27

CONCEPTS OF INFECTIOUS DISEASE EPIDEMIOLOGY

Chapter contributed by Ian Dohoo and Graham Medley

OBJECTIVES

After reading this chapter, you should be able to:

1. Understand why infectious disease data are fundamentally different from other forms of data previously dealt with in this text.

2. Know the terms used to describe infection and disease processes and understand the basic principles of disease transmission.

3. Understand the principles of modelling infectious disease transmission including SIR and SEIR models

4. Understand the concepts of effective contact rate and R_0.

5. Be able to estimate R_0 from epidemiologic data using a variety of approaches including regression and mathematical modelling.

27.1 INTRODUCTION

In previous chapters, we have examined methods for analysing data in which we assumed the observations were independent (Chapters 14-19), or had some dependency (Chapters 20-26). This dependency may have arisen as a function of the observations sharing a common environment, being repeated measures on the same individual or being spatially related. Infectious diseases are different in that the dependency among individuals is related to the state of other individuals in the population. A new infection can only arise if the agent is present in one or more other individuals in the population (or is present in the environment as a result of shedding by other individuals). Thus, the probability that a susceptible individual will become infected in a given time period depends on the number of other individuals in the population shedding the agent.

Infectious diseases have a number of other unique features which must be considered when analysing these data.

- The system is dynamic. The probability of new infections occurring changes over time as the mixture of infectious and susceptible individuals in the population changes. Additionally, the dynamics of the infection process are non-linear. In practical terms, this means that the system does not respond in proportion to changes. For example, vaccinating 50% of a population with an effective vaccine will reduce the incidence of new infections by more than 50%.
- There is heterogeneity among populations due to the stochastic nature of infection transmission. Two populations with exactly the same starting conditions (*ie* population structure and environment, number of infectious individuals *etc*) could have very different disease outbreaks due to the fact that transmission of the agent among individuals is a stochastic process. The probability of transmission depends on the rate of contact between individuals and the probability of transmission occurring if contact is made. However, for any given probability, a range of possible outbreak scenarios is possible. (The impact of this is shown in Section 27.4.4.)
- Threshold effects are present. If one infected individual passes the infection (on average) to less than one other individual, the infection will die out. The number of new infections arising from one individual in a fully susceptible population is called the basic reproductive number (R_0) and this is discussed in more detail in Section 27.4.
- In susceptible populations, the increase in the number of infected individuals is often exponential (*ie* has a constant doubling time) until a substantial portion of the population has been exposed and developed some immunity.
- Infectious agents are continually evolving. For example, pathogenicity of an agent, susceptibility to drugs, *etc* will change over time. Given the relatively short life cycle for many agents, this evolution may take place quite rapidly.

As a consequence of these features, interventions (*eg* vaccination, medication) can dramatically affect the disease process in a population. In particular, interventions targeted at a proportion of a population will have an effect on those individuals not targeted.

It is also important to distinguish between **microparasitic** and **macroparasitic** infections. Microparasites are essentially those infectious agents that are too small to count individually, so that modelling the parasite population explicitly is not practically possible. Thus, when modelling such infections, it is necessary to consider the state of the host (*eg* susceptible,

infected, immune) leading to **compartmental models**. Macroparasites are those infections for which there is sufficient information and reason to consider modelling the parasite population explicitly using **intensity models** where the principal outcome is a measure of the number of parasites (*eg* parasite burden, tick count).

At a modelling level, the distinction between microparasite and macroparasite is essentially pragmatic. Some infectious agents (*eg Theileria* sp.) may be treated as either largely depending on the data available and the problem being addressed. This chapter will focus on compartmental models for microparasitic infectious diseases. For information on macroparasitic infections, the reader is referred to Anderson & May (1991); Cox (1993).

Microparasite infections tend to be associated with the following biological characteristics.

- The agents are generally small and (relatively) antigenically simple (*eg* viruses and bacteria).
- The agents generally multiply rapidly within the host resulting in either death of the host or the production of immunity. Pathogenicity tends to be high. Immunity tends to be strong and long lasting. The duration of infection (*ie* before death or immunity) can be short.
- Sexual reproduction (recombination) is relatively rare and not obligatory; strain variation, for example, in pathogenicity is often considerable.
- If the duration of infection is short, the prevalence of infection is low, and persistence of infection within a population requires continual supply of susceptible hosts (*eg* through births or antigenic change).

In contrast, macroparasitic infections tend to have the following characteristics.

- The agents are large and (relatively) antigenically complex (*eg* nematodes, lice, ticks).
- The agents have complex life cycles involving multiple hosts or free-living stages, and multiply at different rates within the different life-cycle stages.
- The agents have the capacity (and often requirement) for sexual reproduction, and may additionally have asexual multiplication stages.
- Immunity to the agents is often relatively ineffective.
- The agents often have apparent low pathogenicity individually, but the disease impact increases with the numbers (burden) within a host (*eg* gastro-intestinal parasites). The burden of infection is key to the impact that these agents have on the host.
- At the population level, macroparasitic infections are characterised by high prevalence of infections with the number of agents present within the host dependent on external factors that affect the life cycle of the agent (*eg* weather).

These are biological characteristics or general tendencies rather than fixed rules. Examples of archetypal microparasite and macroparasite are rinderpest and *Ascaris*, respectively. Many bacterial and protozoan infections sit somewhere between the 2. For example, bacteria causing bovine intramammary infections may have low pathogenicity (rarely kill the host), generate little immunity in the host and may be prevalent at quite high levels. Nevertheless, they are modelled using compartmental models because it is not possible (or necessary) to quantify the number of bacteria present.)

27.2 INFECTION VS DISEASE

Infection and **disease** must not be confused. While infection is necessary for disease, not all infected animals will become diseased (*ie* develop clinical signs). (**Note** The fact that infection is a necessary cause for disease is a function of the fact that we name diseases etiologically—*eg Salmonella sp* is a necessary cause of salmonellosis. If diseases were named according to their manifestation, specific agents are no longer necessary causes—*Salmonella sp* is not a necessary cause of diarrhoea.). The time courses of infection and disease are shown in Fig. 27.1.

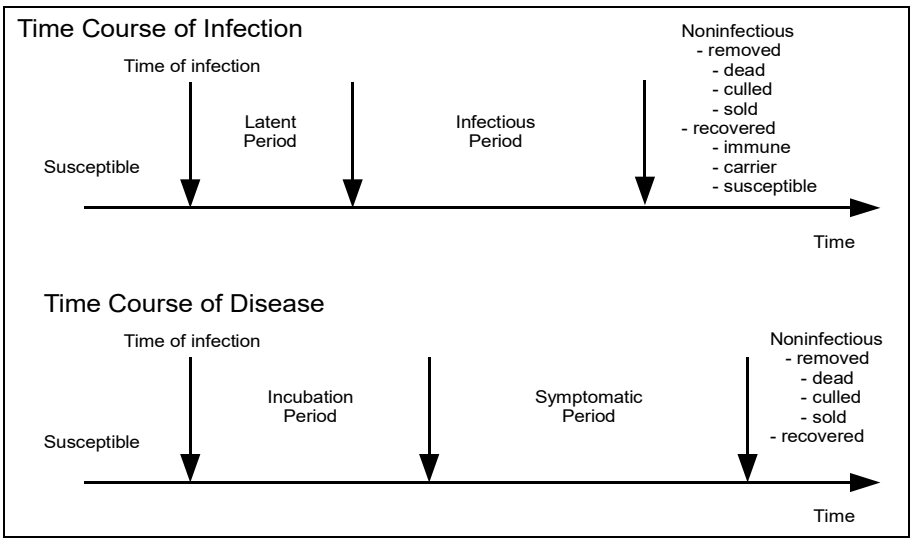

Fig. 27.1 Time courses of infection and disease processes

The states of infection are as follows. A **susceptible** individual becomes infected by transmission (see Section 27.3) of the agent from an infectious individual. The individual then enters a **latent period** when the infectious agent is present but the individual is not capable of transmitting the infection. This latent period may be very short (*eg* approximately one day for avian influenza or very long (*eg* many years for *Mycobacterium avium* subspecies *paratuberculosis*—or MAP for short). (**Note** for non-infectious diseases, the term 'latent period' refers to the time from the onset of detectable changes (*eg* lesions present or changes in biochemical parameters evident) to the onset of clinical signs. This is also referred to as the asymptomatic disease period). The latent period is followed by an **infectious period** during which the host is capable of transmitting the agent. Again, this may be short (*eg* approximately 2 weeks for the infectious salmon anemia virus) or long (years for MAP). The lengths of the latent and infectious periods cannot be observed directly and are usually estimated from experimental infections. The infectious period ends when the host is removed from the population (*eg* dies, culled or sold) or becomes non-infectious through an intervention (*eg* treatment) or the development of immunity which either eliminates the agent from the host or sufficiently suppresses replication to effectively prevent transmission. For some infections, immunity may be long lasting (*eg* canine distemper) while for others the individual may quickly return to a susceptible state (*eg* intramammary infections).

The states of disease are as follows. Once a susceptible individual becomes infected, it enters

the **incubation period** and remains there until signs of disease develop or immunity develops (in which case the individual may never be symptomatic). The length of the incubation period cannot be observed directly unless the timing of infection is known. Methods for estimating the incubation period from outbreak data have recently been reviewed (Cowling *et al,* 2007). The symptomatic period ends when the host is removed (*eg* dies or culled) or recovers (with or without treatment).

The relationship between the **latent period** and the **incubation period** is important from a disease control perspective. If the latent period is shorter than the incubation period (*eg* foot-and-mouth disease), there will be asymptomatic individuals which are shedding the agent and this complicates disease control (it is not adequate to focus solely on symptomatic animals). If, on the other hand, the latent period is longer than the incubation period (*eg* canine heartworm) prompt treatment or removal of animals as soon as they develop signs can reduce/prevent the spread of disease. Individual animals vary considerably in their time course of infection and disease. This variation may be related to environmental factors (*eg* nutrition), host genetics (*eg* breed), agent genetics (*eg* strain), dose of infection *etc*.

27.3 TRANSMISSION

From the viewpoint of the agent, it must transmit in order to survive because no host is immortal. There are a number of routes whereby **transmission** of an agent from an infectious individual to a susceptible one can occur.

- **vertical transmission**—transplacental or perinatal transmission from a mother to her offspring (eg *Neospora caninum* in cows)
- **horizontal transmission**
  - indirect transmission
    - **vector borne**—often by insect vectors but may also be by contaminated needles or other fomite (*eg* mosquito transmission of West Nile virus, transmission of *Staphylococcus aureus* by milking machine).
    - **reservoir species**—transmission of an agent by another species (*eg* rabies transmission from an infected bat to cattle).
  - **direct transmission**
    - **close contact**—transmission requires close contact between 2 individuals (*eg* bovine tuberculosis).
    - **casual contact**—less intimate contact is required for transmission (*eg* agents associated with respiratory disease complex in feedlot calves).
    - **sexual transmission**—requires sexual contact (or spread by artificial insemination) (*eg Trichimonas fetus* in cattle).
    - **air/water borne transmission**—can occur over long distances by the agent being carried by wind or water movement (*eg* foot-and-mouth disease).
    - **contaminative transmission**—the agent can survive for prolonged periods of time in the environment and exposure of susceptibles is from environmental exposure (*eg* anthrax).

The rest of this chapter will focus primarily on diseases with direct transmission, particularly those spread through close or casual contact or sexual transmission.

The probability of direct transmission taking place depends on the frequency of **contacts**

among individuals in the population and the probability of transmission occurring given that contact is made. Contacts are defined as those encounters between individuals that could (but not necessarily would) result in the transmission of the agent if one individual was infectious and the other susceptible. For *Trichimonas fetus,* this requires the mating of a cow by a bull. The **contact rate** (c) is defined as the number of contacts that an animal makes with other animals in one time period. Table 27.1 presents a list of parameters and definitions used throughout this chapter, along with a simple set of calculations, based on the first day of a hypothetical outbreak, showing their relationships.

Table 27.1 Infectious disease parameters

Parameter	Description	Assumption/Calculation
S, I, R, N	the numbers of susceptible, infectious, removed and total number of animals in the population (respectively)	S=996, I=4, N=1000
c	rate of contacts an animal makes with other animals in one time period	10/fish/day
p	probability of transmission of the infection if one animal is infectious and one is susceptible	0.15
cp	rate of 'effective' contacts	1.5/fish/day
I/N	proportion of population that is infectious	4/1000=0.004
$\lambda=cp(I/N)$	rate at which susceptible animals becomes infectious - this is equivalent to our usual definition of I— (incidence rate) (see Section 27.1) - this is also called the 'force of infection' or 'transmission rate per susceptible'	1.5x0.004=0.006 new inf. per fish per day (or 0.006 per fish-day)
$i=\lambda S$	Incidence=rate at which new infections are occurring in the population - this is the population incidence rate (designated i to differentiate it from I used elsewhere in this book)	0.006x996≈6 new inf per day
d	Duration=duration of the infectious period	5 days
cpd	rate of effective contacts per infectious period	1.5x5=7.5/fish/period
$s=S/N$	proportion of the population that is susceptible **Note** In a completely susceptible population $S_0=N$ so $s_0=1$	996/1000≈1
$R_0=cpd$	R_0=basic reproductive number=# of new cases that arise from an infectious individual in a completely susceptible population.	R_0=1.5x5=7.5
$R_t=cpds_t$	R_t=effective reproduction number=# of new cases arising from each infectious individual at time t. **Note** $s_t=S_t/N$ **Note** at t=0, s=1, so $R_0=cpd$	

Note The calculations in this table pertain to the first day of the hypothetical outbreak. On the second day, there will be 10 (6+4) infected fish so λ becomes 1.5*0.01=0.015 and i increases to 0.015*990=14.85.

In animal populations, the nature and frequency of contacts is heavily determined by the structure and management of the population. Many domestic animals are kept in situations of high density, resulting in high contact rates (as a result of casual contacts). How the animals are managed will also affect contact patterns. Cows kept in a closed herd and bred by artificial insemination will not contact *T. fetus*. On the other hand, cows kept on a community pasture with multiple bulls may have many potential contacts. Mixing is considered **homogeneous** if all animals have an equal chance of contacting all other animals. Mixing is considered **heterogeneous** if, for reasons such as behaviour, management *etc*, animals are more likely to contact other animals with certain characteristics than individuals randomly chosen from the population (*eg* contact between cattle varies with stage of the oestrus cycle).

The **probability of transmission** (p) is defined as the proportion of contacts which result in transmission if one of the individuals is infectious and the other susceptible. The probability of transmission depends on the characteristics of the agent, the nature of the contact and the degree of infectiousness of the infected animal. The product of the contact rate and probability of transmission (cp) is called the **effective contact rate**.

27.4 MATHEMATICAL MODELLING OF INFECTIOUS DISEASE TRANSMISSION

Terminology/Acronyms

Throughout this chapter, the term I is used to denote the number of infectious individuals in a population. (It does not represent an incidence rate as used elsewhere in the book). The term R is used to denote the number of 'removed' individuals. R is also used for the reproductive number, but in this case it will always have a subscript R_0 or R_t.

Mathematical models of disease transmission may be developed and used for 3 reasons (Green & Medley, 2002).

- Conceptual models may be used to better understand the effects of interventions in disease control processes. For example Medley *et al* (2008) used conceptual models to evaluate the effects of control programs based on active surveillance.
- Conceptual models may be combined with data derived from experimental research to develop predictive models of disease transmission within populations having set characteristics. These can then be used to estimate the effects of disease control interventions. For example see Tildesley *et al* (2006).
- Conceptual models may be combined with observed data (*eg* from naturally occurring outbreaks) to gain insights into the epidemiology of the disease in those populations and to estimate parameters such as R_0. This chapter focuses on this application of disease models.

Generally, disease models may be **deterministic** or **stochastic**. The former produce a single outcome for a given set of parameters. The latter incorporate the chance nature of events (such as transmission) and produce a probability distribution of possible outcomes (Ball & Neal, 2002; MacKenzie & Bishop, 2001). The mean of the stochastic model is not always the same as the deterministic model, and might show that the outcome has a bimodal distribution; for example when a highly infectious individual is introduced into a susceptible population, deterministically the early outcome is an exponentially increasing epidemic, but the stochastic outcome is bimodal with a peak of probability at no epidemic (if the initial invasion happens to

fade out by chance) and a peak of probability at an epidemic level. In this case, the mean of the stochastic process is equal to the deterministic model.

The simplest model is the **Susceptible-Infectious-Recovered** (**SIR**) model as shown in Fig. 27.2. In an SIR model, susceptible animals (S) are assumed to become infected (and immediately infectious) at a defined rate (λ). Infected animals (I) are assumed to recover (and be immune) at a defined rate (γ). Recovered animals (R) are assumed to remain in that state. The total population size is $N=S+I+R$.

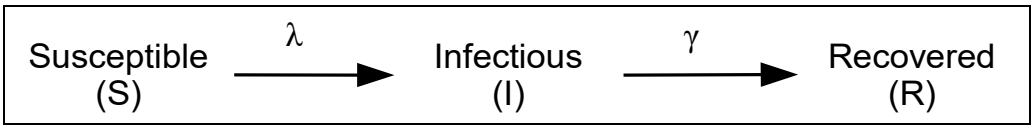

Fig. 27.2 A simple SIR model with two transmission rates

An SIR model can be used to define the key parameters required for modelling infectious diseases. These are all listed in Table 27.1, along with a simple example of their calculation. This example assumes that an infectious agent is introduced into a net pen of completely susceptible Atlantic salmon with the following specific assumptions used in the calculations.

- There are 1,000 fish in a net pen and they mix homogeneously (N=1000).
- The contact rate among fish is 10 contacts per fish per day (*ie* each fish has 10 contacts with other fish on a daily basis) and is homogeneous within the pen.
- A new infectious agent is introduced into the pen and it initially infects 4 fish.
- The probability of transmitting the disease during any one contact is 15% (0.15).
- Fish are infectious for 5 days, at which point they recover and develop sufficient immunity to protect from further infection.

27.4.1 Incidence rates

There are 2 incidence rates presented in Table 27.1. λ corresponds to the usual incidence rate (I) described in Chapter 2. In the example population of 1,000 fish, there are 1,500 effective contacts happening each day, but only 0.4% of them involve infectious fish, so the chance of an individual susceptible fish becoming infected on this first day is only 0.006 (equating to 6 fish becoming infected on that day). From the perspective of a susceptible fish, λ represents their chance of getting infected in a small time period. This is sometimes referred to as the **force of infection**.

As previously noted, the incidence rate (i) in Table 27.1 is different from the use of the term I elsewhere in the book. It is the population incidence rate and represents that rate at which new infections are occurring in the whole population. As the proportion of the population that is susceptible declines, i will fall, even though λ might remain high.

27.4.2 Basic reproductive number (R_0)

As noted in Table 27.1, the basic reproductive number (R_0) represents the number of new infections which arise, on average, from one infected individual when the entire population is susceptible (*ie* at the beginning of an epidemic).

$$R_0 = cpd \qquad \textit{Eq 27.1}$$

R_0 is a product of the rate of effective contacts (cp) and the duration of the infectious period (d). It is a key parameter for understanding infectious diseases. If R_0 >1, then the infection is expected to spread because each infected individual generates, on average, more than one new infection. If R_0 <1, then the infection is expected to die out. However, as noted above, transmission is a stochastic process and there is no guarantee that disease will spread if R_0 >1, or that there will not be additional cases with R_0 <1. In addition, R_0 is an average value for a population and there may very well be clusters of individuals within a population in which the R_0 could be much higher or lower than the average (heterogeneous contact).

As the infection spreads in the population and some individuals move from susceptible to infectious and then to recovered, the population is no longer completely susceptible. The effective reproduction number at time t (R_t) can be computed as:

$$R_t = cpds_t = R_0 s_t \qquad \textit{Eq 27.2}$$

where s_t=S/N (the proportion of the population that is susceptible at time t). This chapter will focus on R_0 as a measure of transmissibility of infection among animals within a population (*eg* pen, herd) but it can also be used at the herd level to quantify transmission between herds (Medley *et al,* 2008; Stegeman *et al*, 1999; Van Nes *et al*, 1998).

27.4.3 R_0 and vaccination

Vaccines are often used to prevent transmission of infection and/or the development of disease in populations. If a vaccine only prevents the development of disease and has no impact on the spread of the infection (*eg* tetanus toxoid), it will have no effect on the dynamics of the agent in the population. We will focus on the situation in which a vaccine has an effect on the transmission of the agent. If a vaccine is 100% effective at preventing infection and is applied to all individuals in the population, then no new infections will occur. However, vaccines are rarely 100% effective and it is often not possible to vaccinate all individuals in a population.

If a proportion (f) of a population is vaccinated with a vaccine which is fully protective, the effective reproductive number (R^*) is:

$$R^* = R_0(1-f) \qquad \textit{Eq 27.3}$$

R^* will be less than one when $f > 1\text{-}1/R_0$, so in this case, the infection will be expected to die out. If you have an estimate of the expected R_0 for an agent in a given population, you can estimate the vaccine coverage that you will need to prevent the spread of infection (sometimes called the **critical percentage** (denoted f_{cp})). For example, if R_0 is 5, then f_{cp}=1-1/5 = 0.8 or 80% of the population needs to be vaccinated with a 100% effective vaccine to prevent the spread of infection. Coleman & Dye (1996) estimated the critical percentage for rabies in dogs (using data from outbreaks in both rural and urban domesticated dogs in 4 countries) to be between 39% and 57%. This is the principle which underlies the concept of **herd immunity**, which states that you do not need to vaccinate every individual in a population in order to prevent an epidemic.

If a vaccine only protects a proportion of individuals (h) (but fully protects those individuals), then R^*=R_0(1-hf) so the proportion of individuals which needs to be vaccinated becomes:

$$f_{cp} = \frac{1 - 1/R_0}{h} \qquad \textit{Eq 27.4}$$

If a vaccine is only partially protective, it may reduce the susceptibility of an individual and hence, reduce the probability of transmission by a factor (designated z_1) from p to z_1p. It may also reduce either the duration of the infectious period by a factor (z_2) and/or the infectiousness of an infected individual by a factor (z_3). If a vaccine has all of these effects, R^* becomes:

$$R^* = R_0 z_1 z_2 z_3 \quad \text{and} \quad f_{cp} = 1 - 1/R^* \qquad \textit{Eq 27.5}$$

For example, the effects of vaccination on transmission parameters for highly pathogenic avian influenza have been investigated (van der Goot *et al,* 2005). These concepts are important not only for modelling the effect of vaccination, but also in planning of field trials to assess effectiveness of vaccination (see Section 11.10).

27.4.4 Limitations of R_0

As noted above, if $R_0 > 1$, you expect the infection to spread in a susceptible population. However, disease transmission is a stochastic process (there is an element of 'chance' as to whether or not transmission will occur). If an infection is introduced into a homogeneous susceptible population, the probability of an outbreak (transmission of the infection past the initial time period) is:

$$p(\text{outbreak}) = 1 - (1/R_0)^{I_0} \quad (\text{assuming } R_0 > 1) \qquad \textit{Eq 27.6}$$

where I_0 is the number of infectious individuals introduced into the population (at t=0) (Keeling, 2005) . Thus, with R_0=7.5 and a single infection introduced, the probability of an outbreak is 87%; but if 4 infections are introduced, this rises to >99%.

However, even if the infection does spread from the initial infectious animals, the size and timing of the resultant epidemics can vary greatly. Similarly, if $R_0 < 1$, there is no guarantee that additional cases will not be observed in the population. Fig. 27.3 shows a range of possible epidemic curves for values of R_0 ranging from 0.5 to 2.0. For each value of R_0, 20 outbreaks were simulated and 5 selected to represent the range of possible outcomes. Each outbreak was based on a population of 100 susceptible individuals with one infectious individual being added and duration (d) is 1 day. The solid line represents the expected course of action (deterministic simulation). Even at R_0=0.5, some limited outbreaks occur. At R_0=2.0, some outbreaks die out quite quickly while in others, most of the population becomes infected.

The consequence of these patterns for estimation of R_0 is that a single outbreak provides relatively little information. (**Note** From a modelling perspective, an outbreak means $R_0 > 1$. From an observed data perspective, an outbreak means that there are more cases of a disease than expected and this can happen with $R_0 < 1$—see panel (a) of Fig 27.3).

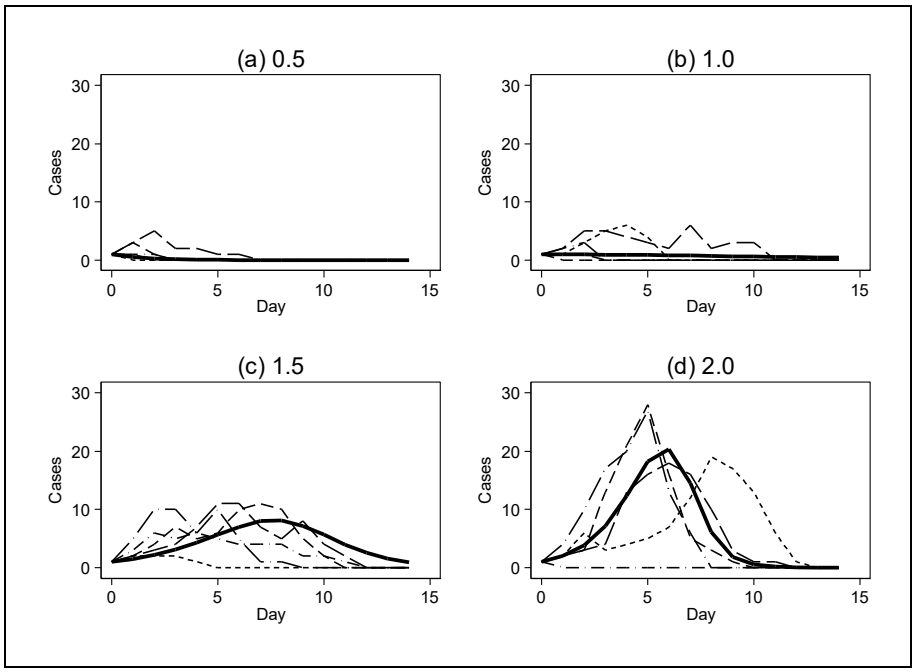

Fig. 27.3 Examples of 6 simulated outbreaks based on R_0 values of 0.5 , 1.0, 1.5 and 2.0 in panels (a) to (d), respectively.

27.5 ESTIMATING R_0 AND OTHER INFECTIOUS DISEASE PARAMETERS

27.5.1 Estimating the Probability of Transmission

Estimation of the contact rate (c) and the probability of transmission (p) is difficult. In outbreak situations, the **secondary attack rate** (**SAR**) is a measure of the probability of transmission. (**Note** SAR is actually a risk not a rate.) The SAR is defined as the probability that a susceptible animal will become infected from the first (primary) case in a population. It is the number of secondary cases (infections derived from the primary case) divided by the number of susceptible animals exposed.

$$\text{SAR} = \frac{\text{number of secondary cases}}{\text{number of exposed individuals}} \qquad \textit{Eq 27.7}$$

SARs are usually calculated from data from small scale outbreaks (*eg* an outbreak in a single pen of cattle) in situations in which the time at which the index case arose is clear. The problem in computing the SAR is deciding which cases are secondary. This requires knowledge about the expected (and minimum and maximum) incubation period of the disease and the latent period of the infection. The time of infection of the primary case (t_0) is estimated based on the expected incubation period and the latent period. Secondary cases are defined as those that occur between t_0 plus the minimum incubation period and t_0 plus the maximum incubation period. This is shown graphically in Fig. 27.4.

An alternative approach to estimating p is to use a **binomial model** of disease transmission.

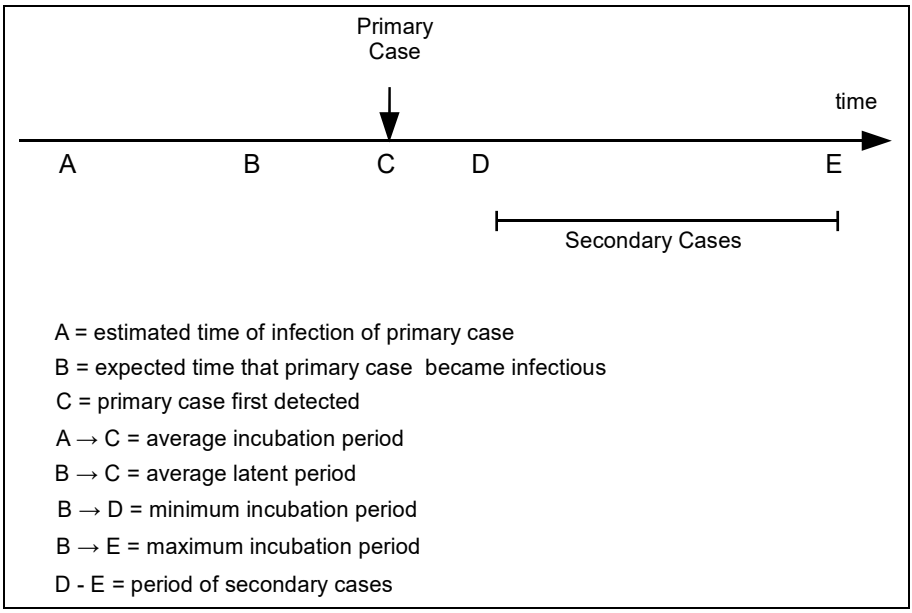

Fig. 27.4 Estimation of a secondary attack rate based on knowledge about expected latent and incubation periods

Rather than basing the probability of transmission on the number of exposed individuals, a binomial model relates the number of new cases to the total number of contacts with infectious individuals. Consequently, it may be appropriate when susceptible animals have multiple contacts with potentially infectious individuals. The probability of transmission during any one contact is p so the probability of avoiding transmission is q=(1-p). The probability of avoiding infection from n potentially infectious contacts is q^n so the probability of becoming infected is $1-q^n=1-(1-p)^n$. These models can be extended over time by assuming that the binomial transmission probability is applicable in discrete time units. These are called **chain binomial models** and include the **Reed-Frost model** and **Greenwood model**. The former assumes that exposure to 2 or more infectious individuals in the same time period are independent events whereas the Greenwood model treats them as a single exposure. Although conceptually appealing for understanding disease transmission, binomial models have not been applied to estimating p.

27.5.2 Estimating R_0

As noted above, R_0 is specific for a given agent in a specified population at a point in time. Nevertheless, estimates of R_0 from some populations can provide some insight into what to expect in new situations. However, given the range of possible outbreaks scenarios that are consistent with a single value of R_0, estimation of R_0 from a single outbreak will be of limited value.

There are a number of approaches to estimating R_0 and some of the more commonly used approaches will be discussed here. See Anderson & May (1991); Becker (1989); Diekmann & Heesterbeek (2000); Dietz (1993) for more complete coverage of these methods.

27.5.3 Estimating R_0 from the exponential phase of an outbreak

During the early stages of an outbreak, the number of new cases occurring is not limited by the availability of susceptibles. This phase is referred to as the **exponential phase** and the **doubling time** (t_d=time it takes for the number of new cases to double) is related to R_0 as follows:

$$t_d = (\ln 2)\, d\, /(R_0 - 1) \approx 0.7\, d\, /(R_0 - 1) \qquad \textit{Eq 27.8}$$

where d is the duration of the infectious period. (This assumes that the outbreak is growing exponentially, which in turn depends on the number of susceptibles not being depleted (Keeling, 2005) . From this, if an estimate of t_d can be derived from the data and an estimate of d is available from experimental research, then:

$$R_0 = 0.7(d\, / t_d) + 1 \qquad \textit{Eq 27.9}$$

One approach to estimating t_d is to fit an exponential regression model, with time as the only predictor, to the data from the exponential phase of an outbreak and:

$$t_d = \ln 2 / \beta \quad \text{and} \quad R_0 = (\beta_d) + 1 \qquad \textit{Eq 27.10}$$

where β is the slope from the regression. If the infection process involves both latent and infectious periods, d is replaced by the **generation interval** which is the sum of the infectious and latent periods. Estimates of t_d (or R_0) using this approach are sensitive to the choice of distribution for the duration of the generation interval (Wallinga & Lipsitch, 2007). This approach was used to estimate R_0 from a subset of 9 cages from infectious salmon anemia (ISA) outbreaks in eastern Canada (Hammell & Dohoo, 2005) (Example 27.1).

27.5.4 Estimating R_0 from the peak of an outbreak

At the peak of an outbreak, where the epidemic curve switches from increasing to decreasing (*ie* the number of new infections in a time period switches from rising to declining) then R_t=1 (*ie* one new infection per existing infection. Because R_t=R_0*s_t=1, then:

$$R_0 = 1 / s_t \qquad \textit{Eq 27.11}$$

where s_t is the proportion of the population susceptible at the time of the peak (called the **critical proportion susceptible**).

27.5.5 Estimating R_0 for endemic diseases

So far, we have considered the situation where an infection is introduced into a population and gives rise to an outbreak or epidemic. However, most infections are endemic (*ie* they survive continuously in a population). Even if there are fluctuations in the incidence over time, the continual existence of the infection means that each infection produces, on a long term average, one other infection (*ie* R_t=1 and therefore s_t=1/R_0). Note that this is the same situation as at the peak of an epidemic. For a disease to remain endemic, there must be a supply of susceptible individuals—the reason that the epidemics in Fig. 27.3 (c) and (d) became extinct was that there were no susceptibles left in the population. The 2 principle sources of susceptible animals are birth and loss of immunity. (The endemic state is actually at a dynamic equilibrium created by the tension between supply and loss (through infection) of susceptibles.)

Example 27.1 Estimation of R_0 from the exponential phase of an outbreak
data = isa_day

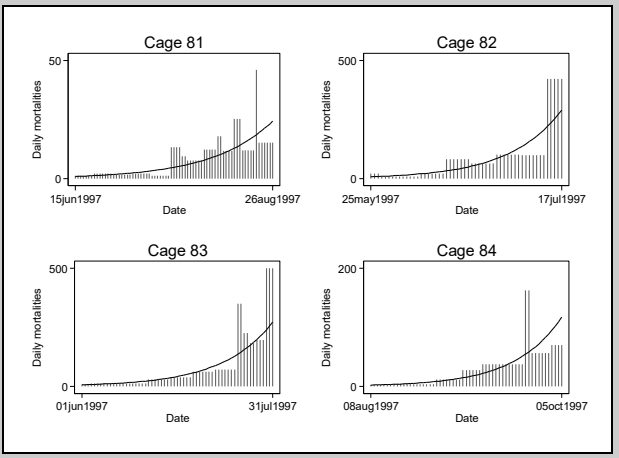

Fig. 27.5 Spikeplots of daily mortalities observed in four cages

These data are a subset (9 cages from one site, but only data from 4 cages are shown in Fig 27.5) of the mortality data that were collected during the initial outbreaks of infectious salmon anemia (ISA) in eastern Canada. As such the population was assumed to be completely susceptible. Dead fish were collected by divers on a periodic basis and the daily mortality was computed as the number of dead fish divided by the interval since the last dive (*eg* if 15 dead fish and the last dive was 3 days ago, it was assumed that the daily mortality was 5 per day). Cages had initial populations at risk of approximately 8,000 to 10,000 fish. The generation interval was estimated to be approximately 21 days and consisted of a seven day latent period and a 14 day infectious period (Mikalsen *et al*, 2001; Moneke *et al*, 2005).

The above graphs are spike plots of daily mortalities with the predicted values from an exponential regression overlaid (for 4 of the 9 cages).

The individual estimates of β, t_d, and R_0 for these 4 cages are shown below. The values of R_0 for all cages (n=9) from this single site ranged from 1.65 to 2.60 with an average of 2.27.

cage	77	78	79	80	81	82	83	84	86	Avg.
β	0.076	0.076	0.070	0.031	0.045	0.068	0.063	0.070	0.045	0.060
t_d (days)	9.17	9.12	9.97	22.53	15.26	10.18	11.05	9.91	15.50	12.52
R_0	2.59	2.60	2.46	1.65	1.95	2.43	2.32	2.47	1.94	2.27

At endemic equilibrium, the average age at which individuals become infected (A) and the average lifespan of individuals (L) can be used to estimate λ (the incidence rate of new infections), as well as R_0 and s^* (the proportion of the population susceptible in that steady state situation).

$$\lambda = 1/A \qquad R_0 = L/A \qquad s^* = A/L \qquad \textit{Eq 27.12}$$

This approach only works when L is a threshold (*ie* all animals hit age L and then die). This would be true for some animals (*eg* those that are slaughtered at a predictable age) but not

generally true (*eg* a dairy herd). The estimate of R_0 is an overall estimate representing an average level of transmission within different age groups and between age groups. Knowledge about contact rates is required to come up with estimates of age-specific estimates of R_0. These assumptions limit the applicability of this procedure for estimating R_0 for endemic diseases but it has been used successfully in situations in which animals are maintained as a distinct group for a considerable portion of their lifespan (Laegreid & Keen, 2004).

The average age of infection can be determined from serologic profiles (*eg* Dietz, 1993), or clinical case data provided most infections produce clinical disease and there is a good system for recording those data. Using serologic data assumes that animals are sampled sufficiently frequently to determine when infection occurred. A related approach for using limited serologic data is based on the final size of the infected (recovered) population (Becker, 1993).

If data about the age-serologic profile of an infection within a population are available, it is possible to derive an estimate of i (population incidence rate). The proportion remaining susceptible at age=a (denoted s_a) can be determined from a survival curve plot (*eg* Fig. 27.6). The slope of this curve at age=a is related to the incidence rate of infections (i) according to the following relationship:

$$\text{slope}(a) = ds/da = -i s_a \quad \textit{Eq 27.13}$$

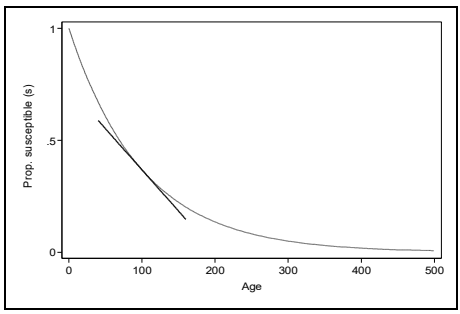

Fig. 27.6 The slope of a survival curve can be used to estimate *i* at any particular age

The approach avoids the assumption that the incidence rate of new infections is constant with age and i can be estimated for different age groups.

Note If only a single age profile at one point in time is available, then time-dependent and age-dependent changes in i are indistinguishable.

For endemic diseases with a stable prevalence, and for which estimates of both the incidence and prevalence are available, the effective contact rate can be estimated. From Table 27.1, $i=\lambda S=cp(I/N)S$, where i is the incidence rate of new infections in the population and I/N is the prevalence of infection. Consequently, $cp=\lambda(N/I)$.

27.5.6 Regression modelling of outbreak data

If data are available from one or more outbreaks, the count of new infections within defined time periods can be modelled using Poisson or negative binomial modelling regressions (see Chapter 18). Given that infectious data usually have extra-Poisson variation (*ie* are more variable than would be predicted by a Poisson process), it is usual to model these data using a negative binomial model.

Fitting a regression model requires 'recreating' the epidemic by 'back calculating' the time at which infections occurred from the observed data. This is shown in Example 27.2 for the ISA data. This example uses point estimates for the duration of the latent and infectious periods but more complex models may assume a distribution for these values (Bos *et al*, 2007).

In order to account for the dependency between the number of new infections at time t and the

state of the population at an earlier time period (in terms of numbers of susceptible and infectious individuals), an **offset** is included in the model. The offset represents the state of the population at the time point when the new infections arose (which may be more than one time period before if the agent has a latent period). Calculation of the offset required an assumption about the nature of the contact pattern within the population. If it is assumed that the offset depends on the proportion of the population that is infectious (I/N) then a **density independent** estimate is computed (SI/N). This situation is applicable when contact is driven by behaviour (*eg* schooling of fish) and does not depend on the size of the population. With this offset $R_0=cpd$. On the other hand, if contact is assumed to be totally random (completely chance encounters) then a **density dependent** estimate (computed as SI) is more appropriate. This is more common for infections for which transmission requires close contact. In this situation $R_0=cpdN$. For a more detailed discussion of this issue, see Begon *et al* (2002); McCallum *et al* (2001).

Once an estimate of β_0 is obtained, R_0 can be computed as follows.

$$R_0 = e^{\beta_0} d \qquad \textit{Eq 27.14}$$

where β_0 is the regression intercept and d is the assumed duration of the infectious period.

The analysis presented in Example 27.2 assumes that all infected fish died. However, this assumption was unrealistic because, in outbreaks which were left to run their course (*eg* Cage 83—Fig. 27.7), the median mortality was only 6.6% (Hammell & Dohoo, 2005). The fact that the outbreak died out after such a small proportion of the fish had died suggested that many fish had become immune and not died.

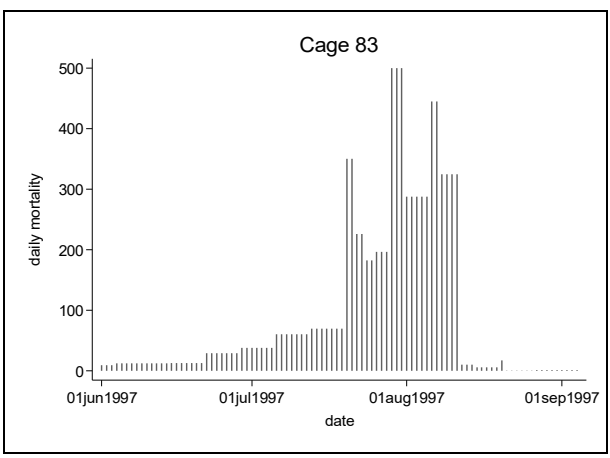

Fig. 27.7 Spike plot of daily mortalities in cage 83

Given the difficulties associated with measuring immunity in fish, it is impossible to determine what proportion of infected fish die and what proportion develop immunity. In this case, the only option was to make an assumption about the proportion of infected fish that die. A sensitivity analysis of this assumption is presented in Example 27.3.

27.5.7 Differential equation models

A very simple SIR model was presented in Fig. 27.2. More complex models can account for more complex interactions between an agent and a host. For example, a newly infected individual may go through a latent period before becoming infectious. This gives rise to an SEIR (Susceptible—Exposed—Infectious—Recovered) model. Similarly, individuals may either die or recover (with immunity) after the infectious period, giving rise to the model shown in Fig. 27.8 (Chowell *et al*, 2007; Keeling & Rohani, 2007; Medley *et al*, 1993). In this model,

Example 27.2 Regression modelling of outbreak data
data = isa_wk

The data are the same as those used in Example 27.1 except that the mortalities were recorded on a weekly basis and it was assumed that infected fish had a latent period of one week followed by an infectious period (*d*) of 2 weeks (Mikalsen *et al* (2001);Moneke *et al* (2005)). Data from 9 cages at one site were used in the analysis.

The outcome of interest for this analysis was the number of new infections (*E*) in a particular week. It was assumed that a single infected fish was added to the population at week -3. This individual fish ultimately generated eight new latent infections in week -2 (*E*) and these fish in turn became infectious fish in weeks -1 and 0 (*Ia* and *Ib*). Ultimately these eight fish died and were the first cases observed (*R*) (week 1). These eight deaths (in week 1) corresponded to new infections in week -2, so the offset for these new infections was computed from the values of *S*, *I* and *N* from the week before (week -3). For the analysis of these data a density independent offset (*SI/N*) was used. Because *N* was large and relatively constant for the duration of these outbreaks, use of a density dependent offset would have produced very similar results. Only records starting from week 0 were used in the regression analysis because this was the first week for which complete information about the number of infectious individuals was available.

cage	Week	Pop. at risk	S	E	Ia	Ib[1]	R	N
81	-3	8120	8119	0	1	0	0	8120
81	-2	8112	8112	8	0	1	0	8120
81	-1	8097	8097	15	8	0	1	8119
81	0	8085	8085	12	15	8	0	8119
81	1	8070	8070	15	12	15	8	8111
81	2	8062	8062	8	15	12	15	8096
81	3	7982	7982	80	8	15	12	8084
81	4				etc			

[1] infectious weeks 1(Ia) and 2(Ib), I=Ia+Ib

Initially, negative binomial models were fit to each outbreak separately.

cage	77	78	79	80	81	82	83	84	86	Avg.
β_0	0.68	0.56	0.90	-0.05	0.05	0.71	0.42	0.72	0.20	0.48
R_0	3.93	3.51	4.92	1.90	2.11	4.07	3.04	4.11	2.44	3.34

The range of values for β_0 was -0.05 to 0.90, which corresponded to values of R_0 of 1.90 to 4.92 with an average value of 3.3 (alternatively, the average β_0 can be converted to R_0 to obtain an average of 3.0).

The data were then compiled into a single dataset and a random effects negative binomial model was fit to the combined data. The overall estimate of β_0 was 0.57 which corresponded to an R_0 of 3.54.

there are only 3 rate parameters (λ γ δ) but there is an additional probability (probability of death) which needs to be estimated. Much more complex models can be found in the veterinary literature, but are beyond the scope of this text. For example, contribution of infections from

Example 27.3 Sensitivity analysis
data = isa_wk

The data from each outbreak were adjusted to allow for only a proportion of the infected fish dying. In the following table, it has been assumed that for every fish that died, there were 4 others which were infected but recovered and became immune. Now the numbers of exposed (*E*) and infectious (*Ia* and *Ib*) fish at each time point was 5 times what it was in Example 27.2.

cage_id	week	par	S	E	Ia [2]	Ib [2]	R	N
81	-3	8133	8132	0	1	0	0	8133
81	-2	8131	8122	10	0	1	0	8133
81	-1	8127	8102	20	10	0	1	8132
81	0	8125	8092	10	20	10	0	8132
81	1	8121	8072	20	10	20	10	8130
81	2	8113	8032	40	20	10	20	8126
81	3	8098	7957	75	40	20	10	8124
81	4				etc			

The outcome of interest is the number of new (latent) infections so, as in Example 2, the offset is based on the total number of infectious fish in the previous week. Random effects negative binomial models were fit to the data. Five models were fit based on the assumption that the ratio of fish infected to fish dying ranged from 1 to 5. These models are summarised in the following table.

infected:dying ratio	1	2	3	4	5
β_0	0.571	0.602	0.633	0.667	0.706
R_0	3.54	3.65	3.77	3.90	4.05

Although there is a general upward trend, the estimates of R_0 are not particularly sensitive to the assumption about the infected:dying ratio. This is because, although the estimated number of new infections goes up substantially as the ratio is increased, so too does the number of infectious fish which generate those infections.

environmental sources has been added for scrapie outbreaks in sheep (Hagenaars *et al*, 2003) and *Salmonella dublin* infections in calves (Nielsen *et al*, 2007). Other factors such as age effects on susceptibility and reproductive status on transmission can be incorporated into complex models (Cherry *et al*, 1998; Hagenaars *et al*, 2003).

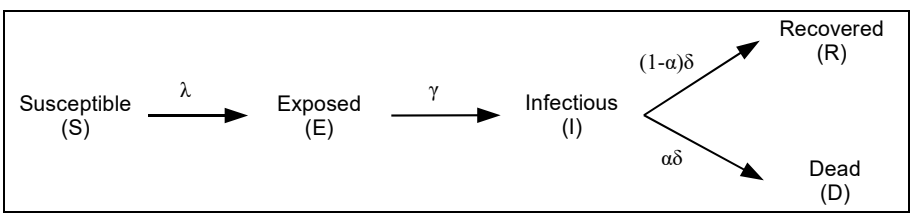

Fig. 27.8 SEIR model that allows for animals to either recover or die

Based on the model for the disease of interest, it is possible to derive estimates of the rate

parameters for the model by solving a set of differential equations. For the model shown in Fig. 27.8 (SEI(RD)) model, the equations to be solved are:

$$\frac{dS}{dt} = -\lambda S \qquad \frac{dE}{dt} = \lambda S - \gamma E \qquad \frac{dI}{dt} = \gamma E - \delta I$$

$$\frac{dR}{dt} = (1-\alpha)\,\delta I \qquad \frac{dD}{dt} = \alpha\,\delta I$$

where α is the probability of an infected fish dying (vs recovering and becoming immune).

These equations define the rates of change in the numbers of fish in each compartment. For example, the change in the number of infectious fish (*ie* dI/dt) is the difference between the rate at which exposed fish become infectious (γE) and the rate at which infectious fish die (δI). The number of recovered fish increased at a rate $(1-\alpha)\delta I$, and the number of dead fish increases at a rate $\alpha\delta I$. Table 27.1 spells out the dependency between λ and I, so that the differential equations for *S, E*, and *I* could be written as:

$$\frac{dS}{dt} = -cp\frac{SI}{N} \qquad \frac{dE}{dt} = cp\frac{SI}{N} - \gamma E \qquad \frac{dI}{dt} = \gamma E - \delta I$$

This indicates more clearly that the rates of change of numbers in each compartment are dependent on the numbers in each compartment—the source of the non-linear nature of epidemics.

The differential equations provide rates of change for each compartment, given the state of each compartment. In order to solve them, we require 'initial conditions' or a starting point. Generally it is impossible to derive an analytical solution for such equations, (*ie* equations of the form $S_t = S_0\, e^{-\lambda t}$) because λ depends on *I* which depends on *S*. Consequently, these equations must be solved numerically. This process follows on from what is presented in Table 27.1—taking a small step in time, one can calculate the new numbers of susceptible, exposed, infected, removed and dead fish, and then repeat this process. To solve these equations accurately requires a small step size (one day is probably too long), so that this only becomes feasible using computer packages, and there are many available for solving such systems of ordinary differential equations (*eg* MatLab, Berkeley Madonna, Mathematica, Maple).

This class of problem is known as an 'initial value problem'—given the model assumptions, parameter values and initial conditions for *S*, *E*, *I*, *R*, and *D* at t=0, what are values of *S*, *E*, *I*, *R*, at t >0? A particularly strong, implicit assumption is that the rates of progress through the *E* and *I* compartments is constant. This is equivalent to a constant hazard rate in survival analysis so, in the ISA example, the assumed duration of stay in classes *E* and *I* is exponentially distributed with means of 7 and 14 days, respectively.

There are a number of ways in which these equations can be fit to the data. The most straightforward is to use non-linear regression to find the values of cp and α that minimise the sum of squared differences between the model and the data (lines and points). R_0 can then be computed as $R_0 = cpd$. This approach was applied to the 9 ISA outbreaks analysed in Examples 27.2 and 27.3 and the results are in Example 27.4. The only known quantity was the number of deaths in each time period and the initial population at risk (at t=0).

As shown in Fig. 27.3, epidemics are stochastic by nature: an individual outbreak rarely

Example 27.4 Differential equation models of ISA outbreaks
data = solutions2.dta

Numerical solutions of the ordinary differential equations arising from the SEI(RD) model were calculated using a standard MatLab function (ode45) to solve the set of equations. Durations in the *E* and *I* compartments were assumed to have exponential distributions with means 7 and 14 days, respectively. Initial conditions at t=0 (S, E, I, R, D) were set to (9969,0,1,0,0). The process was used to generate the estimates of *cp* and α which generate the epidemic curve that best fit the observed data.

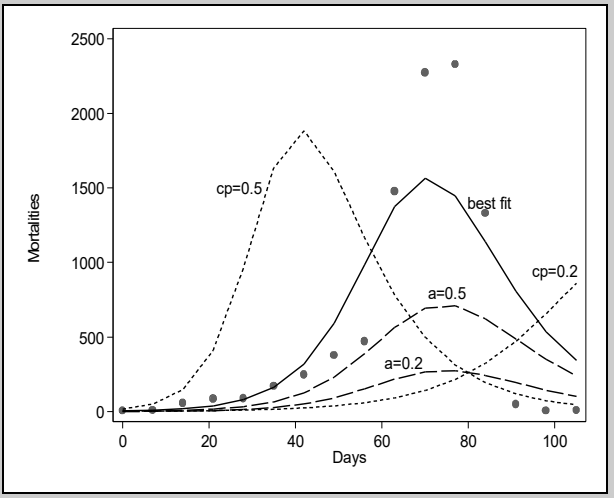

Fig. 27.9 Four poorly fitting curves (see text for legend). Data from cage 83

Four combinations of *cp* and α that generate poorly fitting curves (dashed lines) were plotted along with the observed data (points) and the best fit solution (*cp*=0.3115 and α=1) (solid line). Clearly, the values of *cp* and α (and therefore R_0) have a large influence on the shape of the epidemic.

When applied to all 9 cages from this site, the estimates of *cp*, R_0 and α were as shown below.

cage	77	78	79	80	81	82	83	84	86	Avg.
cp	0.38	0.43	0.60	0.21	0.35	0.40	0.31	0.46	0.29	0.38
R_0	5.34	6.01	8.46	2.89	4.96	5.59	4.36	6.44	4.00	5.34
α	0.23	0.06	0.52	0.25	0.10	1.00	1.00	0.24	0.18	0.40

The estimates of R_0 are somewhat larger than those derived from the negative binomial regression approach. This is not surprising for 2 reasons. First, the mathematical modelling approach assumes that the latent and infectious periods had exponential distributions (with means of 7 and 14 days, respectively) while the regression approach assumed that all fish had periods of exactly 7 and 14 days. Second, the mathematical modelling approach estimates instantaneous rates compared to weekly rates estimated by the regression approach. Given the differences in procedures it is not surprising that the 2 approaches give different results. However, in the long run (with lots of high quality data observed on short time intervals), one would expect the 2 sets of estimates to converge.

An advantage of the mathematical modelling approach is that it allows for α (proportion of infected fish that die) to be freely estimated rather than assuming a fixed value as in the regression approach.

conforms to the expected average. Thus, the *cp* estimated from a single outbreak (as in Fig. 27.9) will not capture the possible values of *cp* that might have created these data. If data from multiple outbreaks are available, the range of estimates will provide more insight into the dynamics of the infectious process. An alternative is to use stochastic models which provide insight into the range of possible values for parameters such as R_0 which are compatible with the single outbreak data. These are beyond the scope of this book, but examples include Bjørnstad *et al* (2002); Ferrari *et al* (2005); Gibson & Renshaw (1998).

REFERENCES

Anderson RM, May RM. Infectious Diseases of Humans. Dynamics and Control. Oxford Univ. Press; Oxford. 1991.

Ball F, Neal P. A general model for stochastic SIR epidemics with two levels of mixing. Math Biosci. 2002; 180: 73-102.

Becker NG. Analysis of infectious disease data. Chapman and Hall; London. 1989.

Becker NG. Martingale methods for the analysis of epidemic data. Stat Methods Med Res. 1993; 2: 93-112.

Begon M, Bennett M, Bowers RG, French NP, Hazel SM, Turner J. A clarification of transmission terms in host-microparasite models: numbers, densities and areas. Epidemiol Infect. 2002; 129: 147-53.

Bjørnstad ON, Finkenstädt BF, Grenfell BT. Dynamics of measles epidemics: Estimating scaling of transmission rates using a Time series SIR model Ecological Monographs. 2002; 72: 169-84.

Bos MEH, Van Boven M, Nielen M, Bouma A, Elbers ARW, Nodelijk G, Koch G, Stegeman A, De Jong MCM. Estimating the day of highly pathogenic avian influenza (H7N7) virus introduction into a poultry flock based on mortality data Vet Res. 2007; 38: 493-504.

Cherry BR, Reeves MJ, Smith G. Evaluation of bovine viral diarrhea virus control using a mathematical model of infection dynamics. Prev Vet Med. 1998; 33: 91-108.

Chowell G, Nishiura H, Bettencourt LMA. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. J R Soc Interface. 2007; 4: 155-66.

Coleman PG, Dye C. Immunization coverage required to prevent outbreaks of dog rabies. Vaccine. 1996; 14: 185-6.

Cowling BJ, Muller MP, Wong IOL, Ho L, Louie M, McGeer A, Leung GM. Alternative methods of estimating an incubation distribution: examples from severe acute respiratory syndrome Epidemiology. 2007; 18: 253-9.

Cox FEG. Modern Parasitology: A Textbook of Parasitology. Blackwell Science; Oxford. 1993.

Diekmann O, Heesterbeek JAP. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Wiley; Chichester. 2000.

Dietz K. The estimation of the basic reproduction number for infectious diseases. Stat Methods Med Res. 1993; 2: 23-41.

Ferrari MJ, Bjørnstad ON, Dobson AP. Estimation and inference of R0 of an infectious pathogen by a removal method. Math Biosci. 2005; 198: 14-26.

Gibson GJ, Renshaw E. Estimating parameters in stochastic compartmental models using Markov chain methods IMA J Math Appl in Med Biol. 1998; 15: 19-40.

Green LE, Medley GF. Mathematical modelling of the foot and mouth disease epidemic of 2001: strengths and weaknesses. Res Vet Sci. 2002; 73: 201-5.

Hagenaars TJ, Donnelly CA, Ferguson NM, Anderson RM. Dynamics of a scrapie outbreak in a flock of Romanov sheep--estimation of transmission parameters. Epidemiol Infect. 2003; 131: 1015-22.

Hammell KL, Dohoo IR. Mortality patterns in infectious salmon anaemia virus outbreaks in New Brunswick, Canada. J Fish Dis. 2005; 28: 639-50.

Keeling MJ. Models of foot-and-mouth disease. Proc Biol Sci. 2005; 272: 1195-202.

Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press; Princeton (NJ). 2007.

Laegreid WW, Keen JE. Estimation of the basic reproduction ratio (R0) for Shiga toxin-producing *Escherichia coli* O157:H7 (STEC O157) in beef calves Epidemiol Infect. 2004; 132: 291-5.

MacKenzie K, Bishop SC. Developing stochastic epidemiological models to quantify the dynamics of infectious diseases in domestic livestock. J Anim Sci. 2001; 79: 2047-56.

McCallum H, Barlow N, Hone J. How should pathogen transmission be modelled? Trends Ecol Evol. 2001; 16: 295-300.

Medley GF, Cooper BS, Green LE. The dynamics of test and clear programmes: interaction between surveillance and control Prev Vet Med 2008, submitted.

Medley GF, Perry BD, Young AS. Preliminary analysis of the transmission dynamics of Theileria parva in eastern Africa. Parasitology 1993; 106 (Pt 3): 251-64.

Mikalsen AB, Teig A, Helleman AL, Mjaaland S, Rimstad E. Detection of infectious salmon anaemia virus (ISAV) by RT-PCR after cohabitant exposure in Atlantic salmon Salmo salar. Dis Aquat Organ 2001; 47: 175-81.

Moneke E, Groman DB, Wright GM, Stryhn H, Johnson GR, Ikede BO, Kibenge FSB. Correlation of virus replication in tissues with histologic lesions in Atlantic salmon experimentally infected with infectious salmon anemia virus. Vet Pathol. 2005; 42: 338-49.

Nielsen LR, van den Borne B, van Schaik G. *Salmonella dublin* infection in young dairy calves: transmission parameters estimated from field data and an SIR-model. Prev Vet Med. 2007; 79: 46-58.

Stegeman A, Elbers AR, Smak J, de Jong MC. Quantification of the transmission of classical swine fever virus between herds during the 1997-1998 epidemic in The Netherlands. Prev Vet Med. 1999; 42: 219-34.

Tildesley MJ, Savill NJ, Shaw DJ, Deardon R, Brooks SP, Woolhouse MEJ, Grenfell BT, Keeling MJ. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK Nature. 2006; 440: 83-6.

van der Goot JA, Koch G, de Jong MCM, van Boven M. Quantification of the effect of vaccination on transmission of avian influenza (H7N7) in chickens. Proc Natl Acad Sci USA. 2005; 102: 18141-6.

Van Nes A, De Jong MC, Buijtels JA, Verheijden JH. Implications derived from a

mathematical model for eradication of pseudorabies virus. Prev Vet Med. 1998; 33: 39-58.

Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci. 2007; 274: 599-604.