Chapter Four

Multi-Pathogen/Multi-Host Models

In this chapter, we follow some recent attempts to develop a theory of "community epidemiology," whereby the traditional framework examining one pathogen spreading through a single host population is extended in two distinct ways: (1) to consider multiple infectious diseases (or strains) spreading through one host species, or (2) a single infectious disease that can be transmitted between different species.

The scientific community is only just beginning to understand the range of complex outcomes when multiple strains compete for a limited supply of susceptibles or when a number of host species share a pathogen. These conceptual extensions of the simple one-pathogen one-host paradigm are becoming increasingly recognized as fundamental to our understanding of key questions in modern epidemiology and public health. Does a reduction in species diversity amplify or reduce the intensity of outbreaks? What role does biodiversity play in the maintenance of pathogens? What are the conditions that determine the spillover of "emerging" pathogens to new host species? How do these community epidemiology issues impact the invasion of species in new environs?

A complete description of models that examine strain dynamics requires a detailed knowledge of the host immune systems, and how different strains interact via the immune response that they elicit. Interest in this area stems from the fact that competition between strains is fundamental to disease evolution. Thus, if we are to understand and ultimately predict the emergence of novel pathogens or, for example, the strain of influenza that will circulate next season, we must better grasp the role of competition between cross-reacting strains of infection.

When one pathogen is transmitted between multiple host species, the models and dynamics are very much akin to those developed to deal with subdivided or structured populations (Chapter 3). The main distinction is that with different species the disease can have dramatically different infectious periods, mortality, and other characteristics, as well as the underlying demographic differences between the species. In general, multi-host models can be partitioned into two different classes: either a secondary obligate host is required for transmission (such as vector-borne diseases), or transmission can occur both within and between the various host species. Vector-borne diseases, such as malaria, dengue fever, and leishmania, are among the most challenging from a public-health perspective, and so models are frequently required to optimize their control. Finally, zoonoses form a special class of multi-host diseases, where animals form a reservoir of pathogens that can be spread to humans. Examples include such high-profile diseases as avian influenza, bubonic plague, hantavirus, Lyme disease, Q-fever, rabies, SARS, toxoplasmosis, trypanosomiasis, West Nile virus, and a range of macro-parasitic infections. In general, the effects of these diseases in humans tends to be quite severe; this is because humans act as a dead end for such infectious diseases and, therefore, there has been little evolutionary pressure for benign infection compared to within the reservoir host species. Although in principle the models for zoonoses have the same structure as other multi-host

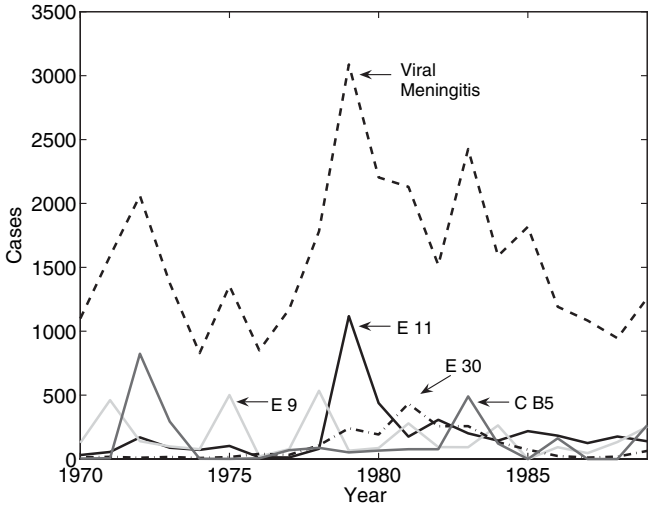


Figure 4.1. Results from the Nonpolio Entrovirus Surveillance scheme in the United States, from 1970–1989. Shown is the total recorded cases of viral aseptic meningitis (dashed line), as well as the constituent cases for echovirus 9, 11, and 30, and coxsackie B5. The interaction between these constituent viral types is complex, but must be understood if we are to successfully understand and predict the incidence of infection (Strikas et al. 1986). (Data from CDC–the Center for Disease Control, USA.)

models, the implications of infection and therefore the need for control measures will be far greater.

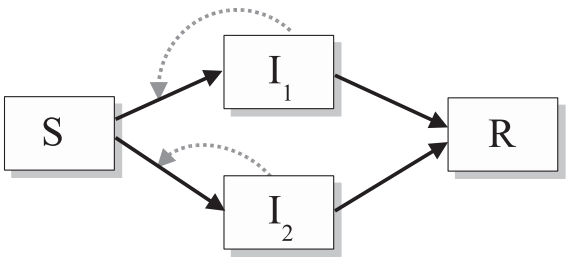
4.1. MULTIPLE PATHOGENS

We first focus on the interaction of two, or more, pathogens within a population. Many infectious diseases that we consider as a single disease are in reality comprised of multiple strains, which interact through the cross-immunity that is invoked within a host. An accurate understanding of such diseases, which include malaria, dengue fever, and influenza, requires the consideration of strain structure (Figure 4.1) (Ferguson et al. 1999a; Gog and Grenfell 2002; Ferguson et al. 2003a). Models incorporating multiple pathogens allow us to investigate questions of disease evolution, from theoretical questions such as understanding current disease behavior in terms of an optimal strategy for transmission, to more applied issues such as predicting the influenza strains for the coming year or understanding the effects of strain-specific control. Finally, multi-strain models offer insights into the increasing prevalence of drug-resistant bacteria, and how to limit their spread (Baquero and Blàzquez 1997; Bohannan and Lenski 2000).

The presence of two or more infectious diseases within a population increases the possible number of compartments into which the population can be subdivided. In the most general form, this formulation should uniquely identify the entire infection history of individuals within each compartment and their immunological status with respect to

the various pathogens/strains under consideration. Such completely general models are more difficult to study due to the large number of degrees of freedom and therefore the difficulty in both parameterization and analysis. Instead, models usually focus on particular assumptions about conferred immunity and the interaction between pathogens. As such, two completely independent infectious diseases provides the simplest model, although it is of little biological interest. We now consider a range of models based on more complex assumptions, where either transmission of, or susceptibility to, one strain is modified due to resistance to another strain. In some cases, it will be assumed that co-infection (simultaneous infection with two pathogens) is so rare that it can be ignored. This simplification is often true for many airborne infections, though for sexually transmitted diseases, infection with one pathogen may increase the susceptibility to others and hence co-infection is promoted (Renton et al. 1998; Chesson and Pinkerton 2000).

4.1.1. Complete Cross-Immunity



For illustrative purposes, we start with a model of two co-circulating strains within a simplified SIR framework, assuming complete cross-immunity. This means infection by either strain confers lifelong immunity to both. The four distinct compartmental classes are: susceptible to both strains, infectious with strain 1, infectious with strain 2, and recovered and therefore immune to both. Mathematically this leads to the following differential equations which are a simple modification to the standard SIR equations:

$$\begin{aligned}
\frac{dS}{dt} &= \nu - \beta_1 SI_1 - \beta_2 SI_2 - \mu S,\\
\frac{dI_1}{dt} &= \beta_1 SI_1 - \gamma_1 I_1 - m_1 I_1 - \mu I_1,\\
\frac{dI_2}{dt} &= \beta_2 SI_2 - \gamma_2 I_2 - m_2 I_2 - \mu I_2,\\
\frac{dR}{dt} &= \gamma_1 I_1 + \gamma_2 I_2 - \mu R.
\end{aligned} \tag{4.1}$$

Here, for generality, it has been assumed that the two strains have different transmission (β), recovery (γ), and mortality (m) rates. The strain-specific basic reproductive ratio is given by $R_0^i = \beta_i/(\gamma_i + \mu + m_i)\,(i = 1, 2)$. It is natural to attempt an understanding of this system by deriving and examining its equilibria. This approach soon leads to the paradoxical situation where the coexistence of both strains requires the fraction susceptible to be at once $1/R_0^1$ and $1/R_0^2$! Closer examination by simulation of equations (4.1) shows that only one strain can persist in such a scenario. Both are competing for the same limited

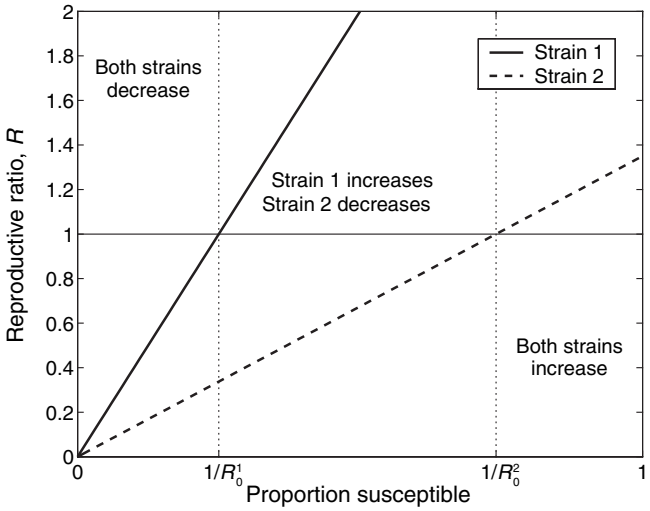


Figure 4.2. The reproductive ratio, $R = S \times R_0$, of two competing strains that offer complete cross-immunity. When the level of susceptibles is low the prevalence of both strains decreases, whereas when a high proportion are susceptible both strains can increase, although this will eventually lead to a decrease in the proportion susceptible, changing the dynamics. In the intermediate region, only strain 1 can increase with the weaker strain 2 being driven to extinction. ($\beta_1 = 4$, $\beta_2 = 1.35$, $\gamma_1 + \mu = \gamma_2 + \mu = 1$, $m_1 = m_2 = 0$.)

resource (susceptibles), and following the ecological tenet, whichever strain utilizes that resource more efficiently will dominate. In particular, the strain with the largest basic reproductive ratio will drive the other to extinction.

When competing strains provide complete protection for each other, the strain with the largest R_0 will force the other strain to extinction.

This competition scenario can be readily understood by considering the growth rate of each strain, captured by the reproductive ratio R^i, as the level of susceptibles varies (Figure 4.2). For convenience, we assume that strain 1 has a higher basic reproductive ratio than strain 2, that is, $R_0^1 > R_0^2$. Standard theory tells us that the fraction of infecteds will increase whenever $S > 1/R_0$ (so that the effective reproductive ratio, $R_0 \times S$, is greater than one), and will decrease whenever $S < 1/R_0$. This behavior leads to three separate regions and two points where one of the strains is at equilibrium ($R_0 S = 1$). If strain 2 is at equilibrium (and therefore $S = 1/R_0^2$), then because the growth rate of strain 1 is positive, this cannot be a stable equilibrium solution for both strains. However, if strain 1 is at equilibrium (and $S = 1/R_0^1$), then because the growth rate of strain 2 is negative, this strain cannot invade and is always forced to extinction. The stable solution is therefore $S = 1/R_0^1$, $I_1 = (\nu R_0^1 - \mu)/\beta_1$, $I_2 = 0$ (assuming births exactly balance deaths).

Although R_0 determines the eventual competitive outcome, pathogens with a more rapid life cycle may be favored in the short term. If two strains invade a highly susceptible population ($S \sim 1$), then the strain with the largest growth rate ($\beta - \gamma - m - d$) will be initially favored. Figure 4.3 shows an example of this type of behavior, starting with two

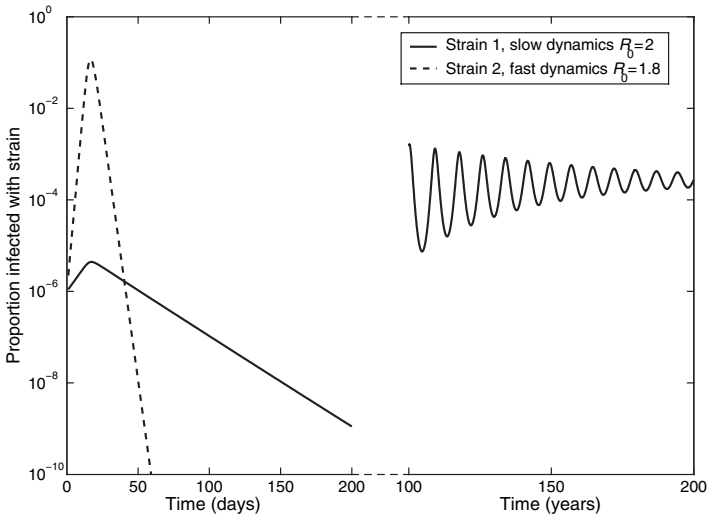


Figure 4.3. The level of two competing strains introduced into a totally susceptible population. During the early epidemic phase (days 0–50), the faster strain 2 dominates, whereas at far later times strain 1 with the largest R_0 is the eventual winner. ($\nu = \mu = 5 \times 10^{-5}$; $\beta_1 = 0.2$, $\gamma_1 = 0.1$, $R_0^1 \approx 2$; $\beta_2 = 1.8$, $\gamma_2 = 1$, $R_0^2 \approx 1.8$; $m_1 = m_2 = 0$. All rates are in days.)

strains invading a totally susceptible population. Although strain 1 has a larger R_0 and therefore eventually dominates (at a time scale of many years), because the life cycle of strain 2 is so much faster it wins out during the initial epidemic lasting about 50 days.

Although the relative R_0 values determine the long-term competitive success, a rapid life cycle may allow short-term dominance.

4.1.1.1. Evolutionary Implications

Now consider what the above results mean in terms of the direction that natural selection would be expected to act. Any mutation that generates a new strain with a larger basic reproductive ratio will be favored, and over time such mutations should accumulate such that R_0 increases. Relating this behavior to more mechanistic parameters implies that both the transmission rate, β, and the infectious period, $1/\gamma$, should increase and where applicable the disease-induced mortality or virulence, m, should decrease (May and Anderson 1983; Bremermann and Thieme 1989). Hence, when no other constrains are present, all infections should become highly transmissible, lifelong infections that are benign or even beneficial to the host (Mann et al. 2003).

Evolution will favor mutants with higher R_0, in theory leading to higher transmission rates, and long-lasting infections associated with a low probability of mortality.

Interestingly, most infectious diseases of public-health concern do not fit this predicted pattern. The reason for this discrepancy has been the subject of much research, generally focused on the existence of trade-offs between the transmission rate of the pathogen and the duration of the infectious period (Bremermann and Thieme 1989; Levin et al. 1999;

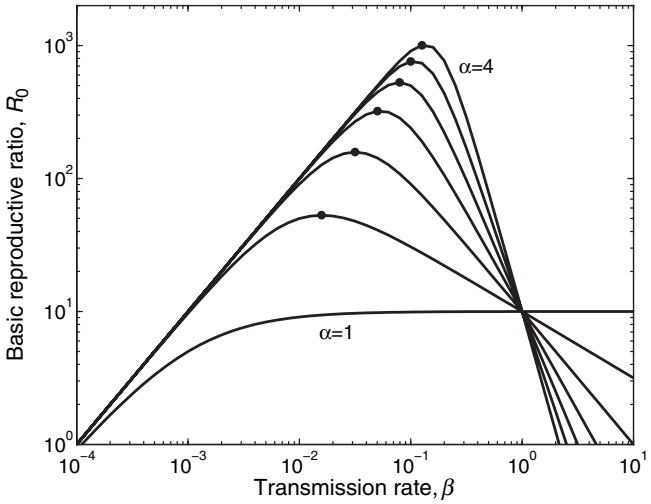


Figure 4.4. Using the power-law trade-off $\gamma + m = k\beta^{\alpha}$, intermediate values of β produce the maximum basic reproductive ratio, R_0. When $\alpha \leq 1$, we predict runaway evolution to ever-larger transmission rates. For $\alpha > 1$, the optimal transmission rate and associated R_0 value (indicated with a dot) increases as α becomes larger. ($k = 0.1$, $\mu = 10^{-4}$, rates given per day.)

Boots and Sasaki 1999). The transmission-virulence trade-off is based upon the notion that any infection producing lots of pathogen particles—while readily transmitted—is likely to be harmful to the host, resulting in rapid host death (Anderson and May 1991; Boots and Sasaki 1999). This trade-off could also apply to the infectious period, such that highly transmissible pathogens are most often of short duration. A crude reflection of this can be seen in the low-transmissibility and long-infectious period of sexually transmitted diseases compared to the generally more transmissible airborne diseases that have much shorter infectious periods. Given a relationship between the duration of infection, $1/(\gamma + m + \mu)$, and the transmission rate, β, the disease will evolve to the set of parameters that maximize R_0:

$$R_0 = \frac{\beta}{\mu + \gamma + m}.$$

The trade-off is frequently assumed to be power-law in shape such that:

$$\gamma + m = k\beta^{\alpha} \qquad \Rightarrow \qquad R_0 = \frac{\beta}{\mu + k\beta^{\alpha}}.$$

Hence, when $\alpha > 1$, R_0 has a well-defined maximum and the transmission rate should evolve toward the intermediate value of $\beta = \sqrt[\alpha]{\mu/(k(\alpha - 1))}$. By applying a trade-off between transmission and infection longevity, we can eliminate runaway evolution and predict intermediate infection parameters and dynamics of the type observed (Figure 4.4).

Trade-offs between transmission rates and duration of infection mean that R_0 is maximized for intermediate values and runaway evolution is prevented.

Although these forms of analysis provide an intuitive understanding of the selection pressures that affect disease evolution, several issues still remain unsolved. Most

immediate is the precise form of the trade-off (Levin 1996). Although the assumed power-law relationship is quite flexible, the true trade-off is likely to be more complex. To date, there is little experimental or observational evidence to support the existence of a trade-off and clearly insufficient data to determine a functional form (Ebert and Bull 2003). For this formulation to be of practical benefit, the trade-off curve associated with every pathogen would have to be determined. Finally, the models implicitly assume that infection parameters occupy a continuous spectrum, whereas in reality their behavior is controlled by a discrete set of genes. However, the mapping from genotype to phenotypic behavior is still a major challenge and we are many years from a detailed empirical understanding.

As well as providing an evolutionary perspective on disease behavior, this theory can also be used to investigate the plausible evolutionary changes in response to social and medical changes. The most applied use is the investigation of antibiotic resistance—when it arises and how its spread can be controlled (Austin and Anderson 1999a; Lipsitch et al. 2000). Consider two bacterial strains, a wild type (I_w) and a resistant strain (I_r), that compete for susceptibles. In general, resistance is assumed to be costly, meaning that the wild type is naturally fitter (has a higher R_0^w). It is this assumption that allows the wild type to dominate in the natural environment. The resistant strain, however, is not eliminated by antibiotics and therefore maintains its R_0^r even when medical treatment is provided. By modeling the effects of treatment on the reproductive ratio of the wild type strain, treatment regimes can be found that minimize the evolution of resistant strains. For example, we might suppose that treatment with antibiotics acts to reduce the infectious period of the wild type strain, allowing individuals to recover more quickly. This leads to the following set of equations describing the dynamics:

$$\begin{aligned}
\frac{dS}{dt} &= \nu - S(\beta_w I_w + \beta_r I_r) - \mu S, \\
\frac{dI_w}{dt} &= \beta_w S I_w - (\gamma + T) I_w - \mu I_w, \\
\frac{dI_r}{dt} &= \beta_r S I_r - \gamma I_r - \mu I_R,
\end{aligned}$$

where $\beta_r < \beta_w$, and treatment with antibiotics, T, acts to increase the recovery rate. The results are shown in Figure 4.5.

As predicted from previous models, it is clear that due to the complete cross-immunity only one strain can persist; when the treatment level is low the wild type dominates, but above a critical treatment level, T_C, the resistant mutant takes over. This critical level is observed when both strains have equal reproductive ratios:

$$\frac{\beta_w}{\gamma + T_c} = \frac{\beta_r}{\gamma} \qquad \Rightarrow \qquad T_C = \gamma \left(1 - \frac{\beta_w}{\beta_r}\right).$$

Minimizing the prevalence of infection in the population occurs for treatment levels just below T_C; therefore, there is a careful balance between reducing infection and not producing conditions favorable to the resistant strains. The greater the difference between wild-type and resistant strains (in terms of R_0), the greater the reduction in prevalence before resistant strains are favored. Two biological factors complicate this picture: (1) compensatory mutations can arise that counteract the reduction in fitness suffered by the resistant strain, and (2) although mutation from wild type to resistant is relatively common, back mutations are far rarer. An additional, though often ignored, subtlety in such

systems is the horizontal transfer of antimicrobial resistance genes mediated by plasmids (Sørensen et al. 2004).

Application of antibiotic treatments requires a careful balance between combating infection and not providing suitable conditions for resistant mutants to out-compete the wild type.

4.1.2. No Cross-Immunity

At the opposite extreme are infectious diseases that result in no cross-immunity. The modeling of such infections only differs from the case of two independent infections in that co-infection is assumed not to occur. Often, this assumption has a strong epidemiological basis because those individuals infected or convalescing from one infection are not mixing with enough individuals to catch subsequent infections. In this instance, the equations become:

$$\begin{aligned}
\frac{dN_{SS}}{dt} &= \nu - \beta_1 N_{SS} I_1 - \beta_2 N_{SS} I_2 - \mu N_{SS},\\
\frac{dN_{IS}}{dt} &= \beta_1 N_{SS} I_1 - \gamma_1 N_{IS} - \mu N_{IS},\\
\frac{dN_{RS}}{dt} &= \gamma_1 N_{IS} - \beta_2 N_{RS} I_2 - \mu N_{RS},\\
\frac{dN_{SI}}{dt} &= \beta_2 N_{SS} I_2 - \gamma_2 N_{SI} - \mu N_{SI},\\
\frac{dN_{RI}}{dt} &= \beta_2 N_{RS} I_2 - \gamma_2 N_{RI} - \mu N_{RI},\\
\frac{dN_{SR}}{dt} &= \gamma_1 N_{IS} - \beta_1 N_{SR} I_1 - \mu N_{RS},\\
\frac{dN_{IR}}{dt} &= \beta_1 N_{SR} I_1 - \gamma_1 N_{IR} - \mu N_{IR},\\
\frac{dN_{RR}}{dt} &= \gamma_1 N_{IR} + \gamma_2 N_{RI} - \mu N_{RR},\\
I_1 &= N_{IS} + N_{IR}, \qquad I_2 = N_{SI} + N_{RI},
\end{aligned} \tag{4.2}$$

where N_{AB} refers to the proportion of the population that are in state A with respect to disease 1 and state B with respect to disease 2, so N_{SI} refers to individuals that are susceptible to disease 1 and infections with disease 2. In this situation, co-existence of the two strains/diseases is generally possible whenever the two basic reproductive ratios are greater than one.

Even when there is no cross-immunity, the absence of multiply infected individuals is epidemiologically plausible, reflecting the reduced number of contacts when ill.

4.1.2.1. Application: The Interaction of Measles and Whooping Cough

One of the clearest applications of this type of model has focused on the dynamics of measles and whooping cough. Measles is caused by a virus and whooping cough is caused by a bacterium, hence we would not expect any specific immune-mediated interaction.

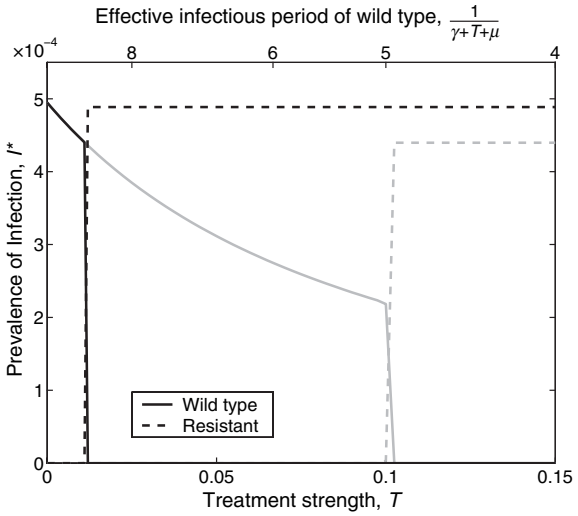


Figure 4.5. The prevalence of wild type, I_w (solid lines), and antibiotic resistant, I_r (dashed lines), infection in the population as differing amounts of treatment, T, are applied. The upper x-axis converts the treatment level into a measure of the infectious period. Due to complete cross-immunity, only one strain can persist. ($\beta_w = 1$, $\gamma = 0.1$, which means that in the absence of treatment, $R_0^w = 10$.) Two resistant scenarios are considered: When the resistant strain is only slightly less fit, $\beta_r = 0.9$, $R_0 = 9$ (black), and when the resistant strain is far less fit, $\beta_w = 0.5$, $R_0 = 5$ (gray). ($\nu = \mu = 5.5 \times 10^{-5}$, all rates are per day.)

However, Rohani et al. (1998) argued that the quarantining that occurs during the later stages of the infectious period and subsequent convalescence means that a fraction of potentially susceptible individuals might be temporarily unavailable to catch any other co-circulating infections. Thus, simple behavioral considerations predict a possible interaction between the epidemics of unrelated infectious diseases. In addition, when pathogens are associated with a considerable probability of mortality following infection, some of those infected become permanently unavailable to contract other diseases (similar to the effects of cross-immunity). This effectively increases the competition for hosts, resulting in "interference" between the epidemics of unrelated infectious diseases.

Rohani et al. (1998) modified equations (4.2) by incorporating a convalescent class and demonstrated that the isolation of individuals during their convalescence period is sufficient to alter the dynamics of competing infectious diseases. They compared the epidemics of measles and whooping cough predicted by two independent single-disease $SEIR$ models to those observed in the modified two-disease model described above. In the two-disease model, the rigidly annual epidemics of whooping cough predicted by the standard $SEIR$ model give way to a range of biennial dynamics, mimicking the epidemics of measles. The presence of whooping cough in the two-disease model results in a reduced effective amplitude of seasonality for measles. As explained by Huang and Rohani (2005), this is because measles has the higher transmission rate, thereby its epidemic patterns become imprinted on whooping cough dynamics as a result of their interaction via the susceptible pool. The most obvious signature of disease interference is, however, out-of-phase epidemics of the two infections when dynamics are multiennial.

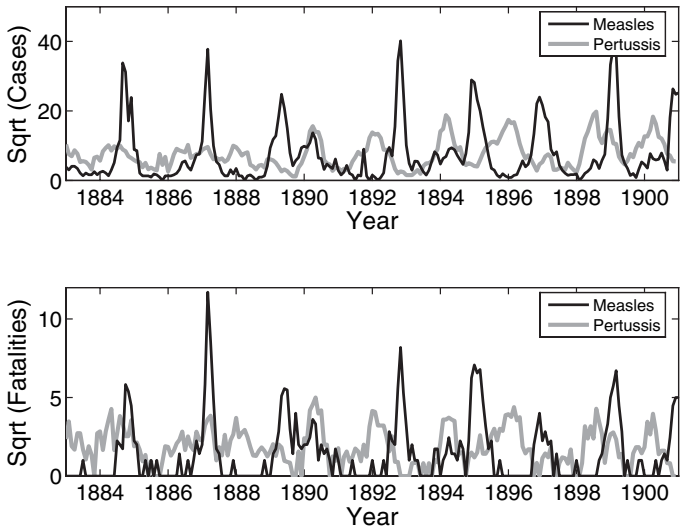


Figure 4.6. Long-term patterns in measles (black lines) and whooping cough (gray lines) epidemics in Aberdeen from 1883 to 1901 (data from Laing and Hay 1902). The top panel contains monthly case notifications; the bottom panel shows monthly case fatalities. The time series for both infections exhibit a strong biennial component, with a striking phase difference between measles and whooping cough outbreaks. (Note that the data is plotted on a nonlinear scale to better illustrate the cyclic nature.)

The most convincing empirical support for this phenomenon has been found in European case fatality data for measles and whooping cough in the early decades of the twentieth century (Rohani et al. 2003). As shown in Figure 4.6, the biennial epidemics of measles and whooping cough are clearly out of phase, particularly from 1891–1902, consistent with the predictions of the two-disease model.

Notice, however, that children are typically exposed to many more infectious diseases than just measles and whooping cough. These other infections include mumps, polio, rubella, or chickenpox, among others. Given the generality of the interaction envisaged by Rohani et al. (1998), would we expect *all* of these infectious diseases to dynamically interact? By extension, ideally, would any modeling work need to include *all* of these pathogens? We believe the answer to this question is likely to be "no." Using simple homogeneous models without age structure, it has been demonstrated that dynamical interference effects are most pronounced between infections with a similar basic reproductive ratio (Rohani et al. 2003, 2007; Huang and Rohani 2006). A complete understanding of this issue will, however, need age-dependence in contact rates to be taken into account. This is because the interference concept relies on "competition" for resources (hosts) between pathogens. For the dynamical effects of this competitive interaction to be noticeable, the pathogens should be infecting largely the same cohort of hosts. Using a two-disease age-structured model, Huang and Rohani (2006) showed that the extent of interference effects is closely determined by the relative distributions of age at infection. Hence, given their similar R_0s, measles and whooping cough are likely to have been strongly "competing" for children in the same age cohorts, whereas their interaction with the other childhood

infections is likely to have been less intense. This logic suggests that infectious diseases such as rubella and polio, or chickenpox and mumps, may also be good potential candidates for the study of interference effects.

4.1.2.2. Application: Multiple Malaria Strains

A second example of such models comes from the epidemiology of *Plasmodium falciparum*, the malaria-causing parasite. Section 4.2.2 of this chapter provides a detailed description of a modeling approach for malaria that includes the transmission via mosquito vectors—however, here we simplify the dynamics by approximating malaria by a directly transmitted infection. The basic reproductive ratio, R_0, of malaria is generally estimated from the increase with age, a, of the proportion of the population seropositive, (i.e., the proportion that is no longer susceptible, $1 - S(a)$). Standard theory (Chapters 2 and 3) predicts that when all ages suffer equal exposure to the pathogen, the level of susceptibles decays exponentially with age:

$$S(a) = \exp\left(-\frac{R_0 - 1}{L}\, a\right), \tag{4.3}$$

where $L = 1/\mu$ is the average life expectancy. This leads to an average age of first infection $A = L/(R_0 - 1)$. Because the average age of infection is around one year old, such calculations lead to estimates of R_0 in the range 50 to 100, making the eradication of malaria extremely difficult. However, these calculations assume that malaria is a single strain infection and that following infection there is complete immunity. However, detailed serological evidence suggests that a diverse range of antigenically distinct strains may cause the disease that is labeled malaria (Day and Marsh 1991).

Following the work of Gupta et al. (1994), we consider several strains of *Plasmodium falciparum*, each of which causes symptoms that are diagnosed as malaria, and suppose that the strains offer no cross-immunity, to each other. A relaxation of this assumption, so that strains confer partial immunity, does not radically change the general conclusions. The proportion of the population that is still susceptible to strain i at age a is given by:

$$S^i(a) = \exp\left(-\frac{R_0^i - 1}{L}\, a\right).$$

However, because there is no cross-immunity, strains act independently. Therefore, the proportion of individuals who are totally susceptible and have no malaria antibodies against *any* strain can be calculated as the independent (and therefore multiplicative) probability of being susceptible to each strain:

$$S^{\text{total}}(a) = \prod_i S^i(a) = \exp\left(-\sum_i (R_0^i - 1)\frac{a}{L}\right). \tag{4.4}$$

Thus, comparing equations (4.3) and (4.4), we see that the value of R_0 that is derived under the assumption of a single malaria strain is related to the sum of the separate R_0 values when multiple strains co-circulate.

$$R_0^{\text{estimated}} = 1 + \sum_i (R_0^i - 1)$$

Hence, the true value of R_0 for each strain is likely to be greatly reduced compared to standard estimates; Gupta et al. (1994) calculated that R_0 may be as low as 6 or 7. So that, instead of having one infectious disease that is very hard to eradicate, this analysis suggests that we have multiple strains, each of which may be substantially easier to eliminate. A control measure that could reduce the overall transmission rate to 10%–15% of its current value would be predicted to eradicate all strains simultaneously. Clearly such results have a profound impact on our understanding of the dynamics and potential control of malaria.

When there is limited cross-immunity, the individual values of R_0 for each strain are lower than estimated from the seropositive level that ignore strain structure.

4.1.3. Enhanced Susceptibility

We now focus on the situation where co-infection with two or more strains is more likely than pure chance would dictate. The classic example here is sexually transmitted infections, where the presence of one infection can increase the *susceptibility* of the host to other infections (Coggins and Segal 1998). This enhanced susceptibility can lead to some surprising results, as discussed below. In addition to these physiological factors, the risk structuring of the population also increases the level of co-infection as the high-risk core group is exposed to a higher force of infection for many sexually transmitted diseases (Chapter 3).

Sexually transmitted infections usually conform to the SIS paradigm, where after treatment infectious individuals are once again susceptible. There are four differential equations correspond to the two possible states (S and I) and the two infections:

$$\begin{aligned}
\frac{dN_{SS}}{dt} &= -\beta_1 N_{SS} I_1 - \beta_2 N_{SS} I_2 + \gamma_1 N_{IS} + \gamma_2 N_{SI} + \gamma_3 N_{II},\\
\frac{dN_{IS}}{dt} &= \beta_1 N_{SS} I_1 - \gamma_1 N_{IS} - \widehat{\beta}_2 N_{IS} I_2,\\
\frac{dN_{SI}}{dt} &= \beta_2 N_{SS} I_2 - \gamma_2 N_{SI} - \widehat{\beta}_1 N_{SI} I_1,\\
\frac{dN_{II}}{dt} &= \widehat{\beta}_1 N_{SI} I_1 + \widehat{\beta}_2 N_{IS} I_2 - \gamma_3 N_{II},\\
I_1 &= N_{IS} + N_{II}, \qquad I_2 = N_{SI} + N_{II}.
\end{aligned} \tag{4.5}$$

These equations assume that those individuals with both infections would be treated for both simultaneously (at rate γ_3); we also make the simplifying assumption that infections are passed on independently, such that those with both infections do not necessarily pass both on to each individual they infect. (The converse assumption does not affect the qualitative results discussed below.) We are now particularly interested in the case where being infected with one disease increases the susceptibility to the other, which translates to $\widehat{\beta}_1 > \beta_1$ and $\widehat{\beta}_2 > \beta_2$.

One interesting feature of such enhanced susceptibility is its effect on the invasion and persistence of the two infections. For either infection to invade a naive, totally susceptible population we require, as usual, that R_0 is greater than one; in particular, for each infection to be able to invade we need:

$$\beta_1 > \gamma_1 \qquad \text{and} \qquad \beta_2 > \gamma_2.$$

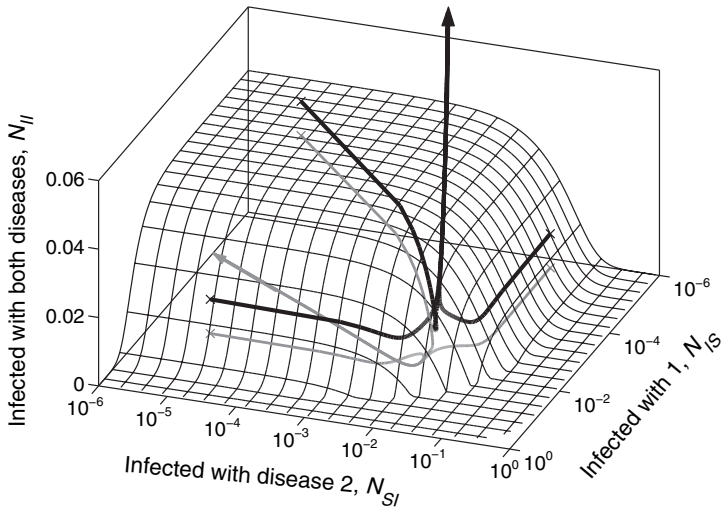


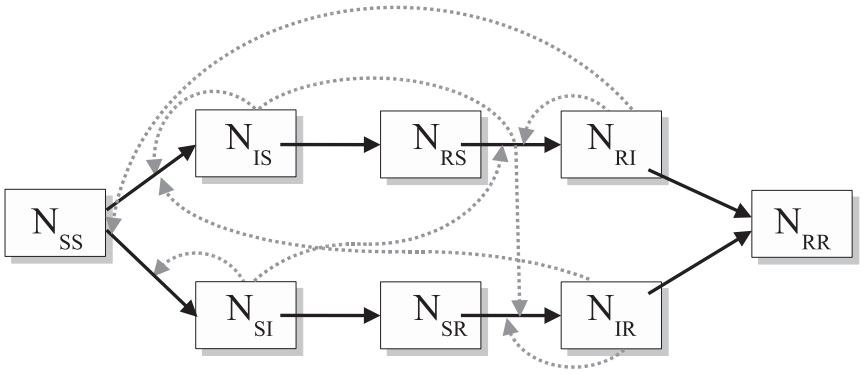
Figure 4.7. Example of six trajectories from the enhanced susceptibility model, equation (4.5), clearly demonstrating the Allee effect. The surface separating persistence from extinction is also shown as a mesh. Gray orbits start just below the surface and lead to extinction, whereas black orbits start just above the surface and tend to a fixed point with a high prevalence of both infections. ($\gamma_1 = \gamma_2 = \gamma_3 = 1$, $\beta_1 = 0.9$, $\beta_2 = 0.85$, $\widehat{\beta}_1 = 8$, $\widehat{\beta}_2 = 7$.)

At invasion, because prevalence is low and hence co-infection very rare, the terms $\widehat{\beta}_1$, $\widehat{\beta}_2$, and γ_3 do not enter into the invasion criterion. However, once the infections are established and co-infection common, these terms can play a pivotal role in maintenance of both infections. In particular, if β_1 and β_2 are small but $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are large, we can experience what is known in ecology as an Allee effect (Courchamp et al. 1999), whereby the infections cannot invade but may persist once they become established. This occurs when the basic reproductive ratio for the infections is below one so that neither can invade a totally susceptible population; however, given a substantial prevalence of infection 1, the average susceptibility of the population to infection 2 is increased and so infection 2 can persist—and vice versa. Figure 4.7 shows an example of this; trajectories (in black) that start above a critical prevalence tend to an endemic equilibrium, whereas orbits (in gray) that start below tend to zero. The mesh separates the regions of persistence and extinction.

Having one sexually transmitted infection can often increase the susceptibility to others, promoting co-infection. In such circumstances the Allee effect may operate, and reducing the prevalence of one infection may lead to a reduction of the other.

The somewhat extreme example of Figure 4.7, where susceptibility is greatly enhanced by infection, is chosen primarily to illustrate the effects of such nonlinear behavior. In practice the increase in susceptibility is generally far less, although the basic results may still hold. The presence of one sexually transmitted infection may make it difficult to eradicate others compared to a combined control strategy; small changes in sexual practices may cause significant changes in prevalence due to the nonlinear nature of the system and Allee effects may be observed. In conclusion, it may be erroneous to study sexually transmitted infections in isolation and a multi-disease approach may frequently be necessary.

4.1.4. Partial Cross-Immunity



Partial cross-immunity refers to the situation where by having experienced and recovered from one infection, or strain, provides some form of limited protection against other infections, or related strains. This form of immunity is the most commonly studied and most widely applicable formulation of the two-strain model. However, a range of differing assumptions can be made about the nature of cross-immunity. Protection can operate either through reduced susceptibility, reduced transmissibility, or a mixture of the two. There is the additional complication that protection might be conferred only to a faction of those individuals, but for the moment we shall assume a homogeneous response, with all recovered individuals experiencing the same reduction. Following a similar notation to before:

This is online program 4.1

$$
\begin{aligned}
\frac{dN_{SS}}{dt} &= \nu - \beta_1 N_{SS} I_1 - \beta_2 N_{SS} I_2 - \mu N_{SS},\\
\frac{dN_{IS}}{dt} &= \beta_1 N_{SS} I_1 - \gamma_1 N_{IS} - \mu N_{IS},\\
\frac{dN_{RS}}{dt} &= \gamma_1 N_{IS} - \alpha_2\beta_2 N_{RS} I_2 - \mu N_{RS},\\
\frac{dN_{SI}}{dt} &= \beta_2 N_{SS} I_2 - \gamma_2 N_{SI} - \mu N_{SI},\\
\frac{dN_{RI}}{dt} &= \alpha_2\beta_2 N_{RS} I_2 - \gamma_2 N_{RI} - \mu N_{RI},\\
\frac{dN_{SR}}{dt} &= \gamma_1 N_{IS} - \alpha_1\beta_1 N_{SR} I_1 - \mu N_{RS},\\
\frac{dN_{IR}}{dt} &= \alpha_1\beta_1 N_{SR} I_1 - \gamma_1 N_{IR} - \mu N_{IR},\\
\frac{dN_{RR}}{dt} &= \gamma_1 N_{IR} + \gamma_2 N_{RI} - \mu N_{RR},\\
I_1 &= N_{IS} + a_1 N_{IR}, \qquad I_2 = N_{SI} + a_2 N_{RI},
\end{aligned}
\tag{4.6}
$$

where α_i is the proportional reduction in the susceptibility to strain i and a_i is the proportional reduction in the transmission of strain i, for individuals who have recovered from the other strain. Clearly, with so many parameters a full description of all possible scenarios is a lengthy undertaking (White et al. 1998); instead we will highlight a few interesting and epidemiologically important points.

Suppose that one strain of the disease has already entered the population and reached equilibrium. The equilibrium fractions of susceptible, infectious, and recovered individuals are those of a single disease in isolation (Chapter 2):

$$N^*_{SS} = \frac{\gamma_1+\mu}{\beta_1}, \qquad N^*_{IS} = \frac{\mu}{\gamma_1+\mu} - \frac{\mu}{\beta_1}, \qquad N^*_{RS} = \frac{\gamma_1}{\gamma_1+\mu} - \frac{\gamma_1}{\beta_1}.$$

Whether or not the second strain can invade can be judged by looking at the growth rate of a small seed of infecteds, seeing whether strain 2 can increase ($\frac{dI_2}{dt} > 0$) when strain 1 is at equilibrium:

$$\begin{aligned}
\frac{dI_2}{dt} &= \frac{dN_{SI}}{dt} + a_2\frac{dN_{RI}}{dt} = \beta_2 N^*_{SS} I_2 + a_2\alpha_2\beta_2 N^*_{RS} I_2 - \gamma_2 I_2 - \mu I_2, \\
&= \beta_2\frac{\gamma_1+\mu}{\beta_1} I_2 + a_2\alpha_2\beta_2\left(\frac{\gamma_1}{\gamma_1+\mu} - \frac{\gamma_1}{\beta_1}\right) I_2 - \gamma_2 I_2 - \mu I_2, \\
&= \beta_2\left[\frac{\gamma_1+\mu}{\beta_1} - \frac{\gamma_2+\mu}{\beta_2}\right] I_2 + a_2\alpha_2\left(\frac{\gamma_1}{\gamma_1+\mu} - \frac{\gamma_1}{\beta_1}\right) I_2.
\end{aligned}$$

So strain 2 can invade if it has a higher reproductive ratio than strain 1, which implies that $\beta_2(\gamma_1+\mu) > \beta_1(\gamma_2+\mu)$; or if the effects of cross-immunity are not particularly strong:

$$a_2\alpha_2 > \frac{\beta_1(\gamma_1+\mu)\left[\beta_1(\gamma_2+\mu) - \beta_2(\gamma_1+\mu)\right]}{\gamma_1(\beta_1-\gamma_1-\mu)\beta_2} \approx \frac{R_0^1(R_0^1-R_0^2)(\gamma_1+\mu)}{R_0^2(R_0^1-1)}. \tag{4.7}$$

By symmetry, we can obtain a similar set of conditions by allowing strain 2 to reach equilibrium and then seeing whether strain 1 can invade. We assume that the strain with the larger R_0 is labeled strain 1; then if condition (4.7) holds, both strains can invade when the other is at equilibrium and hence coexistence is possible. In fact, more detailed analysis shows that under these assumptions any initial condition (where both strains are present) leads to the two strains coexisting at some equilibrium.

Coexistence of strains is possible when their respective R_0 values are close and cross-immunity is weak.

The inclusion of partial immunity bridges the gap between complete cross-immunity where only one strain persists and no cross-immunity where both strains always coexist. Because the coexistence condition (4.7) contains only the product $a_2\alpha_2$, it is irrelevant for a two-strain model, whether partial cross-immunity acts on transmission (a), susceptibility (α), or a mixture of the two. However, when the competition between more than two strains is considered, differences between reduced transmission and reduced susceptibility do occur. To understand these differences, consider a three-strain model, in a triangular arrangement, where strain 1 confers complete immunity against strain 2, strain 2 confers complete immunity against strain 3, and strain 3 confers immunity against strain 1, with

no other interactions. In the reduced susceptibility model, an individual recovered from strain 1 cannot catch strain 2 when challenged, but can later catch strain 3. However, in the reduced transmissibility model, once an individual has recovered from strain 1 he or she can catch (but not transmit) strain 2, which provides immunity against strain 3. This simple conceptual example shows how a detailed understanding of the host's immunological responses is necessary to deal with realistic multi-strain models because the implications of the basic assumptions are profound.

When many cross-reactive strains are circulating within the population, different assumptions about the nature of cross-reactivity can have profound effects on the modeling outcome. Detailed immunological studies are required to clarify the typical behavior.

4.1.4.1. Evolutionary Implications

The partial cross-immunity model can again be considered within an evolutionary setting, by examining competition between strains. In general, most interest has focused on the evolutionary dynamics of influenza, where new strains continually arise and face little herd immunity within the population (Fitch et al. 1997; Earn et al. 2002); models for this type of dynamics therefore must include the interaction of multiple strains (Andreasen et al. 1996; Andreasen et al. 1997; Gomes et al. 2001; Gog and Swinton 2002; Gog and Grenfell 2002; Ferguson et al. 2003a). However, as shown above, this competition does not have to lead to the exclusion of one strain; coexistence is possible. What makes the inclusion of partial cross-immunity so relevant to evolutionary modeling is that when multiple strains are considered, it is plausible to assume that closely related strains have a high level of cross-immunity whereas distantly related strains display little or no cross-immunity. This basic rule has much supporting evidence (de Jong et al. 2000; Earn et al. 2002). Significant jumps in strain structure, due to recombination, may produce strains with high virulence and little population-level immunity; these jumps are epidemiologically important because they may lead to pandemics, as seen 1918 (Gibbs et al. 2001).

The assumption of high cross-immunity for closely related strains has a strong evolutionary implication. The vast majority of new strains, which are genetically and phenotypically close to their parent strain, will face the same level of immunity in the population as their parent and therefore are unlikely to rise to dominance. It is only the rare distant mutation that, facing little conferred immunity, can increase dramatically. Thus, evolution does not lead to the gradual change, but instead progresses by a series of jumps whenever sufficiently distant mutants occur.

The great difficulty with multi-strain models that attempt to track the evolutionary progress of diseases is the rapid proliferation in the number of equations with the number of strains (Gomes et al. 2001). To completely capture the dynamics of n interacting strains within the SIR framework requires 3^n differential equations, or $(n+2)2^{n-1}$ equations if co-infection is ignored. Clearly this sort of exponential growth in model complexity places severe restrictions on the number of strains that can be simulated, even with modern computational technologies. However, substantial simplifications to the model can be made given two reasonable assumptions (Andreasen et al. 1996; Gog and Swinton 2002; Gog and Grenfell 2002). The first is that immunity acts to reduce transmission, not susceptibility; this means that all individuals are equally at risk to any strain irrespective of their epidemic history. Physiologically, this assumption implies that even when an

individual has immunity to a particular strain, if challenged by that strain they still become infected and mount a full immune response, but clear the pathogen before transmission can occur. The second assumption is that partial immunity acts heterogeneously, by making a proportion of recovered individuals totally immune, rather than the homogeneous assumption when all recovered individuals have an equally reduced transmission rate. These two assumptions mean that we need to consider only the proportion of individuals "susceptible to" and infectious with each strain. Hence for a model with n strains there are only $2n$ equations:

$$\begin{aligned} \frac{dS_i}{dt} &= \nu - \sum_j \beta_j c_{ij} S_i I_j - \mu S_i, \\ \frac{dI_i}{dt} &= \beta_i S_i I_i - \gamma_i I_i - \mu I_i, \end{aligned} \tag{4.8}$$

where $0 \leq c_{ij} \leq 1$ provides information on the gain of complete immunity to strain i due to infection (or challenge) with strain j. We would normally insist that $c_{ii} = 1$ so that the single strain dynamics are SIR. S_i refers to the fraction of the population who are not immune and therefore able to transmit strain i. Hence, the susceptibility classes are not mutually exclusive; an individual susceptible to (i.e., able to catch and subsequently transmit) strains 1 and 2 will belong to both S_1 and S_2. Therefore, the population components no longer sum to one.

This change of emphasis, from discrete compartments to overlapping classes, can at first be difficult to grasp. The reasoning behind equation (4.8) can be explained as follows. The I_i equation is relatively straightforward: Only those individuals without immunity to the strain (S_i) can be infectious following transmission. Those individuals who are immune and become infected, but are unable to transmit, do not play a role in the spread of infection and can be ignored. The number of individuals who are susceptible to strain i are not just reduced due to infection with strain i, but other strains can also affect the immunity (captured by c_{ij}). By assuming that *all* individuals can catch *all* strains, we do not need to worry about the immune status of an individual with respect to strain j, just the effect of strain j on the immunity with respect to strain i.

This simple, but powerful, model can now be used to investigate antigenic drift which is evolution driven by the immunity of the population with no discernible change in infection characteristics (so $\beta_i = \beta$ and $\gamma_i = \gamma$). Following Gog and Grenfell (2002) (see also Andreasen et al. 1996), we position strains on a one-dimensional line and assume that immunity is conferred most strongly to nearby strains:

$$c_{ij} = \exp(-A[i-j]^2).$$

It is also presumed that random mutation, at a rate ε, can lead to the spontaneous creation of adjacent strains. This modifies the basic equation:

$$\frac{dI_i}{dt} = \beta S_i I_i - \gamma I_i - \mu I_i - \varepsilon I_i + \frac{1}{2}\varepsilon I_{i+1} + \frac{1}{2}\varepsilon I_{i-1}. \tag{4.9}$$

Figure 4.8 shows an example of the type of drift dynamics predicted by this model. With 100 strains, a traditional approach that tracks all infection histories would require over 6×10^{31} equations—this is computationally infeasible. Two clear features emerge from this model that mimic the observed behavior of influenza (Fitch et al. 1997; Earn

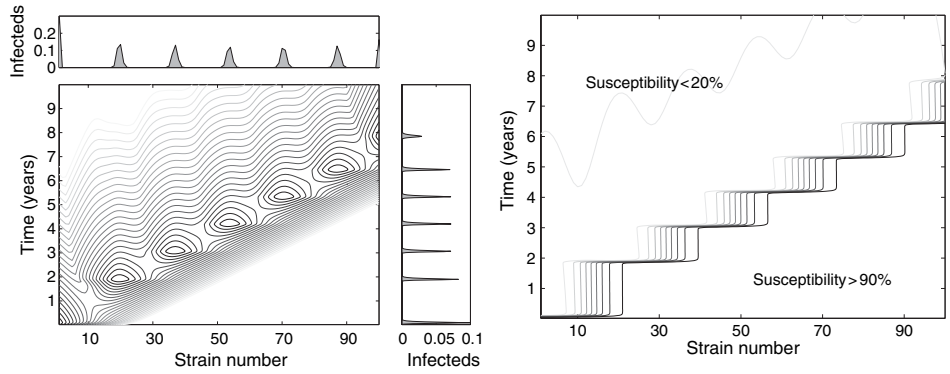


Figure 4.8. Dynamics of a strain-structured model of influenza (equations (4.8) and (4.9)), where partial immunity is conferred to nearby strains and is assumed to act heterogeneously on the transmission rate. The left-hand composite of three graphs show the prevalence of infection within the population. The main graph gives the strain-specific level over time $I_i(t)$ (contours plotted on logarithmic scale, with darkest contours referring to highest prevalences of infecteds); the top graph shows the total prevalence of infection with each strain $\int I_i(t)dt$ over the 10-year simulation; the right graph shows the total infected against time $\sum_i I_i(t)$. The right-hand graph gives the level of "susceptibility" to each strain $S_i(t)$; here the contours are linearly spaced. ($\nu = \mu = 5 \times 10^{-5}$ per day, $1/\gamma = 7$ days, $R_0 = 3$, $A = 0.01$, $\varepsilon = 0.01$ per year.)

et al. 2002). First, there are clear oscillations in the fraction of infecteds (summed over all strains) driven by the strain structure and rate of mutation; in practice, any oscillations in the prevalence of influenza are reinforced by seasonal effects (however, see Dushoff et al. 2004 and Viboud et al. 2006). Second, due to the cross-immunity invoked towards nearby strains, the next epidemic strain must be sufficiently distinct from previous strains; the time taken for these new distant mutants to arise sets the duration between epidemics.

Although such models are a powerful tool for providing insights into the evolutionary dynamics of diseases, the assumed structure of strain space is somewhat naive. A better approximation would be to assume that the space is high-dimensional (Gupta and Maiden 2001), and that, although in general most mutations give rise to nearby strains, some mutations (or recombination events) may cause a departure to more distant regions of strain space where there is little cross-immunity. These large jumps, known as *antigenic shift*, have profound public health implications and are likely to lead to far larger epidemics than normal. Genetic evidence suggests that the 1918 influenza pandemic that killed 10–20 million people was due to the recombination of co-circulating strains (Gibbs et al. 2001). It remains a considerable challenge to determine the precise structure of this strain space and the rate of mutation between various points. The current concerns over H5N1 influenza in poultry emphasizes the need for a detailed genetic understanding of immunity, virulence, and the factors affecting the ability to transmit between humans.

4.1.4.2. Oscillations Driven by Cross-Immunity

Persistent, large amplitude epidemic cycles are generally considered to be a signature of underlying seasonal effects or temporal forcing (Chapter 5), with the majority of unforced models asymptoting to a fixed prevalence. The dynamics of multi-strain models are an

exception to this rule. Consider the behavior seen in Figure 4.8. If the strains were not positioned along a line, but the ends were joined together to make a circle, then an everlasting wave of epidemics could propagate around the circle. By the time one complete revolution of the circle has been made, there will have been sufficient births to increase susceptibility to the initial strain.

Slight variants of equations (4.8) and (4.9) can produce self-sustained oscillations with as few as four interacting strains (Andreasen et al. 1997; Gupta et al. 1998; Gomes et al. 2002; Dawes and Gog 2001). Following the work of Gupta et al. (1998), we consider the dynamics of four interacting strains in a circular arrangement where recovery from one strain offers partial protection to neighboring strains (e.g., recovery from strain 1 offers partial protection against strains 2 and 4, recovery from strain 2 offers partial protection to strains 1 and 3, etc.). Using the notation of Gupta et al. (1998), we have:

This is online program 4.2

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu - S_i \sum_j c_{ij}\lambda_j - \mu S_i, \\
\frac{dP_i}{dt} &= S_i \sum_{j\neq i} c_{ij}\lambda_j - \beta P_i I_i - \mu P_i, \\
\frac{dR_i}{dt} &= (S_i + P_i)\lambda_i - \mu R_i, \\
\frac{d\lambda_i}{dt} &= [S_i + a P_i]\,\lambda_i - \gamma I_i - \mu I_i,
\end{aligned} \tag{4.10}$$

where S_i, P_i, and R_i are those totally susceptible to, partially susceptible to, and fully immune to (either recovered from or infected with) strain i, and λ_i is the force of infection associated with strain i. (c_{ij} is one if i and j are neighboring strains, but zero otherwise.) Compared to the more complete notation given in equation (4.6), the above model assumes that partial immunity decreases an individual's transmissibility ($a < 1$) acting homogeneously, but does not affect its susceptibility ($\alpha = 1$); all strains have identical characteristics. In particular, we can map the terms in equation (4.10) to those in equation (4.6); for strain 1 we have:

$$\begin{aligned}
S_1 &= \sum_{Q\in\{S,I,R\}} N_{SSQS}, \\
P_1 &= \sum_{Q\in\{S,I,R\}} \left[\sum_{A\in\{S,I,R\}} \sum_{B\in\{S,I,R\}} N_{SAQB} - N_{SSQS} \right], \\
R_1 &= \sum_{Q,A,B\in\{S,I,R\}} N_{R,A,Q,B} + N_{I,A,Q,B}, \\
\lambda_1 &= \beta \sum_{Q\in\{S,I,R\}} N_{ISQS} + a\beta \sum_{Q\in\{S,I,R\}} \left[\sum_{A\in\{S,I,R\}} \sum_{B\in\{S,I,R\}} N_{IAQB} - N_{SSQS} \right],
\end{aligned}$$

so that those individuals totally susceptible to strain 1 must not have encountered strains 1, 2, or 4, but could be in any state with respect to strain 3 because strains 1 and 3 do not interact. However, those who are partially resistant to strain 1 must not have encountered strain 1, but must have encountered either strain 2 or 4. Similar arguments can be made to construct the other terms.

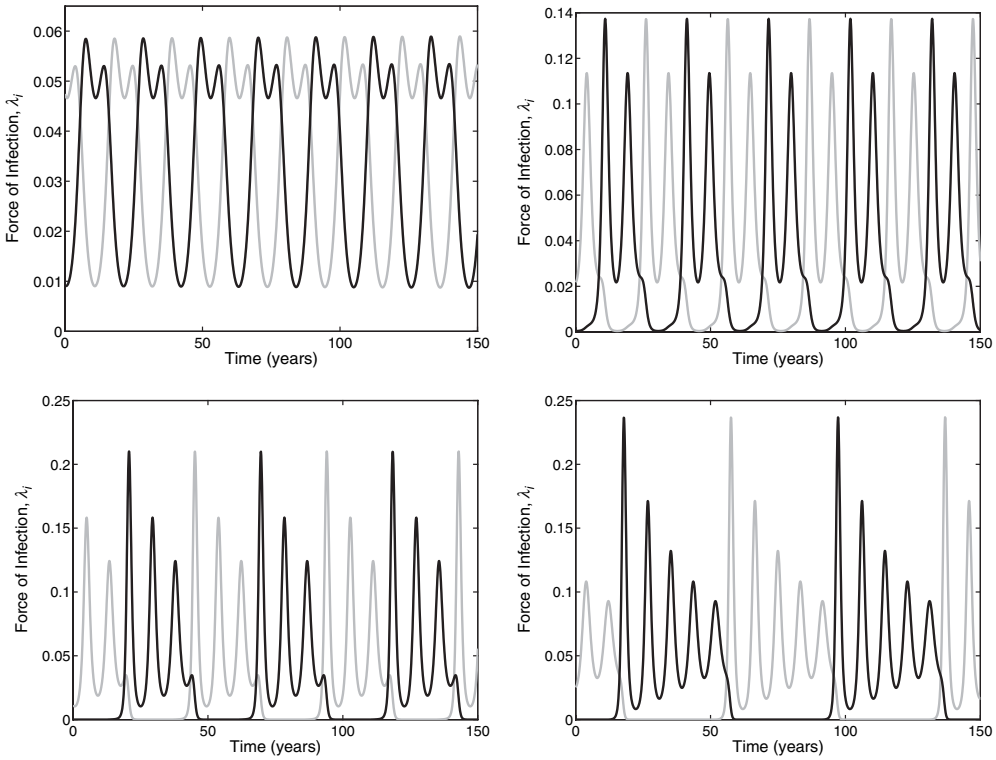


Figure 4.9. The dynamics of strains 1 and 2 of a four-strain system as typified by the force of infection for each strain, λ_i. The level of cross-immunity, a, increases from top left to bottom right ($a = 0.55, 0.6, 0.65, 0.7$). ($\mu = 0.02$ per year, $\gamma = 10$ per year ($1/\gamma = 36.5$ days), $\beta = 40$ per year, hence $R_0 = 4$.)

Figure 4.9 shows that the dynamics of strains 1 and 2 (strains 3 and 4) behave identically as the level of cross-immunity, a, varies. Increasing the level of cross-immunity increases the period of the dynamics and the complexity of the epidemic cycles; values of a less than about 0.53 lead to coexisting equilibrium dynamics where all strains asymptote to the same constant level. It is important to note that sustained regular oscillations are not always the signature of seasonal forcing. When $a = 0.55$ the period is around 21 years, whereas when $a = 0.7$ the period has increased to about 79 years. Although the existence of such cycles is inherently interesting, it is questionable whether such long-term oscillations play any meaningful role in the dynamics of influenza. However, recent work on dengue fever (Wearing and Rohani 2006), respiratory syncytial virus (White et al. 2005), and cholera (Koelle et al. 2005) all show the propensity for multi-strain diseases to exhibit complex cycles. More long-term strain-structured data and a more detailed understanding of the levels of partial immunity are required before the implications of these large-amplitude fluctuations can be practically assessed or their dynamics predicted.

Strain structure and partial cross-immunity between nearby strains can lead to long-period oscillatory dynamics without the need for external forcing.

The interaction between partially cross-immune strains can therefore lead to epidemic cycles in the total infection prevalence (summed across all strains). Two other long-term behaviors of cross-reactive multi-strain models exist, and both are readily achievable using equations (4.8) and (4.9). The first, and most simple, is homogeneous equilibria where all strains asymptote to the equilibrium level of abundance in the population. The second, and more interesting, is heterogeneous equilibria where some strains persist at a much higher prevalence than others; which strains are most abundant depends on the initial conditions. These three scenarios (temporal oscillations, homogeneous abundance, and heterogeneous equilibria) have interesting parallel with Turing patterns (Turing 1952; Murray 1982). Strain-structured models possess local suppression (in terms of local immunity), activation (in terms of susceptibles), and diffusion (in terms of mutation); it is therefore not surprising that Turing-like dynamics can occur in strain space.

Models of strain structure with local immunity and mutation can lead to traveling waves (observed as dynamic oscillations) or large amplitude stationary patterns in strain space, parallelling Turing patterns.

4.1.5. A General Framework

Finally, we present a comprehensive, flexible mathematical framework that incorporates various possible interactions between pathogens (Rohani et al. 2006). The framework is presented for two infectious agents (labeled disease 1 and 2), though extending it to include multiple pathogens is straightforward although lengthy. In developing the model, we envisage a simplified natural history of infection for each disease:

- All newborns are fully susceptible to both infections.
- Upon infection, a susceptible individual enters the exposed (infected but not yet infectious) class, and has a probability of contracting the "competing" infection simultaneously (represented by the cross-immunity parameter ϕ_i, where $i = 1, 2$).
- After the latent period, the individual becomes infectious but is not yet symptomatic and still has a reduced risk (ϕ_i) of becoming co-infected with the other disease.
- Typically, when symptoms appear, the disease is diagnosed and the individual is sent home to convalesce for an average period given by $1/\delta_i$. During convalescence, the competing infection may be contracted, with the transmission rate additionally modulated by the parameter ξ_i, which may represent quarantine or temporary cross-immunity (if less than one) or temporary immuno-suppression (if greater than one).
- Depending upon the disease, host age, and host condition (typically nutritional status), infection may be fatal owing to complications (such as pneumonia and encephalitis, in the case of measles and pertussis). This is represented by per capita infection-induced mortality probabilities ρ_i. It is assumed that mortality occurs at the end of the convalescent period, so that the effects of mortality can be separated from the effects of the infectious and convalescent period. This is a very different assumption to that used in Chapter 2, although it is equivalent under a (complex) change of variables.
- Upon complete recovery, the individual is assumed immune to the infection (disease 1) and reactivates susceptibility to disease 2, if previously not exposed to it. At this stage, we introduce the term α_i to explore the implications of long-lasting immuno-suppression ($\alpha_i > 1$) or cross-immunity ($\alpha_i < 1$) for the susceptibility to disease j following infection with disease i.

The model incorporates a large number of possible mechanisms for interaction among infections. At the immunological level, the parameter α_i represents the long-term cross-immunity or immuno-suppression resulting from becoming infected with infection i. The term ϕ_i represents the extent of cross-immunity to infection j while individuals are experiencing infection i. At the ecological level, the convalescent class and possible subsequent death following infection give rise to competition among infections for susceptibles. Hence, depending on the system of interest, the model can be adapted accordingly.

Mathematically, these assumptions lead to the following set of ordinary differential equations:

This is online program 4.3

$$
\begin{aligned}
\frac{dS}{dt} &= \nu N - (\lambda_1 + \lambda_2)S - \mu S,\\
\frac{dE_1}{dt} &= \lambda_1 S - \phi_2\lambda_2 E_1 - (\sigma_1 + \mu)E_1,\\
\frac{dE_2}{dt} &= \lambda_2 S - \phi_1\lambda_1 E_2 - (\sigma_2 + \mu)E_2,\\
\frac{dI_1}{dt} &= \sigma_1 E_1 - \phi_2\lambda_2 I_1 - (\gamma_1 + \mu)I_1,\\
\frac{dI_2}{dt} &= \sigma_2 E_2 - \phi_1\lambda_1 I_2 - (\gamma_2 + \mu)I_2,\\
\frac{dC_1}{dt} &= \gamma_1 I_1 - \xi_2\phi_2\lambda_2 C_1 - (\delta_1 + \mu)C_1,\\
\frac{dC_2}{dt} &= \gamma_2 I_2 - \xi_1\phi_1\lambda_1 C_2 - (\delta_2 + \mu)C_2,\\
\frac{dR_1}{dt} &= (1-\rho_1)\delta_1 C_1 - \alpha_2\lambda_2 R_1 - \mu R_1,\\
\frac{dR_2}{dt} &= (1-\rho_2)\delta_2 C_2 - \alpha_1\lambda_1 R_2 - \mu R_2,\\
\frac{dR_{12}}{dt} &= (1-\rho_1)(1-\rho_2)\left(\lambda_2\phi_2 E_1 + \phi_2 I_1 + \xi_2\phi_2 C_1 + \lambda_1\phi_1 E_2 + \phi_1 I_2 + \xi_1\phi_1 C_2\right),\\
&\quad +(1-\psi_2\rho_2)\alpha_2\lambda_2 R_1 + (1-\psi_1\rho_1)\alpha_1\lambda_1 R_2 - \mu R_{12},\\
\frac{d\epsilon_1}{dt} &= \lambda_1 S + \phi_1\lambda_1 E_2 + \phi_1\lambda_1 I_2 + \xi_1\phi_1\lambda_1 C_2 + \alpha_1\lambda_1 R_2 - (\sigma_1 + \mu)\epsilon_1,\\
\frac{d\epsilon_2}{dt} &= \lambda_2 S + \phi_2\lambda_2 E_1 + \phi_2\lambda_2 I_1 + \xi_2\phi_2\lambda_2 C_1 + \alpha_2\lambda_2 R_1 - (\sigma_2 + \mu)\epsilon_2,\\
\frac{d\lambda_1}{dt} &= \beta_1\sigma_1\epsilon_1 - (\gamma_1 + \mu)\lambda_1,\\
\frac{d\lambda_2}{dt} &= \beta_2\sigma_2\epsilon_2 - (\gamma_2 + \mu)\lambda_2\,.
\end{aligned}
$$

This model represents an example of *history*-based formulation, rather than *status*-based. The variables require some explanation. All those susceptible to both infections are denoted by S. The variables E_i, I_i, and C_i ($i = 1, 2$) represent those currently exposed, infectious, or convalescing (respectively) after infection with disease i, with no previous exposure to

the other pathogen. The term R_i $(i = 1, 2)$ represents all individuals who have previously experienced infection i and as a result are now only (partially) susceptible to infection j $(j \neq i)$. For bookkeeping purposes, we let ϵ_i and λ_i/β_i represent individuals latent and infectious with disease i $(i = 1, 2)$, irrespective of their status for the other disease. Additionally, R_{12} are all those no longer susceptible to either infection, and may include those who are still exposed or infectious with one or both diseases and expected to fully recover. Thus, in terms of the parameters used earlier: $S = N_{SS}$, $E_1 = N_{ES}$, $I_1 = N_{IS}$, and so forth, $E_2 = N_{SE}$, and so forth, $\epsilon_1 = \sum_Q N_{EQ}$, $\epsilon_2 = \sum_Q N_{QE}$, $\lambda_1 = \beta_1 \sum_Q N_{IQ}$, $\lambda_2 = \beta_2 \sum_Q N_{QI}$. However, this new notation has the distinct advantage that the parameters ϵ_i and λ_i provide a useful shorthand. The total population size (N) is the sum of the first ten variables only $(N = S + \sum_{i=1}^{2}(E_i + I_i + C_i + R_i) + R_{12})$. The full derivation of this model is presented in detail by Vasco et al. (2007). The model's parameters are explained in Table 4.1.

TABLE 4.1.
Description of model parameters. Subscripts refer to disease i $(i = 1, 2)$.

Parameter	*Epidemiological Description*	*Typical Range*
ν	*Per capita* birth rate	0.01–0.5 per year
μ	*Per capita* death rate	0.01–0.5 per year
$1/\sigma_i$	Latent period	1–2 weeks
$1/\gamma_i$	Infectious period	1–3 weeks
$1/\delta_i$	Quarantine period	1–4 weeks
ρ_i	Probability of infection-induced mortality	0–1
ϕ_i	Co-infection probability	0–1
ξ_i	Temporary immuno-suppression/cross-immunity	≥ 0
α_i	Permanent immuno-suppression/cross-immunity	≥ 0
ψ_i	Differential infection-induced mortality	0–1

One intuitively obvious possible consequence of interaction among infections is reduced abundance. Surprisingly, however, detailed analyses have demonstrated that disease interference does not manifest itself by significantly altering infection prevalence; changes in model parameters such as the convalescence period translate into negligible changes in the number of infectives of either infection (Huang and Rohani 2005). Perhaps more surprisingly, epidemiological interference exerts little influence on the coexistence likelihood of pathogens. Defining the basic reproductive ratio of each infection as $R_0^j = \beta_j \sigma_j / (\sigma_j + \mu)(\gamma_j + \mu)$ $(j = 1, 2)$, it is straightforward to show that coexistence requires $R_0^j > 1$ and

$$R_0^j > \frac{R_0^i}{1 + a_i(R_0^i - 1)}, \tag{4.11}$$

where

$$a_i = \frac{1}{\sigma_i + \mu}\left\{\phi\mu + \frac{\sigma_i}{\gamma_i + \mu}\left(\phi\mu + \frac{\gamma_i}{\delta_i + \mu}(\xi\phi\mu + \alpha(1 - \rho_i)\delta_i)\right)\right\}, \tag{4.12}$$

where $i, j = 1, 2$, $j \neq i$, and the diseases are assumed to have symmetric values of ϕ, α, and ξ (such that, for example, $\phi_1 = \phi_2 = \phi$; details provided in Vasco et al. 2007).

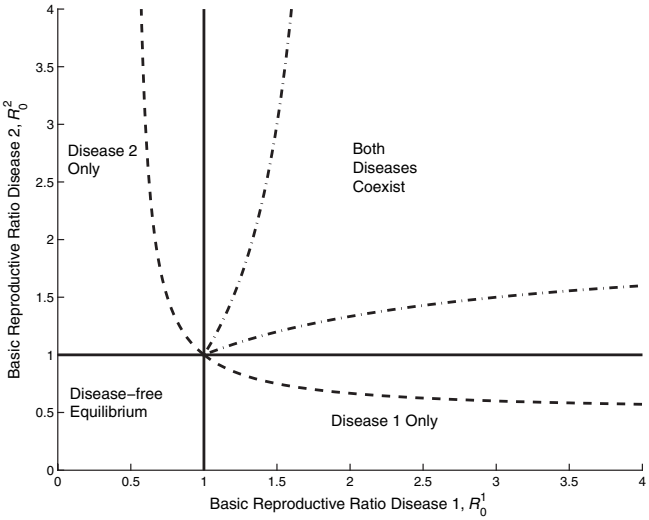


Figure 4.10. The figure demonstrates that coexistence of the two infections can be affected by immuno-suppression and disease interference. In the absence of immune-mediated interactions ($\phi_i = \alpha_i = 1$, $i = 1, 2$), large levels of disease-induced mortality ($\rho_i = 50\%$: dot-dashed line), can cause the region of two-disease coexistence to shrink somewhat. In contrast, strong levels of permanent immuno-suppression ($\phi_i = 1, \alpha_i = 2, \rho_i = 0, i = 1, 2$: dashed line) can expand the coexistence domain. Taken from Vasco et al. 2007. Model parameters were $\mu = 0.02$, $1/\sigma_1 = 1/\sigma_2 = 8$ days, $1/\gamma_1 = 5$ days, $1/\gamma_2 = 14$ days, $\xi = 1$, $1/\delta_1 = 7$ days, and $1/\delta_2 = 14$ days.

In Figure 4.10, we explore the conditions for disease coexistence in this model. In the absence of pathogen-induced mortality ($\rho_1 = \rho_2 = 0$) and with no long-term immunological interactions ($\alpha = 1$), the lack of coinfection alone has little effect on the stable two-disease equilibrium, with the coexistence criterion effectively reducing to R_0^1, $R_0^2 > 1$. It is only after we assume a 50% (dash-dotted line) probability of death following infection that the region of endemic coexistence of both diseases shrinks slightly. On the other hand, if we ignore ecological factors (such as pathogen virulence), immuno-suppression resulting from one infection can facilitate the invasion and persistence of the competing disease even if the invading infection has R_0 lower than one (dashed line).

4.2. MULTIPLE HOSTS

Although many diseases are host-specific, many others can infect multiple and often highly diverse species (Woolhouse et al. 2001b). The models for these types of disease mirror the risk-structured framework developed in Chapter 3, with species being the important risk factor. However, in contrast to risk-structured models, different species may have very different epidemiological and physiological responses to the same infection. In one species a infection may be short-lived and highly virulent, whereas in another species long-term chronic infection may be the norm; the spread of infection between two such host species is a complex problem that can only be understood with mathematical models. We focus on three distinct scenarios that cover a wide spectrum of infections.

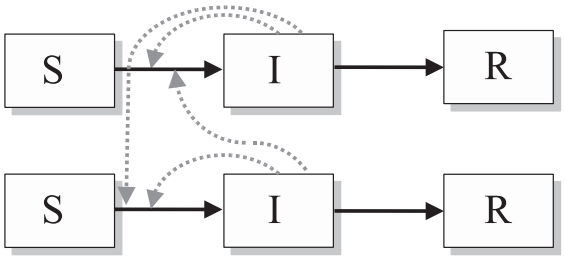
First we consider two (or more) host species and a single disease that can be transmitted both within and between the species. This kind of interaction is especially of interest because, in some cases, the host species do not directly interact. The presence of a shared natural enemy—the infection—gives rise to an indirect or "apparent" competition (Holt 1977). A classic high-profile example of such a system is bovine-tuberculosis, where great attention has been focused on the spread of infection between badgers and cattle. Bovine-tuberculosis also exemplifies the parameterization difficulties that are encountered. In modeling terms, the extent of disease spread from badgers to cattle simply relates to one term in the transmission matrix and yet it is the subject of continual controversy despite many years of research (Krebs 1997; Bourne et al. 2000). Another example of a multi-host-pathogen system is foot-and-mouth virus, which can infect a large variety of livestock species such as cattle, sheep, and pigs, despite their physiological differences. Foot-and-mouth is a major problem for farmers in many areas of the world, and understanding the role that different species play in its transmission and persistence is vital for effective, and often species-specific, control measures.

The second type of multi-host systems are vector-transmitted diseases, such as malaria or dengue fever, which require a secondary "host" to spread infection between primary hosts. For both malaria and dengue fever this secondary host is the female mosquito, which spreads the infecting pathogen as it takes blood-meals from humans or other primary hosts. Vectors are almost always arthropods, and include a range of blood-sucking parasites such as fleas, ticks, lice, and mosquitoes. Unlike the range of infectious diseases considered so far, close contact between infected and susceptible humans is not a requisite of transmission; instead, vectors can spread the infection over a wide range. Infectious vectors can even be carried for thousands of miles in aircraft, promoting public health fears of transporting these vector-born diseases to naive, previously disease-free populations. Models for vector-transmitted diseases follow a similar pattern to standard multi-species models, but the parameterization tends to be simpler due to the absence of within-species transmission; mathematically, the diagonal terms of the transmission matrix are zero. Additionally, the rapid life cycle of the vector compared to epidemic timescales can be used to further simplify the modeling.

Finally we focus on zoonoses; these are infections of animals that can also be transmitted to humans, and therefore form a special class of multi-host model. In general, the animal population is the main reservoir for the infection and human cases are sporadic. Often these infections are also vector-transmitted, so that direct contact between humans and the reservoir species is not required. Examples of zoonoses include such high-profile diseases as bubonic plague, West Nile virus, and Ebola. The existence of an animal reservoir and the sporadic nature of transmission to humans has several implications. Without detailed surveillance of the reservoir species, the pattern of human cases may appear confusing and spontaneous. Additionally, by the time human cases arise and public health agencies are aware of the disease's presence, an epidemic with the reservoir species may be difficult to control. Finally, zoonotic diseases are often (re-)emerging pathogens, so outbreaks are unexpected, and experience of control measures is limited.

In the work that follows, due to the fact that we are dealing with different species whose populations may fluctuate independently, it will be prudent to utilize numbers rather than proportions. Hence we shall work with X for the number of susceptible individuals, rather than the proportion S.

4.2.1. Shared Hosts



The approach to modeling a single infectious disease that can be transmitted within and between two different species has strong parallels with the work on risk-structured models (Chapter 3) where there is transmission within and between two different risk groups. Thus for two species (A and B), the SIR-type dynamics will be given by

$$\begin{aligned}
\frac{dX_A}{dt} &= \nu_A - X_A(\beta_{AA}Y_A + \beta_{AB}Y_B) - \mu_A X_A,\\
\frac{dY_A}{dt} &= X_A(\beta_{AA}Y_A + \beta_{AB}Y_B) - \gamma_A Y_A - \mu_A Y_A,\\
\frac{dX_B}{dt} &= \nu_B - X_B(\beta_{BA}Y_A + \beta_{BB}Y_B) - \mu_B X_B,\\
\frac{dY_B}{dt} &= X_B(\beta_{BA}Y_A + \beta_{BB}Y_B) - \gamma_B Y_B - \mu_B Y_B.
\end{aligned} \tag{4.13}$$

The birth rates for the two species, ν_A and ν_B, may include complex density-dependent terms, such that in the absence of infection the population levels tend to a carrying capacity. The transmission term is not divided by the population size because we are dealing with two separate populations and the interaction is likely to depend on the density of the two species—we have assumed pseudo mass-action transmission. The main distinction from risk-structured models of Chapter 3 is that different species are likely to have different responses to infection and thus differing transmission rates and recovery periods, as well as differing demographic parameters. The effect of the different transmission parameters is to break the symmetry that is usually associated with the transmission matrix, $\boldsymbol{\beta}$, of risk-structured models. This is because although the mixing between species A and species B is the same as the mixing between B and A, the transmission may be far stronger in one direction because one species may shed more pathogen than the other, or may have different physiological responses to infection.

In multi-host models, the transmission matrix is no longer expected to be symmetric, due to species differences.

From the perspective of wildlife species, it is useful to incorporate well-documented density-dependent population regulation into the modeling framework. This is done by Dobson (2004), who derived analytical expressions for R_0 under alternative assumptions of disease transmission (also see Diekmann et al. 1990). He found that host species diversity has an amplifying effect on outbreaks when transmission is density-dependent (pseudo mass-action; see Chapter 4), as might be the case for pathogens with a free living stage or transmitted by aerosol. On the other hand, for vector-borne diseases, where

transmission is frequency-dependent (mass-action), increasing the number of host species has a detrimental effect on pathogen prevalence (see Section 4.2.2.1). Dobson (2004) also demonstrated that pathogen persistence is influenced by the relative strength of between- and within-species transmission, with greater heterospecific transmission leading to greater likelihood of long-term extinction.

4.2.1.1. Application: Transmission of Foot-and-Mouth Disease

The cause and implications of the asymmetry in the transmission matrix can be more readily seen by example. We focus on the spread of foot-and-mouth disease (FMD), which is a highly contagious infection with SIR-type dynamics that is rapidly transmitted between a variety a livestock, especially cattle, sheep, and pigs. Estimates from the 2001 epidemic within the United Kingdom (Keeling et al. 2001b) showed that FMD is transmitted slightly better by cattle than sheep (ratio cattle:sheep = 1.8 : 1) and that cattle are far more susceptible (ratio 15 : 1), although these factors are somewhat offset by the higher numbers of sheep within the United Kingdom (ratio 1 : 4.2). Hence, sheep and cattle respond very differently to this infection, so any attempt at prediction must recognize these differences and model the two species separately. Here, the susceptible and transmission ratios not only incorporate innate differences between the species, but also differences in farming practices and therefore the likelihood of infection being moved on and off a farm. Despite forming the index case, pigs played only a very minor role in the subsequent epidemic and can therefore be ignored.

If we naively assume random mixing between cattle and sheep, then the transmission matrix becomes:

$$\beta_{AB} = bs_A\tau_B \qquad \beta = b\begin{pmatrix} 27 & 15 \\ 1.8 & 1 \end{pmatrix}, \tag{4.14}$$

where s and τ give the susceptibility and transmissibility for the two species, and the parameter b scales the transmission matrix to obtain the observed growth rate of reported cases such that $R_0 \approx 2.5$ (Woolhouse et al. 2001a). In truth, cattle and sheep are aggregated at both the farm and regional level, but to capture these effects requires detailed spatial models (see Chapter 7; Keeling et al. 2001b). Finally, although there is some speculation that sheep may be infected for slightly longer than cattle, partly due to the difficulty with diagnosing infected sheep, we shall assume that both species have an equal infected period $(1/\gamma)$ of around 11 days. This period is composed of around 4 days of latent period, a further 5 days of infectivity before symptoms emerge, and an average of around 2 days before the animals are slaughtered to prevent further transmission (Ferguson et al. 2001a).

Figure 4.11 shows the predicted epidemic dynamics of the foot-and-mouth model (equation (4.13) with the transmission matrix given by (4.14)), starting with 100 infected sheep and the approximate number of susceptible animals in the Cumbria at the start of the 2001 UK epidemic (number of cattle, $N_c \approx 5.12 \times 10^5$; number of sheep, $N_s \approx 2.64 \times 10^6 \Rightarrow b = 1.38 \times 10^{-8}$). The model results are highly reminiscent of those for risk-structured models where the lower-risk group (sheep) is more prevalent. We notice (inset to top graph) that initially the number of infected sheep decreases because the infection cannot be sustained in the sheep population alone and cattle are needed to maintain the transmission. In the absence of cattle, the basic reproductive ratio of this infection in the sheep population is just 0.4; even with cattle included, each sheep infects

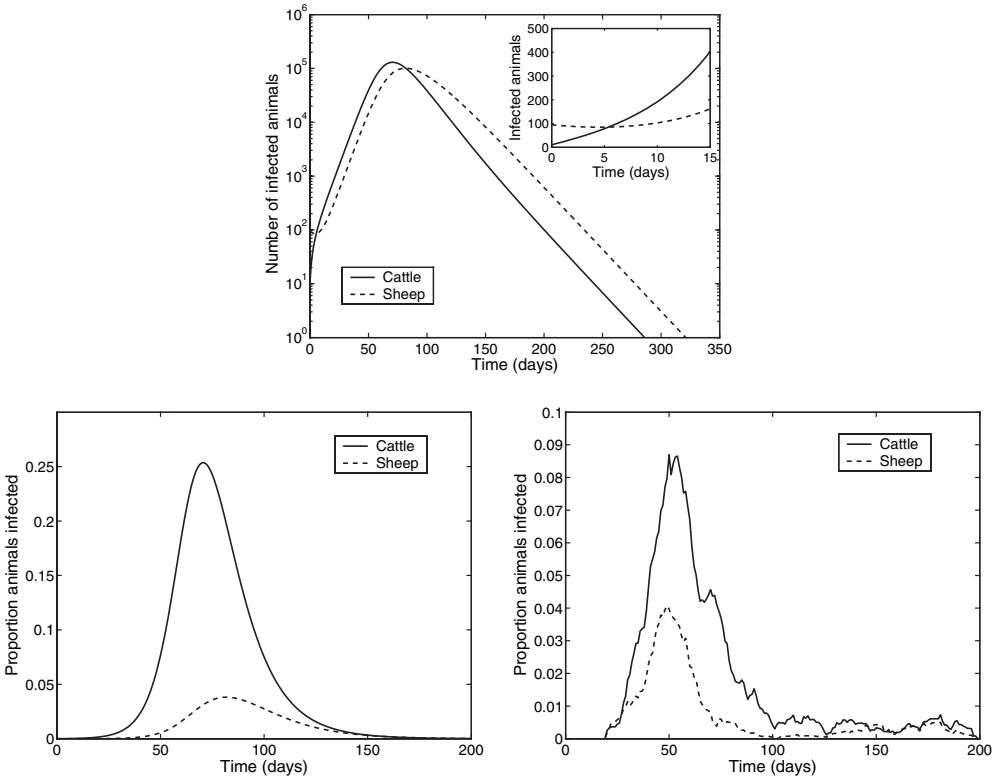


Figure 4.11. Dynamics of a multiple-host model of foot-and-mouth disease, based on the characteristic transmission and susceptibility parameters from the 2001 UK outbreak (Keeling et al. 2001b), and using the animal populations from Cumbria. Despite the much larger number of sheep in the UK, our model predicts that cattle are the main driving force in agreement with more complex simulations. The top graph gives the number of infected cattle (solid line) and sheep (dashed line), respectively, on a logarithmic scale; the inset shows the early epidemic behavior on a linear scale. The bottom graphs show the proportion of animals infected with the virus from the model (left) and the actual reported data from Cumbria in 2001 (right); the Cumbrian data assumes that all animals on a farm are infected. (Birth and deaths are not included in these model results.)

only 1.5 other animals. This strongly indicates that a control measure (such as vaccination) should be primarily focused toward the cattle industry (Tildesley et al. 2006) because without susceptible cattle the disease would soon die out.

The importance of cattle is further illustrated in the lower left-hand graph, showing that throughout much of the epidemic the prevalence of infection in cattle was far higher than it was in sheep; however, toward the end of the epidemic the levels became comparable. The actual data from the 2001 epidemic in Cumbria (lower right-hand graph) supports this basic result, although the difference between cattle and sheep is less pronounced.

Many elements are missing from this simple model, most obviously that livestock do not randomly mix but are aggregated into farms. It may therefore be more appropriate to formulate the "multi-host" model at the farm level, partitioning the population into

different farm types, rather than different species.

$$\frac{dX_F}{dt} = \nu_F - X_F \sum_f \beta_{Ff} Y_f,$$

$$\frac{dY_F}{dt} = X_F \sum_f \beta_{Ff} Y_f - \gamma_F Y_F,$$

where F and f are farm types, X_F and Y_F refer to the number of farms of a particular type, and again $1/\gamma_F \approx 11$ days is the time from infection to slaughter. A plausible partitioning of farms would distinguish between large and small numbers of livestock as well as predominantly sheep, predominantly cattle, or mixed (Ferguson et al. 2001a), leading to six distinct "species" of farms. Although such a high-dimensional model can be parameterized from our knowledge of the 2001 epidemic, a realistic model would also need to account for the complex temporally varying control measures and the intense local spread of infection (Keeling et al. 2001b). Such a data-intensive model is beyond the scope of this book, although see Chapter 8, Box 8.1.

4.2.1.2. Application: Parapoxvirus and the Decline of the Red Squirrel

Since its introduction from America at the start of the twentieth century, the gray squirrel (*Sciurus carolinensis*) has displaced the red squirrel (*S. vulgaris*) from much of its home range in the United Kingdom and mainland Europe (Lloyd 1983; Reynolds 1985). Although gray squirrels have an innate competitive advantage over reds (MacKinnon 1978), this advantage is not sufficient to explain the gray's rapid expansion and the reds' decline (Rushton et al. 1997). The action of a disease, parapoxvirus, has therefore been postulated as a likely cause of red squirrel decline—two-species disease models are therefore needed to understand and predict the likely competitive outcome. This infection has a negligible effect on gray squirrels but causes high mortality in reds, and therefore further decreases the red's competitive ability. Following the work of Tompkins et al. (2003), the following set of equations form a suitable model for the two populations:

$$\begin{aligned}
\frac{dX_G}{dt} &= \left[r_G - \frac{N_G + c_R N_R}{K_G}\right] N_G - \mu_G X_G - (\beta_{GG} Y_G + \beta_{GR} Y_R) X_G, \\
\frac{dY_G}{dt} &= (\beta_{GG} Y_G + \beta_{GR} Y_R) X_G - \gamma_G Y_G - \mu_G Y_G, \\
\frac{dZ_G}{dt} &= \gamma_G Y_G - \mu_G Z_G, \\
\frac{dX_R}{dt} &= \left[r_R - \frac{N_R + c_G N_G}{K_R}\right] N_R - \mu_R X_R - (\beta_{RG} Y_G + \beta_{RR} Y_R) X_R, \\
\frac{dY_R}{dt} &= (\beta_{RG} Y_G + \beta_{RR} Y_R) X_R - m_R Y_R - \mu_R Y_R.
\end{aligned} \tag{4.15}$$

The first term in both of the susceptible equations leads to a density-dependent birth rate, with the parameters c_G and $c_R = 1/c_G$ measuring the competitive effect of gray squirrels on red and visa versa (Begon et al. 1996). The remaining terms are the familiar

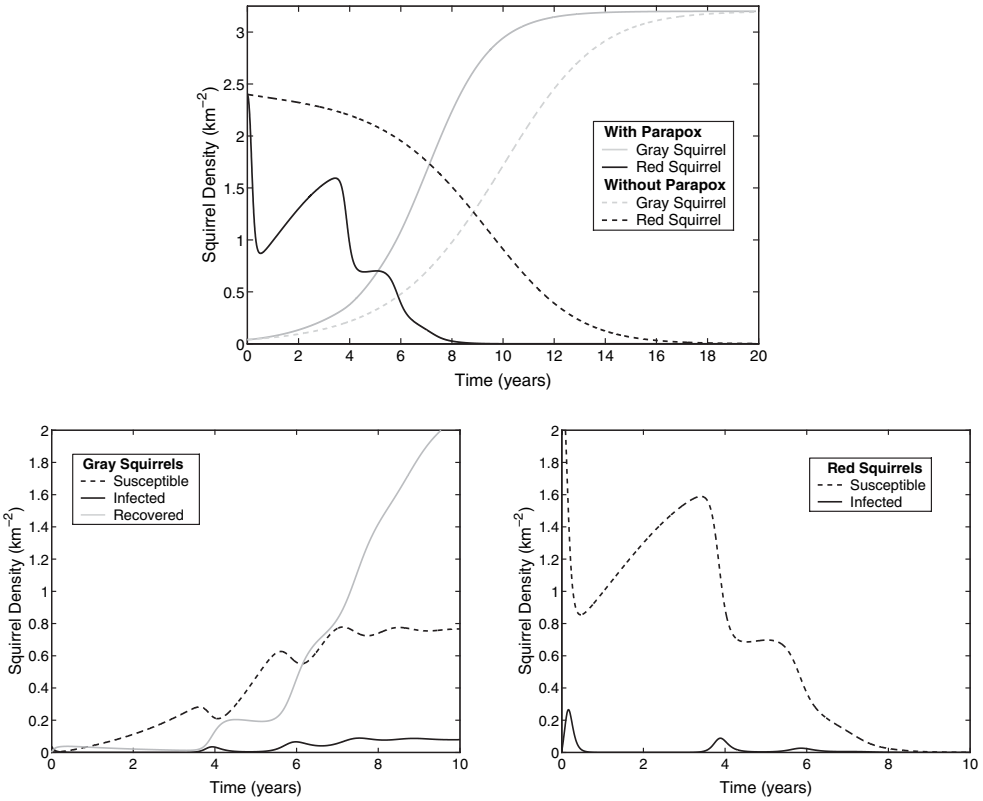


Figure 4.12. Invasion dynamics of gray squirrels into a population of red squirrels. The top graph shows the competitive effects of gray squirrels on the red population in the presence (solid line) and absence (dashed line) of the parapoxvirus. The lower left- and right-hand graphs give the levels of susceptible, infected, and recovered squirrels in the gray and red populations, respectively. The simulations are initialized with the invasion of one gray squirrel (either infected or susceptible) into a 5×5 kilometer area. (Parameters taken from Tomplins et al. (2003) are $r_G = 1.2$, $r_R = 1.0$, $K_G = 4$, $K_R = 4$, $c_R = 0.61$, $c_G = 1.65$, $\mu_G = \mu_R = 0.4$, $\beta_{GG} = \beta_{GR} = \beta_{RR} = \beta_{RG} = 17.5$, $\gamma_G = 13$, $m_R = 26$.)

epidemiological ones, although whereas gray squirrels recover from infection (at rate γ_G), the disease is always fatal to red squirrels (with mortality rate m_R). We have again assumed density-dependent (pseudo mass-action) transmission, in line with the standard paradigm on how directly transmissible wildlife diseases spread.

Figure 4.12 (top graph) shows the competition between red and gray squirrels, following the release of a low number of grays, both when parapox disease is present (solid lines) and absent (dashed lines). With parapoxvirus, during the first year there is a dramatic early decline in the red squirrel population; this is predominantly due to the epidemic dynamics rather than interspecific competition. In general, the addition of this infection into the model leads to the localized extinction of red squirrels within about 8 years, approximately

twice as fast as without the infection and more in keeping with field observations (Rushton et al. 1997; Reynolds 1985).

Although the assumption of random mixing between red and gray squirrels can be justifiable at the scale of an individual wood, when contemplating invasion and extinction at a national scale spatial factors play a more dominant role—with invasion moving in a wave-like fashion across the country (Reynolds 1985). A fully predictive model would need to account for such localized movements, as well as the variablity in habitat quality between regions (see Chapter 7). In addition, when dealing with invasions and extinctions—both of which involve low numbers of individuals—a stochastic element to the model becomes vital (see Chapter 6). Despite these shortcomings, this relatively simple model highlights the importance of pathogens that may be introduced along with an invading and competitive species.

4.2.2. Vectored Transmission

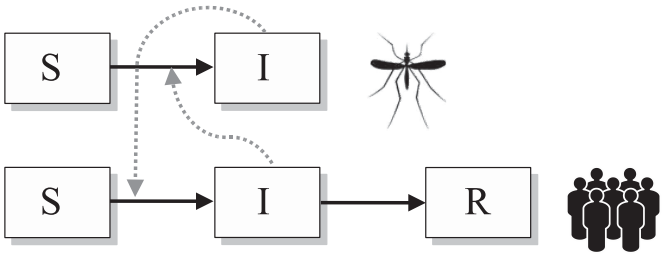
Many infections are transmitted via blood-sucking arthropods known as vectors. Malaria, yellow fever, dengue fever, trypanosomiases, and leishmania are all highly prevalent diseases of tropical and subtropical regions that are spread by this mechanism. Malaria alone is responsible for one million deaths and 300 million acute illnesses per year worldwide, making it one of the most devastating of all diseases. From a public health perspective, models may be crucial in determining which strategies or combinations of strategies are likely to be most successful against these devastating infections. Given the vast scale of these diseases, in some of the poorest areas of the world, cost-effective and optimally targeted controls are vital—well parameterized models, informed by good epidemiology and entomology, can play an important role in assessing the likely success of any policy.

In general, vector-borne diseases cannot be passed between primary hosts (person to person or animal to animal) but only through an intermediate insect host or vector. The natural history of vector-borne diseases therefore follows a standard pattern. An insect vector takes a blood-meal from an infected primary host (human or animal); therefore, with a given probability it becomes infected and is soon infectious. When the insect next feeds on a susceptible (and hence different) host, the pathogen enters the host's bloodstream and infection can occur, again with a given probability. In this way the 2×2 transmission matrix has zero on the diagonal elements, with all the transmission operating through the off-diagonal terms.

For vector-borne diseases, because no transmission occurs between humans (or animals) or between vectors, the diagonal elements of the transmission matrix are zero.

As mentioned in Chapter 1, malaria is caused by a single-celled protozoan, generally *Plasmodium falciparum*, but also *Plasmodium vivax*, *Plasmodium malariae*, and *Plasmodium ovale*. The life cycle associated with malaria is more complex than we have considered thus far, largely because there is sexually reproducing phase within humans and an asexually reproducing phase within *Anopheles* mosquitoes. However, despite these complexities, dynamics within humans are sufficiently fast that the SIR modling paradigm represents a reasonable description.

4.2.2.1. Mosquito Vectors



We now consider a simple model for the spread of infection between humans (or other primary hosts) via mosquitoes following the framework founded by MacDonald (1957). This model is also applicable to other vectors, such as tsetse flies, midges, or ticks, that take a single blood-meal from a host and then move on. First, we focus on the distinguishing feature of these models—the rate, r, at which a particular human is bitten by a particular mosquito:

$$r = \frac{b}{N_H},$$

where b is the bite rate of mosquitoes (the number of bites per unit time) and N_H is the number of humans. This formula therefore assumes that each mosquito bites at a constant rate b and that this is shared among all the human hosts within an area. The equations for the disease dynamics (dealing with numbers of individuals) now become:

This is online program 4.4

$$\begin{aligned}
\frac{dX_H}{dt} &= \nu_H - rT_{HM}Y_M X_H - \mu_H X_H, \\
\frac{dY_H}{dt} &= rT_{HM}Y_M X_H - \mu_H Y_H - \gamma_H Y_H, \\
\frac{dX_M}{dt} &= \nu_M - rT_{MH}Y_H X_M - \mu_M X_M, \\
\frac{dY_M}{dt} &= rT_{MH}Y_H X_M - \mu_M Y_M,
\end{aligned} \tag{4.16}$$

where T_{HM} (≤ 1) is the probability that an infected mosquito biting a susceptible human transmits the infection, with T_{MH} being the probability of transmission in the reverse direction. Although at first this transmission mechanism appears to be density-dependent, the inclusion of the parameter r means that transmission is actually frequency-dependent with respect to the human population. This is because it is assumed that each mosquito bites at a constant rate (irrespective of the number of available humans), whereas the rate at which humans are bitten will increase proportionally to the number (or density) of mosquitoes (Box 4.1). Finally, we note that this formulation assumes that the mosquitoes (or other appropriate vector) can feed only on humans; when other species are part of the menu, a three- or more species model is required, accounting for the different epidemiological parameters associated with each host.

Box 4.1 Minimum Infected Ratio

Measuring the proportion of infected mosquitoes, $I_M = Y_M/N_M$, for wild populations is a difficult and time-consuming task. In general, the fraction of infecteds is low, so testing each individual mosquito would realize a vast number of uninfected mosquitoes for every positive one. Instead, groups of mosquitoes are tested in batches (or pools), thereby increasing the chance that a batch is infected. This leads to the Minimum Infected Ratio (MIR) per thousand mosquitoes, which is defined as:

$$\text{MIR} = 1000\frac{\text{Number of infected batches}}{\text{Total number of mosquitoes tested}}$$

In the ideal scenario, every batch would contain exactly M mosquitoes, which provides a direct link between MIR and the proportion of infected mosquitoes, I_M. Suppose that b batches are tested, then:

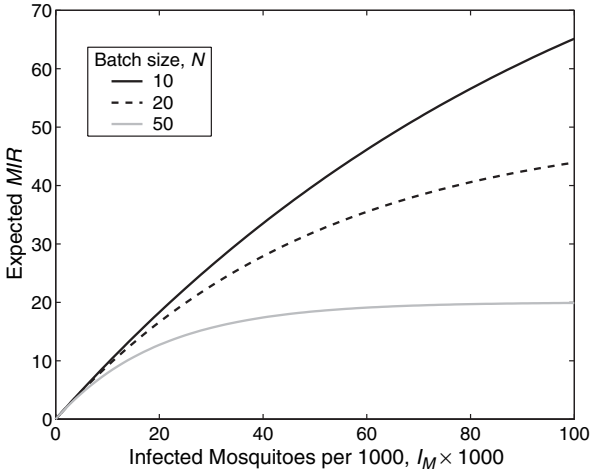
$$\text{MIR} = 1000\frac{b \times [\text{proportion of infected batches}]}{b \times M} = 1000\frac{[1-(1-I_M)^M]}{M}.$$

Thus, when I_M is small, the Minimum Infected Ratio scales almost linearly with increasing prevalence:

$$\text{MIR} = 1000I_M - 500(N-1)I_M^2 + \ldots$$

but as I_M and the batch size, M, increase, MIR becomes an underestimate hence—the name *Minimum* Infected Ratio.

The relationship between the actual prevalence of infection in mosquitoes and that estimated by MIR. Clearly, using smaller batch sizes produces more reliable results, but with the disadvantage that more batches need to be tested for the same total number of mosquitoes.



Simulations show that (assuming a plentiful supply of mosquitoes), the number per batch M should be chosen such that $1-(1-I_M)^M \approx \frac{1}{2}I_M M$. This ensures that I_M can be estimated from MIR with the greatest accuracy; if M is too large or too small, then either most of the batches will be infected or most will be uninfected, reducing the sensitivity of the results.

The disease transmission dynamics are specified in terms of the bite rate of mosquitoes and the probabilities of transmission following a bite.

In terms of a traditional multi-host model, the transmission matrix is given by:

$$\boldsymbol{\beta} = \begin{pmatrix} 0 & rT_{HM} \\ rT_{MH} & 0 \end{pmatrix}.$$

The set of equations (4.16) is for the number of individuals or the density within a given area (and not the proportion); often we can interchange proportions and numbers at will (Chapter 3), but because human and mosquito populations may fluctuate independently, this interchange is no longer viable.

The mosquito parameters, ν_M and μ_M, are likely to vary with climatic conditions. Thus, in regions of the world where there are pronounced climatic variations, strong seasonal effects may dominate the dynamics (Chapter 5). The extreme case of this is temperate regions, where only a low number of mosquitoes successfully overwinter, and hence transmission may be negligible for a significant fraction of the year.

To get a better understanding of the range of dynamics of these vector-transmitted diseases, we shall calculate the basic reproductive ratio, R_0. This can be done from first principles, which provides a more intuitive understanding of the early dynamics. Let us start with one mosquito that has just become infected, then R_0 is the number of secondary infections in mosquitoes that will be generated. First, we calculate the expected number of infected humans from this primary mosquito assuming all humans are susceptible:

$$\text{infected humans} = \frac{rT_{HM}N_H}{\mu_M} = \frac{bT_{HM}}{\mu_M}.$$

Now we calculate the number of mosquitoes infected by an infectious human:

$$\text{infected mosquitoes} = \frac{rT_{MH}N_M}{\gamma_H + \mu_H} = \frac{bT_{MH}N_M}{(\gamma_H + \mu_H)N_H}.$$

Thus, R_0 is given by the product of these two terms:

$$R_0 = \frac{b^2 T_{HM}T_{MH}N_M}{\mu_M(\gamma_H + \mu_H)N_H}. \tag{4.17}$$

Note that each mosquito could infect less than one human on average, and yet R_0 could still be more than one. This definition of R_0, which is used throughout vector-borne epidemiology, does not correspond exactly with the definition from two-species or risk-structured models (Chapter 3). Although both methods agree at the critical point when $R_0 = 1$, the vector approach is the square of the two-species approach because the vector approach includes the multiplication of two transmission steps.

The ratio of mosquitoes to humans is vital in determining both R_0 and the dynamics of infection. When there are many more humans compared to mosquitoes, sustained transmission may be impossible.

R_0 increases with the number (or density) of mosquitoes, but surprisingly decreases with the number (or density) of humans. This is because when there are many humans (and relatively few mosquitoes), the chance of someone being bitten twice in quick succession—once to catch the infection and once to pass it on before recovery—is very small. Therefore, for the infection to successfully spread and invade, the ratio of mosquitoes to humans has

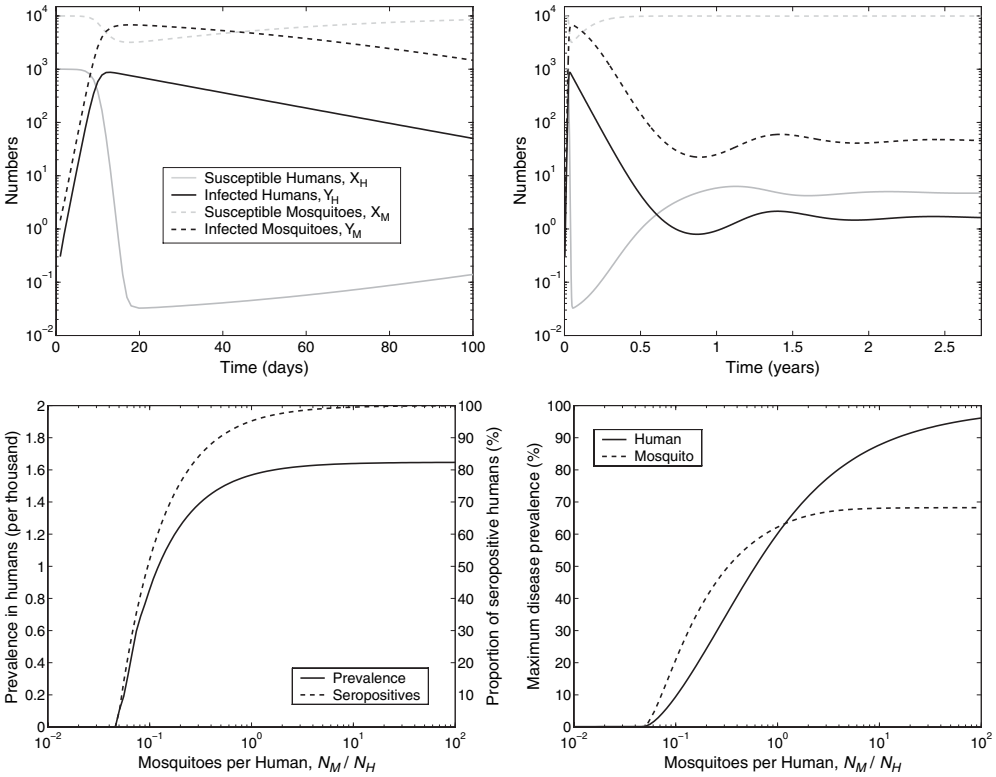


Figure 4.13. The top two graphs show typical dynamics of infected and susceptible humans and mosquitoes against time, starting with a single infected mosquito. The left-hand graph focuses on the early epidemic dynamics, whereas the right-hand graph shows the eventual equilibrium distribution. The lower graphs give the prevalence of infection in humans as the ratio of mosquitoes to humans ($\frac{N_M}{N_H}$) varies. The left-hand graph shows the final equilibrium prevalence in humans (solid line) and the proportion that have recovered from the infection and are seropositive (dashed line). The right-hand graph shows the maximum prevalence of infection in both humans and mosquitoes which occurs during the peak of the initial epidemic. ($N_H = 1000$, $\mu_H = 5.5 \times 10^{-5}$ (50-year human life span), $\nu_H = N_H \mu_H$ (constant human population size), $\mu_M = 0.143$ (1-week mosquito life span), $\nu_M = N_M \mu_M$ (constant mosquito population size), $\gamma_H = 0.033$ (infectious period of one month), $T_{HM} = 0.5$, $T_{MH} = 0.8$, $b = 0.5$. In the top graphs $N_M = 10^4$, which means that $R_0 \approx 200$.)

to be sufficiently large that double bites are common:

$$\text{mosquitoes per human needed for invasion, } \frac{N_M}{N_H} > \frac{\mu_M(\gamma_H + \mu_H)}{b^2 T_{HM} T_{MH}}.$$

Figure 4.13 gives an example of the typical dynamics, starting with a single infectious mosquito and susceptible mosquitoes and humans at equilibrium. Again, as expected from this type of two-species model, the rates of increase during the early epidemic are slaved (Chapter 3). Due to the high value of R_0, a large epidemic occurs on the timescale of a few weeks, after which the prevalence settles toward their equilibrium values, with the

vast majority of humans having experienced infection. The lower two graphs consider the dynamics because the ratio of mosquitoes to humans varies. The critical ratio of around 0.048 mosquitoes per human, when $R_0 = 1$, is clear in both graphs as the point where an epidemic is just possible. Above this level, the equilibrium prevalence and proportion of seropositive individuals (left-hand graph) rapidly increase to their asymptotic values. In contrast, the peak human prevalence during the initial epidemic shows much weaker saturation as the ratio of mosquitoes to humans increases. Therefore, although ratios in excess of 1:1 have little impact on the equilibrium prevalence, when faced with the invasion of a infection the precise ratio has a significant impact of the scale of the human epidemic.

In most situations, we expect the number of mosquitoes to far exceed the number of humans (or other hosts). The only notable exception is in subtropical or temperate regions where the number of vectors may be low during the colder winter months and hence disease transmission may be negligible. Such regions often form the boundary to areas where these infections are endemic, and thus understanding their dynamics is crucial if we wish to predict the spread of infection and the epidemiological implications of global warming. Including temporal (climatic) forcing into these vector-based models requires a detailed understanding of the entomology and vector ecology of these species and the models will rely on the techniques developed in Chapter 5. The example of West Nile Virus in Section 4.2.3.2 illustrates how this climatic forcing could be included.

Given that the life cycle of mosquitoes is much faster than both the epidemic and human timescales, each mosquito effectively experiences a constant level of human infection during its lifetime. Mathematically, we can use this fact to produce quasi-equilibrium calculations (Box 4.2) which assume that the mosquito population rapidly converges to an equilibrium state that depends exclusively on the current host population levels.

Due to the rapid life cycle of mosquitoes, a quasi-equilibrium approach can be used wherein mosquito populations are assumed to rapidly converge to equilibrium levels that are functions of the human population.

The quasi-equilibrium solution shows that the force of infection to humans rapidly saturates with increasing levels of human infection. This contrasts with the linear behavior of directly transmitted infections.

The quasi-equilibrium relationship between infection prevalence in mosquitoes and its prevalence in humans provides a deeper understanding of the behavior of vector-borne diseases (Figure 4.14). We observe that the quasi-equilibrium prevalence in mosquitoes begins to saturate with increasing prevalence in the host (solid line). This allows a comparison between the dynamics of vector-born and directly transmitted infections. The dashed line shows the expected linear behavior for a directly transmitted infection with a similar basic reproductive ratio, $R_0 \approx 200$; in contrast, the saturation of the vector-based infection curve, Y_M^*, shows that the force of infection (to humans) saturates. Therefore, relatively less transmission occurs when the prevalence in the host is high, and hence the equilibrium level of seropositives is lower in a vector-borne disease compared to a directly transmitted infection with the same R_0. The converse argument is also true; for a given level of seropositives in the human population, R_0 is larger for a vector-borne infection and the infection is more difficult to eradicate than results based on directly transmitted pathogens would suggest.

Box 4.2 Fast Vector Dynamics

One difficulty with such vector-based transmission of diseases is that there are double the number of equations compared to standard single-species models: a set for both the host and the vector. However, we can take advantage of the rapid vector life cycle (often 1 to 2 weeks) to simplify the equations. By assuming that the dynamics of the vector are fast compared to those of the host, we can find (quasi-)equilibrium vector abundances for any host population levels by setting the vector rates of change equal to zero:

$$\frac{dX_M}{dt} = \nu_M - rT_{MH}Y_H X_M - \mu_M X_M = 0,$$
$$\frac{dY_M}{dt} = rT_{MH}Y_H X_M - \mu_M Y_M = 0,$$

which implies that

$$X_M^*(X_H, Y_H) = \frac{\nu_M}{rT_{MH}Y_H + \mu_M}, \qquad Y_M^*(X_H, Y_H) = \frac{rT_{MH}\nu_M Y_H}{(rT_{MH}Y_H + \mu_M)\mu_M}.$$

These (quasi-)equilibrium vector population levels, which depend on the current host population, can then be substituted into the host equations to give a smaller, but more complex, set of differential equations:

$$\frac{dX_H}{dt} = \nu_H - T_{HM}b\frac{bT_{MH}\nu_M Y_H}{(bT_{MH}Y_H + \mu_M N_H)\mu_M N_H}X_H - \mu_H X_H,$$
$$\frac{dY_H}{dt} = bT_{HM}\frac{bT_{MH}\nu_M Y_H}{(bT_{MH}Y_H + \mu_M N_H)\mu_M N_H} - \mu_H Y_H - \gamma_H Y_H.$$

These equations give *exact* equilibrium solutions; however, the dynamic approach to equilibrium is an approximation and will be affected by the quasi-equilibrium assumption for the vector.

These results emphasize the crucial point that although the vector-borne infections generally spend the vast majority of their time in the primary (human or animal) host, the role of the vector cannot be simply ignored. The nonlinear behavior due to the obligatory role of the vector in transmission between hosts can have a pronounced effect on our understanding and parameterization of such disease models.

4.2.2.2. Sessile Vectors

Some blood-sucking arthropods, such as fleas and lice, tend to remain with a host for several generations. In such cases, the bite rate of the vector has little relevance to infection transmission because the vector is unlikely to have left the host. A more plausible model is therefore to consider transmission through a pool of free-living infected vectors (Y_V) that are in search of a new host:

$$\begin{aligned}
\frac{dX_H}{dt} &= \nu_H - rT_{HV}N_H Y_V \frac{X_H}{N_H} - \mu_H X_H,\\
\frac{dY_H}{dt} &= rT_{HV}N_H Y_V \frac{X_H}{N_H} - \mu_H Y_H - \gamma_H Y_H - m_H Y_H,\\
\frac{dY_V}{dt} &= T_{VH}(\mu_H + m_H + l)Y_H K_V - rN_H Y_V - \mu_V Y_V,
\end{aligned} \tag{4.18}$$

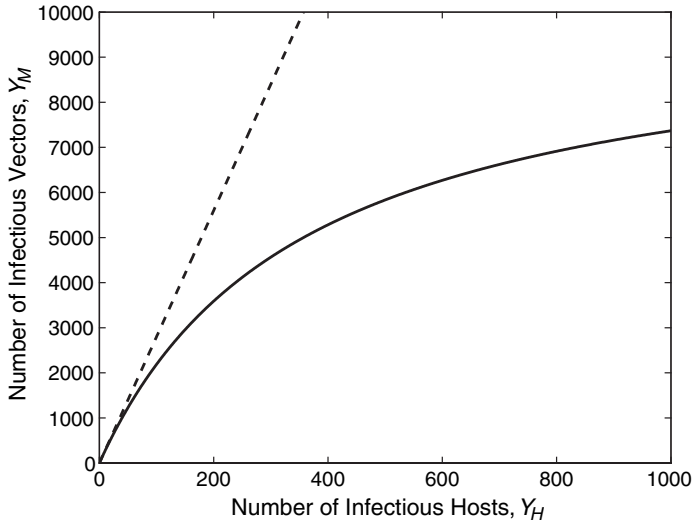


Figure 4.14. The number of infected mosquitoes (solid line) as calculated from the quasi-equilibrium equations. The dashed line shows the linear function that would be needed for the dynamics to approximate host-to-host transmission. All parameters are the same as Figure 4.13, with $N_M = 10{,}000$.

where the subscripts V and H refer to vector and host. Whenever an infected host dies (either naturally at rate μ_H or due to disease-induced mortality at rate m_H), the vectors leave the dead host in search of another live one at rate r. It is assumed that the vector life cycle is rapid so that each host on average supports a population of K_V vectors, and that a proportion T_{VH} of these are infected. Additionally, vectors may leave a living infected host at rate l, thus increasing transmission. The free-living infected vectors then encounter a new host at rate rN_H, and if this host is susceptible ($\frac{X_H}{N_H}$), they may transmit the infection with the probability T_{HV}. Finally, free-living vectors have a natural death rate, μ_V.

Figure 4.15 gives typical equilibrium-level dynamics for this type of vector. Not surprisingly, as the number of vectors per host increases, so does prevalence in the population (left-hand graph)—this increase in the number of vectors will have a linear effect on the overall transmission. For the parameters used in the model, each host must support more than approximately 25 vectors before transmission is possible. The dynamics with respect to mortality is more complex (right-hand graph). With directly transmitted infections, an increase in the mortality rate, m_H, decreases the infectious period and therefore decreases R_0. However, for this class of vector-borne diseases, the death of an infected host actually releases infected vectors into the environment. Thus as the mortality increases, the number of new cases (gray line), and the number of infected vectors in the environment (dashed lines), also increases.

- For infections spread by ticks, fleas, or lice, a high disease mortality may lead to greater transmission (despite the shorter infectious period) because it increases the rate at which vectors leave the host.

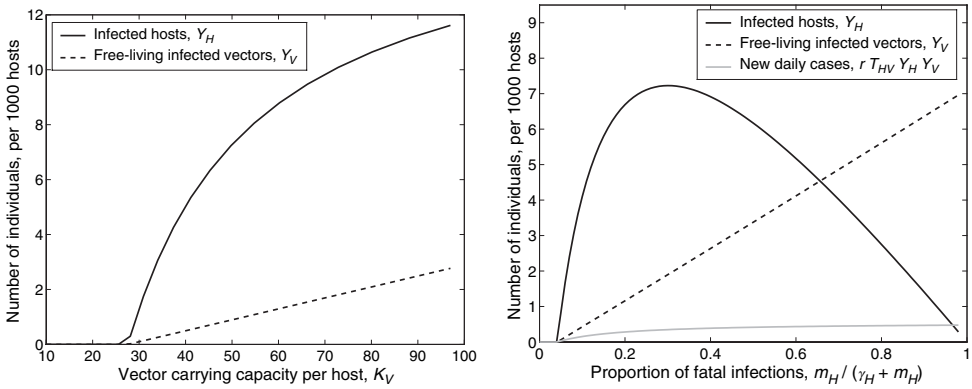


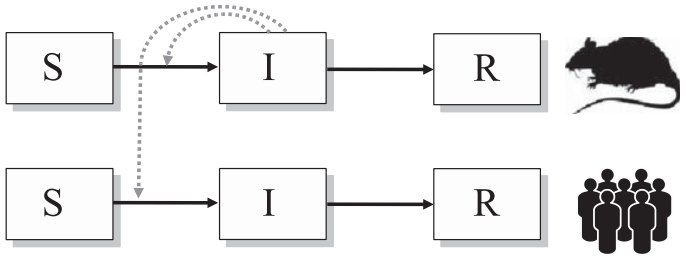
Figure 4.15. The left-hand figure shows the equilibrium number of infected hosts and free-living infected vectors, because the average number of vectors per host, K_V, is varied. The right-hand figure considers the effects of changes in the disease mortality; it also shows the equilibrium number of infected hosts and free-living infected vectors, but additionally gives the number of daily cases. ($N_H = 1000$, $\mu_H = 5.5 \times 10^{-4}$ (5-year host life span), $\nu_H = N_H \mu_H$ (constant host population size), $\mu_V = 0.071$ (2-week life span of free-living vectors without a host), $\gamma_H = 0.033$ (infectious period of one month), $T_{HM} = 0.5$, $T_{MH} = 0.8$, $r = 10^{-3}$, $l = 2.7 \times 10^{-3}$ (once per year). Left-hand graph $m_H = 0$, right-hand graph $K_V = 10$ (an average of 10 vectors per host).)

Again, some understanding is gained by calculating R_0 from the first principles starting with one free-living infected vector:

$$\begin{aligned} R_0 &= \text{Probability that a vector finds a host} \times \text{Probability of infection} \\ &\quad \times \text{Number of infected vectors released during host's infectious period} \\ &= \frac{rN_H}{rN_H + \mu_V} \times T_{HV} \times T_{VH} \frac{\mu_H + m_H + l}{\mu_H + m_H + \gamma_H} K_V. \end{aligned} \tag{4.19}$$

Although complex, this formula explains the model behavior because K_V and m_H are varied.

4.2.3. Zoonoses



Zoonotic diseases are defined as those that can be passed from animals to humans. In general, the animal host is the main reservoir for the infection, and humans contribute little to the overall transmission. From the disease's perspective, human cases are an

irrelevance, but are obviously the focus of public health interest. This dichotomy is also present in the models, where the animal population controls the infection dynamics but the human population determines the disease's impact. In this way, models of directly transmitted zoonoses are much simpler than standard multi-species models, because the 2×2 transmission matrix has zeros in one column (representing the lack of transmission from humans). Models of vector-borne zoonoses follow the standard template for all vector-borne diseases but include the additional transmission to humans. These two distinct classes are dealt with separately.

4.2.3.1. Directly Transmitted Zoonoses

Zoonoses are a ubiquitous challenge to human health, and are associated with a wide range of reservoir species (Acha and Szyfres 1989; Frank and Jeffrey 2001). It has been estimated that three-quarters of emerging human pathogens are zoonotic (Woolhouse 2002), thus a better understanding of their dynamics is likely to play an important role in public health planning. Many prominent zoonoses are associated with reservoirs in household pets (e.g., toxoplasmosis in cats) and livestock (e.g., brucellosis in cattle); this prominence reflects the greater mixing, and therefore transition, between these species and humans, rather than any epidemiological characteristics of the infections. Although the model given below (equation (4.20)) is generic, and could be applied to many directly transmitted zoonoses with appropriate parameterization, several zoonotic diseases deserve special mention due to their epidemiological importance, and scientific and public interest:

- Anthrax is a bacterial infection that affects a wide range of species, especially herbivores. Whereas early modeling focused on the transmission of infection between animals (Furniss and Hahn 1981; Hahn and Furniss 1983), more recent attention has concerned its use as a bioterrorism weapon (Webb and Blaser 2002; Wein et al. 2003), although the risk of subsequent transmission (and therefore R_0) in such situations is very low.
- Brucellosis is a coccobacilli that can be transmitted to humans from cattle, pigs, sheep, and dogs (Corbel 1997). It was once a major public health concern, and although veterinary efforts have dramatically reduced the number of cases in the United States and Europe, there are still several hundred thousand cases per year in humans worldwide. Modeling interest has primarily focused on the effect of the disease on the natural bison population, due to conservation issues, rather than the implications for human health (Peterson et al. 1991; Dobson and Meagher 1996).
- Ebola is one of the most notorious zoonotic diseases, causing a rapid onset hemorrhagic fever and very high mortality. The need for very close contact, and the severity of the symptoms soon after infection, means that outbreaks have been locally isolated. The animal reservoir species for Ebola is still unknown, despite much research.
- Hantavirus is primarily associated with rodent hosts, with each viral type having its only preferred host species. The "Four-Corners" outbreaks in Arizona, Colorado, New Mexico, and Utah in the 1990s have been linked to an infectious reservoir in the deer mouse (*Peromyscus maniculatus*) (Mills et al. 1999). Infection leading to

Hantavirus Pulmonary Syndrome is frequently fatal in humans, unless early treatment is given. Modeling of hantavirus has focused on the role of fluctuating rodent reservoir populations in an attempt to understand the observed large amplitude spatial and temporal variability (Abramson and Kenkre 2002; Abramson et al. 2003; Sauvage et al. 2003; Buceta et al. 2004).

- Rabies has been the focus of much mathematical modeling, due to its public health importance and the long-term spatiotemporal data that has been available. Many species can act as a reservoir for rabies, with foxes (in Europe), raccoons (in United States), and dogs (worldwide) being the primary sources for human infection. Early characteristics of rabies in humans occur 1 to 3 months after infection and are nonspecific and flu-like, with rapid progression to neurological symptoms including anxiety, confusion, slight or partial paralysis, excitation, hallucinations, agitation, hypersalivation, and hydrophobia. Rabies is almost inevitably fatal once symptoms emerge. Modeling efforts can be partitioned into two overlapping groups: Following the lead of Murray and co-workers (Kallen et al. 1985; Murray et al. 1986), many models consider the spatial spread of rabies in a wavelike manner (Moore 1999; Smith et al. 2002), whereas more applied models focus on the impact of specific control mechanisms for either preventing epidemic invasion (Smith and Harris 1991) or reducing the impact where the disease is endemic (Tischendorf et al. 1998; Rhodes et al. 1998; Suppo et al. 2000; Bohrer et al. 2002; Kitala et al. 2002; Smith and Wilkinson 2003). All of these models focus on the infection dynamics within the host reservoir, with little quantitative consideration given to the number of human cases.
- Toxoplasmosis is one of the most well-known zoonotic infections in the developed world. Its natural reservoir is the domestic cat (although sheep and other livestock are often infected), which explains its high prevalence in humans. Generally the symptoms of toxoplasmosis are mild and flu-like, but if caught during pregnancy the effects on the unborn child may be severe (Dubey 1988). Due to its usually benign nature, little quantitative informative is known about the epidemiology and transmission rates of this infection, which limits the amount of predictive modeling that is feasible (Ades and Nokes 1993).

We now consider a very general model for the dynamics of a zoonotic disease in both its animal reservoir and in humans. Assuming SIR-type dynamics in both the human (H) and animal (A) reservoir populations, the equations for directly transmitted zoonoses are:

$$\begin{aligned}
\frac{dX_A}{dt} &= \nu_A - \beta X_A Y_A - \mu_A X_A, \\
\frac{dY_A}{dt} &= \beta X_A Y_A - \mu_A Y_A - \gamma_A Y_A - m_A Y_A, \\
\frac{dX_H}{dt} &= B_H - \varepsilon\beta X_H Y_A - \mu_H X_H, \\
\frac{dY_H}{dt} &= \varepsilon\beta X_H Y_A - \mu_H Y_H - \gamma_H Y_H - m_H Y_H,
\end{aligned} \tag{4.20}$$

where ε is generally small and measures the trickle of infection from the animal population into the human one. The birth and death rates for the animal population (ν_A and μ_A) may be quite complex; seasonal factors, density dependence, and stochastic variation may all impact the dynamics. In fact, it is often our lack of quantitative knowledge of the basic ecology of the reservoir species that limits our modeling of the zoonoses. Because all transmission events are assumed to be from infectious animals, we have again adopted a density-dependent approach; this means that large fluctuations in the wildlife population can greatly increase the risk of an epidemic. Outbreaks of hantavirus are thought to arise via such a mechanism.

In a purely deterministic setting, the number of human cases parallels the number of animal cases ($\beta X_A Y_A \propto \varepsilon\beta X_H Y_A$), although there will be far fewer human cases. However, due to the low numbers involved, it is often far better to use a stochastic approach (Chapter 6). As such, $\varepsilon\beta X_H Y_A$ is the probabilistic rate of new cases in humans, and the probability of detecting at least one human case within time-interval t_1 to t_2 is:

$$P(t_1, t_2) = 1 - \exp\left(-D\varepsilon\beta X_H \int_{t_1}^{t_2} Y_A dt\right), \tag{4.21}$$

where D is the probability of successful diagnosis. For many zoonoses, it is difficult (and time consuming) to monitor the infection within the animal population, therefore the onset of human cases is usually the only indicator of a major epidemic within the animal population and therefore an elevated risk to humans. The problem is then a statistical one, determining whether the increase in human cases is merely a statistical fluctuation or the signature of an underlying epidemic.

For zoonotic diseases when human cases are rare, it may be difficult to separate the observation of a few chance human cases and the start of a larger-scale outbreak.

The typical dynamics of a zoonotic infection where humans play a negligible role in transmission are shown in Figure 4.16. The epidemic in the animal population has the characteristic shape that we have come to expect from such simple epidemics, and is unaffected by the behavior of the human population. As seen in the left-hand figure, the chance of observing human cases increases with number of animals infected so far, the relative transmission rate to humans ε, the size of the susceptible human population X_H, and the detection rate D. In particular, the size of the detected human epidemic can be estimated as:

$$\text{Number suscept. humans} \times \text{prob. of infection} \;=\; X_H(0)\left[1 - \exp(-\varepsilon D R_0 R_\infty)\right]$$

where, using standard notation, R_∞ is the proportion of animals infected. Hence, if R_0 is significantly larger than 1, and therefore R_∞ is close to one, we expect to see human cases if $\varepsilon D R_0 X_H > 1$.

The right-hand graph of Figure 4.16 considers the probabilistic nature of human infection in more detail, assuming that at least one human case is diagnosed and hence the epidemic is identified. The gray line shows the expected number of observed human cases after the initial diagnosis; clearly, this increases rapidly with the scaling factor, $\varepsilon D X_H$, and shows remarkably little stochastic variation. In contrast, the number of infected animals when the first human case is observed shows much more variation and, in general,

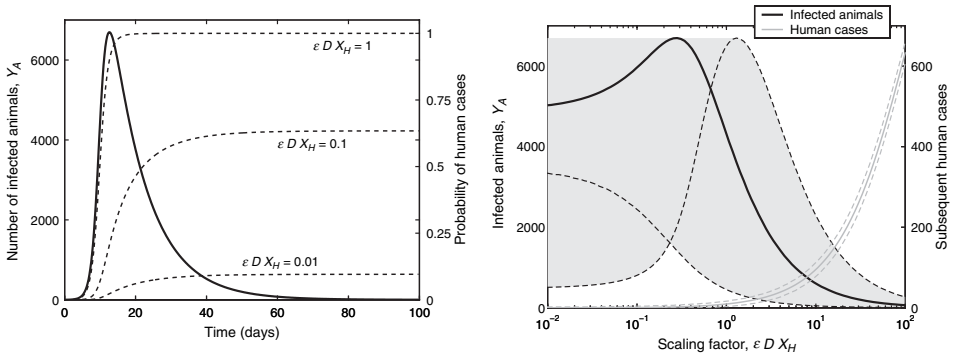


Figure 4.16. The left-hand figure shows a typical animal epidemic (solid line) ($N_A = X_A + Y_A + Z_A = 10^4$, $\gamma_A = 0.1$, $m_A = 0$, $R_0 = 10$, $\mu_A = 2.74 \times 10^{-4}$ (life expectancy of around 10 years), $\nu_A = \mu_A N_A$), and the probability of there being at least one human case detected (dashed line), (scaling factor $\varepsilon D X_H = 0.01, 0.1, 1$). The right-hand graph uses the same animal epidemic, but is conditional on there being at least one human case detected. The black line shows the expected number of infected animals, $Y_A(t)$, when the first human case is detected. The dashed lines correspond to values of Y_A with early and late detection (based on 95% confidence intervals for the timing of the first human case), and the shaded area shows the corresponding range of Y_A. The gray lines give the number of detected human cases (and 95% confidence intervals), assuming no control and $X_H = 1000$.

decreases as the scaling factor increases. Given that there is at least one human case, the first case is expected to occur at time T_1, such that:

$$\frac{P(0, T_1)}{P(0, \infty)} = \frac{1 - \exp\left(-\varepsilon D \beta X_H \int_0^{T_1} Y_A dt\right)}{1 - \exp\left(-\varepsilon D X_H R_0 R_\infty\right)} = \frac{1}{2},$$

where P is defined in equation (4.21) as the probability of identifying at least one case within a given time interval. The solid black line gives the number of infected animals at this time, $Y_A(T_1)$. Similarly, times can be found when the conditional probability is 0.05 and 0.95; the number of infected animals at these times are shown as dashed lines and the range incorporated is shaded.

Two opposing public health implications are associated with the results of this simple model. Although a low scaling factor ($\varepsilon D X_H$) means that few human cases will arise, it also implies that the animal epidemic is likely to be large before cases are detected and therefore difficult to control. Conversely, a high scaling factor should mean that the epidemic is detected far sooner allowing for easier control, but the cost to human health for not controlling the disease is more severe. Intermediate values of the scaling may present the greatest challenge; there is a potential for many human cases so the epidemic must be controlled, but detection is often delayed, making control much more difficult. All these problems become exacerbated if early cases are not quickly diagnosed, as tends to be the case with emerging zoonoses.

When a zoonoses is identified only by rare human cases, an epidemic within the animal hosts can be large before it is discovered. In such cases the epidemic may be difficult to control.

4.2.3.2. Vector-Borne Zoonoses: West Nile Virus

There are many infectious diseases that have a primary animal host, but that can be spread to humans via an insect vector. Examples include Chagas' disease (which infects dogs and is spread by the *Triatoma* and *Rhodnius* species of the True Bugs or Heteroptera family), Lyme disease (which infects rodents and dogs and is spread by *Ixodes* ticks), Q fever (which infects birds, rodents, and a range of household pets and is spread by ticks), leishmaniasis (which infects dogs and is spread by mosquitoes), and bubonic plague (which infects rats and other rodents and is spread by fleas). In recent years, West Nile virus (WNV, which infects birds and is spread by mosquitoes) has hit the headlines due to a significant number of deaths in the United States. Here we will concentrate on developing a model of WNV as an illustration of the general methods and complexities involved with understanding vector-borne zoonoses.

West Nile virus (WNV) provides an encompassing example of all that has been discussed in this chapter. It is a vector-borne zoonoses that has multiple host reservoirs, and during the 1990s a new strain (Lineage 1) emerged that has been associated with increased virulence and a range expansion. These elements make WNV a major health concern (especially in the United States but increasing in Europe), negate standard epidemiological rules-of-thumb which are based on experience from directly transmitted single-species pathogens, and make the formulation and parameterization of a detailed model extremely complex.

West Nile virus was first identified in 1937 in the West Nile region of Uganda, hence its name. It is a flavivirus commonly found in Africa, West Asia, and the Middle East. The natural host reservoir for WNV is birds (of many different species), with infection vectored by mosquitoes; occasionally an infected mosquito will bite a human, leading to infection. For the vast majority of human cases, symptoms are mild and flu-like with most individuals not even realizing that they have been infected. However, in a small proportion of cases, the infected person can develop meningoencephalitis, which can be fatal.

West Nile virus made international headlines in 1999 when it was responsible for a number of deaths in New York state, echoing an increasing trend for severe human cases and a high rate of avian mortality (Hubalek and Halouzka 1999; Petersen and Roehrig 2001; Campbell et al. 2002). In subsequent years this infection has spread to cover the majority of the United States, has invaded Canada and the Caribbean, and the death toll has continued to rise (see Figure 4.17). In the United States in 2002 there were, 3,873 clinical cases and 246 deaths, and data from New York City suggests that around 80% of the cases are subclinical (asymptomatic). Despite the shocking number of severe cases and fatalities, the actual incidence in the human population is very low. Levels of seroprevalence in Queens (New York) after the 1999 outbreak were estimated at only 3%. This is in direct contrast to the data from areas of Africa where the infection is endemic (and probably Lineage 2), where seroprevalence levels are about 50% in children and 90% in adults.

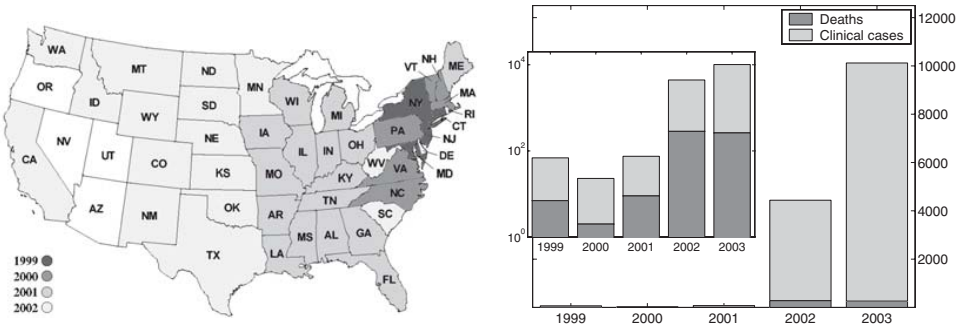


Figure 4.17. Data from the spread of West Nile Virus in the USA from 1999 to 2003. From its initial focus in New York State, the left-hand graph shows the dramatic range expansion that has occurred as the invading wave spreads west. The right-hand graph gives the number of reported clinical cases (light gray) and deaths (dark gray) due to WNV over the same period; the inset graph is plotted on a logarithmic scale to improve the clarity of the early data.

Modeling of West Nile virus within the United States (and its possible spread to other areas) is complicated by a variety of factors:

1. WNV has been found in 138 bird species within the United States, with susceptibility, transmissibility, and infectious period varying between species. House sparrows (*Passer domesticus*) may be a major reservoir due to their long infectious period and high-level of exposure to the virus—up to 60% (Komar et al. 2001). In contrast, the American crow (*Corvus brachyrhynchos*) is considered a sentinel species, due to its high level of mortality—crows comprised over 70% all the dead antibody positive birds reported. Sentinel species may be pivotal in providing an early warning of increasing incidence within the bird population (Eidson et al. 2001).
2. Multiple mosquito vectors may be responsible for transmission. Some, such as *Culex restuans*, feed predominantly on birds (ornithophilic) and therefore are responsible for amplification of the infection within the bird population but cause few human cases. In contrast, more opportunistic mosquito species (such as *Cx pipiens*) that feed on both birds and mammals may generate more human cases.
3. The temperate climate in the United States means that mosquito (and bird) populations fluctuate throughout the year—introducing temporal forcing into the model. The persistence of WNV from one year to the next relies on the successful overwintering of infected mosquitoes.

We have a very limited quantitative knowledge of the basic ecology of the species concerned and the epidemiological parameters and characteristics of their infection. However, a plausible attempt can be made at defining the basic structure of a full model for West Nile virus, after which simplifications can be made that will allow us to parameterize and simulate its dynamics. There are four basic components to the full dynamics of West Nile virus: birds, ornithophilic mosquitoes, opportunistic mosquitoes, and humans, with each being subdivided into a number of species. We start with the transmission matrix $\boldsymbol{\beta}$,

which can be partitioned into a number of nonzero components:

$$\beta = \left(\begin{array}{c|c|c|c} 0 & \begin{array}{c}\text{to birds from}\\ \text{ornithophilic}\\ \text{mosquitoes}\end{array} & \begin{array}{c}\text{to birds from}\\ \text{opportunistic}\\ \text{mosquitoes}\end{array} & 0 \\ \hline \begin{array}{c}\text{to ornithophilic}\\ \text{mosquitoes}\\ \text{from birds}\end{array} & 0 & 0 & 0 \\ \hline \begin{array}{c}\text{to opportunistic}\\ \text{mosquitoes}\\ \text{from birds}\end{array} & 0 & 0 & 0 \\ \hline 0 & 0 & \begin{array}{c}\text{to humans from}\\ \text{opportunistic}\\ \text{mosquitoes}\end{array} & 0 \end{array}\right).$$

Given the 29 mosquito species and 138 bird species that are known to have been infected with WNV within North America, the transmission matrix is 168×168 with over 8,000 parameters to be estimated. This is clearly impractical; we therefore focus on a much reduced model, which contains five basic elements: house sparrows (as the reservoir bird species), crows (as a sentinel bird species), ornithophilic and generalist (opportunistic) mosquitoes, and finally humans. With five interacting groups, parameterization will still be difficult, although some progress can now be made. The formulation of the equations follows the same mechanisms as elsewhere in this chapter, just with a greater number of components. We subdivide the populations into X, W, Y, and Z corresponding to susceptible, exposed, infectious, and recovered, and use the subscripts S, C, O, G, and H, to refer to sparrows, crows, ornithophilic mosquitoes, generalist mosquitoes, and humans:

$$\begin{aligned}
\frac{dX_b}{dt} &= \nu_b - (r_O T_{bO} Y_O + r_G T_{bG} Y_G) X_b - \mu_b X_b, \\
\frac{dW_b}{dt} &= (r_O T_{bO} Y_O + r_G T_{bG} Y_G) X_b - \sigma_b W_b - \mu_b W_b, \\
\frac{dY_b}{dt} &= \sigma_b W_b - \gamma_b Y_b - m_b Y_b - \mu_b Y_b, \\
\frac{dZ_b}{dt} &= \gamma_b Y_b - \mu_b Z_b \qquad \text{where } b \in \{S, C\}, \\
\frac{dX_m}{dt} &= \nu_m - (r_m T_{mS} Y_S + r_m T_{mC} Y_C) X_m - \mu_m X_m, \\
\frac{dW_m}{dt} &= (r_m T_{mS} Y_S + r_m T_{mC} Y_C) X_m - \sigma_m W_m - \mu_m W_m, \\
\frac{dY_m}{dt} &= \sigma_m W_m - \mu_m Y_m \qquad \text{where } m \in \{O, G\},
\end{aligned}$$

$$\begin{aligned}
\frac{dX_H}{dt} &= \nu_H - r_G T_{HG} Y_G X_H - \mu_H X_H,\\
\frac{dW_H}{dt} &= r_G T_{HG} Y_G X_H - \sigma_H W_H - \mu_H W_H,\\
\frac{dY_H}{dt} &= \sigma_H W_H - \gamma_H Y_H - m_H Y_H - \mu_H Y_H,\\
\frac{dZ_H}{dt} &= g_H Y_H - \mu_H Z_H,\\
r_O &= \tfrac{b_O}{N_S+N_C} \qquad r_G = \tfrac{b_G}{N_S+N_C+N_H}.
\end{aligned} \tag{4.22}$$

Here we explicitly assume that sparrows always recover from infection ($m_S = 0$), crows always die of the disease ($\gamma_C = 0$), mosquitoes can catch WNV only from birds, and humans can catch WNV only from generalist (opportunistic) mosquitoes. As with many wildlife diseases and vector-borne infections, the birth and death rate of the birds and mosquitoes may well be seasonal and density dependent.

The results of the model for West Nile virus given by equation (4.22) are shown in Figure 4.18. Although this model is much reduced in complexity from one that includes all species and all possible interactions, many of the parameters are largely a matter of speculation, and a rich variety of dynamics are possible. The parameters we have chosen mean that after the initial epidemic the level of seroprevalence in the sparrow population is around 50%, which is in agreement with observations (Komar et al. 2001). We also find from this model that the peak numbers of infectious generalist mosquitoes occur in late August early September, which is slighter later than the peak in mosquito numbers and agrees with the times when humans are most at risk.

This model demonstrates that using the methodology developed within this chapter, we can readily create models for a large number of interacting species. The primary difficulty comes from parameterizing such models, because data on the individual constituent mechanisms is difficult and time consuming to obtain—our knowledge of the basic ecology of both birds and mosquitoes is still too poor to allow a detailed parameterization of this model. However, such models can still be used to consider a variety of control measures (such as the use of insecticides) in order to limit the disease dynamics, with the ultimate aim of minimizing human cases. However, great care must be taken to ensure that the results are robust to the uncertainties in parameter values. To fully predict the complete behavior of West Nile virus however, would necessitate a model that can capture the heterogeneities at a variety of scales, from the local patchiness of mosquito breeding grounds to the national-scale spread of infection. Such spatial models are explored in Chapter 7.

4.3. FUTURE DIRECTIONS

In the coming decades it is likely that far more genetic, molecular, and immunological data will become available. The challenge will be to integrate this knowledge with the evolutionary disease models developed earlier in this chapter. Currently, models of disease evolution are in their infancy—far more work is required to integrate the type of models outlined in this chapter with the novel immunological models that are currently being developed (Nowak and May 2005). A comprehensive understanding of disease evolution would

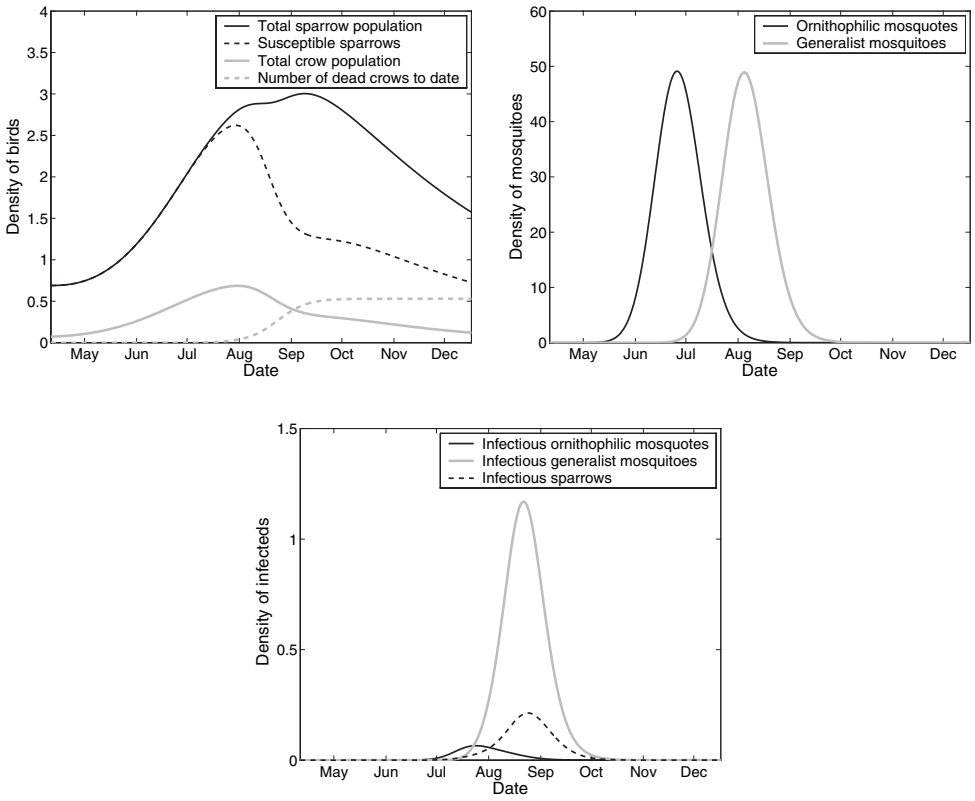


Figure 4.18. Model results for the spread of West Nile virus between species, using the formulation given in equation (4.22). Some parameters are easily found from the literature, $1/\mu_G \approx 1/\mu_O \approx 7$ days, $T_{SO} \approx T_{CO} \approx T_{SG} \approx T_{CG} \approx 0.4$ (Goddard et al. 2002), $T_{OS} \approx T_{OC} \approx T_{GS} \approx T_{GC} \approx 0.5$ (range 0.3–0.8), $1/\sigma_S \approx 1/\sigma_C \approx 1/\gamma_S \approx 1/m_C \approx 3.5$ days (Komar et al. 2001), $1/\sigma_H \approx 1/\gamma_H \approx 5$ days, $m_H \approx 1.3 \times 10^{-3}$ (Campbell et al. 2002), $1/b_O \approx 1/b_G \approx 3$ days. Other parameters can be found by matching to the observed proportion of seropositive birds, the minimum infectious ratio of mosquitoes, and the number of human cases. In addition, the number of bird and mosquito births are assumed to have a Gaussian distribution $\nu_Q = \hat{\nu}_Q \exp(-\frac{1}{2}(t - t_Q)^2/V_Q)$, where $Q \in \{S, C, O, G\}$. ($1/\hat{\nu}_S = 20$ days, $1/\hat{\nu}_C = 60$ days, $1/\hat{\nu}_O = 1/\hat{\nu}_G = 0.127$ days, $t_S = t_C = 190$, $t_O = 170$, $t_G = 210$, $V_S = V_C = 1800$, $V_O = V_G = 150$.)

allow us predict the short-term behavior of influenza, including the probable strains for the next season as well as the likelihood of a pandemic. In addition, a more complete knowledge of viral and bacterial genetics may allow us to predict with greater accuracy methods of preventing the evolution of drug-resistant strains and their spread through the population.

A second area where substantial advances are required is in the parameterization of multi-species models, where the number of parameters usually grows quadratically with the number of species. However, it is often our understanding of the basic ecology of the host species that is lacking, and only detailed field work can resolve many of the issues.

From a modeling perspective, it is important that we ascertain how sensitive models are to these unknown ecological factors, so that field work can be directed toward the key factors that can shape an epidemic.

4.4. SUMMARY

This chapter addressed two contemporary but very different issues, the competition and evolution of infections/strains in a single host population and the spread of a single infection between multiple host species. Both of these modeling issues have important applied implications to public health. Understanding disease evolution would enable us to predict and prepare for future epidemics. The study of infections in multiple hosts has a more immediate impact, because it concerns a range of high-profile diseases (such as malaria) that are responsible for millions of deaths worldwide every year.

The main findings about competing and evolving infections can be summarized as follows:

➤ When competing strains provide complete protection for each other, the strain with the largest R_0 will force the other strain to extinction, although a rapid life cycle may allow short-term dominance.

➤ Evolution will favor mutants with higher R_0, leading to higher transmission, life-long infections with low mortality. However, trade-offs between transmission rates and duration of infection mean that R_0 is maximized for intermediate values and runaway evolution is prevented.

➤ Application of antibiotic treatments requires a careful balance between combating infection and not providing suitable conditions for resistant mutants to outcompete the wild type.

➤ Even when there is no cross-immunity, the absence of multiply infected individuals is epidemiologically plausible, reflecting the reduced number of contacts when ill. This is believed to be why cases for measles and whooping cough are often out of phase.

➤ Research into malaria strains shows that when there is limited cross-immunity, the individual values of R_0 for each strain are lower than estimated from seropositive levels that ignore strain structure, reducing R_0 for each malaria strain to as low as 6 or 7.

➤ Having one sexually transmitted infection can often increase the susceptibility to others, promoting coinfection. In such circumstances the Allee effect may operate, and reducing the levels of one infection may lead to a reduction of the other.

➤ Coexistence of competing strains is possible when their respective R_0 values are close and the level of cross-immunity is weak.

➤ Models of strain structure with immunity to genetically close strains and mutations can lead to both traveling waves or large amplitude patterns in strain space.

Multiple-host models have much in common with the risk-structured models of Chapter 3, with species playing the role of risk. A number of general issues of importance are:

➤ In multi-host models (unlike risk-structured models), the transmission matrix is nolonger expected to be symmetric due to species differences. However, we still expect

to see very early dynamics determined by the initial conditions before the behavior of infection in all the hosts becomes slaved, increasing with an exponent determined by R_0.

➤ For vector-borne diseases, such as malaria, because there is no transmission between humans (or animals) and no transmission between vectors, the diagonal elements of the transmission matrix are zero—which dramatically simplifies the calculation of R_0.

➤ The ratio of mosquitoes to humans is vital in determining both R_0 and pathogen dynamics. When there are many more humans compared to mosquitoes, sustained transmission may be impossible because humans rarely experience two bites—one to infect the human and one to infect subsequent mosquitoes.

➤ Due to the rapid life cycle of mosquitoes, a quasi-equilibrium approach can be used where mosquito populations are assumed to rapidly converge to equilibrium levels that are functions of the human population. The quasi-equilibrium solution shows that the force of infection to humans rapidly saturates with increasing levels of human infection. This contrasts with the linear behavior of directly transmitted infections.

➤ For infections spread by ticks, fleas, or lice, a high disease mortality may lead to greater transmission (despite the shorter infectious period) because it increases the rate at which vectors leave the host.

➤ For zoonotic diseases (those spread from animals to humans) when human cases are rare, it may be difficult to separate the observation of a few chance human cases and the start of a larger scale outbreak. When the zoonoses is identified only by rare human cases, an epidemic within the animal hosts can be large before it is discovered, making the epidemic difficult to control.

➤ For zoonotic diseases such as West Nile virus, the vast number of animal hosts and mosquito vectors makes parameterization of even the simplest model very difficult—a greater understanding of host and vector ecology is needed.