviridae*, only replicates in dividing cells. HIV has some unfortunate unique properties even within this retrovirus family such as using the mRNA processing of the cell it invades to synthesise its own viral RNA. Although studies (Ho et al. 1995) have shown the dynamics of viral replication is very high *in vivo* the immune system can counteract this replication from 5 to 10 years or more depending on the initial infection. Cases of haemophiliacs who have been given contaminated blood have succumbed in a matter of months.

Infection by the virus HIV-1, the most common variety, has many highly complex characteristics, most of which are still not understood. The fact that the disease progression can last more than 10 years from the first day of infection is just one of them. Another is that while most viral infections can be eliminated by an immune response, HIV is only briefly controlled by it. HIV primarily infects a class of white blood cells or lymphocytes, called CD4 T-cells, but also infects other cells such as dendritic cells. The virus has a high affinity for a receptor present on the cell surface of each of these cells which guides the virus to their location *in vivo*. When the CD4 T-cell count, normally around 1000/μL, decreases to 200/μL or below, a patient is characterized as having AIDS. There are very specific clearly defined clinical categories (Morb Mort Week Report 42 (No. RR-17), Table 308-1 and Table 308-2, December 18, 1992) which are used to diagnose the AIDS; the CD4 T-cell count is not the only factor. The categories are regularly updated. These are used by the Centers for Disease Control for surveillance purposes. For example, if a patient with the virus has a CD4 T-cell count greater than 500/μL but has, or has had one of a variety of diseases then a formal diagnosis is made and registered. The reason for the fall in the T-cell count is unknown. T-cells are normally replenished very quickly in the body, so the infection may affect the source of new T-cells or the life span of preexisting ones. Although HIV can kill cells that it infects, only a small fraction of CD4 T-cells are infected at any given time. Because of the central role of CD4 T-cells in immune regulation, their depletion has widespread deleterious effects on the functioning of the immune system as a whole and this is what leads to AIDS.

Since the mid-1980's, numerous models, deterministic and stochastic, have been developed to describe the immune system and its interaction with HIV. It is a highly

controversial area. Stochastic models aim to account for the early events in the disease when there are few infected cells and a small number of viruses. Nowak et al. (1996; see earlier references there) look at the effects of variability among viral strains but this and earlier work has been commented on critically by Stilianakis et al. (1994) and Wein et al. (1998).

Most models have been deterministic such as those by McLean and Nowak (1992), Perelson et al. (1993), Essunger and Perelson (1994), Frost and McLean (1994), Stilianakis et al. (1994), Kirschner and Webb (1997) and Wein et al. (1998). Deterministic models, which attempt to reflect the dynamic changes in mean cell numbers, are more applicable to later stages of the process when the population is large. These models typically consider the dynamics of the CD4 cells, latently infected cells and virus populations as well as the effects of drug therapy.

Because of the ethics, among other things, of doing experiments on humans, fundamental information has been lacking about the dynamics of HIV infection. For example, since the disease takes an average of 10 years to develop it was widely thought that the components of the disease process would also be slow. A combination of mathematical modelling and experiment has shown this is not the case by showing that there are a number of different timescales in HIV infection, from minutes to hours and days to months. The current understanding of the rapidity of HIV infection has totally changed the manner in which HIV is treated in patients and has had a major impact in extending peoples' lives; see the review paper by Perelson and Nelson (1999).

Figure 10.4 shows a typical course of HIV infection. Immediately after infection the amount of virus detected in the blood, V, increases rapidly. After a few weeks to months the symptoms disappear and the virus concentration falls to a lower level. An immune response to the virus occurs and antibodies against the virus can be detected in the blood. A test, now highly refined, to detect these antibodies determines if a person has been exposed to HIV. If the antibodies are detected, a person is said to be HIV-positive.

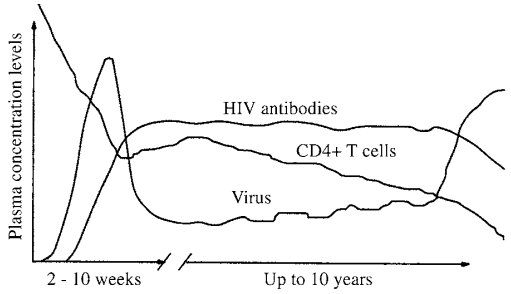


Figure 10.4. Schematic time course of a typical HIV infection in an infected adult. The viral concentration, the level of antibodies and the CD4 T-cells are sketched as a function of time. The early peak corresponds to the primary infection which leads to a period of latency. Note the typical gradual decline in the level of CD4 T-cells over the years. Eventually the symptoms of full-blown AIDS start to appear. (From Perelson and Nelson 1999 and reproduced with permission)

The level the virus falls to after the initial infection has been called the *set-point*. The viral concentration then seems to remain at a quasi-steady state level during which the concentration of CD4 T-cells measured in blood slowly declines. This period in which the virus concentration stays relatively constant but in which the T-cell count slowly falls is typically a period in which the infected person has no disease symptoms.

A key question then is what is going on during this asymptomatic period. Many believed that the virus was simply quiescent or latent during this period, as seen in other viral diseases, such as herpes. One method of determining whether or not the virus is active is to perturb the host–virus system during the asymptomatic period. In the mid-1990's work started on new antiretroviral drugs, the protease inhibitors. With their introduction it became possible to perturb the host–virus system during the asymptomatic period. In 1994, David Ho (Aaron Diamond AIDS Research Center) ran an experiment which examined the response of 20 patients infected with HIV to the protease inhibitor, ritonavir. The results were dramatic. Figure 10.5 shows the amount of virus measured in blood plasma fell rapidly once the drug was given. Alan Perelson (Los Alamos National Laboratory) and his colleagues then developed a model system which was applied to the patient data and estimations of crucial parameters were obtained. The work is reported in Ho et al. (1995).

Before discussing a model which includes protease inhibitor treatment, we first describe an early model by Anderson et al. (1986) for pedagogical reasons since it is a common way of constructing an epidemic model using a flow chart. It is much less specific and less directly related to current HIV thinking than the one we discuss below in relation to the data and qualitative behaviour of the virus as shown in Figures 10.4 and 10.5. A nice review of the state of AIDS modelling at the time is given by Isham (1988).

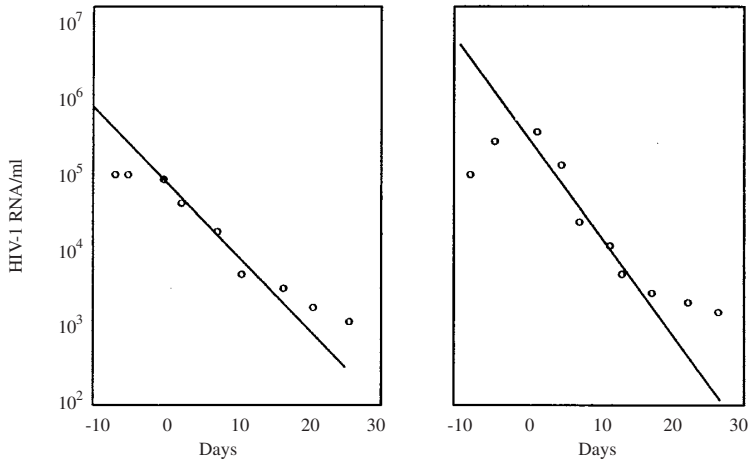


Figure 10.5. After treatment started at $t = 0$ with a protease inhibitor the plasma viral load declined rapidly. The data are from 2 of the 20 patients studied in Ho et al. (1995): all 20 patients exhibited similar rapid declines. (Reproduced with permission)

Basic Epidemic Model for HIV Infection in a Homosexual Population

Here we are interested in the development of an AIDS epidemic in a homosexual population. Let us assume there is a constant immigration rate B of susceptible males into a population of size $N(t)$. Let $X(t)$, $Y(t)$, $A(t)$ and $Z(t)$ denote respectively the number of susceptibles, infectious males, AIDS patients and the number of HIV-positive or seropositive men who are noninfectious. We assume susceptibles die naturally at a rate μ; if there were no AIDS, the steady state population would then be $N^* = B/\mu$. We assume AIDS patients die at a rate d: $1/d$ is of the order of months to years, more often the latter. Figure 10.6 is a flow diagram of the disease on which we base our model.

As in previous models we consider uniform mixing. A reasonable first model system, based on the flow diagram in Figure 10.6, is then

$$\frac{dX}{dt} = B - \mu X - \lambda c X, \quad \lambda = \frac{\beta Y}{N}, \tag{10.28}$$

$$\frac{dY}{dt} = \lambda c X - (v + \mu) Y, \tag{10.29}$$

$$\frac{dA}{dt} = pvY - (d + \mu) A, \tag{10.30}$$

$$\frac{dZ}{dt} = (1 - p) v Y - \mu Z, \tag{10.31}$$

$$N(t) = X(t) + Y(t) + Z(t) + A(t). \tag{10.32}$$

Here B is the recruitment rate of susceptibles, μ is the natural (non-AIDS-related) death rate, λ is the probability of acquiring infection from a randomly chosen partner ($\lambda =$

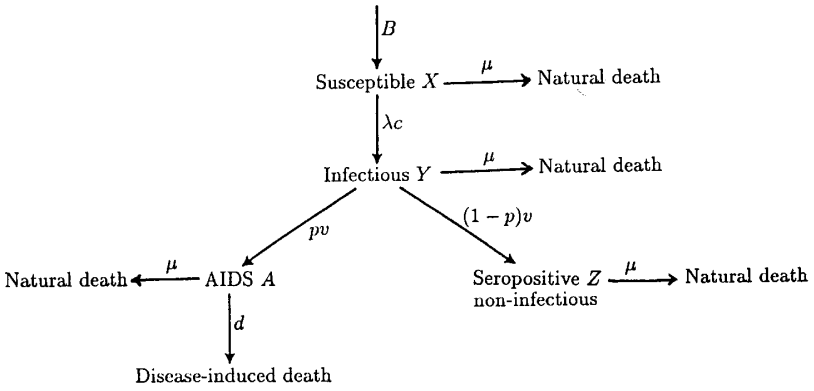


Figure 10.6. The flow diagram of the disease as modelled by the system (10.28)–(10.32). B represents the recruitment of susceptibles into the homosexual community. The rate of transferral from the susceptible to the infectious class is λc, where λ is the probability of acquiring infection from a randomly chosen partner and c is the number of sexual partners. A proportion of the infectious class is assumed to become noninfectious with the rest developing AIDS. Natural (non-AIDS induced) death is also included in the model. Parameters are defined in the text.

$\beta Y/N$, where β is the transmission probability), c is the number of sexual partners, d is the AIDS-related death rate, p is the proportion of HIV-positives who are infectious and v is the rate of conversion from infection to AIDS here taken to be constant. $1/v$, equal to D say, is then the average incubation time of the disease. (Actually λ here is more appropriately $\beta Y/(X+Y+Z)$ but A is considered small in comparison with N.) Note that in this model the total population $N(t)$ is not constant, as was the case in the epidemic models in Section 10.2. If we add equations (10.28)–(10.32) we get

$$\frac{dN}{dt} = B - \mu N - dA. \tag{10.33}$$

An epidemic ensues if the basic reproductive rate $R_0 > 1$: that is, the number of secondary infections which arise from a primary infection is greater than 1. In (10.32) if, at $t = 0$, an infected individual is introduced into an otherwise infection-free population of susceptibles, we have initially $X \approx N$ and so near $t = 0$,

$$\frac{dY}{dt} \approx (\beta c - v - \mu)Y \approx v(R_0 - 1)Y \tag{10.34}$$

since the average incubation time, $1/v$, from infection to development of the disease, is very much shorter than the average life expectancy, $1/\mu$, of a susceptible; that is, $v \gg \mu$. Thus the approximate threshold condition for an epidemic to start is, from the last equation,

$$R_0 \approx \frac{\beta c}{v} > 1. \tag{10.35}$$

Here the basic reproductive rate R_0 is given in terms of the number of sexual partners c, the transmission probability β and the average incubation time of the disease $1/v$.

When an epidemic starts, the system (10.28)–(10.32) evolves to a steady state given by

$$\begin{aligned} X^* &= \frac{(v+\mu)N^*}{c\beta}, \quad Y^* = \frac{(d+\mu)(B-\mu N^*)}{pvd} \\ Z^* &= \frac{(1-p)(d+\mu)(B-\mu N^*)}{pd\mu}, \quad A^* = \frac{B-\mu N^*}{d}, \\ N^* &= \frac{B\beta[\mu(v+d+\mu)+vd(1-p)]}{[v+\mu][b(d+\mu)-pv]}. \end{aligned} \tag{10.36}$$

If we linearise about this steady state it can be shown that (X, Y, Z, A) tends to (X^*, Y^*, Z^*, A^*) in a damped oscillatory manner with a period of oscillation given in terms of the model parameters; the method to obtain this is exactly the same as described in Chapter 3 but the algebra is messy. With typical values for the parameters at the time (Anderson et al. 1986) the period of epidemic outbreaks was of the order of 30 to 40 years. It is unrealistic to think that the parameters characterising social behaviour associated with the disease would remain unchanged over that time span. The life expectancy

of people with HIV has dramatically increased since then, due mainly of course, to new medicines such as AZT and protease inhibitors.

We can get some interesting information from an analysis of the system during the early stages of an epidemic. Here the population consists of almost all susceptibles and so $X \approx N$ and the equation for the growth of the infectious, that is, HIV-positive, Y-class is approximated by (10.34), the solution of which is

$$Y(t) = Y(0)\, e^{v(R_0-1)t} = Y(0)\, e^{rt}, \tag{10.37}$$

where R_0 is the basic reproductive rate, $1/v$ is the average infectious period and $Y(0)$ is the initial number of infectious people introduced into the susceptible population. The intrinsic growth rate, $r = v(R_0 - 1)$, is positive only if an epidemic exists ($R_0 > 1$). From (10.37) we can obtain the doubling time for the epidemic, that is, the time t_d when $Y(t_d) = 2Y(0)$, as

$$t_d = r^{-1} \ln 2 = \frac{\ln 2}{v(R_0 - 1)}. \tag{10.38}$$

We thus see that the larger the basic reproductive rate R_0 the shorter the doubling time.

If we substitute (10.37) into equation (10.30) for the AIDS patients, we get

$$\frac{dA}{dt} = pvY(0)\, e^{rt} - (d+\mu)A.$$

Early on in the epidemic there are no AIDS patients, that is, $A(0) = 0$, and so the solution is given by

$$A(t) = pvY(0)\frac{e^{rt} - e^{-(d+\mu)t}}{r+d+\mu}.$$

Estimates for the parameter r were calculated by Anderson and May (1986) from data from 6875 homosexual and bisexual men who attended a clinic in San Francisco over the period 1978 to 1985: the average value is $0.88\,\mathrm{yr}^{-1}$. Crude estimates (Anderson and May 1986, Anderson et al. 1986) for the other parameter values are $R_0 = 3$ to 4, $d + \mu \approx d = 1 - 1.33\,\mathrm{yr}^{-1}$, $p = 10$ to 30% (this is certainly very much higher), $v \approx 0.22\,\mathrm{yr}^{-1}$, $c = 2$ to 6 partners per month. With these estimates we then get an approximate doubling time for the HIV-positive class as roughly 9 months.

Numerical simulations of the model system of equations (10.28)–(10.31) give a clear picture of the epidemic development after the introduction of HIV into a susceptible homosexual population. Figure 10.7 shows one such simulation: the model predicts that HIV incidence reaches a maximum around 12 to 15 years after the introduction of the virus into the population. It should be kept in mind that this is an early (and now more a pedagogical) model. It is interesting to compare these predictions of the mid-1980's with the situation in 2000.

In spite of the simplicity of the models, the results were in line with observation in homosexual communities. More realistic, and not always more complex models, have

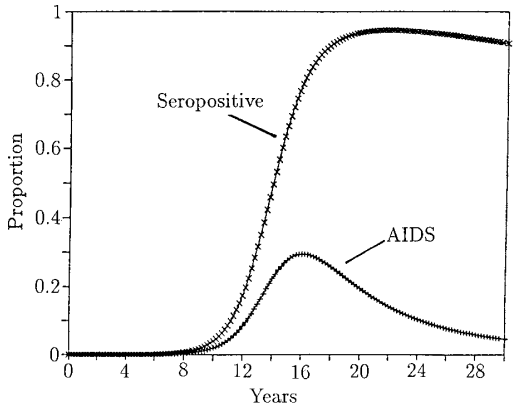


Figure 10.7. Numerical solution of the model system (10.28)–(10.31) with initial conditions $A(0) = Z(0) = 0$, $S(0) + Y(0) = N(0) = 100{,}000$. Parameter values: $B = 13333.3\ \text{yr}^{-1}$, $v = 0.2\ \text{yr}^{-1}$, $\mu = (1/32)\ \text{yr}^{-1}$, $d = 1\ \text{yr}^{-1}$, $p = 0.3$, the basic reproductive rate of the epidemic $R_0 \approx \beta c/v = 5.15$. The graphs give the proportion of those HIV-positive (seropositive) and the proportion who develop AIDS. (After Anderson et al. 1986)

been proposed such as those discussed below. A review of some of the current mathematical models for the transmission dynamics of HIV infection and AIDS is given by Perelson and Nelson (1999). With the accumulation of more data and information of the epidemic, even more sophisticated models will no doubt be required in the normal progression of realistic modelling. A practical use of good models at any stage is that, among other things, it poses questions which can guide data collection and focus on what useful information can be obtained from sparse or less than complete data. Estimates of epidemic severity doubling time, and so on, are in themselves of considerable interest and use. The model here is for a homosexual population. Now that the epidemic is very much heterosexual other models are required. The approach described here is a reasonable starting point. The models we now discuss take a very different approach to HIV infection in that we deal with the actual viral population and not human populations. As such they can be more closely tied to *in vivo* data.

10.6 HIV: Modelling Combination Drug Therapy

This section is in part based on the work of Nelson (1998). We start with the simple, but experimentally based model, proposed by Perelson et al. (1996). We then develop a more complex nonlinear model which includes treatment for HIV infection with a protease inhibitor and a reverse transcription inhibitor such as AZT.

The Ho et al. (1995) model was a simple linear first-order equation which accounted for viral production and viral decline; namely,

$$\frac{dV}{dt} = P - cV, \tag{10.39}$$

where P represented a source of viral peptides and c was the viral clearance rate. While many factors play a role in the clearance of viral peptides such as immune cells, fluid flow and absorption into other cells, c did not distinguish between them. After introduction of the protease inhibitor (the specific type of drug used on the patients) it was assumed that the drug would be completely effective, or in other words, the drug would block all viral production after being introduced. Hence $P = 0$, and we are left with the simple equation

$$\frac{dV}{dt} = -cV \ \Rightarrow\ V(t) = V_0 e^{-ct}, \tag{10.40}$$

where V_0 is measured as the mean viral concentration in the plasma before treatment. Plotting $\ln V$ against t and using linear regression to determine the slope (see Figure 10.5) gave an estimate for c and hence for the half-life of the virus in the plasma; namely, $t_{1/2} = \ln 2/c$. The mean for the half-life was $t_{1/2} = 2.1 \pm 0.4$ days; see Ho et al. (1995) for the complete data. The experimentalists then assumed that the patients were in a quasi-steady state before treatment: that is, the levels of viral load measured in the plasma remained fairly constant. With this assumption, and knowing the value for c and the initial viral concentration, V_0, they were able to compute the viral production before therapy by solving $P = cV$. While these results were minimal estimates, based on the assumption of a perfect drug (with no delays), they still provided an estimate of over 1 billion viral particles being produced daily. This important result was contrary to the belief that the viral dynamics during this latent period was close to dormant.[9] It is an excellent example where even simple, mathematically trivial, models can be of immense help in extracting crucial information from patient data. Another example which changed the way patients with liver disease were assessed for (toxic) medication levels is given in Connor et al. (1982a,b).[10]

Due to these results of Ho et al. (1995) many more models have been developed to study the HIV; see Perelson and Nelson (1999) for a comprehensive review. In the rest of this section we examine several models, in particular one which looks at combination drug therapy and briefly discuss another which includes a delay.

Protease inhibitors are drugs which target the protease enzymes in the cell and cause newly produced viruses to be noninfectious. To date there is no single drug (nor even a combination of them) which completely kills the HIV infection because of the ability of the virus to mutate into a drug resistant form. It takes time, however, for a

[9] This result, based on an incredibly simple mathematical model, did much to boost the usefulness of mathematical models in the medical community, a consequence of which is that many more laboratories are now looking for theoreticians to help in the modelling process.

[10] The model consisted of a two-compartment model which results in a pair of coupled linear ordinary differential equations which can be solved simply analytically with patient-based initial conditions. I set it as a modelling exercise for first-year mathematics students in Oxford but the question was not well described so it was not clear exactly what was required. One of the college tutors, dealing with the difficulties his students were having in understanding what was both going on and required, said I must have done it deliberately to simulate what it was like talking to doctors.

new form to evolve. The idea behind combination drug treatment is when the virus is presented with two quite different antiviral drugs the time it takes for a mutiple-drug resistant strain to emerge is much longer than if the virus had to contend with only one toxic drug. This is also discussed in the paper by Perelson and Nelson (1999). The use of multiple drug treatments, such as protease inhibitors together with AZT, has already had a major effect (in the devloped world) in significantly slowing down the progression from HIV infection to full-blown AIDS. It has not, however, effected a cure for the disease. Already there is reemergence of drug-resistant strains of HIV in homosexuals in San Francisco who have been taking the combination drug cocktail.

We consider each drug to be less then perfect, which thus allows for viral mutation to a resistant form if administered independently. Let n_p be a measure of the effectiveness of a protease inhibitor or combination of protease inhibitors in blocking production of infectious virions so this will affect the viral dynamics directly and the T-cells indirectly. Other commonly used drugs are reverse transcriptase inhibitors, of which AZT is perhaps the best known. After the development of the protease inhibitors, a combination, or cocktail, therapy which included multiple drugs was prescribed. For instance, patients would take a combination of three drugs made of up of a protease inhibitor and two reverse transcriptase inhibitors. This combination was dramatic initially in reducing the number of viral peptides detectable in the patient and it was thought that this might be the cure for the AIDS virus. Unfortunately, with a virus as complex as the HIV it was only a matter of time before the emergence of resistant viruses. While the combination treatment is still showing promise for prolonging the lives of infected patients, it is too early (2001) to say whether or not the virus is even permanently controlled, far less cured.

We develop a four-species model which includes an equation for uninfected T-cells, T, productively infected T-cells, T^* (not all infected T-cells produce the virus), infectious viruses, V_I and noninfectious viruses, V_{NI}. The model consists of the following equations which we motivate in turn below.

$$\begin{aligned}
\frac{dT}{dt} &= s + pT\left(1 - \frac{T}{T_{\max}}\right) - d_T T - kV_I T,\\
\frac{dT^\star}{dt} &= (1 - n_{rt})kV_I T - \delta T^\star,\\
\frac{dV_I}{dt} &= (1 - n_p)N\,\delta T^\star - cV_I,\\
\frac{dV_{NI}}{dt} &= n_p N\,\delta T^\star - cV_{NI}.
\end{aligned} \tag{10.41}$$

In the T-cell equation we consider the cells to be destroyed proportional to the number of infected viruses and cells with clearance parameter k. In the absence of infection there is a nonzero steady state, T_{s1}, so we have a quadratic polynomial in T for the uninfected T-cell dynamics: s, p, $T_{\max}$, d_T and k are positive constants. The specific form of the T-cell kinetics, namely, with a logistic form plus another source (s) and a clearance term ($-d_T T$), is because of the form of T-cell recovery after therapy as indicated by patient data. With the reverse transcriptase (RT) drug like AZT, the RT-inhibitor acts on the source term for productively infected T-cells with $0 \le n_{rt} \le 1$ the measure of its efficacy; if $n_{rt} = 1$ it is completely effective and prevents all production of infected

T-cells while if $n_{rt} = 0$ it implies no RT-inhibitor is given. In the T^* equation the effect of the RT-inhibitor is to reduce the production of the infected cells. These cells also have a natural death with a rate parameter, δ. The protease inhibitor acts on the source of the virus and so appears in the V_I equation with n_p a measure of its efficacy. The specific appearance in the equations for the effects of the drugs is due to the cellular mechanisms of each drug and the stage at which they aim to target during infection. When a drug is completely effective we set $n_p = 1$ or $n_{rt} = 1$. In the infected virus V_I equation there is a factor N which is the bursting parameter for the viral production after lysis (essentially the breaking up, or death, of the cell due to its penetration by the infected virus and subsequent generation of a large number of viruses); it is of the order of 480 virions/cell (a virion is a complete virus with all its coating, proteins and so on). The infected viruses are considered to die naturally at a rate c. Finally the noninfectious viruses are produced with a rate dependent on the protease drug and we assume they die off at the same rate as the infected ones. This model lets us explore the effect of the drugs on the HIV by varying, in particular, the parameters n_{rt} and n_p. For example, if $n_p = 0$ we are using only the reverse transcriptase, or RT-inhibitors. We now analyse this system in several ways and compare the results with patient data.

Some idea of the values of the dependent variables are (from Ho et al. 1995): $T \sim 180$ cells/mm^3, $T^* \sim 2\%$ T-cells, $V_I \sim 134 \times 10^3$ virions/ml, $V_{NI} = 0$ virions/ml. Available parameter estimates are: the viral activity rate $k \sim 3.43 \times 10^{-5}$ virions/ml (Ho et al. 1995), death rate of infected cells $\delta \sim 0.5$/day (Perelson et al. 1996), viral production by the bursting cell $N \sim 480$ virions/cell (Perelson et al. 1996), clearance rate of the virus $c \sim 3$/day (Perelson et al. 1996), T-cell source $s = 0 - 10$ cells/mm^3/day (Kirschner and Webb 1996) and death rate of targeted cells $d_T \sim 0.03$/day (McLean and Mitchie 1995).

T-Cell Recovery

Some models have assumed that the T-cells do not change dynamically during the first weeks of treatment and hence set $T =$ constant $= T_0$. However, after antiretroviral therapy is initiated some recovery of T-cells is observed and patient data presented by Ho et al. (1995) suggest that over a period of weeks the recovery of T-cells can be described by either a linear or exponential function of time, with no statistically significant difference between the two functions over that time period. After therapy is initiated $V_I(t)$ falls rapidly. For a perfect protease inhibitor, namely, $n_p = 1$, the solution of the fourth equation of (10.41) is $V_I(t) = V_0 e^{-ct}$ and so after a few days (depending on c of course) the term $-kV_I T$ could be negligibly small in the equation for T-cells. T-cell replacement can be due to the source s, which incorporates the generation of new cells in the thymus, their export into the blood and the transport of already created T-cells in tissues to the blood, or to proliferation of cells. It was previously thought that the adult thymus no longer produced T-cells but with the significant advances in the study of HIV dynamics some believe this to be incorrect. If the source s is the major mechanism of T-cell replacement, we can then approximate the T-cell dynamics by

$$\frac{dT}{dt} = s - d_T T \quad \text{or} \quad T(t) = T_0 + at,$$

where a is a rate constant.

If we now consider the effect of only protease inhibitor drugs, that is, $n_{rt} = 0$, which relates directly to the patient data of Ho et al. (1995), and further assume the above linear T-cell growth in line with the patient data, (10.41) become

$$\begin{aligned} \frac{dT^\star}{dt} &= kV_I(T_0 + at) - \delta T^\star, \\ \frac{dV_I}{dt} &= (1 - n_p)N\,\delta T^\star - cV_I, \\ \frac{dV_{NI}}{dt} &= n_p N\,\delta T^\star - cV_{NI}, \end{aligned} \tag{10.42}$$

which is a nonautonomous system but which can be trivially made into one. To do it we simply replace the $T_0 + at$ in the first equation by T and add to the sytem the differential equation $dT/dt = a$ with initial condition $T(0) = T_0$. Typical solutions of this system are shown in Figure 10.8 with the estimated parameter values given in the legend. In Figure 10.9 we show how the solutions compare with specific patient data. The comparison is quantitatively very good.

We now consider the full nonlinear model given by (10.41), which, as is easily shown, has two steady states one of which is the noninfected steady state $(T_{s1}, 0, 0, 0)$ with preinfected T-cells; we are only interested in $T_{s1} \geq 0$ of course. A little algebra shows that this uninfected state is

$$T_{s1} = \frac{T_{\max}}{2p}\left[(p - d_T) + \sqrt{(p - d_T)^2 + \frac{4sp}{T_{\max}}}\right]. \tag{10.43}$$

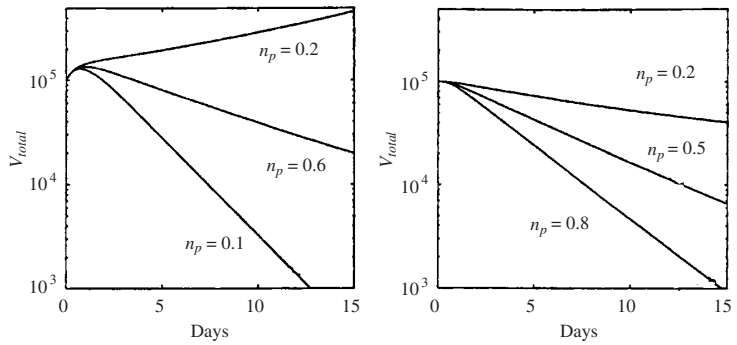


Figure 10.8. Solutions for the total virus population, $V_I + V_{NI}$, of the system (10.42) which assumes linear T-cell growth $T(t) = T_0 + at$ and monotherapy with a protease inhibitor; that is, there is no AZT-like drug so $n_{rt} = 0$. Note the change in viral output as a function of the level of drug effectiveness. (**a**) The viral decay assuming a pretreatment steady state value with $c = NkT_0$ and varying n_p. (**b**) Viral decay after treatment without a pretreatment value for c: the critical efficacy here is $n_c = 0.33$. Parameter values: $N = 480$, $k = 3.43 \times 10^{-5}$/day, $\delta = 0.43$/day, $T_0 = 180$, $a = 1$ cells/day, $c = 2$/day. (From Nelson 1998)

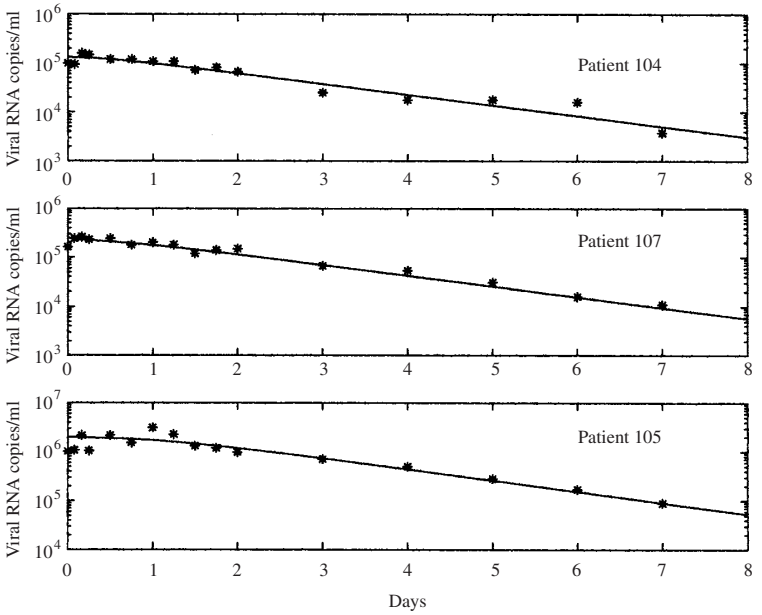


Figure 10.9. Solutions for the total virus population, $V = V_I + V_{NI}$, plotted in terms of the RNA, of the system (10.42), which assumes constant T-cell growth and monotherapy with a protease inhibitor and their comparison with three patient data of Ho et al. (1995). These are typical of the other patients. The parameters for each patient were obtained from a best fit using figures equivalent to those in Figure 10.5. Parameter values: $\delta = 0.5$/day then, patient 104: $T_0 = 2\,\text{mm}^{-3}$, $\alpha = 1.5$, $V(0) = 52 \times 10^3$, $c = 3.7$/day, total viral production rate: $P = 2.9 \times 10^9$/day; patient 105: $T_0 = 11\,\text{mm}^{-3}$, $\alpha = 10.18$, $V(0) = 643 \times 10^3$, $c = 2.1$/day, total viral production rate: $P = 32.1 \times 10^9$/day; patient 107: $T_0 = 412\,\text{mm}^{-3}$, $\alpha = 2.64$, $V(0) = 77 \times 10^3$, $c = 3.1$/day, total viral production rate $= 3.0 \times 10^9$/day. (From Perelson and Nelson 1999 and reproduced with permission)

We are interested in the stability of this steady state if perturbed with the introduction of HIV. We examine the stability in the usual way, exactly as we did, for example, in Chapter 3 by looking at the eigenvalues of the perturbed linear system. After somewhat more algebra we find the eigenvalues are

$$\begin{aligned} \lambda_1 &= p\left(1 - \frac{2T_{s1}}{T_{\max}}\right) - d_T, \quad \lambda_4 = -c, \\ \lambda_{2,3} &= -\frac{c + \delta}{2} \pm \frac{1}{2}\sqrt{(c + \delta)^2 - 4c\delta + 4\delta NkT_{s1}(1 - n_c)} \\ n_c &= 1 - (1 - n_{rt})(1 - n_p), \end{aligned} \tag{10.44}$$

where n_c represents the effectiveness of the combination treatment. For stability we require the eigenvalues to be negative; they are all real here. The only non obvious

negative eigenvalues are λ_1, which requires $T_{s1} > (1/2p)(p - d_T)T_{\max}$, a condition that is obviously satisfied from (10.43), and λ_2. The eigenvalue $\lambda_2 < 0$ is satisfied if

$$\lambda_2 = -\frac{c+\delta}{2} + \frac{1}{2}\sqrt{(c+\delta)^2 - 4c\delta + 4\delta NkT_{s1}(1-n_c)} < 0;$$

that is,

$$c + \delta > \sqrt{(c+\delta)^2 - 4c\delta + 4\delta NkT_{s1}(1-n_c)}.$$

So, the uninfected steady state is stable if

$$n_c = 1 - (1-n_p)(1-n_{rt}) < \frac{c}{NkT_{s1}} \Rightarrow n_c > 1 - \frac{c}{NkT_{s1}}. \tag{10.45}$$

This means that if the drug treatment is strong enough the virus will be eliminated. (Actually, the virus is never eliminated but it does fall below detectable levels.) We can estimate the required effectiveness of treatment from this condition. Under the assumption of a pretreatment steady state, with $T = T_0$, the second and third of (10.41) imply $c = NkT_0$. By way of example, if we set $n_{rt} = 0$, the stability condition (10.45) then becomes

$$n_p > 1 - \frac{T_0}{T_{s1}}.$$

Healthy individuals have T-cell counts of about 1000/mm^3 so we can assume $T_{s1} = 1000$. Hence for a patient with a pretreatment T-cell count of say, $T_0 = 200$, we find n_p needs to be greater than 0.8. For a less advanced patient with a T-cell count of $T_0 = 500$, n_p need only be greater than 0.5. Thus, this analysis supports the notion that patients should be started on antiretroviral drug therapy as early as possible (Perelson and Nelson 1999). On the other hand if we have both drugs administered the condition is then

$$(1-n_p)(1-n_{rt}) < 1 - \frac{T_0}{T_{s1}}$$

and with $T_0 = 200$ we need only have, for example, an efficacy of 0.55 for each of n_p and n_{rt}.

The second steady state, the infected steady state, is obtained, after some algebra, from (10.41) as

$$T_{s2} = \frac{c}{Nkn_c}, \quad \bar{V}_I = \frac{s}{kT_{s2}} + \frac{1}{k}\left[p\left(1 - \frac{T_{s2}}{T_{\max}}\right) - d_T\right],$$
$$\bar{T}^\star = \frac{c\bar{V}_I}{\delta N(1-n_p)}, \quad \bar{V}_{NI} = \frac{n_p\bar{V}_I}{1-n_p},$$

where overbars denote steady state quantities and as before $n_c = (1-n_{rt})(1-n_p)$. In the absence of treatment, $n_c = 1$ but here we are concerned with studying the effects of less than perfect drugs so $0 < n_c < 1$.

This steady state is relevant only if $\bar{V}_I > 0$; that is,

$$\frac{s}{T_{s2}} + p - d_T - p\frac{T_{s2}}{T_{\max}} > 0. \tag{10.46}$$

If the inequality (10.46) is replaced by an equality and the equation $\bar{V}_I = 0$ solved for T_{s2}, we obtain an expression identical to the expression for T_{s1}. Thus, at $\bar{V}_I = 0$ the uninfected and infected steady states merge. Further, as T_{s2} decreases the left-hand side of (10.46) increases. So, for $\bar{V}_I > 0$, an infected steady state exists, $0 < T_{s2} < T_{s1}$, which of course makes biological sense since in the infected steady state the system should have fewer T-cells than in the uninfected state.

Substituting the expression for T_{s2} into the steady state equation for V_I gives a necessary condition for the infected steady state to exist; namely,

$$\bar{V}_I = \frac{sN(1-n_c)}{c} + \frac{1}{k}\left[p\left(1 - \frac{c}{NkT_{\max}(1-n_c)}\right) - d_T\right] > 0. \tag{10.47}$$

If we look at a limiting case where $s = 0$, from (10.47), certainly if

$$Nk < \frac{c}{T_{\max}(1-n_c)} \Rightarrow V_I < 0.$$

Let us now consider the stability of this infected steady state by calculating the eigenvalues. The Jacobian matrix, evaluated at the infected steady state, is

$$\begin{pmatrix} p(1-\frac{2\bar{T}}{T_{\max}}) - d_T - k\bar{V}_I & 0 & -k\bar{T} & 0 \\ (1-n_{rt})k\bar{V}_I & -\delta & (1-n_{rt})k\bar{T} & 0 \\ 0 & \delta N(1-n_p) & -c & 0 \\ 0 & \delta N n_p & 0 & -c \end{pmatrix},$$

where $\bar{T} = T_{s2}$.

The characteristic equation immediately gives one eigenvalue as $\lambda_4 = -c < 0$. The other three eigenvalues, λ, are determined by solving the cubic

$$\left[p\left(1 - 2\frac{\bar{T}}{T_{\max}}\right) - d_T - k\bar{V}_I - \lambda\right][(c+\lambda)(\delta+\lambda) - k\bar{T}\delta N(1-n_c)]$$
$$- k\bar{V}_I k\bar{T}\delta N(1-n_c) = 0$$

which, using the steady state value for $\bar{T}$, simplifies to

$$\left[p\left(1 - \frac{2\bar{T}}{T_{\max}}\right) - d_T - k\bar{V}_I - \lambda\right][\lambda^2 + (\delta+c)\lambda] - kc\delta\bar{V}_I = 0;$$

that is,

$$\lambda^3 + A\lambda^2 + B\lambda + C = 0,$$

where

$$A = \delta + c + \frac{2p\bar{T}}{T_{\max}} - (p - d_T) + k\bar{V}_I,$$
$$B = (\delta + c)\left[\frac{2p\bar{T}}{T_{\max}} - (p - d_T) + k\bar{V}_I\right], \quad C = c\delta k\bar{V}_I.$$

We do not need the actual expressions for the eigenvalues, only the sign of their real part. The Routh–Hurwitz conditions (see Appendix B) state that, if $A > 0, C > 0$ and $AB - C > 0$ then the eigenvalues have negative real parts. By inspection, $C > 0$. At steady state,

$$s + (p - d_T)\bar{T} - \frac{p\bar{T}^2}{T_{\max}} = k\bar{V}_I\bar{T}.$$

Since $s > 0$,

$$(p - d_T)\bar{T} - \frac{p\bar{T}^2}{T_{\max}} < k\bar{V}_I\bar{T}$$

or

$$p - d_T < \frac{p\bar{T}}{T_{\max}} + k\bar{V}_I,$$

from which it follows that $A > 0$. The remaining condition necessary for stability of the infected steady state is $AB - C > 0$. Let us write $A = (\delta + c + B_1)$ with B_1 defined by the expression for A and note that B can then be written as $B = (\delta + c)B_1$. Exploiting this form and noting that B_1 contains the term $k\bar{V}_I$, it can then be simply shown that $AB = B_1(\delta + c)^2 + B_1^2(\delta + c) > \delta ck\bar{V}_I = C$. Hence the infected steady state, if it exists, is stable.

As noted above if the infected steady state exists, $T_{s2} < T_{s1}$, which we can rewrite as

$$c < NkT_{s1}(1 - n_c).$$

To summarise, if $c > NkT_{s1}(1 - n_c)$ then the only nonnegative steady state is the uninfected steady state and it is stable. Conversely, if $c < NkT_{s1}(1 - n_c)$ then the uninfected state is unstable and the infected state exists and is stable. This is equivalent to saying that there is a transcritical bifurcation when $c = NkT_{s1}(1 - n_c)$. We can express these conditions in a different way in terms of the model parameters. The critical treatment efficacies, for example, are related to the model parameters by

$$\left[(1 - n_p)(1 - n_{rt})\right]_{\text{critical}} = \frac{c}{2skN}\left[\sqrt{(p - d_T)^2 + \frac{4sp}{T_{\max}}} - (p - d_T)\right]. \quad (10.48)$$

10.7 Delay Model for HIV Infection with Drug Therapy

We now touch on some recent work in which a discrete delay is added to the model to account for the time lag between the time a cell becomes infected and the time at which the infected cell starts producing virus. Work in this area has shown that including a delay of this form affects the estimated values derived from kinetic experiments for the half-life of productively infected cells and viruses. Here we only give a very brief description of the current work with delay models.

Time Lags in the HIV Infection Process

The virus life cycle plays a major role in disease progression during HIV infection. The binding of a viral particle to a receptor on the CD4 T-cell, or other targeted cell, begins a chain of events that can eventually lead to the CD4 T-cell becoming productively infected, that is, producing new viruses. Most previous models consider this process to occur instantaneously. In other words, it is assumed that as soon as virus contacts a targeted cell the cell begins producing viruses. However, biologically there is a measurable time delay between initial viral entry into a cell and subsequent viral production. Recently there have been models which examine this effect and they have shown that this delay needs to be taken into account to determine accurately the half-life of a free virus from drug perturbation experiments. If the drug is assumed to be completely efficacious, the delay does not affect the estimated rate of decay of viral producing T-cells (Herz et al. 1996, Mittler et al. 1998, 1999). If the assumption of the drug being completely effective is not assumed (to date no drug is 100% effective) the introduction of the delay then affects the estimated value for the infected T-cell loss rate (Nelson 1998, Nelson et al. 2000).

Such a delay model, based on the one in the last section, is given and discussed in detail by Nelson et al. (2000). In it we incorporate the intracellular delay by considering the generation of virus-producing cells at time t to be due to the infection of targeted cells at time $t - \tau$, where the delay, τ, is taken to be a constant. Of course in reality the delay is a distributed function (refer to Chapter 1). We also assume uninfected T-cells remain constant; that is, $T = T_0$. Model equations describing this scenario are (compare with (10.41))

$$\begin{aligned}
\frac{dT^\star}{dt} &= kT_0V_I(t - \tau) - \delta T^\star, \\
\frac{dV_I}{dt} &= (1 - n_p)N\delta T^\star - cV_I, \\
\frac{dV_{NI}}{dt} &= n_pN\delta T^\star - cV_{NI},
\end{aligned} \quad (10.49)$$

where the term $V_I(t-\tau)$ allows for the time delay between contact and viral production. The average life span of a virus is $1/\delta$. With delay we are saying that the average life span of a cell from time of infection to death is $\tau + (1/\delta)$. We do not have a precise value for τ but estimates of 1 to 1.5 days (Perelson et al. 1996, Mittler et al. 1998, 1999). So, to a first approximation, T-cells infected with HIV-1 might live on average 2 to 3 days

rather an average of 1 to 2 days. The rate constant for infection, k, is assumed constant because the drug we are modelling, namely, a protease inhibitor, does not affect k. If a reverse transcriptase inhibitor were being used then the appropriate model would have k in the $dT^\star/dt$ equation replaced by $[1-n_{rt}(t-\tau)]k(t-\tau)$. Here the V_{NI} equation is uncoupled from the T^* and V_I equations and so can be solved independently once the solution for the first two equations is known. Analysis of a more general form of this delay model, which included uninfected T-cells and nonlinearities are given in Nelson (1998). The method of analysis is similar to that discussed in detail in Chapter 7.

This model has been used, among other things, to analyze the change in parameters associated with the decay rate seen in data from patients undergoing antiviral treatment. It has also helped in getting better estimates for crucial parameters from patient data. The main conclusions from the analysis of the model, with experimentally estimated parameters, is that when the drug efficacy is less than 100%—the case *in vivo* at present—the rate of decline of the virus concentration in the plasma primarily depends on the efficacy of the therapy, the death rate of the virus producing cells and the length of the delay. These are all to be expected. The main point of the model and its analysis is that the results quantify these effects in terms of the measurable (and experimentally changeable) parameters.

10.8 Modelling the Population Dynamics of Acquired Immunity to Parasite Infection

Gastrointestinal nematode parasite infections in man are of immense medical importance throughout the developing world. An estimated 800 to 1000 million people are infected with *Ascaris lumbricoides*, 700 to 900 million with the hookworms *Ancyclostoma duodenale* and *Nector americanus* and 500 million with the whipworm *Trichuris trichiura* (Walsh and Warren 1979). To design optimal control policies, we must have an understanding of the factors which regulate parasite abundance and influence the size and stability of helminth populations. So, in this section we present a model for the immunological response by the host against gastrointestinal parasites which was proposed and studied by Berding et al. (1986). We show that such relatively simple modelling can have highly significant implications for real world control programmes.

Parasites invoke extremely complex immunological responses from their mammalian hosts. We still do not know exactly how these come about but current experimental research provides some important pointers which form the basis for the mathematical model. Also the modelling in this section demonstrates how to determine some of the parameter estimates from a combination of theory and experiment which would be difficult to obtain from experiment alone.

Let us first summarise the relevant biological facts starting with a brief review of key experiments. Laboratory experiments in which mice are repeatedly exposed to parasite infection at constant rates can provide a suitable test for mathematical models of helminth population dynamics. Experiments relevant for our model (Slater and Keymer 1986) involve two groups of 120 mice, which are fed on artificial diets containing either 2% ('low protein') or 8% ('high protein') weight for weight protein. Both groups were subdivided into 4 groups of 30 mice, which we denote by (a), (b), (c) and (d), which were subjected to repeated infection with larvae of the nematode *Heligmosoides polygyrus*. The subgroups were infected at different rates: group (a) with 5 larvae/mouse/two weeks, group (b) with 10 larvae/mouse/two weeks, group (c) with 20 larvae/mouse/two weeks and group (d) with 40 larvae/mouse/two weeks. So, we have a total of 8 subgroups of 30 mice differing either in their infection rates or in the protein diets they were fed on. It is known that protein deprivation impairs the function of the immune system so this scheme lets us compare parasite population dynamics, under various infectious conditions, in the presence and the absence of an acquired immune response.

The most important experimental observations were the temporal changes in the mean worm burden, M, namely, the total number of adult worms divided by the total number of hosts. Every two weeks throughout the experiment a sample of 5 mice from each group was examined for the presence of adult parasites. The number of parasites present in each mouse was determined by postmortem examination of the small intestine. The main experimental results are shown in Figure 10.10 which displays the mean

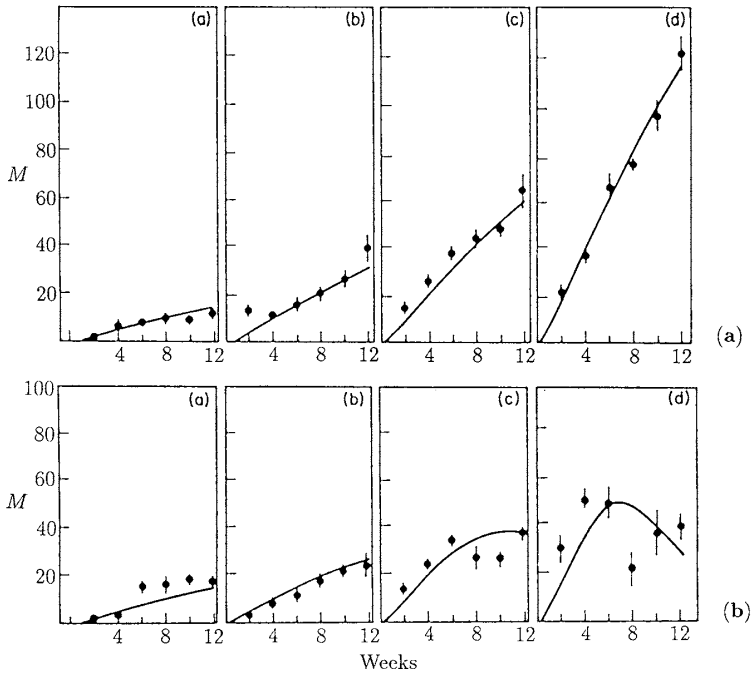


Figure 10.10. Change in mean adult worm burden, M, in mice hosts fed on a protein diet for a repeated infection over a 12-week period: (**a**) low protein diet; (**b**) high protein diet. The infection rates are (a) 5, (b) 10, (c) 20, (d) 40 larvae/mouse/2 weeks. The periods are the experimental points from Slater and Keymer (1986). The continuous lines are solutions of the mathematical model; how these were obtained is described in the text in the subsection on the population dynamics model and analysis. (From Berding et al. 1986)

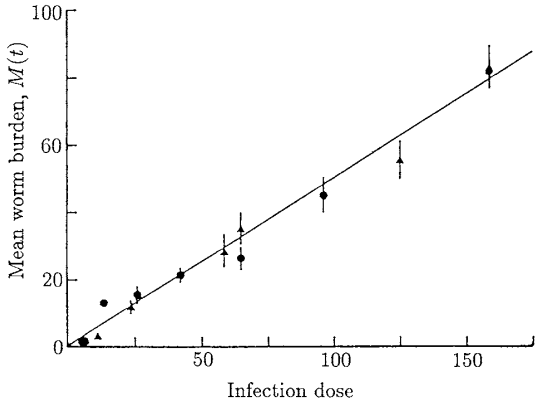


Figure 10.11. These experimental results show the worm survival after a single infection of larvae. There is a linear relationship between larval dose and adult worm burden; these results are after 14 days from the infection. The circles represent mice fed on a low protein diet and the triangles are for mice fed on a high protein diet. The solid line is a best fit linear description of the data; the gradient is 0.64, from which we deduce that 64% of the larvae survive. (From Berding et al. 1986)

worm burden as a function of time for the low and the high protein groups, respectively. The letters (a), (b), (c) and (d) refer to infection rates of 5, 10, 20 and 40 larvae per two weeks.

Other experiments were carried out to quantify parasite establishment and survival in primary infection. Here only a single dose of larvae was given unlike the repeated infection in Figure 10.10. The results shown in Figure 10.11 give the mean worm burden as a function of the infection dose. From this figure we estimate that approximately 64% of the infective larvae survive to become adult worms.

The survival of adult worms in a single infection is summarised in Figure 10.12: it again shows the mean worm burden as a function of time. In this situation the worm population remains free from effects of the host's immune system: we use this figure to estimate the natural death rate of the adult worms.

So as to be able to construct a realistic model, let us summarise these and related experimental observations.

(A1) Infective parasite larvae, after ingestion by the host, develop into tissue dwelling larvae which become adult worms found within the lumen of the alimentary canal. This invokes a distinct immunological response from the host. Typically, the tissue dwelling larvae are in the most immunogenic stage in the parasite life cycle. We thus assume that the immune system is triggered according to the larval burden experienced by the host.

(A2) Many experiments point to the presence of delay, that is, memory, effects in immune response. Some of these effects can be accounted for by including delay in the models.

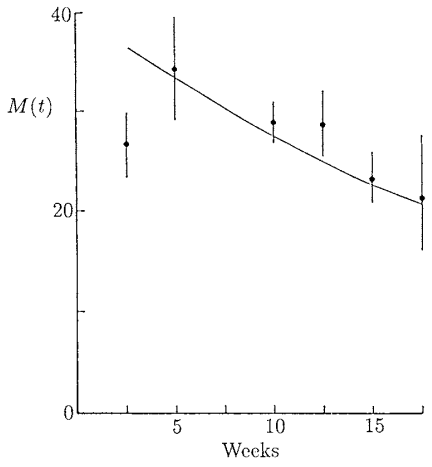


Figure 10.12. Survival of adult worms which follows a single infection of 50 larvae/mouse on day zero. The results are for mice fed on a low protein diet. The continuous line is a best fit for an exponential survival model with a constant death rate $\delta = 5.6 \times 10^{-3}\,\text{day}^{-1}$: that is, we assume the worms die proportional to their mean population (see equation (10.53) below). (From Berding et al. 1986)

(B1) The experimental results shown in Figure 10.10 suggest, importantly, that the strength of the immune response is very much dependent on the nutritional status of the host. We interpret the differences in the dynamics of infection in mice feeding on low and high protein diets, that is, Figures 10.10(a) and (b) respectively, as a consequence of a relationship between the nutritional status and immunological competence.

(B2) The similarity between Figures 10.10(a) (a),(b) and Figures 10.10(b) (a),(b) and the differences between Figures 10.10(a) (c),(d) on the one hand and Figures 10.10 (b),(c),(d) on the other, clearly indicate a threshold behaviour of the immune system. Biologically this means that the full activation of the immune system requires a certain threshold of exposure to parasite infection.

(B3) Available evidence on the effectiveness of the immune response, which we define here as the per capita rate of limitation in parasite establishment and survival, in relation to its stimulus, that is, increased exposure to infection, suggests the following scenario. After an initial increase, the activity of the immune response to the parasites saturates at a maximum level. Further stimulation does not seem to increase the subsequent effectiveness of acquired immunity. So, we assume here that the activity of immune response saturates at a defined maximum level.

(C) The immunological response may act against several stages in the parasite life cycle. In some strains of mice it directly kills tissue-dwelling larvae. However, in others the immune response is not capable of preventing larval development. In these, larvae subjected to immunological attack emerge as stunted adults, with a correspondingly high mortality rate. So, to reflect these experimental findings we model immunological competence by an increased mortality rate of the adult parasite.

Let us now construct the model on the basis of these assumptions, all firmly based on experimental observations, in the following three main steps.

(i) We introduce a variable, E, for the immune system, which takes into account assumptions (A1) and (A2), by

$$E = \int_{t-T}^{t} L(t')\,dt', \tag{10.50}$$

where $L(t)$ denotes the mean number of tissue-dwelling larvae in a host at time t with T the time-span over which the immune system retains memory of past infections. So, E is a measure of the number of larvae in the host during the time interval $(t - T, t)$. Note that with the form (10.50) different situations (for example, a small infection persistent for a long time and a large infection persistent for a short time) can lead to the same values of E.

(ii) To account in a simple way for the biological facts in (B1) through (B3) we introduce an expression to describe the immune system's activity; namely,

$$I \equiv I_{\alpha\beta}(E) = \frac{\alpha E^2}{\beta + E^2}, \tag{10.51}$$

where E is the input variable (10.50), α is the maximum functional activity of the host's immune response and β provides a measure of the sensitivity of the immune system. (Recall the predation response in the budworm model dynamics in Chapter 1.) According to (B1), α also reflects the nutritional status of the host being considered; we can think of α as a monotonic increasing function of the nutritional status. β also may be host specific since it seems likely that β also has a direct biological interpretation in genetic terms since different strains of mice differ in their immune response against parasitic infections.

(iii) Finally we have to incorporate (10.50) with (10.51) into a dynamical model for the complete host–parasite community. According to assumption (C), and independent of the specific dynamical situation under consideration, the activity of the host's immune system simply leads to an increase in the mortality of the adult parasites. This requires an additional loss term in the dynamical equations for the mean worm burden, $M(t)$, of the form $-IM(t) < 0$, where I, the strength of immunological response, plays the role of a death rate for parasites; it depends on the level of infection.

Population Dynamics Model and Analysis

From the above, mice fed on low protein diets appear to have little or no immune response; we refer to these as the low protein diet group (LPG) and investigate the dynamics of their mean worm burden by a simple immigration–death model. On the other hand, hosts feeding on a high protein diet are expected to show an immune response.

Low Protein Model

Let us start by considering the parasite dynamics of the LPG. The parasites, harboured by a host population of constant size, are subdivided into two categories: larvae in the wall of the small intestine, and adult worms in the gut lumen. We model the dynamics of the mean number of larvae, L, per host by

$$\frac{dL}{dt} = \lambda_i - \mu D L, \quad i = 1, 2, 3, 4, \tag{10.52}$$

where λ_i, $i = 1, 2, 3, 4$, refer to the experimentally controlled infection rates, for example, 5, 10, 20 and 40 larvae per mouse per 2 weeks as in the experiments recorded in Figure 10.10. Here $1/D = C_L$ denotes the proportion of larvae developing into adult worms after a developmental time delay t_L, here denoted by $1/\mu$. For the parasite *Heligmosoides polygyrus*, $t_L = 1/\mu \approx 8$ days and from Figure 10.11 we estimate $C_L = 0.64$. We can now evaluate the net loss rate of the larval population per host as $\mu D \approx 0.195\,\text{day}^{-1}$ which implies (i) an effective life span of a larval worm of $1/(\mu D) \approx 5.12$ days, and (ii) the natural larval mortality rate $\mu_0 = \mu(D - 1) \approx 0.07\,\text{day}^{-1}$.

We model the dynamics of the mean adult worm burden, M, by

$$\frac{dM}{dt} = \mu L - \delta M, \tag{10.53}$$

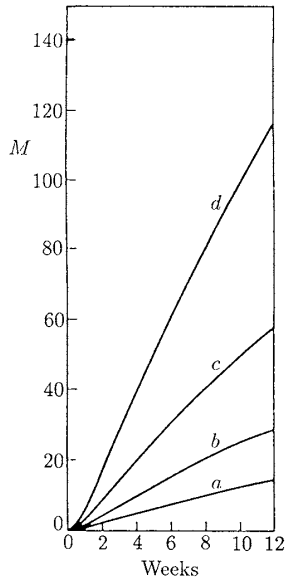


Figure 10.13. Mean worm burden $M(t)$ for mice on the low protein diet (LPD) obtained from the analytical solution (10.54) of the model (10.52) and (10.53). The curves correspond to the different larvae infection rates λ_i, $i = 1, 2, 3, 4$: (*a*) 5, (*b*) 10, (*c*) 20, (*d*) 40 larvae per mouse per 2 weeks. Parameter values: $\mu = 0.125\,\text{day}^{-1}$, $D = 1.56$, $\delta = 5.6 \times 10^{-3}\,\text{day}^{-1}$. These curves correspond to those superimposed on Figure 10.10(a).

where δ denotes the natural death rate of the adult worms in the absence of competitive or immunological constraints. We estimate $\delta = 5.6 \times 10^{-3}\,\text{day}^{-1}$ from the experimental results of a single infection shown in Figure 10.12, which implies an adult worm life span of approximately 25 weeks.

Solutions of the linear equations (10.52) and (10.53), with the initial conditions $L(0) = M(0) = 0$, are simply

$$\begin{aligned} L(t) &= \frac{\lambda_i}{\mu D}(1 - e^{-\mu D t}) \\ M(t) &= \frac{\lambda_i}{D}\left\{\delta^{-1}(1 - e^{-\delta t}) \right. \\ &\quad \left. + (\mu D - \delta)^{-1}(e^{-\mu D t} - e^{-\delta t})\right\}, \quad i = 1, 2, 3, 4. \end{aligned} \tag{10.54}$$

Figure 10.13 plots $M(t)$ for $i = 1, 2, 3, 4$ for the first 12 weeks using the above estimates for the parameter values. These are the curves which are superimposed on the experimental results in Figure 10.10(a); there is very good quantitative agreement.

High Protein Model

With this diet the host's immune system comes into play and so we have to incorporate its action into the dynamical equation (10.53) for the worm burden. In line with the observation (iii) above, this equation now takes the form

$$\frac{dM}{dt} = \mu L - (\delta + I)M, \tag{10.55}$$

$$I = \frac{\alpha E^2}{\beta + E^2}, \quad E = \int_{t-T}^{t} L(t')\,dt', \tag{10.56}$$

where I is the cumulative effect of increased mortality of the worms by the immune response.

The larvae equation is still taken to be (10.52) since we assume the immune response does not principally alter the larvae dynamics. The infection pattern in the laboratory situation is then given by (10.54) as

$$L(t) = \begin{cases} 0, & t < 0 \\ \dfrac{\lambda_i}{\mu D}(1 - e^{-\mu D t}), & t > 0 \end{cases}, \tag{10.57}$$

where λ_i, $i = 1, 2, 3, 4$ are the different larval infection rates. This generates the immune system input function E given by (10.56); integration gives

$$E(t) = \begin{cases} \dfrac{\lambda_i}{\mu D}\left\{t - \dfrac{1}{\mu D}(1 - e^{-\mu D t})\right\}, & 0 < t < T \\ \dfrac{\lambda_i}{\mu D}\left\{T - \dfrac{1}{\mu D}e^{-\mu D t}(1 - e^{\mu D t})\right\}, & t > T \end{cases} \tag{10.58}$$

which, as $t \to \infty$, asymptotes to the constant $\lambda_i T/\mu D$.

The high protein model consists of (10.52) and (10.55) and to solve it we must first obtain estimates for the immune system parameters T, α and β, respectively the memory time from past infections, the maximum mortality contribution from the immune system and the worm burden at which the immune response is switched on. Accurate estimates of immunological memory time T are not available. Some data (Rubin et al. 1971) indicate that some mice retain active immunity against *Heligmosoides polygyrus* for at least 30 weeks after infection. On the basis of this we assume T is at least larger than the experimental duration time of 12 weeks of experiments; see Figure 10.10.

Consider now the parameter α, which characterises maximum functional activity of the host immune response and also reflects the nutritional status of the host. We can estimate it from the asymptotic steady state value $M(\infty) = M_\infty$ of the worm burden. Let us consider the highest infection rate λ_4, then from (10.58) in the limit $t \to \infty$, we have

$$E = \frac{\lambda_4 T}{\mu D},$$

which on substituting into (10.56) gives

$$I \approx \alpha \quad \text{for} \quad T \gg \frac{\mu D\sqrt{\beta}}{\lambda_4}. \tag{10.59}$$

The experimentally observed saturation, as described in (B3), ensures the validity of this assumption on T. We can use (10.59) with (10.55) at the steady state to determine α, to get

$$\alpha M_\infty = \mu L(\infty) - \delta M_\infty \quad \Rightarrow \quad \alpha = \frac{\lambda_4}{D M_\infty} - \delta. \tag{10.60}$$

Since, within the experimental observation time, the system does not reach its final steady state, we use (10.60) to predict M_∞ as a *function of* α, namely,

$$M_\infty = \frac{\lambda_4}{D(\alpha + \delta)}. \tag{10.61}$$

Finally we use the experimental data given in Figures 10.10(b)(d), which correspond to the highest rate of infection, to determine the sensitivity of the immune system as measured by β. Note there that the mean adult worm burden rises to a maximum value M^* at a time t^* and then declines under the influence of host immunity, despite continual reinfection, to settle at the asymptotic steady state value M_∞. For the maximum point (M^*, t^*), $M^* = M(t^*)$, equations (10.55) and (10.56) give

$$0 = \mu L - \delta M^* - \frac{\alpha E^2 M^*}{\beta + E^2}. \tag{10.62}$$

Since, in the laboratory situation, t^* satisfies

$$\frac{1}{\mu D} \ll t^* < T, \tag{10.63}$$

we use the first of (10.58) to get E and then solve (10.62) for β to get

$$\beta = \frac{E^2(\alpha M^* - \mu L + \delta M^*)}{\mu L - \delta M^*}, \tag{10.64}$$

where $L(t)$ and $E(t)$ are their values at $t = t^*$; as before $M^* = M(t^*)$. We estimate the values for $M^*(\approx 50\,\text{worms})$ and $t^*(\approx 7\,\text{weeks})$ from the experimental data in Figures 10.10(b)(d) and in turn use (10.64) in subsequent calculations to determine the sensitivity β, which is measured in worm^{2}day^2, as a function of α. So, we have used the experimental data and the fact that the laboratory situation is in the regime $t < T$ to determine the respective parameters α and β by using (10.60) and (10.64).

We can now analyze the complete nonlinear immigration–death model

$$\begin{aligned} \frac{dL}{dt} &= \lambda_i - \mu DL, \quad i = 1, 2, 3, 4, \\ \frac{dM}{dt} &= \mu L - (\delta + I)M, \end{aligned} \tag{10.65}$$

where the acquired immune response function I is given in terms of E and L by (10.56). Numerical integration of (10.65) gives the solution for the mean adult worm burden as a function of time; the results are plotted in Figure 10.14 for the first 12 weeks. The different curves again represent the different infection rates.

Here we have chosen α equal to $0.5\,\text{day}^{-1}$, which implies $\beta \approx 6.1 \times 10^6$ worms^2-day^2 (the units are dictated by the form of the immune function I in (10.56)). With these, the solutions give a very satisfactory fit to the experimental data in Figure 10.10. Since α is related to M_∞ by (10.61), we thus predict that a continuation of the present experimental setting eventually leads to an asymptotic steady state of $M_\infty = 4$ worms. If we restrict ourselves to the same genetic type of hosts and the same dietary conditions, the model can also be used to investigate more realistic situations such as when the hosts are subjected to natural infection. We briefly discuss this below.

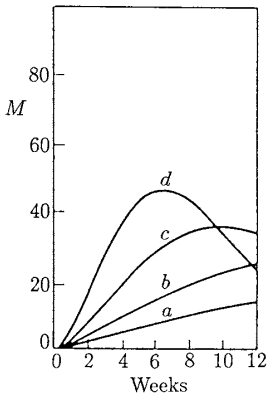


Figure 10.14. The time evolution in the mean worm burden, $M(t)$, in mice hosts fed on a high protein diet for a 12-week period of repeated infection, obtained from a numerical integration of equations (10.65), which govern the population dynamics in the presence of host immune response. These curves are the ones used to compare with the experimental data in Figure 10.10(b). Parameter values: $\mu = 0.125\,\text{day}^{-1}$, $D = 1.56$, and $\delta = 5.6 \times 10^{-3}\,\text{day}^{-1}$ as in Figure 10.13, and $\alpha = 0.5$ (which imply $\beta = 0.1 \times 10^6$) for the maximum functional activity of the immune system.

For even more general applications of (10.65), such as to arbitrary nutritional conditions or different strains of mice, further experiments are necessary to clarify: (i) the detailed functional dependence of the maximum functional activity, α, on the nutritional status of the hosts, (ii) the specific relationship of the sensitivity, β, to various strains of mice and (iii) the size of the memory time, T. With these the system (10.75) can be used to predict the time-evolution and the final steady state of the mean worm burden dependence on the nutritional status and the genetic properties of the hosts being considered.

Among the goals of any mathematical modelling in epidemiology are: (i) to provide a proper mechanistic description of the field situation and (ii) to provide a sound basis for making practical predictions. Usually, however, a major difficulty is the practical estimation of the many parameters which are involved in the models. Controlled laboratory experiments, which study particular aspects of the complete dynamics, while keeping all other parts of the system under experimental control, have proved very useful in this respect. The experiments described here have specifically highlighted the role of the immune response. As a result we have been able to develop and exploit a simple but realistic mathematical model, which admits a full *quantitative* description of the population dynamics in the presence of host immune response.

At this point a few cautionary remarks should be made. First, the model as it stands does not, nor was it intended to, give a full picture of the underlying delicate biochemical and biocellular processes. It does, however, provide a quantitative picture of the macroscopic features of immune response: the per capita rate of limitation in parasite survival can be related quantitatively to the antigenic stimulus (that is, the exposure to infection). Second, the choice of the input function E for the immune system in (10.50) and in particular (10.51) is, of course, not unique; it seems, however, a very plausible one in view of the biological observations listed. In fact the qualitative features of the experimental data are reproduced even with a linear function $I(E)$ in place of the immune activity function in (10.56). However, numerical simulations show that this latter model assumption gives a more satisfactory, *simultaneous* fit of the four graphs corresponding to the four different infection rates, (Figures 10.10(b) (a) to (d)), than a linear version of (10.51). In summary then, the model is supported by the following facts: (i) it is in keeping with the biological observations, (ii) it provides a quantitative fit for the experimental data used to test it and (iii) the parameters introduced are biologically meaningful and can be estimated.

The importance of an acquired immune response in human infection with several species of helminth parasites has also been shown, for example, in the immunological and epidemiological studies of Butterworth et al. (1985). They describe the immune response of 'resistant' and 'susceptible' Kenyan school-children to infection with the blood fluke *Schistosoma mansoni*. The role of human immunity in controlling other worm infections is similarly well established. There is an urgent need for fieldwork studies: basic mathematical models of the type described and used here can be of enormous help in their design and interpretation. In addition, extension of the modelling technique to the 'real world' can provide a cheap and effective way of testing the efficiency of various parasite control programmes, without resort to lengthy and expensive field trials. Further modelling on the lines described in this section have been carried out by Berding et al. (1987) for further laboratory studies in which there is a genetically

heterogeneous host population and in which there is natural transmission of the parasite. As before the mice populations had different protein diets. They also discuss the significance of the results from a real world medical viewpoint.

A direct practical (and commercial) application of the concepts and modelling techniques in this section was given by Parry et al. (1992). They applied it to coccidial infection in chickens with emphasis on vaccinating the chickens by delivering oocysts (early stage coccidia) in their feed: this induced an immune response at a much lower level of parasite burden.

What is already abundantly clear is that in real world practical terms, the nutritional status of the host is an important factor in the population dynamics of parasite infections, and must not be ignored in the design of optimal health control policies.

10.9 Age-Dependent Epidemic Model and Threshold Criterion

In many diseases the chronological age of the individual is an important factor in assessing their vulnerability and infectiousness. For example, the interesting data quoted by Bernoulli (1760) on the incidence and severity of smallpox with age is a vivid illustration: vulnerability and mortality go down markedly with age. A variety of age-dependent models was discussed, for example, in the book by Hoppensteadt (1975). Dietz (1982), for instance, proposed such a model for river blindness (onchocersiasis) and used it to compare various possible control strategies.

Age may also be interpreted as the time from entry into a particular population class such as the susceptibles, infectives or the removed group in a basic *SIR* model. The two interpretations of age are often the same. With the specific case we analyse in the following section, on a drug use epidemic model, age within a class, the users, is the relevant interpretation. Another more relevant and practical example involving bovine tuberculosis is discussed in detail in Section 10.11.

Consider the population we are interested in can reasonably be divided into susceptibles, $S(t)$, and infectives, $I(a,t)$, where a is the age from exposure to the disease so we are considering an SI age-dependent model. The number of susceptibles decreases through exposure to the disease. The removal rate of susceptibles is taken to be

$$\frac{dS}{dt} = -\left[\int_0^\tau r(a')I(a',t)da'\right]S, \quad S(0) = S_0. \tag{10.66}$$

That is, the removal due to infectives is weighted with an age-dependent function $r(a)$ which is a measure of the infectiousness of the infectives. Since the infective is only infectious for a limited time, τ, this is the upper limit in the integral.

To get the equation for the infective population $I(a,t)$ we use a conservation approach. In a time Δ there is an advance in chronological age and in infective class age from (t,a) to $(t+\Delta, a+\Delta)$. Conservation then says that the change in the number of infectives in a time Δ must be balanced by the number removed. We thus have, in time Δ,

$$I(a+\Delta, t+\Delta) - I(a,t) = -\lambda(a)I(a,t)\,\Delta,$$

where $\lambda(a)$ is the age-dependent removal factor. In the limit as $\Delta \to 0$ we then get, on expanding in a Taylor series, the partial differential equation

$$\frac{\partial I}{\partial t} + \frac{\partial I}{\partial a} = -\lambda(a)I. \tag{10.67}$$

At time $t = 0$ there is some given age-distributed class of infectives $I_0(a)$. At $a = 0$ there is recruitment from the susceptible class into the infectives. Since all new infectives come from the susceptibles, the 'birth rate' $I(0,t)$ is equal to $-dS/dt$. Thus the boundary conditions for (10.67) are

$$I(a,0) = I_0(a), \quad I(0,t) = -\frac{dS}{dt}, \quad t > 0. \tag{10.68}$$

The integrodifferential equation model now consists of (10.66)–(10.68), where $I_0(a)$ and S_0 are given. We assume the functions $r(a)$ and $\lambda(a)$ are known, at least qualitatively, for the disease and in control procedures can be manipulated as is often the case.

An infection will not spread if the number of susceptibles expected to be infected by each infective drops below one. If the number exceeds one then the infection will spread and we have an epidemic. The number γ of initial susceptibles expected to be infected by each infective is

$$\gamma = S_0 \int_0^\tau r(a) \exp\left[-\int_0^a \lambda(a')\,da'\right] da. \tag{10.69}$$

As in (10.66), $r(a)$ here is the infective capability of an infective. It is weighted with an exponential function which is the probability of an initial infective surviving to age a: $\lambda(a)$ is the same as in (10.67). The threshold value for an epidemic is $\gamma = 1$ above which the infection spreads. We now show how the severity of the epidemic, as measured by the ratio $S(\infty)/S_0$, depends on γ. Clearly from (10.66) since $dS/dt \le 0$, $S(t) \to S(\infty)$, where $0 \le S(\infty) \le S_0$.

We solve the mathematical problem (10.66)–(10.68) using the method of characteristics. (A similar procedure was used in Chapter 1, in the single population growth model with age distribution.) The characteristics of (10.67) are the straight lines.

$$\begin{aligned} \frac{dt}{da} = 1 \quad \Rightarrow \quad a &= t + a_0, \quad a > t \\ &= t - t_0, \quad a < t, \end{aligned} \tag{10.70}$$

where a_0 and t_0 are respectively the age of an individual at time $t = 0$ in the given original population and the time of birth of an infective; see Figure 10.15.

The characteristic form of (10.67) is

$$\frac{dI}{da} = -\lambda(a)I \quad \text{on} \quad \frac{dt}{da} = 1,$$

and so, with Figure 10.15 in mind, integrating these equations we get

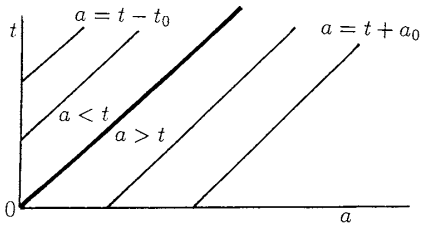


Figure 10.15. Characteristics for the infectives equation (10.67). On $t = 0$, $I(a, 0) = I_0(a)$, which is given, and on $a = 0$, $I(0, t) = -dS/dt$, $t > 0$.

$$\begin{aligned} I(a,t) &= I_0(a_0)\exp\left[-\int_{a_0}^{a}\lambda(a')\,da'\right], \quad a > t \\ &= I(0,a_0)\exp\left[-\int_0^a \lambda(a')\,da'\right], \quad a < t. \end{aligned}$$

Thus, from (10.70),

$$\begin{aligned} I(a,t) &= I_0(a-t)\exp\left[-\int_{a-t}^{a}\lambda(a')\,da'\right], \quad a > t \\ &= I(0,a-t)\exp\left[-\int_0^a \lambda(a')\,da'\right], \quad a < t. \end{aligned} \tag{10.71}$$

From (10.66) the solution S(t) is

$$S(t) = S_0\exp\left[-\int_0^t\left\{\int_0^\tau r(a)I(a,t')da\right\}dt'\right]. \tag{10.72}$$

Using (10.71) for $I(a,t)$, in the ranges $a < t$ and $a > t$,

$$\begin{aligned} \int_0^\tau r(a)I(a,t')da &= \int_0^t r(a)I(0,t'-a)\exp\left[-\int_0^a\lambda(a')\,da'\right]da \\ &\quad + \int_t^\tau r(a)I_0(a-t')\exp\left[-\int_{a-t'}^{a}\lambda(a')\,da'\right]da. \end{aligned} \tag{10.73}$$

Since the time of infectiousness is τ, the last integral vanishes if $t > \tau$; we can think of it in terms of $r(a) = 0$ if $a > \tau$. For $S(t)$ in (10.72) we have, using (10.68) and (10.73),

$$\begin{aligned} &\int_0^t\int_0^\tau r(a)I(a,t')\,da\,dt' \\ &\quad = -\int_0^t\int_0^{t'} r(a)\exp\left[-\int_0^a\lambda(a')\,da'\right]\frac{dS(t'-a)}{dt'}\,da\,dt' \\ &\quad\quad + \int_0^t\int_{t'}^\tau r(a)I_0(a-t')\exp\left[-\int_{a-t'}^{a}\lambda(a')da'\right]da\,dt'. \end{aligned}$$

Interchanging the order of integration in the first integral on the right-hand side we get

$$\begin{aligned} &\int_0^t\int_0^\tau r(a)I(a,t')da\,dt' \\ &\quad = -\int_0^t r(a)\exp\left[-\int_0^a\lambda(a')da'\right](S(t-a)-S_0)\,dt + m(t), \end{aligned} \tag{10.74}$$

where

$$m(t) = \int_0^t\int_{t'}^\tau r(a)I_0(a-t')\exp\left[-\int_{a-t'}^{a}\lambda(a')da'\right]da\,dt'. \tag{10.75}$$

Substituting (10.74) into (10.72) we then get

$$S(t) = S_0\exp\left\{-m(t) + \int_0^t r(a)\exp\left[-\int_0^a\lambda(a')da'\right](S(t-a)-S_0)da\right\}. \tag{10.76}$$

If we now let $t \to \infty$, remembering that $r(a) = 0$ for $a > \tau$, we get, using γ defined in (10.69),

$$F = e^{-m(\infty)+\gamma(F-1)}, \quad F = \frac{S(\infty)}{S_0}. \tag{10.77}$$

We are interested in the severity of the epidemic as measured by F, that is, the fraction of the susceptible population that survives the epidemic, and how it varies with γ. For given $r(a)$, $I_0(a)$ and $\lambda(a)$, (10.75) gives $m(t)$ and hence $m(\infty)$. If $0 < m(\infty) = \varepsilon \ll 1$, Figure 10.16 shows how F varies with γ. For each value of γ there are two roots for F but, since $S(\infty) \le S_0$, only the root $F = S(\infty)/S_0 \le 1$ is relevant. Note how the severity of the epidemic is small for ε small as long as $\gamma < 1$ but it increases dramatically; that is, $S(\infty)/S_0$ decreases (from $S(\infty)/S_0 \approx 1$) quickly for $\gamma > 1$. For example, if $0 < \varepsilon \ll 1$ and $\gamma \approx 1.85$, $S(\infty)/S_0 \approx 0.25$.

Suppose a single infective is introduced into a susceptible population of size S_0. We can approximate this by writing $I_0(a) = \delta(a)$, the Dirac delta function, then

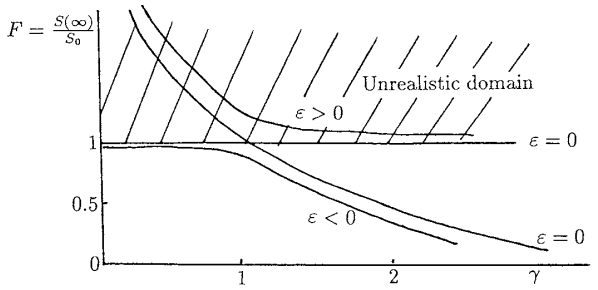


Figure 10.16. Dependence of the epidemic severity $F = S(\infty)/S_0$, that is, the fraction of the susceptible population who survive the epidemic, on the threshold parameter γ from (10.77), namely, $F = \exp[-\varepsilon + \gamma(F-1)]$. The only realistic values, of course, are $F = S(\infty)/S_0 \le 1$.

$$\int_0^\tau I_0(a)\,da = 1.$$

In this case, from (10.75),

$$\begin{aligned} m(t) &= \int_0^t \int_{t'}^\tau r(a) I_0(a-t') \exp\left[-\int_{a-t'}^a \lambda(a')da'\right] da\,dt' \\ &= \int_0^t \int_{t'}^\tau r(a)\delta(a-t') \exp\left[-\int_{a-t'}^a \lambda(a')da'\right] da\,dt' \\ &= \int_0^\infty r(t') \exp\left[-\int_0^{t'} \lambda(a')da'\right] dt' \\ &= \frac{\gamma}{S_0} \end{aligned}$$

from (10.69). Thus (10.77) becomes

$$F = \exp\left[\gamma\left(F - 1 - \frac{1}{S_0}\right)\right], \quad F = \frac{S(\infty)}{S_0}. \tag{10.78}$$

Since $1/S_0 \ll 1$ in general the solutions for F in terms of γ are typically as given in Figure 10.16. Thus $\gamma > 1$ need *not* be large for a severe epidemic to occur. Therefore, it is the estimation of the parameter γ in (10.69) that is critical in the epidemiology of age-dependent models. This we do in the following section for a very simple and primitive model of drug use.

10.10 Simple Drug Use Epidemic Model and Threshold Analysis

The spread of the use of self-administered drugs, therapeutic and illicit, is in some cases a result of the enthusiastic proselytising by a user in the initial stages of use. We describe here a simple illustrative model discussed by Hoppensteadt and Murray (1981) for the etiology of such a drug and show how to determine the threshold parameter γ. This entails the evaluation of the infectiousness which we relate to the response of the user to the drug. The novel feature of the epidemic model studied here is the inclusion of the user's personal response to the drug. The model is a pedagogical one: we do not have a specific drug in mind.

Suppose the drug is introduced into the blood stream in dosages $d(t)$ and let it be removed at a rate proportional to $c(t)$, the drug concentration in the blood; that is, a first-order kinetics removal. The governing equation for the blood concentration $c(t)$ is then

$$\frac{dc}{dt} = d(t) - kc, \quad c(0) = 0, \tag{10.79}$$

where $k > 0$ is constant and $t = 0$ is the time the individual is first recruited as a user. In drug abuse, the dosage $d(t)$ tends to be oscillatory or approximately periodic with a

progressively decreasing period. The solution of (10.79) is

$$c(t) = e^{-kt} \int_0^t e^{kt'} d(t')\,dt'. \tag{10.80}$$

For many drugs the body has specific sites and it is the binding of these sites which evokes a response in the user. Denote the number of free sites, that is, active or unbound, by $A(t)$, the number of bound, that is, inactive, sites by $B(t)$ and the total number by N. We assume that no new sites are being created so $A(t) + B(t) = N$. We take as a site binding model the very simple system

$$\begin{aligned} \varepsilon \frac{dA}{dt} &= \alpha B - \beta c A, \quad A(0) = N, \\ \varepsilon \frac{dB}{dt} &= \beta c A - \alpha B, \quad B(0) = 0, \end{aligned} \tag{10.81}$$

where α, β and ε are positive constants: the inclusion of ε here is for later algebraic convenience when we take it to be small. We are thus assuming that the rate of binding of active sites is proportional to the amount of the drug $c(t)$ in the body and the number of active sites available: that is, $\beta c A/\varepsilon$. There is also a replenishment of the active sites proportional to the number of bound sites: that is, $\alpha B/\varepsilon$. With $A + B = N$ the equation for B is then given by the second of (10.81).

Suppose now that the reaction, $r(t)$, to the drug is proportional to the blood concentration and the number of free sites. We thus take it to be

$$r(t) = Rc(t)A(t), \tag{10.82}$$

where $R > 0$ is a measure of the individual's response to the drug.

If the rate of binding is very fast, that is, α and β are $O(1)$ and $0 < \varepsilon \ll 1$ in (10.81), the number of free and bound receptors reaches equilibrium very quickly. Then, using $A + B = N$,

$$B = \frac{\beta c A}{\alpha} \quad \Rightarrow \quad A = \frac{\alpha N}{\alpha + \beta c}, \quad B = \frac{\beta N c}{\alpha + \beta c}, \tag{10.83}$$

and the individual's response is

$$r = \frac{R\alpha N c}{\alpha + \beta c}, \tag{10.84}$$

which is a Michaelis–Menten (cf. Chapter 6, Section 6.2) type of response which saturates to $r_{\max} = R\alpha N/\beta$ for large blood concentration levels c. Note that with B as in (10.83) the response $r = R\alpha B/\beta$; that is, the response is proportional to the number of bound sites.

If ε in (10.81) is $O(1)$ we can incorporate it into the α and β; this is equivalent to setting $\varepsilon = 1$. Now with $B = N - A$ the equation for $A(t)$ from (10.81), with $\varepsilon = 1$, is

$$\frac{dA}{dt} = \alpha N - A(\alpha + \beta c), \quad A(0) = N$$

which has solution

$$\begin{aligned} A(t) = N \exp\left[-\int_0^t \{\alpha + \beta c(t')\}\, dt'\right] \\ + \alpha N \int_0^t \exp\left[-\int_{t'}^t \{\alpha + \beta c(\tau)\} d\tau\right] dt', \end{aligned} \tag{10.85}$$

with $c(t)$ from (10.80).

If $d(t)$ is known we can carry out the integrations explicitly to get $c(t)$ and $A(t)$: it is algebraically rather complicated for even a simple periodic $d(t)$. Since the algebraic details in such a case initially tend to obscure the key elements we consider here the special case $d(t) = d$, a constant, and assume that the recovery rate of active sites from their bound state is very small: that is, $\alpha \approx 0$. Then, from (10.80) giving $c(t)$ and the last equation giving $A(t)$, we have

$$\begin{aligned} c(t) &= \frac{d}{k}(1 - e^{-kt}), \\ A(t) &= N \exp\left[-\frac{\beta d}{k}\left\{t + \frac{1}{k}(e^{-kt} - 1)\right\}\right] \end{aligned} \tag{10.86}$$

and the response $r(t)$ from (10.82) is

$$r(t) = RcA = \frac{RNd}{k}(1 - e^{-kt}) \exp\left[-\frac{\beta d}{k}\left\{t + \frac{1}{k}(e^{-kt} - 1)\right\}\right]. \tag{10.87}$$

Figure 10.17 illustrates the form of $c(t)$ and $r(t)$ from (10.86) and (10.87).

It is interesting to note that even with this very simple illustrative model, the response of an individual does not just increase with dosage: after an initial stage of increasing response it actually decreases with time.

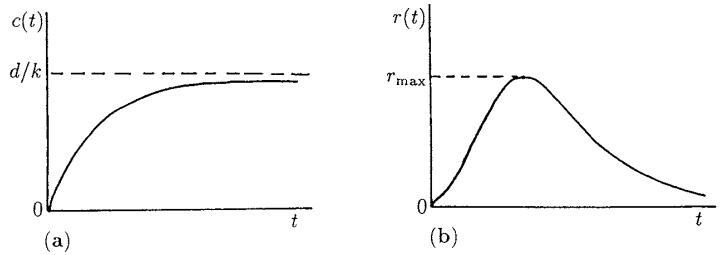


Figure 10.17. (**a**) The blood concentration $c(t)$ of the drug: from (10.86) it saturates to d/k after a long time. (**b**) The body's response to the drug from (10.87). Note the initial increase before it tails off with continuous drug use.

Now consider the possibility of an epidemic of drug use appearing in a population S_0 of nonusers after the introduction of a single user. We assume $1/S_0 \ll 1$, as is reasonable, and so $F = S(\infty)/S_0$ is given by the solution $F < 1$ in Figure 10.16 for the appropriate γ, which we now evaluate.

Here age is measured from the first time of using the drug. There is no time limit for infectiousness so in the definition (10.69) for γ we set $\tau = \infty$. From Figure 10.17(b) the response $r(t) \to 0$ as $t \to \infty$; that is, the infectiousness, or proselytising fervour, becomes less effective with time. For simplicity we assume the probability factor in (10.69) has λ constant and so

$$\gamma = S_0 \int_0^\infty r(t) e^{-\lambda t}\, dt. \tag{10.88}$$

We can now evaluate γ for various limiting situations in terms of the parameters α, β, γ and k in the user model (10.79)–(10.82).

In the case $d(t) = d$, a constant, we get Table 10.1 after some elementary algebra. It gives the user's response $r(t)$ and the corresponding epidemiological threshold parameter γ. For example, in the case $0 < \varepsilon \ll 1$, (10.84) holds if $\alpha \ll \beta$, $r(t) \approx RN\alpha/\beta$, a constant, and (10.88) gives $\gamma \approx S_0 RN\alpha/(\lambda\beta)$. On the other hand if $0 < \varepsilon \ll 1$ and $\beta \gg \alpha$ then, from (10.83), $r(t) = RNc(t)$ and, with $c(t)$ from (10.86), γ is given, from (10.88), by

$$\gamma = S_0 \int_0^\infty \frac{RN\, d(1 - e^{-kt})}{k} e^{-\lambda t}\, dt = \frac{S_0 RN\, d}{\lambda(\lambda + k)}.$$

A similar type of asymptotic approach results in the other forms in Table 10.1.

In the case of most self-administered drugs $0 < \varepsilon \ll 1$; that is, the response is very fast. The possibility of an epidemic depends on the relative magnitude of the various parameters in a simple way. This case is covered by (i)–(iv) in Table 10.1. For example if the rate of freeing of bound sites is much slower than the binding rate, $\beta \gg \alpha$ (case (ii)) then, since most sites will be bound, the user's reaction is small. This reduces the user's 'infectiousness' and hence the epidemic risk.

If we increase the cure rate, that is, increase λ, there is a reduction in γ. Decreasing the individual's response, such as by education or chemotherapy, also reduces γ and

Table 10.1. The case $d(t) = d$, a constant. Here $r(t)$ is a measure of the drug user's response and γ is the epidemic infectious (or recruitment) rate. (Table from Hoppensteadt and Murray 1981)

Case		$r(t)$	γ/S_0
(i)	$\varepsilon \ll 1, \alpha \ll \beta$	$RN\alpha/\beta$	$RN\alpha/\lambda\beta$
(ii)	$\varepsilon \ll 1, \beta \gg \alpha$	$RNc(t)$	$RN\,d/[\lambda(\lambda + k)]$
(iii)	$\varepsilon \ll 1, k \gg 1$	$RN\,d/k$	$RN\,d/k\lambda$
(iv)	$\varepsilon \ll 1, k \ll 1,\ \alpha/\beta\, d \ll 1$	$RN\,dt/[1+(\beta\,dt/\alpha)] \sim RN\alpha/\beta$	$RN\alpha/\beta\lambda$
(v)	$\varepsilon = 1, k \ll 1$	$N\,dRt \exp[-2\,d\beta t]$	$RN\,d/(2\,d\beta + \lambda)^2$
(vi)	$\varepsilon = 1, k \gg 1$	$R\,d \exp[-d\beta t/k]/k$	$RN\,d/k\lambda$

hence reduces the possibility of a severe epidemic. The results are in line with a heuristic common sense approach.

If we define the critical population S_c by

$$S_c = \left\{ \int_0^\infty r(a) \exp\left[-\int_0^a \lambda(a')\, da' \right] da \right\}^{-1}, \tag{10.89}$$

then if $S_0 > S_c$, which implies $\gamma > 1$, an epidemic occurs, whereas if $S_0 < S_c$ it does not. The sensitivity of S_c to the parameters can only really be determined if r and γ are known with some confidence.

The type of drug use models we have described and analysed here, but without age dependence, have been very useful in their application to certain aspects of chronic alcohol misuse and even in trying to come up with a better breathalyser. Ethanol metabolism, associated with alcohol eradication in the body, is very different in normal subjects as compared to alcoholics. Smith et al. (1993) used such a model for the study of ethanol metabolism to try to understand the difference between normal users and abusers of alcohol and compared the results and predictions with subject data. An even simpler model, essentially $dc/dt = d(t) - k$ where $c(t)$ is the blood alcohol level, $d(t)$ is the alcohol intake and k is the metabolic decay rate, was used by Lubkin et al. (1996) in a study to try and determine whether it was possible to have a more sophisticated model for alcohol breath exhalation which would make roadside breathalysers more accurate: basically the answer was 'no.'[11]

10.11 Bovine Tuberculosis Infection in Badgers and Cattle

Bovine tuberculosis infection is an insidious disease, which often does not become apparent until it has reached an advanced stage in cattle, badgers and also swine. Investigations carried out suggest that in the southwest of England, for example, badgers constitute a significant reservoir of the bovine Tb, *Mycobacterium bovis* (*M. bovis*) and that badgers, because of their population density, could be a major factor in its spread. Conditions in these affected areas, and as mentioned, the social organisation of badgers, not only favour the transmission of the disease from one infected badger group to another but also from badgers to cattle and vice versa.

Within specific regions in England and Wales, badger habitats are usually intimately intermeshed with intensively used cattle pastures (Neal 1986, MAFF Report 1987; see also 1994). Field studies conducted over a period of about 10 years in such regions confirm that the foraging activities of badgers on cattle pasture with their preferred food items (earthworms, insects and fruits), which are exploited alternatively because they show marked seasonal fluctuations, cause a high frequency of urination and defecation as a direct consequence of their eating habits (MAFF Report 1987). Therefore, diseased badgers tend to contaminate the environment heavily with bacilli, through their feeding habits and suppurating bite wounds, for prolonged periods. Even though a majority of bacilli may be killed early by exposure to direct sunlight, some do survive in the microhabitat for periods of several weeks depending on the prevailing climatic conditions. Studies by MacDonald (1984) indicate that in the wild, the risk of infection depends partly on the viability of the bacilli. In bronchial pus, these survive in appreciable numbers for up to four weeks in winter and one week in summer, in urine for seven days and three days respectively, and in cattle dung for five months and two months respectively. In general, warm, dark, moist locations appear optimal for bacterial survival on the soil surface (MacDonald 1984).

[11] When Washington State Trooper Sgt. Rod Gullberg, a co-author on the paper, first phoned me to see if he could come and talk about the problem, he volunteered to come to the campus. I naively said, 'Yes, of course, but the parking problem on the campus is absolutely horrendous' to which he calmly replied 'I don't think I'll have a problem.' He arrived in his enormous police car and parked it right in front of the main entrance to the building beside what I had always taken to be the equivalent of about 10 solid yellow lines with your car and you being whisked off in a matter of seconds. He then came into the building, in uniform, bristling with all the police accoutrements of baton, gun and so on, and asked, 'Where can I find Professor Murray?' He was followed upstairs with intense curiosity. People felt I must have another very different secret life.

Cattle are most likely to become infected in several ways: they might inhale bacilli during an encounter with badgers with severe pulmonary and kidney lesions or they might graze or sniff at grass contaminated with infectious badger products (sputum, pus from lungs and bite wounds, faeces and urine). Thus a criss-cross infection may arise when cattle come into contact with the bacilli either directly from the environment or indirectly from infectious badgers. Certain farm practices, namely, allowing badgers access to cattle sheds, salt licks and water troughs could also contribute to disease transmission. There is therefore a significant probability for badger-to-cattle and cattle-to-badger disease transmission.

In this section we describe a criss-cross epidemic model for bovine tuberculosis infection between badgers and cattle that Dr. D.E. Bentil and I developed in the mid-1990's and deduce some analytical results. The main objective is to use these results in the following section to study the dynamics of immunization programmes and suggest how certain practical control measures could be adopted with the ultimate aim of minimizing the spread of infection from badgers to cattle and vice versa, should an epidemic occur.

Criss-Cross Model System for Bovine Tb

When dealing with two populations—here badgers and cattle—we require an epidemic system for each population and then couple the systems through infection of susceptible cattle by infected badgers and susceptible badgers via infected cattle. With an *SEIR* model such as discussed in detail by Bentil and Murray (1993) this would result in a model with 8 coupled partial differential equations if we include age structure as we should. In principle models should be developed from the simple to the complex. Here we have to choose between considering only time-dependent populations, without age structure, or consider fewer subpopulations and include age structure. Here we adopt the latter strategy and consider two subpopulations in each of the badgers and the cattle, that is, an *SI*-type age-structured criss-cross epidemic model to study the disease transmission dynamics between them. So, we consider a model involving two distinct populations (badgers and cattle) and an infection which is communicated between them. We investigate a simple, age-structured, criss-cross model which describes the rate at which cub and adult badgers and cattle go through two different—susceptible

and infectious—states. Here, one of the basic assumptions is that badgers endure a prolonged illness once infected: for example, 12 naturally infected badgers held in captivity survived for between 165 and 1305 days (MacDonald 1984). It is during this prolonged illness that it is assumed they contaminate cattle pasture with bacilli. The mortality due to *M. bovis* infection in both badgers and cattle is low (Cheeseman et al. 1988) so it is not unreasonable to assume that disease-induced death is negligible as compared with normal death. We also assume constant death rates for both badgers and cattle. Other forms of death rates could be used but at this stage add unnecessarily to the complexity of the analysis. It is useful and important to get in the first instance some general guidelines. The contraction of *M. bovis* infection does not confer immunity so we assume that infected badgers either die or recover temporarily and become susceptible again. We assume a similar disease transmission dynamics for cattle. The flow diagram of the disease transmission dynamics in terms of the two distinct interacting populations, namely, badgers and cattle, is schematically shown in Figure 10.18.

We take the total number of cub and adult badgers and cattle at risk of infection to be constant and equal to N and $\tilde{N}$ respectively. We have also assumed that infected cattle recover at a rate $\tilde{r}$ which is proportional to $\tilde{W}$, the infected cattle, and infected badgers recover at a rate r proportional to W, the infected badger population. Cattle appear to develop symptoms much more readily so we assume $\tilde{r} \gg r$. Cattle are newly infected at rates $\tilde{\beta}_1$, $\tilde{\beta}_2$ which are proportional to the product of the number of susceptible cattle, $\tilde{U}$, and the sum of infectious cattle, $\tilde{W}$, and badgers, W. Similarly, newly infected badgers occur at rates β_1, β_2 which are proportional to the product of the number of susceptible badgers, U, and the sum of infected cattle and badgers, namely, $\tilde{W}$ and W. The parameters β_1, β_2 and $\tilde{\beta}_1$, $\tilde{\beta}_2$ are the disease transmission coefficients for badgers and cattle respectively.

Figure 10.18 is certainly basic and contains many simplifying assumptions. With these caveats we write the model system as

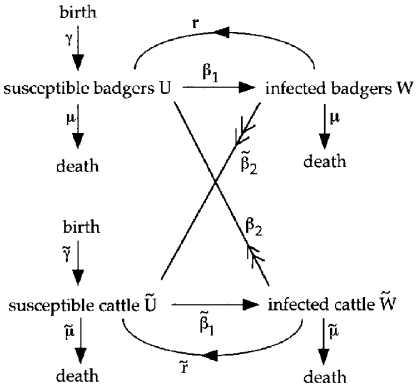


Figure 10.18. Diagrammatic flow chart of a criss-cross model for an infection between badgers and cattle. Each class is a disease host for the other. Here we have divided the badger population into susceptibles, U, and infectious, W. The cattle population is treated similarly with the susceptible and infectious cattle population denoted by $\tilde{U}$ and $\tilde{W}$. The contraction of *M. bovis* infection does not confer immunity and so an infected animal becomes susceptible again after recovery.

$$\begin{aligned}
\frac{\partial U}{\partial t} + \frac{\partial U}{\partial a} &= -\lambda_1 U + rW - \mu U,\\
\frac{\partial W}{\partial t} + \frac{\partial W}{\partial a} &= \lambda_1 U - rW - \mu W,\\
\frac{\partial \tilde{U}}{\partial t} + \frac{\partial \tilde{U}}{\partial a} &= -\tilde{\lambda}_1 \tilde{U} + \tilde{r}\tilde{W} - \tilde{\mu}\tilde{U},\\
\frac{\partial \tilde{W}}{\partial t} + \frac{\partial \tilde{W}}{\partial a} &= \tilde{\lambda}_1 \tilde{U} - \tilde{r}\tilde{W} - \tilde{\mu}\tilde{W},
\end{aligned} \tag{10.90}$$

where the force of infection for the respective populations is given by

$$(B) \qquad \lambda_1(t) = \beta_1 \int_0^\infty W(t,a)\,da + \beta_2 \int_0^\infty \tilde{W}(t,a)\,da, \tag{10.91}$$

$$(C) \qquad \tilde{\lambda}_1(t) = \tilde{\beta}_1 \int_0^\infty \tilde{W}(t,a)\,da + \tilde{\beta}_2 \int_0^\infty W(t,a)\,da \tag{10.92}$$

which are partial contributions from both badgers (B) and cattle (C). The initial age distribution of the respective classes at $t = 0$ is given by

$$\begin{aligned}
U(0,a) &= U_0(a), \quad \tilde{U}(0,a) = \tilde{U}_0(a),\\
W(0,a) &= W_0(a), \ \tilde{W}(0,a) = \tilde{W}_0(a),
\end{aligned} \tag{10.93}$$

and the renewal (boundary) conditions

$$\begin{aligned}
N(t,0) &= \gamma \int_0^\infty N(t,a)\,da = \gamma N(t), \ t > 0,\\
\tilde{N}(t,0) &= \gamma \int_0^\infty \tilde{N}(t,a)\,da = \tilde{\gamma} N(t), \ t > 0.
\end{aligned} \tag{10.94}$$

The absence of a birth term in the model equations is because the only input into the host population is into the class of age zero and so appears as a boundary condition. If we hold to the assumption that all newborn badgers and cattle are susceptible in constant populations where the birth rates, γ, $\tilde{\gamma}$ are set equal to the death rates, μ, $\tilde{\mu}$, for badgers and cattle respectively, then the boundary conditions in (10.94) for the various groups are

$$\begin{aligned}
U(t,0) &= N(t,0) = \gamma N(t), \ W(t,0) = 0,\\
\tilde{U}(t,0) &= \tilde{N}(t,0) = \tilde{\gamma}\tilde{N}(t), \ \tilde{W}(t,0) = 0.
\end{aligned} \tag{10.95}$$

Here, for example, $\gamma N(t)$ is the number of births of badgers at age 0 for all t, and $N(t)$ is the total badger population. At any time t, the age distribution of both badgers and cattle can be expressed as

$$\begin{aligned}
N(t,a) &= \gamma N(t) \exp\left(-\int_0^a \mu(s)\,ds\right) = \gamma N(t) m(a),\\
\tilde{N}(t,a) &= \tilde{\gamma}\tilde{N}(t) \exp\left(-\int_0^a \tilde{\mu}(s)\,ds\right) = \tilde{\gamma}\tilde{N}(t)\tilde{m}(a),
\end{aligned} \tag{10.96}$$

which define the survival probability $m(a)$ and $\tilde{m}(a)$ functions. For example, $m(a)$ is the probability that a badger will live to age a.

We now rescale the problem to make the system nondimensional. This introduces dimensionless groupings which highlight certain ecological facts. We first factor out the death rate in the model system (10.90) by making the substitutions

$$\begin{aligned} u(t,a) &= \frac{U(t,a)}{N(t,a)},\ w(t,a) = \frac{W(t,a)}{N(t,a)}, \\ \tilde{u}(t,a) &= \frac{\tilde{U}(t,a)}{\tilde{N}(t,a)},\ \tilde{w}(t,a) = \frac{\tilde{W}(t,a)}{\tilde{N}(t,a)}. \end{aligned} \tag{10.97}$$

If we choose reference scales for $u(t,a)$, $w(t,a)$, $\tilde{u}(t,a)$, $\tilde{w}(t,a)$, a and t and scale these variables by the maximum values they can realistically obtain (the maximum value for $u(a)$ occurs at $u(0)$) and we scale the time and chronological age by setting $r = rt$, $\alpha = ra$ we obtain the nondimensional system

$$\begin{aligned} \frac{\partial u}{\partial \tau} + \frac{\partial u}{\partial \alpha} &= -\frac{1}{r}\lambda_1 u + w, \\ \frac{\partial w}{\partial \tau} + \frac{\partial w}{\partial \alpha} &= \frac{1}{r}\lambda_1 u - w, \\ \frac{\partial \tilde{u}}{\partial \tau} + \frac{\partial \tilde{u}}{\partial \alpha} &= -\frac{1}{\tilde{r}}\tilde{\lambda}_1 \tilde{u} + \tilde{w}, \\ \frac{\partial \tilde{w}}{\partial \tau} + \frac{\partial \tilde{w}}{\partial \alpha} &= \frac{1}{\tilde{r}}\tilde{\lambda}_1 \tilde{u} - \tilde{w}. \end{aligned} \tag{10.98}$$

The boundary conditions become

$$u(\tau,0) = 1,\ \ w(\tau,0) = 0,\ \ \tilde{u}(\tau,0) = 1,\ \ \tilde{w}(\tau,0) = 0, \tag{10.99}$$

and initial conditions are given by

$$\begin{aligned} u(0,\alpha) &= u_0(\alpha),\ w(0,\alpha) = w_0(\alpha), \\ \tilde{u}(0,\alpha) &= \tilde{u}_0(\alpha),\ \tilde{w}(0,\alpha) = \tilde{w}_0(\alpha). \end{aligned} \tag{10.100}$$

The force of infection for the respective populations is given by

$$\begin{aligned} (B)\quad \lambda_1(\tau) &= \frac{\beta_1}{r}\int_0^\infty w(\tau,\alpha)N(\tau,\alpha)\,d\alpha + \frac{\beta_2}{\tilde{r}}\int_0^\infty \tilde{w}(r,\alpha)\tilde{N}(\tau,\alpha)\,d\alpha, \\ (C)\quad \tilde{\lambda}_1(\tau) &= \frac{\tilde{\beta}_1}{\tilde{r}}\int_0^\infty \tilde{w}(\tau,\alpha)\tilde{N}(\tau,\alpha)\,d\alpha + \frac{\tilde{\beta}_2}{r}\int_0^\infty w(\tau,\alpha)N(\tau,\alpha)\,d\alpha. \end{aligned} \tag{10.101}$$

The force of infection determines whether or not an epidemic will occur. We saw in the simple models we discussed in earlier sections that there are threshold conditions which must be obtained if the number of infected animals is going to increase. So, the evaluation of the λ's is an essential part of the study of the spread of a disease. In its

simplest form if the force of infection is greater than 1 it means that more than one susceptible will be infected by one infective. In the case of the *SEIR* age-dependent model discussed by Bentil and Murray (1993) the conditions for an epidemic were reduced to determining whether or not a function of λ, obtained from the expression for the force of infection analogous to (10.101) had a solution $\lambda > 1$. With this, threshold values of parameters and populations for an epidemic to ensue were obtained.

The mathematical problem posed by (10.98)–(10.101) is not easy to solve in general. At an equilibrium state, however, we can obtain solutions relatively easily. After a long time we assume an equilibrium is reached, that is, where all $\partial/\partial\tau$ terms are set equal to zero and the various classes are only functions of age a. The λ's in (10.101) are constants since the integrals do not involve τ ($\tau \to \infty$ at equilibrium). The equations in (10.98) are then a set of 4 linear ordinary differential equations uncoupled into two pairs, one for $u(\alpha)$ and $w(\alpha)$ and the other set for $\tilde{u}(\alpha)$ and $\tilde{w}(\alpha)$. The respective fractions of infective and susceptible badgers and cattle at equilibrium are easily derived. For example, with the first two equations in (10.98), on adding and using the boundary conditions $u(0) = 1, w(0) = 0$, we get a linear first-order equation in $u(\alpha)$ which is trivially solved. With these solutions we then have, after some elementary algebra, the equilibrium forces of infection, denoted by λ_2 and $\tilde{\lambda}_2$ (we use the subscript 2 to distinguish them from the time-dependent forces of infection) as

$$\begin{aligned} (B)\qquad \lambda_2 &= \frac{\beta_1\lambda_2\gamma N}{r(\lambda_2+r)}\int_0^\infty m(\alpha)\left[1-\exp\left(-\frac{\lambda_2}{r}-1\right)\alpha\right]d\alpha \\ &\quad + \frac{\beta_2\tilde{\lambda}_2\tilde{\gamma}\tilde{N}}{\tilde{r}(\tilde{\lambda}_2+\tilde{r})}\int_0^\infty \tilde{m}(\alpha)\left[1-\exp\left(-\frac{\tilde{\lambda}_2}{\tilde{r}}-1\right)\alpha\right]d\alpha, \\ (C)\qquad \tilde{\lambda}_2 &= \frac{\tilde{\beta}_1\tilde{\lambda}_2\tilde{\gamma}\tilde{N}}{\tilde{r}(\tilde{\lambda}_2+\tilde{r})}\int_0^\infty \tilde{m}(\alpha)\left[1-\exp\left(-\frac{\tilde{\lambda}_2}{\tilde{r}}-1\right)\alpha\right]d\alpha \\ &\quad + \frac{\tilde{\beta}_2\lambda_2\gamma N}{r(\lambda_2+r)}\int_0^\infty m(\alpha)\left[1-\exp\left(-\frac{\lambda_2}{r}-1\right)\alpha\right]d\alpha, \end{aligned} \tag{10.102}$$

where $m(\alpha)$ and $\tilde{m}(\alpha)$, the survival probabilities, are defined by (10.96). We can go no further with the analysis until we specifiy these functions. If we assume the death rate $\mu(a)$ is a constant, then $m(\alpha) = e^{-\mu a}$ and $\tilde{m}(\alpha) = e^{-\tilde{\mu}a}$ and we can then easily evaluate the integrals in (10.102). We then get coupled transcendental equations to determine the forces of infection in the badgers and the cattle. In general these have to be solved numerically for given parameter values.

By way of illustration let us assume that the contributions from within the respective animal populations are negligible and only a cross-type of infection prevails; that is, $\beta_1 = 0 = \tilde{\beta}_1$ and the death rates are constant. In this situation, after some algebra, we get

$$(B) \qquad \lambda_2 = \frac{\beta_2 \tilde{\lambda}_2 \tilde{\gamma} \tilde{N}}{\tilde{\lambda}_2 + \tilde{r}} \left[\frac{1}{\tilde{\mu}} (1 - e^{-\tilde{\mu} L}) - \frac{1}{\tilde{\lambda}_2 + \tilde{r} + \tilde{\mu}} (1 - e^{-(\tilde{\lambda} + \tilde{r} + \tilde{\mu}) L}) \right],$$

$$(C) \qquad \tilde{\lambda}_2 = \frac{\tilde{\beta}_2 \lambda_2 \gamma N}{\lambda_2 + r} \left[\frac{1}{\mu} (1 - e^{-\mu L}) - \frac{1}{\lambda_2 + r + \mu} (1 - e^{-(\lambda_2 + r + \mu) L}) \right], \tag{10.103}$$

where L is the life expectancy and (B) and (C) refer to badgers and cattle respectively. In both cases as $L \to 0$, $\lambda_2 \to 0$ and $\tilde{\lambda}_2 \to 0$ as they should.

For large L, from (10.103) we have for the badgers and cattle respectively

$$\frac{\lambda_2}{\tilde{\lambda}_2} = \frac{\beta_2 \tilde{\gamma} \tilde{N}}{\tilde{\mu}(\tilde{\lambda}_2 + \tilde{r} + \tilde{\mu})}, \quad \frac{\tilde{\lambda}_2}{\lambda_2} = \frac{\tilde{\beta}_2 \gamma N}{\mu(\lambda + r + \mu)}, \tag{10.104}$$

and the following inverse proportionality relation is obtained

$$1 = \frac{\beta_2 \tilde{\beta}_2 \gamma \tilde{\gamma} N \tilde{N}}{\mu \tilde{\mu} (\lambda_2 + r + \mu)(\tilde{\lambda}_2 + \tilde{r} + \tilde{\mu})} \tag{10.105}$$

or

$$\frac{\beta_2 \gamma N}{\mu(\lambda_2 + r + \mu)} = \left[\frac{\tilde{\beta}_2 \tilde{\gamma} \tilde{N}}{\tilde{\mu}(\tilde{\lambda}_2 + \tilde{r} + \tilde{\mu})} \right]^{-1}. \tag{10.106}$$

These are closely related to the conditions we found for epidemics to exist in the discussion on venereal disease models in Section 10.3. To interpret the results we must now determine parameter estimates.

Parameter Estimation

We know some of the key parameters that influence the demography of badgers and cattle in the absence of *M. bovis* infection; see, for example, Anderson and Trewhella (1985) and Brown et al. (1994). However, it is extremely difficult to get convincing field data for criss-cross disease spread between badgers and cattle: those available are somewhat inconsistent and address specific epidemiological parameters while other relevant parameters are chosen arbitrarily. This lack of adequate information and inconsistent data estimates make it difficult to obtain reliable disease transmission rates from the expressions in (10.102). From a modelling point of view, the choice of parameter values is a crucial factor in determining the level of prevalence of the disease. We therefore used numerical techniques, and particularly the Logical Parameter Search (LPS) Method developed by Bentil and Murray (1993) to generate appropriate parameter values to mimic the observed trend when no field data were available. The LPS method is an online search procedure that scans given parameter ranges and generates parameter sets that satisfy some given logical conditions. To apply it to this criss-cross model, for example, we partly used field data obtained from the literature (see Table 10.2) to set up realistic parameter ranges. The procedure then scanned consecutively the various parameter

Table 10.2. Parameter values used for the basic criss-cross model of bovine tuberculosis between badgers and cattle. LPS estimates were found as described in the text. (From Anderson and Trewhella 1985, BTEC 1987)

Parameter	Symbol	Value	LPS Estimates
Total population (Cattle)	$\tilde{N}$	—	10 cattle km^{-2}
Total population (Badgers)	N	2–5 badgers km^{-2}	3 badgers km^{-2}
Death rate (Cattle)	$\tilde{\mu}$	—	0.25 $year^{-1}$
Death rate (Badgers)	μ	0.25 $year^{-1}$	0.125 $year^{-1}$
Birth rate (Cattle)	$\tilde{\gamma}$	—	0.05 $year^{-1}$
Birth rate (Badgers)	γ	0.125 $year^{-1}$	0.02 $year^{-1}$
Average removal rate (Cattle)	$\tilde{r}_1$	—	2 $year^{-1}$
Avererage removal rate (Badgers)	r_1	2 $year^{-1}$	1 $year^{-1}$
Disease transm. coef. (Cattle–cattle)	$\tilde{\beta}_1$	—	2.0 km^2 $year^{-1}$
Disease transm. coef. (Cattle–badgers)	$\tilde{\beta}_2$	—	1.0 km^2 $year^{-1}$
Disease transm. coef. (Badgers–badgers)	β_1	1.54 km^2 $year^{-1}$	1.54 km^2 $year^{-1}$
Disease transm. coef. (Badgers–cattle)	β_2	—	3.5 km^2 $year^{-1}$
Average life expectancy (Cattle)	$\tilde{L}$	—	10 years
Average life expectancy (Badgers)	L	3.5–5.5 years	10 years

ranges for suitable parameter sets, which satisfied some given logical conditions (for example, criteria for disease prevalence that was obtained from the model analysis). We cross-checked the generated parameter sets with the threshold conditions (that is, the

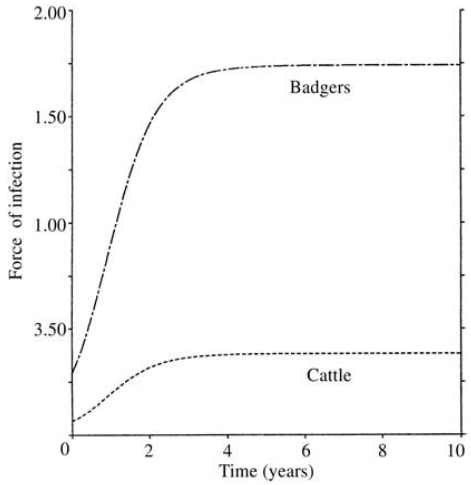


Figure 10.19. Graphical representation of the force of infection corresponding to the disease transmission dynamics for badgers and cattle. Here, a primary assumption is that most of the badger groups sleep in communal huddles in the setts with the environmental conditions that greatly enhance the spread of Tb among them, which in part accounts for an increasingly higher force of infection for badgers; it stabilises after some time.

disease incidence increases after the introduction of an infected group) to make sure that all other requirements, namely, conditions for disease incidence and prevalence had been satisfied. The model equations (10.90) were solved by finite difference schemes with parameter values as in Table 10.2. The initial conditions were set by solving the ordinary differential equations obtained by dropping the time derivatives from which we obtained stable age distributions determined by the age-specific birth and death rates and perturbing the whole system by shifting 10% of susceptible badgers and 5% of susceptible cattle into the infective classes.

Numerical Results and Predicitions

The time-dependent forces of infection in the badger and cattle populations are given by (10.101) which can be evaluated only by solving the full system. This was done with the parameter values given in Table 10.2 and the results are shown in Figure 10.19. As we saw earlier, we could evaluate the integrals and obtain algebraic relations, namely, (10.104)–(10.106), between the two forces of infection for the equilibrium state where life expectancy, L, is long, the death rate a constant (giving an exponential survival probability) and a criss-cross type of infection is the main route by which infection may occur. These imply that the ratio of the force of infection of badgers to cattle is inversely proportional to the ratio of the force of infection of cattle to badgers. The implication here is that if the spread of bovine tuberculosis remains unchecked it may be possible to predict the dynamics of disease spread within badgers for different age groups by studying that for cattle alone (and vice versa).

The model predictions, as illustrated in Figure 10.20 indicate that the number of susceptible badgers and cattle declines while there is a gradual increase in the number of infected badgers and cattle, and much more so within badger populations. This suggests that should a criss-cross type of infection occur the impact of the disease could be felt much more within badger populations. This confirms our assumption that badgers endure a prolonged illness once infected and that it is during this prolonged period of illness that they contaminate cattle pasture with bacilli.

The basic age-structured criss-cross model we have discussed here is based on the assumption of horizontal transmission by bite wounding, aerosol infection, infection contracted through grazing on pastures and so on. Vertical transmission (mother to cub) may be important but we did not take this into account. Broadly speaking, cattle cannot be regarded as a reliable sentinel for the prevalence of infection in badgers everywhere because of the variation in the degree of contact. The proposed models therefore reflect the epidemiology of the disease in areas with good habitats where both species coexist.

As we have mentioned, it is difficult to establish the actual force of infection especially within various badger groups where, for instance, age is determined by weight, size and dental structure as opposed to precise observed trends in cattle. In any event, with the implementation of the LPS method, we were able to make various predictions using the model equations. We speculate that cattle are more or less kept under more hygienic conditions in farms and thus the tendency of high levels of infection is markedly reduced. There is no oscillatory trend in disease incidence between the two distinct groups but, among badgers, some observations indicate a possible cyclic trend in disease incidence (see Cheeseman et al. 1989 and Bentil and Murray 1993). This

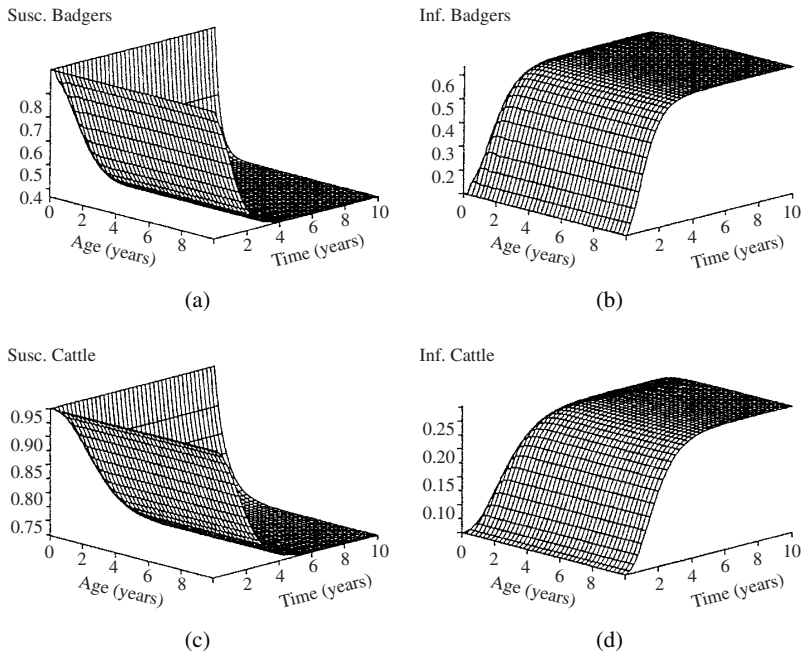


Figure 10.20. Numerical solution of the criss-cross model showing patterns for susceptible and infectious badgers and susceptible and infectious cattle with respect to age and time (horizontal axis) after an initial infection. The vertical axis denotes the corresponding fractions of the various subpopulations. For the chosen parameter values (using LPS estimates in Table 10.2) the number of infectious badgers (b) increases to about 60% while the number of infectious cattle (d) increases to about 30% before stabilizing.

may seem to be the case from our model predictions and makes the study of a possible (hypothetical) criss-cross type of infection all the more relevant.

Results of this study indicate that it is possible to estimate the age-specific equilibrium values of the force of infection knowing which survival functions to use. A constant death rate for badgers and cattle gives, for example, an inverse proportionality relationship which makes it easier to predict the disease transmission dynamics within different age groups. It may be possible to analyse the model behaviour for a step function death rate but the results will be much more difficult to obtain, other than numerically.

A major motivation for the comparative study of an age-structured model for this recurrent disease is the evaluation of control measures for the eradication of the disease as pointed out, for example, by Dietz and Schenzle (1985), Anderson and May (1985) and Murray et al. (1986). The main objective of the above modelling is to use the results to study the dynamics of immunization programmes and suggest how certain control measures could be adopted with the ultimate objective of minimizing the spread of infection from badgers to cattle and vice versa, should an epidemic occur. We discuss this in the following section.

10.12 Modelling Control Strategies for Bovine Tuberculosis in Badgers and Cattle

Bentil and Murray (1993) developed and analysed models for the dynamics of bovine tuberculosis (*Mycobacterium bovis*) infection in the wild badger population. Because of the possibility (I believe, high probability) of badgers being the reservoir for the disease in cattle in the southwest of England in the last section we proposed and analysed a simple criss-cross model. As pointed out, the eradication of the disease when there is a feral infected animal population, such as in the southwest of England and in New Zealand, has not yet been successful in spite of the implementation of an intensive national tuberculosis eradication campaign. The eradication, or rather acceptable control, of bovine tuberculosis by testing and slaughtering programmes has been successful in many countries, as we have said, but total elimination has not been achieved. In the U.S.A., all cattle are systematically tested and those reactors are slaughtered. As a result, reactor rate was reduced from about 5% to 0.03% (USDA Report 1982a,b).

Badgers occupy a variety of habitats, especially woodland areas interspersed with arable and pasture land (Clements et al. 1978, Kruuk 1988). Such habitats are usually intimately intermeshed with intensively used cattle pastures which makes the likelihood for badger-to-cattle and cattle-to badger disease transmission all the more possible. The analysis of the model in the last section for the dynamics of a hypothetical criss-cross infection provides some guidelines concerning the likely impact of the disease between the two distinct populations. Results from that criss-cross model suggest that it may be possible to predict the disease transmission dynamics for one group, namely, cattle, if we know that for badgers and vice versa.

In Britain, for example, programmes for the control of bovine tuberculosis in areas of frequent herd infection have been centred on the reduction of badger density by removal of entire groups of badgers (usually by gassing, which particularly incenses the English) where one or more individuals were thought to be infected. The MAFF control policy (MAFF report 1987) assumed, wrongly as it turned out, that a single intensive intervention to remove all infected groups of animals would suffice to eliminate infection from contaminated areas for long periods of time. As mentioned above the MAFF (1994) control was more selective and probably no more effective. In this section we discuss a new approach that Dr. D.E. Bentil and I developed in the mid-1990's. We model the dynamics of specific immunization programmes and suggest how certain control measures could be adopted with the ultimate objective of minimizing the spread of infection from badgers to cattle and vice versa should an infection occur. We compare several vaccination strategies and deduce a cost benefit criteria for them. The model is in effect a spatial one in that we present a discrete approach to the study of the problem. The discrete approach uses a cellular automaton model which could easily be understood by nonspecialists. We shall also show how a characteristic empirical response to the vaccination policies could be achieved.

Criss-Cross Model with Immunization

Based on the age-dependent criss-cross model in the last section we examine two aspects of the impact of immunization which we expect will reduce the net rate of disease

transmission between the two populations by decreasing the per capita force of infection. In particular we consider the following aspects of immunization:

(i) its effect on the steady state or equilibrium conditions, that is, the state towards which a population may converge in the long term under the influence of an immunization programme;

(ii) its effect in the short term on the temporal dynamics of the infection within the various groupings as they move to a new steady state following the initiation of an immunization programme.

Suppose the age-specific rates of immunization are $c(a)$ and $\tilde{c}(a)$ for badger and cattle populations respectively. This introduces a further removal term in the criss-cross model for the susceptible population dynamics. We modify the model (10.90) for badger–cattle disease transmission dynamics to read

$$\begin{aligned}
\frac{\partial U}{\partial t} + \frac{\partial U}{\partial a} &= -[\lambda_1 + c(a)]U + rW - \mu U,\\
\frac{\partial W}{\partial t} + \frac{\partial W}{\partial a} &= \lambda_1 U - rW - \mu W,\\
\frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial a} &= c(a)U - \mu Z,\\
\frac{\partial \tilde{U}}{\partial t} + \frac{\partial \tilde{U}}{\partial a} &= -[\tilde{\lambda}_1 + \tilde{c}(a)]\tilde{U} + \tilde{r}\tilde{W} - \tilde{\mu}\tilde{U},\\
\frac{\partial \tilde{W}}{\partial t} + \frac{\partial \tilde{W}}{\partial a} &= \tilde{\lambda}_1 \tilde{U} - \tilde{r}\tilde{W} - \tilde{\mu}\tilde{W},\\
\frac{\partial \tilde{Z}}{\partial t} + \frac{\partial \tilde{Z}}{\partial a} &= \tilde{c}(a)\tilde{U} - \tilde{\mu}\tilde{Z},
\end{aligned} \tag{10.107}$$

where we have introduced another subpopulation in both the badgers and cattle, namely, the immune classes Z and $\tilde{Z}$ respectively. As before U and W are respectively the susceptible and infectious badgers with similar definitions for $\tilde{U}$ and $\tilde{W}$. The force of infection for the respective populations are again given by (10.91) and (10.92). The initial age distribution of the respective classes is given by (10.93) but with the addition of initial conditions for the immune classes, so

$$\begin{aligned}
U(0,a) &= U_0(a), \quad \tilde{U}(0,a) = \tilde{U}_0(a),\\
W(0,a) &= W_0(a), \quad \tilde{W}(0,a) = \tilde{W}_0(a),\\
Z(0,a) &= Z_0(a), \quad \tilde{Z}(0,a) = \tilde{Z}_0(a),
\end{aligned} \tag{10.108}$$

and represent the preimmunization equilibrium distributions. The renewal (boundary) conditions for the various groups are given by (10.95) with the addition of those for the immune classes, namely,

$$\begin{aligned}
U(t,0) = N(t,0) = \gamma N(t), \quad W(t,0) = Z(t,0) = 0,\\
\tilde{U}(t,0) = \tilde{N}(t,0) = \tilde{\gamma}\tilde{N}(t), \quad \tilde{W}(t,0) = \tilde{Z}(t,0) = 0,
\end{aligned} \tag{10.109}$$

where again $\gamma N(t)$ is the number of births of badgers at age 0 for all t, with the birth rate γ assumed constant and $N(t)$ the total badger population. So, at any time t, the age distribution of both badgers and cattle is again given by (10.96).

We again rescale the problem in the same way as we did in the last section by writing

$$\begin{aligned} u(t,a) &= \frac{U(t,a)}{N(t,a)},\ w(t,a) = \frac{W(t,a)}{N(t,a)},\ z(t,a) = \frac{Z(t,a)}{N(t,a)}, \\ \tilde{u}(t,a) &= \frac{\tilde{U}(t,a)}{\tilde{N}(t,a)}, \tilde{w}(t,a) = \frac{\tilde{W}(t,a)}{\tilde{N}(t,a)},\ \tilde{z}(t,a) = \frac{\tilde{Z}(t,a)}{\tilde{N}(t,a)}, \end{aligned} \tag{10.110}$$

and again rescaling the time and chronological age by setting $\tau = rt, \alpha = ra$ (badgers) $\alpha = \tilde{r}a$ (cattle) we get the nondimensional system

$$\begin{aligned} \frac{\partial u}{\partial \tau} + \frac{\partial u}{\partial \alpha} &= -\frac{1}{r}(\lambda_1 + c)u + w, \\ \frac{\partial w}{\partial \tau} + \frac{\partial w}{\partial \alpha} &= \frac{1}{r}\lambda_1 u - w, \\ \frac{\partial z}{\partial \tau} + \frac{\partial z}{\partial \alpha} &= \frac{1}{r}cu, \\ \frac{\partial \tilde{u}}{\partial \tau} + \frac{\partial \tilde{u}}{\partial \alpha} &= -\frac{1}{\tilde{r}}\tilde{\lambda}_1 \tilde{u} + \tilde{w}, \\ \frac{\partial \tilde{w}}{\partial \tau} + \frac{\partial \tilde{w}}{\partial \alpha} &= \frac{1}{\tilde{r}}(\tilde{\lambda}_1 + \tilde{c})\tilde{u} - \tilde{w}, \\ \frac{\partial \tilde{z}}{\partial \tau} + \frac{\partial \tilde{z}}{\partial \alpha} &= \frac{1}{\tilde{r}}\tilde{c}. \end{aligned} \tag{10.111}$$

Now let vaccination be given so that fractions $f, \tilde{f}, g, \tilde{g}$ of susceptible badgers and cattle become immune at ages T_1 and $T_2 > T_1$, say. Then the conditions at ages T_1 and T_2 depicting the relationship at the points of discontinuity of u and $\tilde{u}$ take the form

$$\begin{aligned} u(T_1 + 0) &= (1 - f)u(T_1 - 0); \quad u(T_2 + 0) = (1 - g)u(T_2 - 0), \\ \tilde{u}(T_1 + 0) &= (1 - \tilde{f})\tilde{u}(T_1 - 0); \quad \tilde{u}(T_2 + 0) = (1 - \tilde{g})\tilde{u}(T_2 - 0). \end{aligned} \tag{10.112}$$

In such circumstances we can deduce the susceptible populations consecutively for the specified 'immunization age intervals' (see, for example, Hethcote 1983) to get the susceptible fractions (in nondimensional terms) for the badger and cattle populations at equilibrium in the age intervals $[0, T_1)$, $[T_1, T_2)$, $[T_2, \infty)$. To do this we need the equilibrium solutions.

The initial values $u(0, \alpha) = u(\alpha)$, $\tilde{u}(0, \alpha) = \tilde{u}(\alpha)$ represent the preimmunization fractional equilibrium distributions given by the time-independent solutions of the set of ordinary differential equations given by (10.112) excluding the immunization terms and with all $(\partial/\partial t)$-terms set equal to zero. These are routinely found to be

$$\begin{aligned} u(\alpha) &= \frac{1}{\lambda_2 + r}\left[r + \lambda_2 \exp\left(-\frac{\lambda_2}{r} - 1\right)\alpha\right], \\ \tilde{u}(\alpha) &= \frac{1}{\tilde{\lambda}_2 + \tilde{r}}\left[\tilde{r} + \tilde{\lambda}_2 \exp\left(-\frac{\tilde{\lambda}_2}{\tilde{r}} - 1\right)\alpha\right], \end{aligned} \tag{10.113}$$

where the forces of infection λ_2 and $\tilde{\lambda}_2$ are those for the equilibrium state and given, respectively, by

$$\begin{aligned} (B) \qquad \lambda_2 &= \frac{\beta_1}{r}\int_0^\infty w(\alpha)N(\alpha)\,d\alpha + \frac{\beta_1}{\tilde{r}}\int_0^\infty \tilde{w}(\alpha)\tilde{N}(\alpha)\,d\alpha, \\ (C) \qquad \tilde{\lambda}_2 &= \frac{\tilde{\beta}_1}{\tilde{r}}\int_0^\infty \tilde{w}(\alpha)\tilde{N}(\alpha)\,d\alpha + \frac{\tilde{\beta}_2}{r}\int_0^\infty w(\alpha)N(\alpha)\,d\alpha. \end{aligned} \tag{10.114}$$

Using these solutions (10.113) we get, from (10.112),

$$u(\alpha) = \begin{cases} \frac{1}{\lambda_2 + r}[r + \lambda_2 \exp(-\frac{\lambda_2}{r} - 1)\alpha], & 0 \le \alpha < rT_1; \\ \frac{(1-f)}{\lambda_2 + r}[r + \lambda_2 \exp(-\frac{\lambda_2}{r} - 1)\alpha], & rT_1 \le \alpha < rT_2; \\ \frac{(1-f)(1-g)}{\lambda_2 + r}[r + \lambda_2 \exp(-\frac{\lambda_2}{r} - 1)\alpha], & rT_2 \le \alpha, \end{cases} \tag{10.115}$$

with a similar relation for susceptible cattle when f, g, r, λ are replaced by $\tilde{f}, \tilde{g}, \tilde{r}$ and $\tilde{\lambda}$ in (10.115).

At equilibrium the basic reproductive rate ρ_0 is related to the total susceptible fraction, u_e, by

$$\rho_0 u_e = 1. \tag{10.116}$$

In this context, the equilibrium disease incidence can then be determined from the relation

$$\rho_0 \int_0^\infty u(a)N(a)\,da = 1, \tag{10.117}$$

with a similar relation holding for the cattle population. For example, if we assume a constant death rate, μ, as we did in the previous section, we have

$$\frac{\rho_0 \gamma}{r^2}\int_0^\infty u(\alpha)e^{-(\mu/r)\alpha}\,d\alpha = 1 \tag{10.118}$$

which on integrating, using (10.115), gives

$$\begin{aligned} &\frac{\rho_0 \gamma}{\lambda_2 + r}\left\{\frac{r}{\mu}[1 - fe^{-\mu T_1} - (1 - f)ge^{-\mu T_2}]\right. \\ &\quad \left. + \frac{\lambda_2}{\lambda_2 + r + \mu}[1 - fe^{-(\lambda_2 + r + \mu)T_1} - (1 - f)ge^{-(\lambda_2 + r + \mu)T_2}]\right\} = 1. \end{aligned} \tag{10.119}$$

Suppose that vaccination takes place only at one age, T_1 say; then $g = 0$ and we get, from the last equation

$$\frac{\rho_0\gamma}{\lambda_2 + r}\left\{\frac{r}{\mu}[1 - fe^{-\mu T_1}] + \frac{\lambda_2}{\lambda_2 + r + \mu}[1 - fe^{-(\lambda_2+r+\mu)T_1}]\right\} = 1. \qquad (10.120)$$

The disease incidence in cattle satisfies the relation

$$\frac{\tilde{\rho}_0\tilde{\gamma}}{\tilde{\lambda}_2 + \tilde{r}}\left\{\frac{\tilde{r}}{\tilde{\mu}}[1 - \tilde{f}e^{-\tilde{\mu}T_1}] + \frac{\tilde{\lambda}_2}{\tilde{\lambda}_2 + \tilde{r} + \tilde{\mu}}[1 - \tilde{f}e^{-(\tilde{\lambda}_2+\tilde{r}+\tilde{\mu})T_1}]\right\} = 1. \qquad (10.121)$$

From the last two equations, λ_2 and $\tilde{\lambda}_2$ corresponding to the equilibrium forces of infection within badgers and cattle after the initiation of an immunization programme can be determined by estimating ρ_0, and $\tilde{\rho}_0$ from prevaccination epidemiological data and using (10.120) and (10.121) to calculate λ_2 and $\tilde{\lambda}_2$ in terms of f and T_1, which characterise the immunization programme. The effective reproductive rate, ρ, (that is, the generation of secondary cases where a proportion is immune) under the mass action assumption of disease spread and transmission, ρ is related to ρ_0 by

$$\rho = \rho_0 u_e = \rho_0(1 - f), \qquad (10.122)$$

where u_e is the fraction of susceptibles and f is the proportion that is temporarily immune.

To be able to eradicate the disease by adopting an appropriate vaccination or treatment coverage, it is necessary to create a level of herd immunity such that the effective reproductive rate, ρ, is reduced to a value less than unity. Herd immunity means that the fraction of the population that is susceptible is sufficiently small that an outbreak would not result if one animal suddenly became infective or if an imported infective were introduced into the environment. It can also be considered as an indirect protection of unvaccinated susceptibles by high levels of vaccination amongst the remaining segments of the population. This protection is a consequence of the reduction in disease transmission brought about by the removal of vaccinated animals from the susceptible class. It is through the effects of herd immunity that it is possible to eradicate a disease without vaccinating every single susceptible (Fox et al. 1971). Formally, the critical level of vaccination coverage corresponds to the limit $\lambda_2 \to 0$, $\tilde{\lambda}_2 \to 0$ in (10.120) and (10.121). In this way, each primary case will generate less than one secondary case as is evident from the ensuing relation

$$\rho_0[1 - f\exp(-\mu T_1)] \leq 1. \qquad (10.123)$$

This means, therefore, that we require the immune proportion of badgers and cattle to exceed a critical value

$$f_c = \left(1 - \frac{1}{\rho_0}\right)\exp\left(\frac{T_1}{L}\right); \quad \tilde{f_c} = \left(1 - \frac{1}{\tilde{\rho}_0}\right)\exp\left(\frac{T_1}{L}\right). \qquad (10.124)$$

Table 10.3.

Demographic	Epidemiological	Operational	Technical (Efficacy)
Population Growth (birth/death rates)	Prevalence level in initial situation	Population coverage	Clinical
	Implementation intensity	Effective contact rate	Waning
		Eligibility criteria	

So, a proportion greater than f_c of each new cohort of cub (calves) at or near birth, or at age T_1 should be immunized. If vaccination is given to very young cubs (calves) only, then $T_1 \approx 0$ and $\rho_0(1 - f) \leq 1$. Eradication is easier if animals are vaccinated at the earliest feasible age, T_1, and essentially impossible if at a later stage.

Case notification records show that generally, there has been a very low level of *M. bovis* infection over several decades (MAFF report 1987, 1994, Cheeseman et al. 1988, 1989). An important question concerns the level of coverage that should be aimed at to eradicate an infection should it occur. In the attempt to choose or adopt an effective strategy for the eradication of the disease, we should note that control measures differ in their effectiveness according to different situations. In Table 10.3 we list some aspects which exert different influences on the relative effectiveness of adopted measures.

Control Programme and Its Implementation

We make the following clinical assumptions.

(i) The development of *M. bovis* in infected badgers and cattle is purely an endogenous process.

(ii) The protective efficacy of vaccines is assumed to wane at a constant rate of 5% per annum and gives 66% protection (Waaler et al. 1969). This means that a vaccination coverage of about 95% amounts to transferring 66% of the noninfected group into a vaccine-protected group.

Studies by Stuart et al. (1988) on the development of diagnostic tests for, and vaccination against, tuberculosis in badgers suggest that badgers mount a weak antibody response to conventional antigens when compared with laboratory rabbits. However, it was found that cell-mediated immunity seems to be enhanced by vaccination and leads to prolonged survival of badgers and delayed excretion of tubercle bacilli.

As a means of reducing the force of infection and hence the number of infectives, in our approach we adopt chemotherapy in the form of oral vaccination to control *M. bovis* infection within badgers and suggest vaccination as a method to fight the disease in cattle. A combination of both strategies where vaccines are administered in mixed food items as well as actual vaccination of groups of animals may be helpful, although this will only really be effective for cattle since badgers, unlike cattle, are not confined to specific areas and sometimes move about randomly within, and sometimes away from, their neighbourhoods (see, for example, Rogers et al. 1998).

We divide the population into two, that is, cubs (calves) and adults. Considering badgers, for example, the fraction of cubs becoming immune at age T_1 (1 year) is f and the fraction of adults becoming immune at age T_2 (5 years) is g. If $\rho_0 u_e \leq 1$, then the disease will eventually die out and herd immunity achieved. From a practical point of view, we may conclude that a policy of 66% vaccination at age one will reduce the yearly incidence of *M. bovis* infection for approximately five years. Thereafter the yearly incidence will be higher if we adopt a no-vaccination policy at all. It is reasonable to suggest a two-66% vaccination policy: one at the beginning, or more precisely, a year after birth and the other after 5 years. This double campaign could reduce *M. bovis* infection for about nine years before any possible epidemic ensues.

An alternate, and we believe better, method is a modification of a vaccination policy which was proposed by Frerichs and Prawda (1975) and adopted for the control of rabies in Colombia. We call it the Preferred Vaccination Policy (PVP). Here, assuming there is a potential outbreak within neighbouring regions, each targeted subregion should be ranked according to the potential contribution it would make to the incidence of *M. bovis* infection. From case notification records, the risk, $R_{i,t}$, of badger/cattle contributing to new infection cases is calculated for each subregion within the specified area as

$$R_{i,t} = C_i U_{i,t} + \frac{1}{5} \sum_{j=1}^{5} C_{i(j)} U_{i(j),t}, \tag{10.125}$$

where

- $R_{i,t}$ is the index for *M. bovis* risk for subregion i at time t. It is calculated from case notification records;
- C_i is the proportion of badger/cattle in subregion i relative to that of neighbouring subregions;
- $U_{i,t}$ is the number of susceptible badger/cattle in subregion i at time t;
- $i(j)$ is a subscript denoting neighbouring subregions j, $j = 1, 5$ surrounding i.

Since there is the possibility for infection from neighbouring subregions, an assumption is made that the sum of the values $C_{i(j)} U_{i(j),t}$ in each of the five neighbouring subregions or social groups is equally as important to the subsequent generation of *M. bovis* infection as the value of $C_i U_{i,t}$ in subregion i itself. All values of $R_{i,t}$ are ranked from highest to lowest. This control policy continuously employs vaccinating teams who are sent to the highest ranked subregion and they remain there until the required proportion of susceptibles is vaccinated. Thereafter, they are sent to the next highest ranked subregion at different time periods and information updated as they visit various subregions.

Cellular Automaton Model for Practical Implementation

We present a discrete approach to the implementation of the PVP using cellular automaton models. Such models have the advantage of providing a visual representation of the main qualitative features of the processes and the results of simulations for various parameter sets. The cellular automaton models are as follows. First, we model the

disease incidence and spread within badger and cattle populations (Rule A) and second, we consider a situation where a proportion becomes immune due to the introduction of a vaccination policy, namely, the Preferred Vaccination Policy (PVP) (Rule B).

The first model consists of a rectilinear grid of cells which represent the contagion of the disease between badgers and cattle (Rule A below). The model includes parameters D_c and D_b for the duration of disease in cattle and badgers respectively. We consider v as a status variable which increases by one unit for each time-step until a maximum of D steps is reached after which v reverts to 0, that is, the animal becomes susceptible again. Here, we have assumed that there is no immune class. The status of the animal changes with time depending on the status of the animal itself and the status of its four contiguous neighbouring cells in accordance with a set of rules for simulating the contagion of the disease within badger populations. The rules are:

Rule A

An animal can become infected if it is in a susceptible state and if it comes into contact with an infected animal; that is, at least one of its four neighbours is infected. When contact occurs, the probability of a susceptible individual becoming infected depends on the parameters P_{cc}, P_{cb}, P_{bb} and P_{bc}. These are the probabilities of disease transmission per unit time from infected cow to susceptible cow, from infected cow to susceptible badger, from infected badger to susceptible badger, and from infected badger to susceptible cow respectively. At each time-step, each cell in the array is evaluated. If the cell contains an infected animal, we determine whether each of its four neighbours is in a susceptible state. If a neighbour is susceptible or exposed and has little or no resistance to infection, it becomes infected with probability p. Once infected the animal remains infective for a specified number of time-steps after which time it dies or becomes susceptible again.

The second model simulates the disease incidence and prevalence in which an immune class has been introduced (Rule B below). An immune class is introduced by sending vaccination teams to highest ranked subregions. This is the Preferred Vaccination Programme (PVP).

Rule B

Using a sparsely populated rectilinear grid, we consider an extension of Rule A to include an immune state for the status variable v. After remaining infected for D time units, an animal may become immune in a subregion with a certain probability. An animal in the immune state remains immune for a specified number of time-steps and dies.

Results and Control Implementation

A direct vaccination coverage to specific age groups is difficult in the case of badgers so oral vaccination could be administered by way of sprinkling food mixed with vaccine. (This method is remarkably efficient in dispensing vaccine to foxes in the case of rabies.) Vaccination is assumed to be given randomly to all badgers while with cattle those in the same age group or those found by a serological test to be susceptible could be vaccinated. The proposed strategy for vaccination should be one year after birth and four years afterwards.

To obtain some measure for the cost, C, of such a programme we exploit (10.120) and (10.121) to propose a cost–benefit criterion. One measure of the effort for a particular strategy to be implemented is

$$C = fe^{-\mu T_1}[1 + e^{-(\lambda_2+r)T_1}] + ge^{-\mu T_2}[1 + e^{-(\lambda_2+r)T_2}], \tag{10.126}$$

and similarly for the cattle population

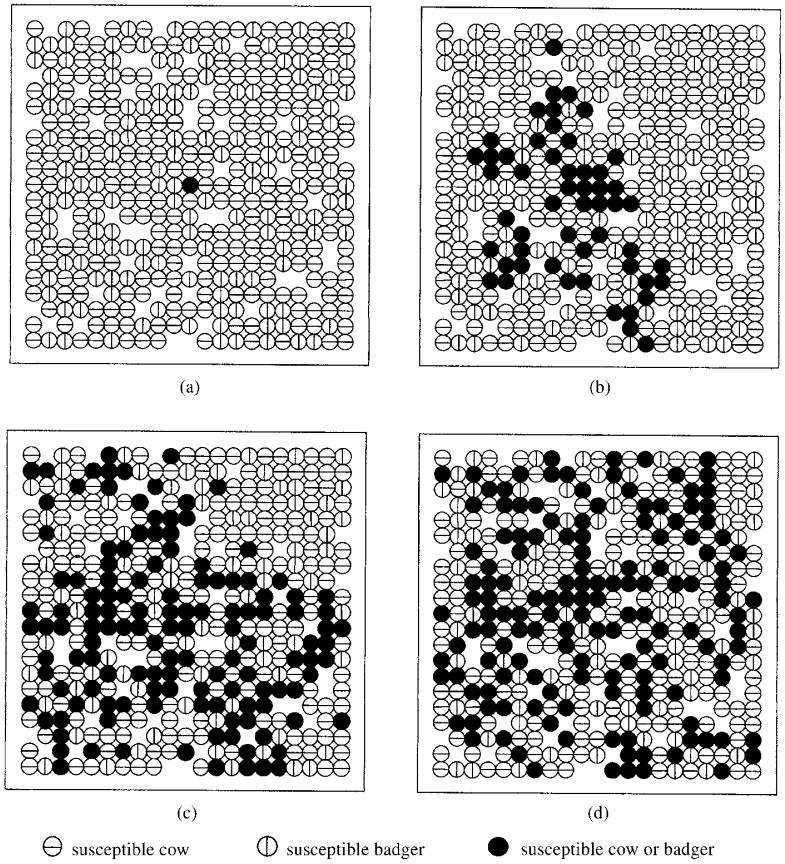


Figure 10.21. Cellular automaton simulation for disease spread between badgers and cattle. We consider a sparsely populated region (white regions imply uninhabited regions). Parameter values: normalised badger density = 0.60, cattle density = 0.40, duration of disease = 3 months. Transmission probabilities: infected cow to susceptible cow = 0.25, infected cow to susceptible badger = 0.1, infected badger to susceptible cow = 0.75. No development of immunity. (**a**) One infected badger introduced in the sett at $t = 0$. (**b**), (**c**) and (**d**) depict patterns of infection at $t = 20$ months, $t = 40$ months, $t = 60$ months. Initial condition: one infected badger was introduced at the centre of the array.

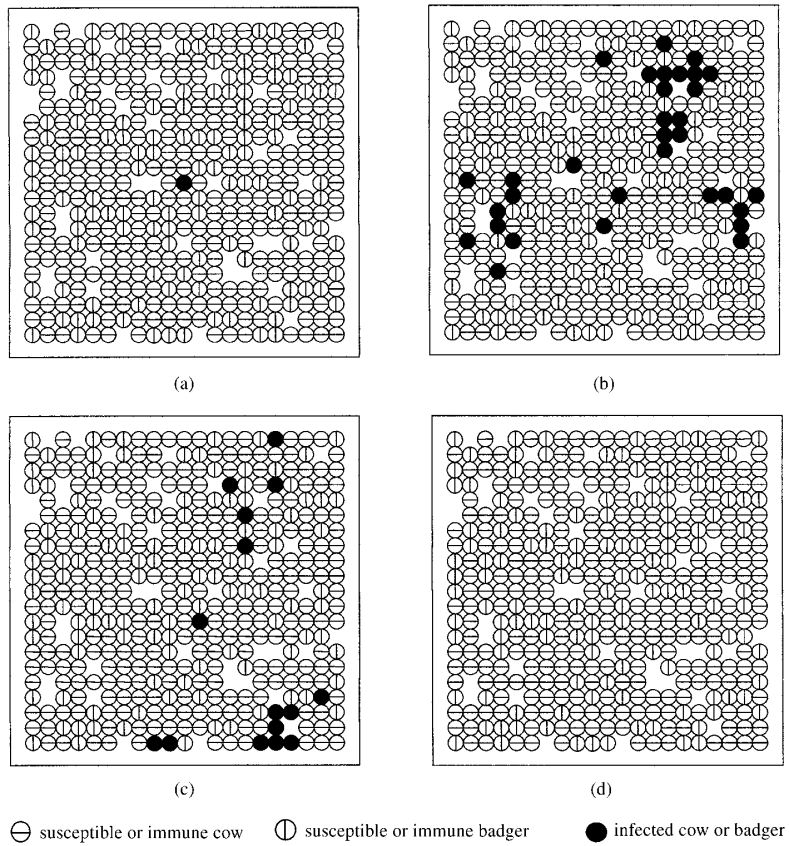


Figure 10.22. Empirical response to a vaccination policy. In comparison with the spread of infection between badgers and cattle (Figure 10.21), with the conferment of immunity due to the introduction of the PVP, the simulations suggest a possible reduction in disease incidence and prevalence. Parameter values: badger density = 0.60, cattle density = 0.40, duration of disease = 3 months. Transmission probabilities: infected cow to susceptible cow = 0.25, infected cow to susceptible badger = 0.1, infected badger to susceptible cow = 0.75. Immunity is 6 months. (**a**) One infected badger introduced in the sett at $t = 0$. (**b**), (**c**) and (**d**) depict patterns of infection at $t = 20$ months, $t = 40$ months, $t = 60$ months. Initial condition: one infected badger was introduced at the centre of the array.

$$\tilde{C} = \tilde{f}e^{-\tilde{\mu}T_1}[1 + e^{-(\tilde{\lambda}_2+\tilde{r})T_1}] + \tilde{g}e^{-\tilde{\mu}T_2}[1 + e^{-(\tilde{\lambda}_2+\tilde{r})T_2}], \tag{10.127}$$

where in each case we consider the force of infection in cattle, say, has an influence on the disease transmission dynamics of badgers and vice versa. Even though the Preferred Vaccination Policy could cost more than the other vaccination policies, the cumulative number of mean infected cattle over a 10-year planning period would be reduced to a significantly lower level. The indicator for the programme's success should therefore not

be cost-per-badger or cattle vaccinated but cost-per-infected badger or cattle prevented from becoming infected.

The cellular automaton models provide a visual representation of the main qualitative features of disease prevalence in badgers and cattle and the impact of a Preferred Vaccination Policy. Figure 10.21 shows a cellular automaton model for a typical criss-cross infection involving two distinct populations, that is, cattle and badgers. Figure 10.22 shows the characteristic empirical response to the described PVP for sparsely populated unit cells of badgers and cattle and an infection between them.

We are interested in the long term effect on the disease prevalence as a consequence of implementing a vaccination policy. So, by way of example, we used similar cellular automaton models to study the disease prevalence in badgers for the following control policies: (i) no vaccination, (ii) approximately 66% initial vaccination, and (iii) 66% initial + 66% revaccination (after five years). The vaccination was assumed to be administered by sprinkling food mixed with vaccines. Figure 10.23 compares results of the possible control experiments with that for the Preferred Vaccination Policy over a 10-year period.

As mentioned at the beginning of this section the MAFF control policy in Britain assumed that such a single and intensive intervention to remove all infected groups of animals would suffice to eliminate infection from contaminated areas for long periods of time (MAFF Report 1987). The MAFF (1994) report resulted in a more selective culling: badgers in Britain are now legally protected. The USDA eradication policy attempted to liquidate all infected animals in the herd with indemnities paid, as available, to help compensate owners for their losses, or hold herds under quarantine and tested

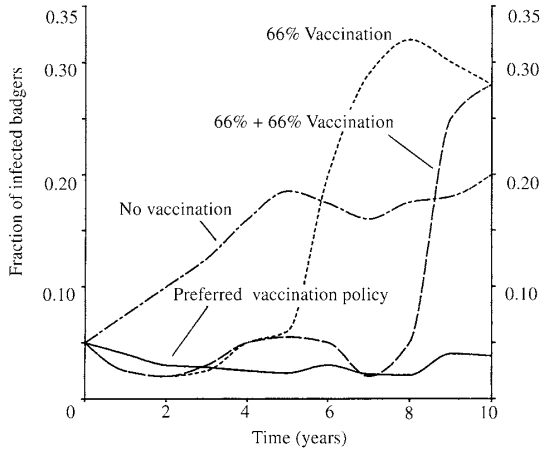


Figure 10.23. Computer simulation of possible scenarios for fractions of infected badgers over a 10-year period under various control strategies. The graphs show an empirical response to the control strategies described above namely: (i) no vaccination, (ii) 66% vaccination, (iii) two 66% vaccination campaigns, one at the beginning and the other after 5 years, (iv) PVP. It is evident that the fraction of infected badgers will be much lower for the PVP, giving an indication that the disease prevalence will be minimal in cattle too.

until all evidence of infection was eliminated (USDA Report 1982a,b) may not be the only approaches to take, from our model predictions. We note that chemotherapy in the form of oral vaccination (sprinkling of food mixed with vaccine) could be an effective way of controlling the spread of the disease in badgers as opposed to gassing, trapping and killing, since badgers, unlike cattle, are not confined to specific regions and move about randomly within and sometimes away from their setts. Again as already mentioned such a procedure has been very successful in the control of fox rabies in parts of Europe. We have also shown that eradication of the disease is easier if animals are vaccinated at the earliest feasible year (one year) rather than wait until there is an infection, in which case it would be almost impossible to eradicate the disease.

Early mass vaccination programmes predict a reduction of the effective reproduction rate of infection within communities, and hence raise the average age at infection amongst those animals which experience the disease. If, however, the force of infection changes with age, the tendency of mass immunization to increase the average age at infection may mean that the rate of exposure of susceptibles to infection in older classes may well differ from the rates acting in the age classes in which the susceptibles would have typically acquired infection prior to immunization. The potential consequence of such a change is to reduce the predicted level of vaccination coverage required to eradicate the infection below a certain level.

It seems that of the three vaccination strategies considered, leaving aside the cost–benefit criteria, the Preferred Vaccination Policy which targets highest ranked 'infectious' regions and vaccinates susceptibles is the best of the control programmes for bovine tuberculosis infection between badgers and cattle if an epidemic occurs. It should be noted, however, that the models we have discussed in this section and the previous one are still fairly basic. Not only that, we have not taken into account the spatial movement of the badgers. Relevant data on badger movement is given by Rogers et al. (1998) for 36 social groups in Gloucestershire in England over a period of 18 years. They show that the movement of badgers within groups varies and with this variation there is a variation in the incidence of bovine tuberculosis. The spatial aspects of disease spread are extremely important. We discuss an example of this later in Chapter 13, Volume II when we discuss the spatial spread of a rabies epidemic.

An interesting and very different new approach to the control of bovine tuberculosis is given by Kao et al. (1997). They develop a herd-based model which involves 'test and slaughter' combined with herd isolation and vaccination and they apply it to the situation in New Zealand. The model system consists of ordinary differential equations, relating movement from one state to another (such as from latent to infected) and an integral equation which gives the number of infected herds.

The question as to what is the best strategy for control is highly complex and clearly species-(and geographically) dependent. In the HIV modelling above we could incorporate two drugs into the models and thereby compare the efficacies of the different treatments. In the case of badgers and Tb, vaccination is now the preferred method of control in England. Rabies, the vector for which is the red fox (*Vulpes vulpes*) in the present epidemic in western Europe and, as mentioned, is controlled by vaccination, does not (yet) exist in Britain where domestic animals are not vaccinated for the disease. (Fox hunting in England is an amazingly inefficient way of keeping down the fox population.) It would be interesting to construct a model which included various meth-

ods of control and to compare the consequences of vaccination, culling, contraception and so on. From the point of culling a new selective culling was introduced in 1994 (MAFF Report 1994): without more knowledge which could come from further studies and realistic modelling it simply fuels the controversy. Hancox (1995) presents the various arguments in the debate. He concludes that there may be cattle reservoirs of bovine TB and so it is the cattle who are continually infecting the badger, a much more appealing scenario for the British with their sentimental view of badgers and who view the slaughtering of badgers with much indignation. The vaccination scenario we discuss in this section does not distinguish which of the badgers or cattle is the reservoir.

BSE (Bovine Spongiform Encephalopathy) and Creutzfeldt–Jacob (CJ) Disease

Although we do not do any modelling it is important to briefly mention bovine spongiform encephalopathy (BSE), or 'mad cow' disease. It was first diagnosed in England in 1986 and by the summer of 1997 there were around 167,000 cases confirmed with undoubtedly many more undetected.[12] The epidemic was severe in both size and in particular the human consequences since it is has given rise to the emergence of a new human disease, namely, a variant of Creutzfeldt–Jacob (CJ) disease.

CJ disease is a particularly horrifying neurodegenerative disease that affects the brain and is always fatal. It is caused by prions, which are very small particles—badly folded proteins—that are particularly tenacious; they cannot be broken down nor killed easily. Unlike viruses they contain no genetic material and so provoke no immune response. These prions accumulate in the brain and make spongelike holes. Prior to death the victims suffer from insomnia, depression, anxiety, memory loss, loss of bodily function control, coordination and blindness. Since prions are only in infected tissue they can easily be missed in an autopsy, which is why CJ is difficult to detect. It is becoming clear that BSE in the cattle was a result of contaminated feed associated with the equivalent disease in sheep and is directly the cause of variant CJ in humans. BSE comes under the category of a *transmissable spongiform encephalopathy* or TSE.

By the end of 2000 only 87 cases in Britain have been found since 1994. Compared with malaria or HIV it is negligible. There are several rather frightening reasons for the panic, particularly in France, because of the chilling list of facts about BSE and how easily it is passed from one infected animal to another. The pathogen is very tenacious and is resistant to heat, boiling, alcohol, ionizing radiation and so on. Surgical instruments which were in contact with the infected tissue can remain contaminated even after normal sterilisation. The pathogen can survive being buried for years, could end up in landfills and possibly passed on to grazing animals. BSE can be passed on by a cow ingesting as little as a few grammes of infected tissue. An animal can harbour the disease wihout showing any symptoms but it can pass it on to another animal. With the variant CJ disease it may be possible for one human, who has it but shows no symptoms, to pass it on to another.[13] The disease, in many ways is like being bitten by a dog that was possibly rabid but without the benefit of any subsequent vaccine, with a gestation period that could last for years with the knowledge that the disease can be passed (possibly trivially) on to others without any knowledge of having done so.

[12]The story of how the U.K. government dealt with the problem is an example of astonishing incompetence if not irresponsibility. As of 2000 the question of the export of British beef—now reputed to be free of BSE—was a matter of litigation between France on one side and Britain and the European Union on the other. The French felt that the evidence that the British beef was now safe was not unequivocal. As of the end of 2000 an epidemic in France, and other European countries, is causing serious concern and not just in view of the connection to Creutzfeldt–Jacob disease.

Another concern about infected animals is that most of a slaughtered animal is used for purposes other than beef for human consumption. It is frequently used in cosmetics, pet chow, beauty preparations and so on; the choreographer, George Balanchine, who died of CJ disease is believed to have contracted it from using a bovine glandular product to preserve youthful looks. The first French case was of a bodybuilder who used a muscle-boosting preparation. One of these currently (2000) available, according to Dr. Michael Hansen of the U.S. Consumers' Union, contains dried bovine brain, spleen, pituitary glands and eye tissue. Another possibility of contracting the disease comes from vaccines which are cultivated in bovine serum as was the case in Britain until 1993; the vaccines were only withdrawn from use in November 2000.

Since it is unknown how easy or difficult it is to contract CJ in humans, how long the gestation period is and so on, it is very difficult at this stage to come up with a model that has any credence as regards prediction. Nevertheless it is important to try and get some idea of the progress of both BSE in cattle and CJ disease in humans. Estimates range from several hundred thousand to (according to Dominique Gillot, the French Minister of Health) several dozen: the latter is clearly ridiculous. The increase in CJ disease in Britain is becoming alarming. Although the numbers are still small, it is the rate of increase that is crucial as we know from the material in this chapter. For example, 14 people died in 1999 and 14 contracted it in the first six and a half months of 2000 by when a total of 74 had died. By the end of 2000, a further 13 had died. The long incubation period of the human form of the disease and the fact that it is probable that several million people were exposed to contaminated beef in the 1980's imply that over the next 25 to 35 years several hundred thousand people could die of CJ disease.

Some modelling has been carried out by Donnelly et al. (1997), who set the demographic scene and discussed control strategies while Ferguson et al. (1997) presented and analysed an age-structured model for the transmission dynamics: unfortunately these have had to be based on the very limited available data about the etiology of the disease. The model includes infection obtained from feed, the primary source of BSE in cattle, as well as from direct horizontal and maternal transmission. They estimate parameters from the data and use back-calculation to reconstruct the past temporal pattern infection. Such back-calculation was used by Murray et al. (1986) in their study of the spatial spread of rabies. A review of this back-calculation methodology associated with parameter estimation in HIV-infection rates has been given by Bacchetti et al. (1993). Ferguson et al. (1997) carried out a sensitivity analysis of the parameters, gave estimates and predictions and discussed some of the implications. As mentioned, with the large number of unknowns any model predictions must be treated with considerable reserve.

In the case of any disease, the ultimate aim of epidemiologists is to eradicate it, or in other words make the virus, bacterium or whatever, become extinct. On the other

[13]Long before the BSE epidemic and the variant CJ disease, corneal transplantation was implicated in one case of human-to-human transmission of Creutzfeldt–Jacob disease (Duffy et al. 1974).

hand ecologists view extinction of a species, decline of their habitat, in fact generally a decline in biodiversity as a disaster. Although epidemiologists and ecologists have opposite goals the mathematical model equations share similar forms and analytical analyses. The paper by Earn et al. (1998) reviews some of the differences and similarities between these two important fields and discusses some of the recent work on their spatial aspects, chaotic behaviour and synchrony.

Exercises

1 Consider the dynamics of a directly transmitted viral microparasite to be modelled by the system

$$\frac{dX}{dt} = bN - \beta XY - bX, \quad \frac{dY}{dt} = \beta XY - (b+r)Y, \quad \frac{dZ}{dt} = rY - bZ,$$

where b, β and r are positive constants and X, Y and Z are the number of susceptibles, infectives and immune populations respectively. Here the population is kept constant by births and deaths (with a contribution from each class) balancing. Show that there is a threshold population size, N_c, such that if $N < N_c = (b+r)/\beta$ the parasite cannot maintain itself in the population and both the infectives and the immune class eventually die out. The quantity $\beta N/(b+r)$ is the *basic reproductive rate* of the infection.

2 Consider an epidemic outbreak of a lethal disease in which the infectious period and the incubation period of the disease are different. Denote the number of susceptibles by $S(t)$, those incubating the disease by $E(t)$, the population who are infectious by $I(t)$ and those that have died by $R(t)$. During the epidemic assume the population is constant, equal to N. If a susceptible can be infected by someone who is incubating the disease but less easily than by an infected person, justify the following *SEIR* model,

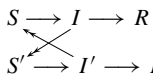
$$\frac{dS}{dt} = -\frac{\beta S}{N}(I + rE), \quad \frac{dE}{dt} = \frac{\beta S}{N}(I + rE) - bE,$$
$$\frac{dI}{dt} = bE - cI, \quad \frac{dR}{dt} = cI,$$

where β, r, b and c are positive constants. What does each of these parameters measure?

Suppose that in the early stages of the epidemic only a few (relative to the total population) individuals, E_0, become infected all at the same time and so are incubating the disease: they do not become infectious for a time of the order of $1/b$. Is this a reasonable presumption? During this time $S(t) \approx N$. Use this to solve for $E(t)$ as a function of t.

With the full system examine the stability of the disease-free steady state and hence determine the conditions for it to be unstable. Hence deduce that the basic reproductive rate $R_0 = (\beta/bc)(b + cr)$.

3 In a criss-cross venereal infection model, with the removed class permanently immune, the infection dynamics is represented by



with the usual notation for the susceptibles, infectives and the removed class. Briefly describe the assumptions made for its model system to be

$$\frac{dS}{dt} = -rSI', \quad \frac{dS'}{dt} = -r'S'I,$$
$$\frac{dI}{dt} = rSI' - aI, \quad \frac{dI'}{dt} = r'S'I - a'I',$$
$$\frac{dR}{dt} = aI, \quad \frac{dR'}{dt} = a'I',$$

where the parameters are all positive. The initial values for S, I, R, S', I' and R' are S_0, I_0, 0 and S_0', I_0', 0 respectively.

Show that the female and male populations are constant. Hence show that $S(t) = S_0 \exp[-rR'/a']$; deduce that $S(\infty) > 0$ and $I(\infty) = 0$ with similar results for S' and I'. Obtain the transcendental equations which determine $S(\infty)$ and $S'(\infty)$.

Show that the threshold condition for an epidemic to occur is at least one of

$$\frac{S_0 I_0'}{I_0} > \frac{a}{r}, \quad \frac{S_0' I_0}{I_0'} > \frac{a'}{r'}.$$

What single condition would ensure an epidemic?

4 Consider a population of haemophiliacs who were given infected blood and so were all infected with HIV at the same time $t = 0$. Denote by $y(t)$ the fraction of the population who have AIDS at time t, and by $x(t)$ the fraction who are HIV-positive but do not yet have AIDS. Let $v(t)$ be the rate of conversion from infection to AIDS. Show that a simple model for the dynamics with relevant initial conditions is then

$$\frac{dx}{dt} = -v(t)x, \quad \frac{dy}{dt} = v(t)x,$$
$$x(0) = 1, \quad y(0) = 0.$$

Assume that the patient's immune system is progressively impaired from the time of infection and so $v(t)$ is an increasing function of time. Examine the system when $v(t)$ varies: (i) linearly with time and sketch the rate of change in the population who develop AIDS and (ii) faster than linearly.

[Peterman et al. (1985) present data on 194 cases of blood transfusion-associated AIDS. With $v(t) = at$ the solution of the model system with $a = 0.237yr^{-1}$ applied to these data gives the rate of increase, dy/dt, in AIDS patients which compares very well (depressingly so) with the data.]

5 For the drug use epidemic model in Section 10.9 show that the values given for the threshold parameter γ/S_0 in cases (iii) and (iv) in Table 10.1 are as given.