CHAPTER 15

MATHEMATICAL MODELING OF CRIMEAN-CONGO HEMORRHAGIC FEVER TRANSMISSION

BEN S. COOPER, PH.D.

Statistics, Modelling and Bioinformatics Department, Centre for Infections, Health Protection Agency, London, UK. E-mail: ben.cooper@hpa.org.uk

Sensibly used, mathematical models are no more, and no less, than tools for thinking about things in a precise way. (Anderson and May 1991) [1]

This chapter is divided into five sections. Section 15.1 discusses the rationale for using mathematical models. Section 15.2 considers the specific areas where models may be useful in studying Crimean-Congo hemorrhagic fever (CCHF). Section 15.3 reviews work on modeling the dynamics of tick-borne diseases and considers the relevance of this work for CCHF. Section 15.4 considers the problem of modeling the nosocomial transmission of CCHF. Section 15.5, lastly, suggests future directions for CCHF modeling work.

15.1. WHY USE MATHEMATICAL MODELS?

Mathematical models of infectious diseases represent simplified representations of known processes and their interactions. Typically these processes are transmission, disease progression, birth, death and recovery, acquisition and loss of immunity, and immigration or emigration. Other processes appropriate for certain applications include boosting of immunity, vector dynamics, vaccination, and other control measures. By using these models we hope to capture important aspects of the behavior of the whole system and gain a full understanding of the role of the individual processes and their interactions in determining this behavior. The properties of these models and their predicted behavior under different scenarios are usually investigated either by solving the equations (in the case of deterministic models, where chance is assumed to play no part, and which can often be considered to model the average behavior of the system), or by simulating many epidemics from stochastic models (an approach known as Monte Carlo simulation because the role of chance in determining the course of

O. Ergonul and C. A. Whitehouse (eds.), Crimean-Congo Hemorrhagic Fever, 187–203.

the epidemic is explicitly accounted for). In both cases, the analysis usually makes use of a computer (though the original work with both types of models was done without one). This *in silico* modeling approach is directly comparable with the use of in vitro or in vivo models in the laboratory; all three approaches work by choosing model systems that are, it is hoped, in some important sense analogous to the system we really want to find out about. We may hope that the model is sufficiently similar to the real system in the ways that matter to give the same answers to the questions we ask, but we can seldom be sure. A cautious and critical approach to interpreting results from all models is therefore needed. Mathematical models are no exception to this rule. Nonetheless, as with other models, their use has the potential to lead to important advances in our understanding.

At the simplest level, mathematical models are just translations of reasoning from natural language into a precise mathematical formulation. The translation process has the virtue of forcing (or at least encouraging) us to think clearly, and tends to makes hidden assumptions explicit. Flaws in the reasoning of simple verbal arguments often become apparent once the models have been constructed. Models also allow the intuition that the verbal arguments rest on to be thoroughly tested. In this sense, models can be thought as formal encapsulations of hypotheses. Sometimes the resulting models support the verbal arguments and we may be encouraged that our intuition was correct, but frequently models show that our loose verbal reasoning is wrong. These models produce counterintuitive results. However, even when model results confirm our intuition, by allowing a fuller and more rigorous analysis of the consequences of the assumptions that make up the model they can broaden our understanding of the system. Often models will make new predictions that can be directly tested. If they pass the test, our confidence in the truth of the hypothesized mechanism the model represents will be strengthened. There is also a danger that new hidden assumptions are made in constructing the model. Careful modeling work should aim to highlight these assumptions and ideally to explore the sensitivity of the results to structural uncertainties in the model as well as to uncertainty in parameter values.

Models are not, as has been suggested, simply substitutes for experiments [13]. Instead, a primary use of models is to broaden our understanding, synthesize information, and to show how diverse outcomes can be understood as the result of similar underlying processes. In this way mathematical models are frequently used to help interpret experimental findings. Models themselves often suggest certain experiments or observations, and such experiments may in turn lead to revisions to the models. Nonetheless, it is true that in all the sciences where experimental manipulation is either difficult or impossible – astrophysics, economics, climatology, geology, ecology (of which infectious disease epidemiology can be considered to be one branch) – mathematical models play a prominent and sometimes central role.

The modern use of models in infectious disease epidemiology dates to the pioneering work of Ronald Ross (though the mathematical study of epidemics can

in fact be traced back to work on smallpox by the great French mathematician Daniel Bernoulli [2]). Ross' work on the biomathematics of malaria originally led him to the conclusion that control of the vector would be the most efficient means of fighting malaria [18]. By the 1970s the use of mathematical models to study infectious diseases was increasing faster than exponentially, and substantial growth continues today. The range of analytical tools available to modelers has also increased considerably. In particular, the ready availability of ever-faster computers have made new types of models possible, and in the last few years new developments in statistics coupled with increased computational power have enabled more detailed and more accurate model-based analysis of epidemiological data.

Such approaches have only been applied to hemorrhagic fevers very recently [4, 9], however, and these methods have not so far been applied to CCHF. The aim of this chapter is to review the potential value of mathematical models for studying CCHF, describe the basic theory and key predictions from simple models, and to highlight some of the most important analytical techniques likely to be of value in studying CCHF. Here, we are exclusively concerned with understanding the system at the population level. Mathematical models also have a central role to play in understanding within-host progression of infectious diseases [15], but this is beyond the scope of this chapter.

15.2. THE USE OF MODELING APPROACHES FOR CCHF

There are at least four reasons why we might want to use mathematical models to study CCHF. First and foremost, they can help us to gain a qualitative understanding of the dynamics of the disease and, in this way, help us to improve our intuition. This is likely to be particularly important for studying nosocomial outbreaks where stochastic (chance) effects will be important. As casino owners know well, most people have rather poor intuition about chance, and the importance of such effects in epidemics in small populations comes as a surprise to many. Second, by highlighting key uncertainties and gaps in our knowledge models may suggest observational or experimental studies that would improve our understanding of key aspects of the whole system. This is likely to be particular important in the area of understanding the vector dynamics where major uncertainties exist for CCHF. Since models can also be considered to be hypotheses about the systems, by confronting models with data we can effectively choose between competing hypotheses. Third, models can help in the selection and evaluation of control policies. This can be done by employing models as statistical tools used to estimate the effect of interventions that have been made (and, equally importantly, to quantify the uncertainty in these estimates). Having quantified the effect of individual interventions, we can then go on to use models to ask "what-if" questions, using models predictively to determine the expected effect of hypothetical combinations of interventions. More generally, by enabling us to identify the most critical parameters affecting the

behavior of the system, models can help us in setting priorities and identifying the most cost-effective control policies. Indeed, the use of dynamic models is essential for accurate economic analyses of control measures for infectious diseases [3]. Fourth, there is the potential to use models for forecasting the future of epidemics. Though popularly imagined as one of the main uses of models (perhaps by analogy to the models used to derive weather forecasts) this is one of the least developed areas in the infectious disease modeling literature, although there is increasing interest in this application.

15.3. MATHEMATICAL MODELS OF TICK-BORNE DISEASE TRANSMISSION

In this section we describe a basic framework for modeling tick-borne infections. This is adapted from the seminal work on the transmission dynamics of tick-borne infections by Medley et al. [12]. We show how this approach enables us to assess the magnitude of interventions needed to control CCHF in different regions and how interventions aimed at controlling the tick population could increase as well as decrease the risk to workers exposed to potentially infected animals. We also describe how this basic framework can be extended to address a wider class of questions.

The model presented by Medley et al. related specifically to the tick-borne transmission of *Theileria parva* in eastern Africa. The modeling framework described was, however, quite general, and with only minor modifications we can apply it to the study of CCHF. We begin by showing how seroprevalence data can be used to estimate the rate at which animals become infected with the virus. We assume that the rate at which an animal of age a is infected with the virus is $\beta(a)$ (mathematically this means that the chance an uninfected animal of age a is infected in a short time interval δt is approximately $\beta(a)\delta t$, the approximation becoming exact in the limit as δt approaches zero. The (a) following the β indicates that this rate is a function of age, and not necessarily constant). From this we can immediately write down a differential equation describing how the number of susceptible animals (those which have never been infected) changes with the age of the animals:

$$\frac{\mathrm{d}X(a)}{\mathrm{d}a} = -\beta(a)X(a) \tag{1}$$

Here we use $X(a)$ to represent the proportion of animals of age a who have not been infected with the virus and who are therefore seronegative. $\mathrm{d}X(a)/\mathrm{d}a$ is the rate at which $X(a)$ changes with age, so the equation specifies the slope of the graph plotting numbers seronegative ($X(a)$) against age. To obtain the model predictions for the actual relationship between $X(a)$ and a we solve Equation (1) by integration. This gives

$$X(a_i) = \exp\left(-\int_0^{a_i} \beta(a)\,\mathrm{d}a\right). \tag{2}$$

Two possible functional forms for $\beta(a)$ are $\beta(a) = b$, i.e. the infection rate is constant with age; and $\beta(a) = ba + c$, i.e. the infection rate starts from some baseline c and increases linearly with age. Substituting these into Equation (2) and solving gives $X(a_i) = \exp(-ba_i)$ and $X(a_i) = \exp(-ca_i - b^2 a_i / 2)$, respectively. Equivalently these models predict that the numbers seropositive by a given age, which we call $S(a)$, are given by $S(a) = 1 - \exp(-ba)$ and $S(a) = 1 - \exp(-ca - ba^2/2)$, respectively. Many other functional forms for $\beta(a)$ are possible (e.g. the infection rate might saturate with increasing age), but these two are the simplest and have found to be adequate for explaining many age-seroprevalence profiles. Using these expressions for $S(a)$ we can estimate the infection rate by fitting the curves to age-seroprevalence profiles. Such profiles could be obtained by a longitudinal study, repeatedly sampling from the same animals over time, or – providing it was reasonable to assume the system was in equilibrium – by using data from a cross-sectional, age-stratified survey of animals.

Figure 15-1 below illustrates how the seroprevalence would be expected to change with age for both functional forms of $\beta(a)$. It also shows simulated data from a hypothetical cross-sectional study. In practice $\beta(a)$ would be estimated from such data by fitting curves for $S(a)$ corresponding to different functional

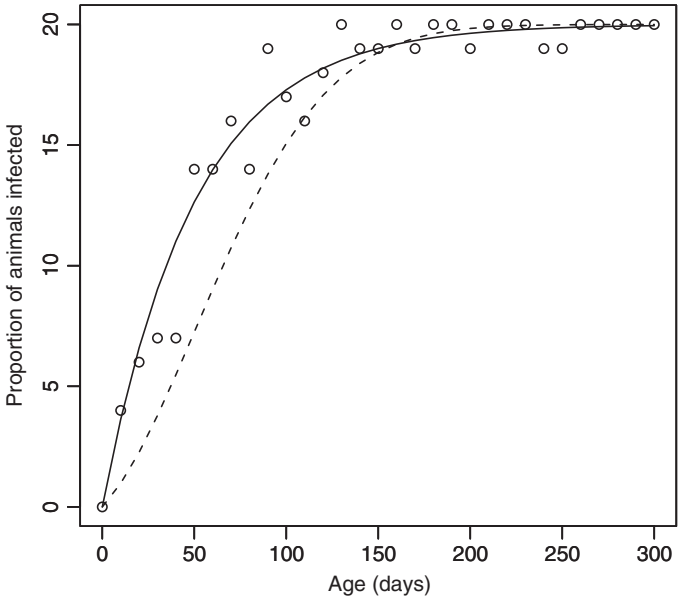

Fig. 15-1. Seroprevalence with age from a hypothetical cross-sectional study with 20 animals in each age group (0, 10, 20, . . ., 300 days). Solid line shows expected seroprevalence assuming $S(\text{age}) = 1 - \exp(-b \times \text{age})$ where $b = 0.02$. Dashed line shows expected seroprevalence when $S(\text{age}) = 1 - \exp(-c \times \text{age} - b \times \text{age}^2 / 2)$, with $b = 0.0002$ and $c = 0.004$. Dots illustrate how typical data from such a study might look when seroprevalence increases with age according to the first functional form (solid line) and is generated from this model assuming binomially distributed errors.

forms of $\beta(a)$ to the data. This can be done using standard maximum likelihood methods assuming binomially distributed data, allowing model parameters and their confidence intervals to be estimated. The best-fitting functional form of $\beta(a)$ can be selected with a likelihood ratio test.

The rate of infection, $\beta(a)$, is equal to the product of the rate of tick attachment as a function of age, $T(a)$, the probability of a tick being infected with the virus in an endemically stable environment, r^*, and the probability of transfer of infection from an infected tick to a host it is attached to, q. Rearranging this gives $T(a) = \beta(a)/(qr^*)$, so once estimates of $\beta(a)$ have been obtained using the method described above, if we also have estimates of q and r^* (which should be relatively easy to obtain), it becomes possible to derive estimates of the tick attachment rate as a function of host age.

When we also know the latent and infectious periods of the virus in a particular animal host, we can construct a dynamic transmission model of the course of infection in the population. Figure 15-2 gives a schematic illustration of the structure of such a model, which forms the basis of many disease transmission models. Each host is assumed to belong to one of four compartments: susceptible to infection (S), latently infected with the organisms (i.e. exposed) but not yet infectious (E), infectious (I), and recovered and immune (R).

As discussed above, the rate at which hosts become infected, $\beta(a)$, can be assumed to increase linearly with the proportion of ticks that are infected, and that proportion in turn would be expected to increase as the proportion of infected hosts increased. A full dynamic model of the system is needed to account for this feedback and a model incorporating both the tick and host dynamics would therefore allow $\beta(a)$ to change over time. However, when the system is in equilibrium (i.e. when the size of the host population and the amount of infection in that population is neither increasing nor decreasing with time, apart from small chance fluctuations) the proportion of ticks infected will not change over time and $\beta(a)$ will also not change over time. Under these circumstances it is possible to write down a system of ordinary differential equations that describes how the proportion of hosts in each compartment changes with the age of the hosts:

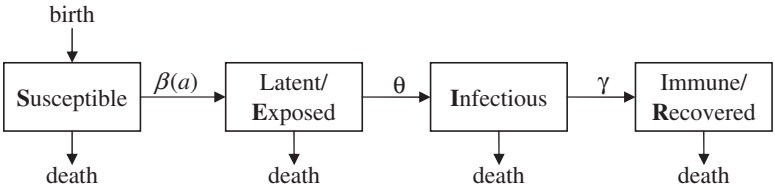

Fig. 15-2. Schematic diagram of a dynamic compartmental SEIR model.

$$
\begin{aligned}
\frac{\mathrm{d}S(a)}{\mathrm{d}a} &= -\beta(a)S(a) - \mu S(a) \\
\frac{\mathrm{d}E(a)}{\mathrm{d}a} &= \beta(a)S(a) - \theta E(a) - \mu E(a) \\
\frac{\mathrm{d}I(a)}{\mathrm{d}a} &= \theta E(a) - \gamma I(a) - \mu I(a) \\
\frac{\mathrm{d}R(a)}{\mathrm{d}a} &= \gamma I(a) - \mu R(a)
\end{aligned} \tag{3}
$$

Here θ is that rate of progression from the latent to infectious compartment ($1/\theta$ gives the mean latent period) and γ is the rate of recovery from infection (the mean infectious period is then $1/\gamma$), and μ is the all-cause mortality rate (in general, this will be a function of age, but for simplicity we take it as a constant here implying that life expectancy is $1/\mu$ days). Essentially, these equations describe the flows between the compartments in Fig. 15-2: hosts leave the S compartment at rate $\beta(a)S(a)$ due to infection, and at rate $\mu S(a)$ due to death. Those becoming infected first flow into the E compartment at the same rate they leave the S compartment. They can leave the E compartment due to death (at rate $\mu E(a)$) or by progressing from the latently infected stage to the infectious stage (at rate $\theta E(a)$). Other terms in the above equations can be explained in a similar manner. Typical output for such a model when $\beta(a)$ is taken as a constant b is illustrated in Fig. 15-3.

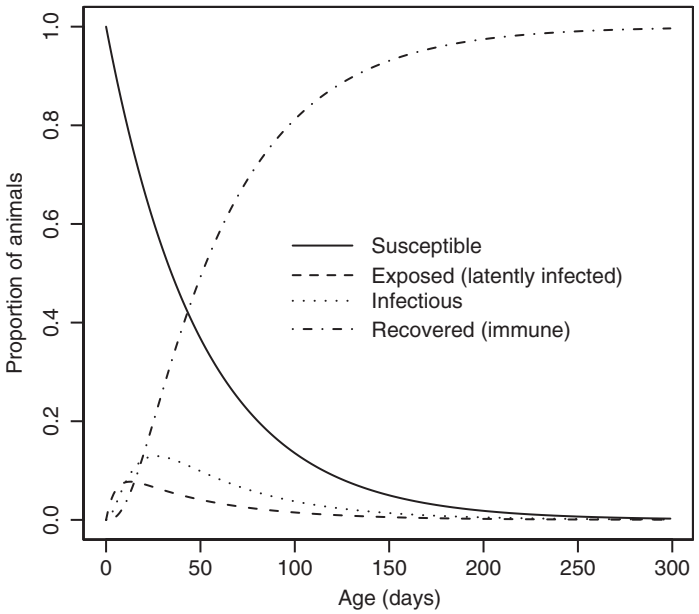

Fig. 15-3. Output from the SEIR model when mean infectious period ($1/\gamma$) is 10 days, mean latent period ($1/\theta$) is 5 days, $\mu = 0.0005$ and $b = 0.02$.

This shows the proportion of animals that are expected to be susceptible, latently infected, infectious, and immune under representative parameter values. It is noteworthy that with the relatively short infectious period assumed here (10 days), only a very small proportion of animals in most age groups are infected. This is despite the fact that infection rate, b, is high enough in this scenario to ensure that almost all animals surviving for a year or more will have been infected. Thus, despite the fact that the infection is highly endemic and would be hard to eradicate, the very low prevalence in all but the youngest age groups implies that the risk to workers handling potentially infected animals will be small provided that contacts are with animals more than about 150 days old. In general, the age profile of infectiousness will depend on all the model parameters (those used here are entirely arbitrary). Knowledge of the pattern of infection with age in different settings, however, could inform risk assessments and control policies aimed at minimizing exposures to infected animals.

A corollary of this observation is that control policies that aim to reduce the total tick population and hence the infection rate $\beta(a)$ (taken to be a constant, b, here) could, under some circumstances, increase the risk to humans. Such a perverse outcome could arise if people in high-risk occupations (veterinarians, slaughterhouse workers, etc.) were preferentially exposed to animals above the age at which most animals became infected in the absence of interventions. If the infection rate, b, was initially high, even large reductions in the tick population could have little effect on the total number of animals escaping infection, but could dramatically affect the ages at which animals became infected (Fig. 15-4). Progressive reductions in b have the effect of substantially increasing the likelihood that older animals are infected. Thus, for the highest value of b in this scenario, there are almost no infected animals which are older than 200 days. As b is reduced there becomes an appreciable chance that animals in these age groups will be infectious, putting those who have contact with them at risk. Such risks should be considered when evaluating the likely benefits of any control measure that aims to reduce infection but that is unlikely to lead to overall control of the disease.

The above equations provide a simple but general description of the infection process in an endemically stable environment (i.e. when disease incidence is neither increasing nor decreasing, and the infection rate $\beta(a)$ does not change with time). If we are interested in studying the temporal evolution of a system that is changing over time (perhaps due to the implementation of a control measure) we need to modify the approach. Most importantly, we expect the probability that a tick is infected, r, to change with the number of animal hosts infected. A simple and biologically plausible assumption is that r will increase linearly with the prevalence of infected hosts. This can be represented mathematically as $r(t) = cY(t)/N$ where c represents the probability that infection passes from an infectious hosts to an uninfected tick and N is the total number of hosts. The infection rate can then be expressed as a function of time (rather than age) as $\beta(t) = qTr(t - \varphi)$, where φ is the time a newly infected tick takes to develop and become infectious. If we

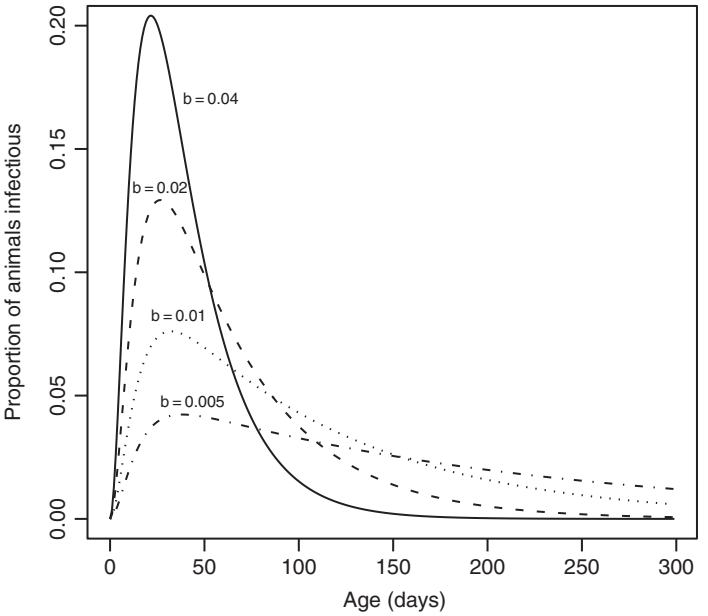

Fig. 15-4. Proportion of animals which are infectious by age of animal. Results from the SEIR model under the same assumptions as Fig. 15-3 are shown, except that the infection rate, $\beta(a) = b$, is progressively reduced by factors of 2.

replace $\beta(a)$ in Equations (3) above with this expression for $\beta(t)$, add a term μN to right-hand side of the Equation (1) to represent births (assumed to balance deaths from all compartments), and replacing a with t, we will have a model suitable for studying the nonequilibrium situation.

One of the most useful concepts in infectious disease epidemiology is the case reproduction number: the average number of secondary cases caused by a primary case. It is useful to distinguish between two reproduction numbers: the basic reproduction number (R_0) and the effective reproduction number (R_t). R_0 is defined as the average number of secondary cases produced by a typical primary case in an otherwise fully susceptible population in the absence of control measures. R_t is defined similarly, except that the population is not required to be fully susceptible and there may be control measures in place. In the absence of control measures R_t can be calculated as the product of R_0 and the proportion of the population that is susceptible. The value of R_t is therefore always less than or equal to R_0.

R_0 is of central importance because its value determines whether or not an epidemic is possible; only when it takes a value greater than one, so that each primary cases generates on average more than one secondary case, can the chain reaction that constitutes an epidemic proceed. When R_0 is less than 1 there may still be chains of disease transmission, and occasionally these may even be quite

long, but these will be self-limiting and have no chances of leading to the self-sustaining chain reaction of a full-blown epidemic that affects thousands.

R_t is important since it determines the rate of epidemic growth at a given point in time. During the course of an epidemic R_t will decline from its initial value of R_0 as the number of susceptibles decreases. It may also decline as result of interventions designed to control the disease. When R_t is equal to 1 the epidemic is neither growing nor falling. If there is no supply of new susceptibles, R_t will proceed to fall below 1 as the pool of susceptibles decreases further. If there is a sufficient supply of new susceptibles (through birth, loss of immunity, or immigration) it is possible to reach an endemic equilibrium state where the generation of new susceptibles matches the loss of susceptibles due to infection. In this case R_t is maintained at a value of 1, and the amount of infection neither increases nor decreases over time. Cyclic behavior can occur when the epidemic causes R_t to fall well below 1, and a longer-term increase in susceptibles due to births eventually restores R_t to a value above 1 permitting another epidemic. This is the cause of the cyclic pattern of childhood diseases such as measles [8].

For a tick-borne disease, we can define R_0 as the average number of secondary infections in hosts from one primary infected host (when all hosts and ticks are susceptible) or, equivalently, as the average number of infectious ticks that arise from a single infected tick in a susceptible population. This definition immediately leads to an expression for R_0 in terms of parameters we have already introduced:

$$R_0 = \frac{qTc\theta}{(\gamma + \mu)(\theta + \mu)} \tag{4}$$

This can be derived by multiplying the probability that the first tick infects its host, q, the probability that the host survives to the infectious state, $\theta/(\theta + \mu)$, the mean number of ticks attaching to the infectious host (which is equal to product of the average duration of the host's infectious period, $1/(\gamma + \mu)$, and the mean rate of tick attachment, T), and the probability that each tick that attaches becomes infected, c. Using this formula (and estimates for the parameters) the impact of interventions on R_0 can be derived in terms of expected impacts of intervention on different aspects of the system. In particular, it is possible to calculate the reduction of the tick attachment rate, T, that would be needed to reduce R_0 to below 1, resulting in local eradication of the disease. An alternative formulation expresses R_0 in terms of the equilibrium infection rate β^*: $R_0 = 1 + \beta^*/\mu$. This holds only when β^* is greater than zero and, therefore, when R_0 is greater than 1. It is useful because an estimate of β^* is relatively easily obtained from age-stratified seroprevalence data as described above, and enables a simple assessment of the likely effort needed to eradicate the virus in a given population.

The above notes provide only a very broad outline of a framework for modeling tick-borne infections, but one that could easily be applied to CCHF if supported by appropriate field research. Such an approach can, and – in the context

of other pathogens – has been extended in a variety of ways; for example, to address the particular biological details of different pathogens, to more fully describe the dynamics of the tick population [16], and to provide economic analyses of control policies [14]. In particular, when modeling interventions that are likely to affect the tick population, models that describe details of the tick life cycle (and how these are affected by the intervention) will usually be required. One interesting recent example of this is due to Ogden et al. [17], who developed a dynamic population model of the tick *Ixodes scapularis*, dividing the tick population into 12 developmental stages, with the aim of investigating the effects of climate (and predicting the effects of climate change) on the range and seasonality of the tick.

Though such complex models are becoming quite common and are, in many cases, entirely appropriate, the degree of detail that should be included in a model will vary according to the application. The general question of how complex models should be was succinctly addressed by Albert Einstein, who said that models should be as simple as possible, but no simpler. This applies as well to infectious disease epidemiology as it does to astrophysics: models should be only as complex as is required to address the questions at hand. Many unfamiliar with mathematical models naively believe that more complex models will provide better answers to most questions. In fact, the reverse is usually true, and for most purposes, including prediction, surprisingly simple models tend to perform better.

15.4. MODELING THE NOSOCOMIAL TRANSMISSION OF CCHF

The approach described above for modeling the tick-borne transmission of CCHF virus in animal hosts used a deterministic formulation. When the populations under investigation are large, this is a reasonable approach: the time evolution of the system is likely to be quite predictable and individual chance events (e.g. whether or not one animal gets infected, how long it takes another to recover) are unimportant. Just as casinos may lose on individual bets but are sure to win in the long run, in a large population when many are infectious there are so many individual unpredictable events that the eventual outcome becomes highly predictable. In small populations, such as hospital units, and at the beginning of epidemics in both large and small populations when only small numbers are infected, this deterministic approach fails badly. The details of the random events that make up an epidemic become important. For example, if the first infected person happens to die before he has a chance to infect anyone an epidemic will not occur, even if it had the potential to (i.e. even if R_0 was greater than 1). A deterministic model would predict a major epidemic every time, which is clearly unrealistic.

The importance of such chance events is illustrated in Figs. 15-5–15-7. These show results from Monte Carlo simulations from a stochastic susceptible-exposed-infectious-recovered (SEIR) model which has a structure similar to that

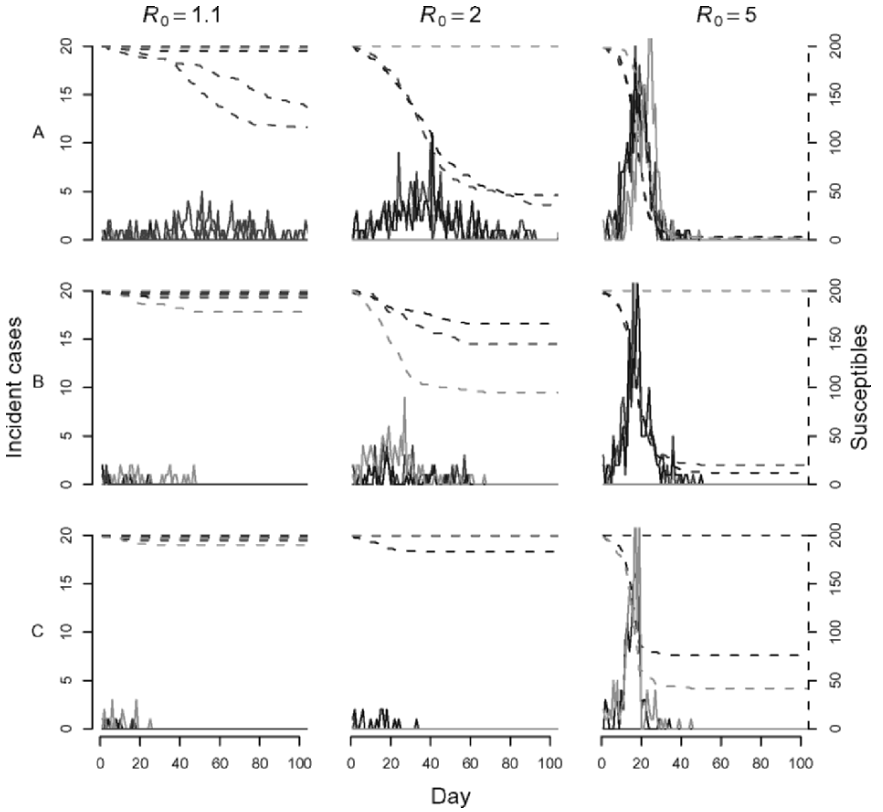

Fig. 15-5. Incident cases (*solid lines*) and number of susceptibles (*broken lines*) from stochastic simulations of an SEIR epidemic model when there is no intervention (row A), an intervention to reduce the probability of transmission from each case by 50% after 20 days (row B), and an intervention to reduce this probability by 90% (row C). Columns (*left* to *right*) show results for initial R_0 values of 1.1, 2, and 5. Three runs from each scenario are shown, except when R_0 is 1.1, when 10 runs are shown (since in this case in most simulations the epidemic does not take off). A mean incubation period of 5 days is assumed (daily probability of progressing from latent to infectious is 0.2) and a mean infectious period of 3.3 days (daily probability of ceasing to be infectious is 0.3), and at day 0 there are 200 susceptibles and one latent case.

shown in Fig. 15-2 (except that births and nondisease-related deaths can be ignored because we are interested in short timescales over which it is reasonable to assume a fixed population size). This simple model could be considered to provide an approximate description of transmission in a small population such as a hospital unit that has stopped admitted new patients (when new patients continue to be admitted the constant supply of susceptibles leads to rather different dynamics [5]). The model assumes that each day each infectious person has some fixed chance of infecting each susceptible person and some fixed probability, γ, of ceasing to be infectious. When infected, individuals enter a latent compartment, with

a daily probability, θ, of progressing to the infectious compartment. The basic reproduction number, R_0, in this model is equal to the mean infectious period ($1/\gamma$) multiplied by the probability that a given susceptible person is infected by one infectious case on a given day, multiplied by the initial number of susceptibles. Figure 15-5 (top row) indicates typical model outcomes in a population of 200 initially susceptible people when no intervention to control the epidemic is made. When R_0 is greater than 1 there is a chance of an epidemic. This chance increases with R_0 (Fig. 15-6). In contrast to the deterministic model, there is also a real chance that that the epidemic dies out almost immediately. This can be seen in Fig. 15-5, where even though R_0 is greater than 1 in some of the simulation runs the epidemic fails to take off and the susceptible population stays near to its initial value of 200. Figure 15-6 shows that even when $R_0 = 2$, there is no secondary transmission at all in about one third of the simulations. As R_0 increases above 1 there is an increasing chance that if there is any secondary transmission a large number of infections will result. This gives rise to bimodal distribution when R_0 is above 1 (Fig. 15-6). It is also noticeable that as R_0 increases, when the epidemic does take off it tends to peak earlier and affects more people, though the precise course of the epidemic is not predictable (Fig. 15-5).

If an intervention is able to reduce the effective reproduction number, R_t, below 1 by reducing the probability of transmission from each case then the

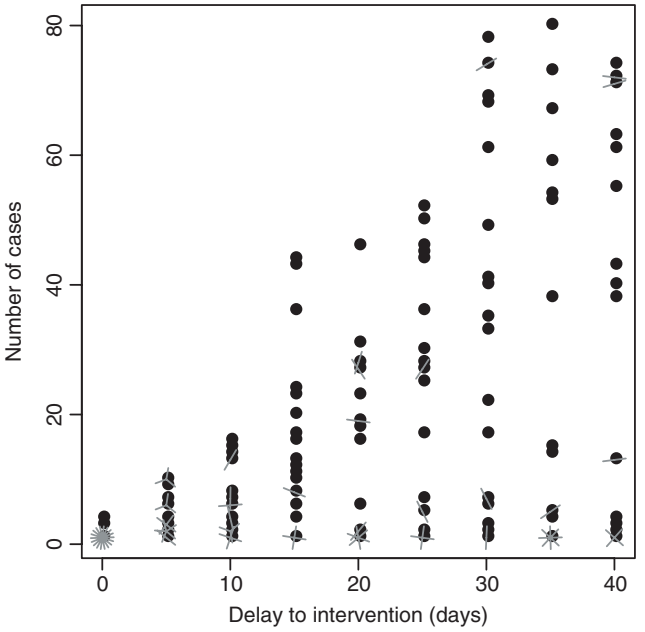

Fig. 15-6. Distributions of outbreak sizes for different values of the basic reproduction number, R_0. Results from 1,000 simulation runs are shown for each value of R_0. Model parameters as in Fig. 15-5.

epidemic will be controlled. This is shown in rows B and C of Fig. 15-5. When transmission probability is reduced by 50% after 20 days (row B) R_t is reduced to approximately half the initial R_0 value. In the first column this is sufficient to control the epidemic (because $R_t \approx 0.55$ at day 20), and though some transmission persists the epidemic comes to an end soon. When R_0 is 2 a 50% reduction in transmission gives an $R_t \approx 1$ at day 20, which is enough to permit prolonged transmission at a low level, but not to allow a large epidemic. In contrast, when R_0 is 5 a 50% transmission reduction brings R_t down to about 2.5. This is not enough to control the epidemic and most of the susceptibles go on to be infected (though fewer than would have been without the intervention). In contrast, the 90% reduction in transmission in row C is sufficient to reduce R_t below 1 in all cases, and in all simulations runs the epidemics are quickly brought under control. Nevertheless, the large number of infected cases by day 20 when R_0 is 5 is sufficient to ensure continued transmission for some time after the intervention.

Figure 15-7 shows the effect of delays in an intervention that is able to bring about control of the epidemic but unable to prevent all transmission. In this

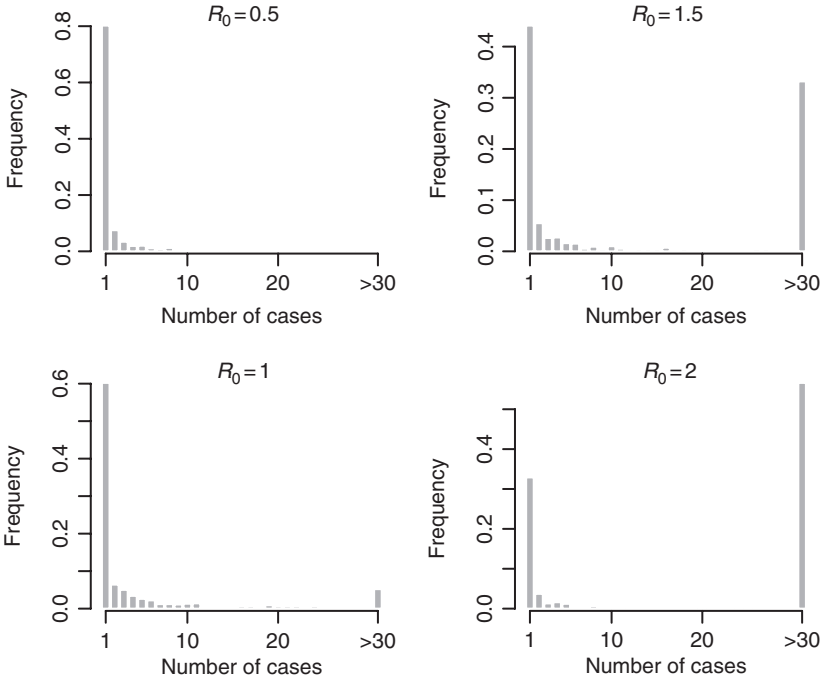

Fig. 15-7. Effect of delays in making an intervention that reduces transmission by 90%, assuming $R_0 = 2$. Twenty simulations were performed for each delay (measured in days since the first infectious case). Where two or more simulation runs have given the same result, the number of "petals" on each point of the sunflower plot indicates the number of simulation runs represented.

scenario if an intervention able to reduce the probability of transmission by 90% is made as soon as the first case is detected there is a high probability that there almost all secondary spread can be eliminated. Delays of 5 days or more lead to significant chances that much larger numbers will be affected, though there remains a substantial chance that the epidemic will die out of its own accord and affect few people. This accounts for the bimodal distribution in the number of cases that starts to become apparent when interventions are delayed by 20 days or more.

15.5. FUTURE DIRECTIONS FOR USING MODELS TO HELP UNDERSTAND CCHF

Future work modeling the spread of CCHF, whether tick-borne or nosocomial, will depend on good quality field data to enable necessary parameter estimates to be obtained and an appropriate model structure selected. Precise details of models will depend to a large extent on the questions that are being asked.

The simple SEIR modeling framework illustrated here can readily be extended to account for the fact that not all individuals infected will themselves become infectious. This can be done by constructing models that allow people to move with some probability straight from the exposed to the recovered compartment. It is also relatively easy to further modify this basic model structure to better capture observed distributions of the latent and infectious periods and to account for variable infectivity with time since infection. Appropriate modifications can also be made to account for more complex mixing patterns (e.g. probabilities of patients infecting health-care workers and other patients may differ).

Models have also proved to be useful for evaluating the role of contact tracing, quarantine, and isolation in the control of infectious diseases. This work has shown how the epidemiological characteristics of different diseases (the basic reproduction number, the length of the latent and infectious periods, and the amount of transmission that occurs prior to the onset of symptoms) largely determine the likely success of control using these measures [7]. As one might expect, the properties of CCHF put it well within the region where these measures can be expected to be effective. The severe acute respiratory syndrome (SARS) epidemic also highlighted the high degree of variability in the number of secondary cases produced by each primary case [10]. When detailed contact tracing data are examined it turns out that this pattern is seen in many infectious diseases [11]. This variability may greatly exceed that assumed in the simple stochastic models presented here. It is, nevertheless, a simple matter to account for such variability in models. Its main impact would be to make major epidemics rather less likely for a given R_0, and to make the role of chance even more dominant.

In many cases it is useful to estimate R_0. The estimate will tell us how close we are to a risking a major epidemic (if R_0 is currently below 1) and allow precautionary measure to be taken. If R_0 is greater than 1 the estimate tells us how

much an intervention would have to do to bring about control or eliminate the chance of a major epidemic. An approach that has proved useful for other diseases is to estimate R_0 from the distribution of the number of cases from clusters of transmission [6]. This is possible because the probability of 0, 1, 2, . . . secondary cases varies with the value R_0 (Fig. 15-6), so the likelihood of different R_0 values can be derived from the distribution of the number of cases.

When more detailed surveillance data are available other approaches can be used to provide much better estimates of R_0 and to assess the impact of interventions. The best of these uses computationally intensive Markov chain Monte Carlo algorithms to estimate the basic reproduction number and other model parameters and quantify the uncertainty in these estimates. Such an approach has recently been used to estimate the basic reproduction number for Ebola and to evaluate the role of interventions in reducing transmission [9]. This method could be adapted relatively easily to study CCHF transmission, and detailed data from outbreaks used to evaluate the evidence of effectiveness for different interventions. Much simpler approaches based on deterministic approximations are also possible [4], but these methods appear not to accurately characterize the uncertainty and may therefore be inappropriate for assessing the evidence of effectiveness of different interventions [9].

REFERENCES

1. Anderson RM, May RM (1991) Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford
2. Bernoulli D (1766) Essai d'une nouvelle analyse de la mortalité causée par la petite vérole. Mém Math Phys Acad R Sci Paris 1–45; English translation by Bradley, L. in Smallpox Inoculation: An Eighteenth-Century Mathematical Controversy, Adult Education Department, Nottingham, 1971
3. Brisson M, Edmunds WJ (2003) Economic evaluation of vaccination programs: the impact of herd-immunity. Med Decis Making 23:76–82
4. Chowell G, Hengartner NW, Castillo-Chavez C, Fenimore PW, Hyman JM (2004) The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. J Theor Biol 229:119–126
5. Cooper BS, Medley GF, Scott GM (1999) Preliminary analysis of the transmission dynamics of nosocomial infections: stochastic and management effects. J Hosp Infect 43:131–147
6. De Serres G, Gay NJ, Farrington CP (2000) Epidemiology of transmissible diseases after elimination. Am J Epidemiol 151:1039–1048
7. Fraser C, Riley S, Anderson RM, Ferguson NM (2004) Factors that make an infectious disease outbreak controllable. Proc Natl Acad Sci USA 101:6146–6151
8. Keeling MJ (1997) Modelling the persistence of measles. Trends Microbiol 5:513–518
9. Lekone PE, Finkenstädt BF (2007) Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. To appear in Biometrics. Available online at: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1541–0420.2006.00609.x
10. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D, Murray M (2003) Transmission dynamics and control of severe acute respiratory syndrome. Science. 300:1966–1970
11. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. Nature 438:355–359

12. Medley GF, Perry BD, Young AS (1993) Preliminary analysis of the transmission dynamics of *Theileria parva* in eastern Africa. Parasitology 106:251–264
13. Meltzer MI, Norval RAI (1993) Mathematical models of tick-borne disease transmission. Parasitology Today 9:277–278
14. Mukhebi AW, Chamboko T, O'Callaghan CJ, Peter TF, Kruska RL, Medley GF, Mahan SM, Perry BD (1999) An assessment of the economic impact of heartwater (*Cowdria ruminantium* infection) and its control in Zimbabwe. Prev Vet Med 39:173–189
15. Nowak M, May RM (2000) Virus Dynamics. Oxford University Press, Oxford
16. O'Callaghan CJ, Medley GF, Peter TF, Perry BD (1998) Investigating the epidemiology of heartwater (*Cowdria ruminantium* infection) by means of a transmission dynamics model. Parasitology 117:49–61
17. Ogden NH, Bigras-Poulin M, O'Callaghan CJ, Barker IK, Lindsay LR, Maarouf A, Smoyer-Tomic KE, Waltner-Toews D, Charron D (2005) A dynamic population model to investigate effects of climate on geographic range and seasonality of the tick *Ixodes scapularis*. Int J Parasitol 35:375–389
18. Ross R (1911) The Prevention of Malaria. Murray, London