4

Binomial and Stochastic Transmission Models

4.1 Overview

How we think about the transmission dynamics of an infectious agent within a host population influences how we design, analyze, and interpret vaccine studies. It can influence our choice of interventions. In this chapter and the next we introduce transmission models necessary for estimating and understanding the effects of vaccination. In this chapter, we present the binomial model and the chain binomial model. These models are central in the formulation of statistical models for estimating transmission parameters and vaccine efficacy parameters. They form the basis of the models in Chapters 10 through 11. The binomial model is also the basic building block of the small- and large-scale stochastic simulation models of vaccination interventions in populations, that can also be used to produce data for design of vaccine studies. In a stochastic model, whether an event occurs is random, depending on a number produced by a random number generator described later.

In Chapter 5 we present simple differential equation transmission models that are generally deterministic. That is, every time the equations are solved, the same answer is obtained. This approach is essential to understanding large complex models of the population effects of vaccination programs, but less relevant to our purposes in this book. Historically, much of theoretical discussion of the effect of vaccination on the basic reproductive number R_0 stems from the solution of differential equations models, so the chapter also contains further discussion of R_0 and the effects of vaccination.

Without getting too formal, all of the models in this and the following chapters assume that people can be in discrete states, such as susceptible, infected but latent, infected and infectious, or recovered. The binomial models in this chapter are discrete event models, in that whole individuals become infected or recover. They are particularly interesting for analyzing data because the likelihood functions for the discrete events can be easily formulated. Binomial models can be formulated in discrete time or in continuous time as we shall show.

In contrast, in the differential equations models, the number of people flowing from one state to another, such as from susceptible to infected, is continuous. That is, there can be 450.75 people in the infected compartment. We consider only differential equations models formulated in continuous time, though discrete time versions are sometimes used.

For all transmission models, whether for estimating parameters of interest or for simulating vaccine interventions, the underlying assumptions about how people mix and contact each other is central. We begin this chapter with a general introduction to mixing structures and population dynamics.

4.2 Contact processes and mixing structures

4.2.1 Contact processes

People make contacts in a population before an infectious agent enters the population. How to think about the contact process in a population can depend on the infectious agent of interest. The contacts of interest may be through the air or casual touching. Some models assume that people behave like gas molecules with the rate of contacts being determined by density. If people are pressed more closely together, as in an urban environment, they contact each other more often than if they were less densely distributed, as in a rural environment. Hence, for disease spread by air, droplet, or casual touching, such as measles, influenza, or mumps, population density plays a role in determining the value of R_0. Alternatively, for diseases spread by contacts made by choice, such as in sexual contacts or injection of intravenous drugs, the contacts may be determined more by social behavior. In many cases, both density and social choice will play a role in determining contact rates and mixing patterns.

4.2.2 Random mixing

Under the assumption of random mixing, every person in the transmission unit is assumed to make contact equally with every other person. Thus, an infective person will equally expose every other person in the transmission unit. In a model of the United States based on random mixing, every infective person in the population will expose every susceptible. In a model with households as the basic transmission unit, the assumption of random mixing implies that each person in the household makes contact with the others equally. We denote by c the constant contact rate that does not change over time in a randomly mixing population.

Most populations do not mix randomly. We consider a few approaches to nonrandom mixing.

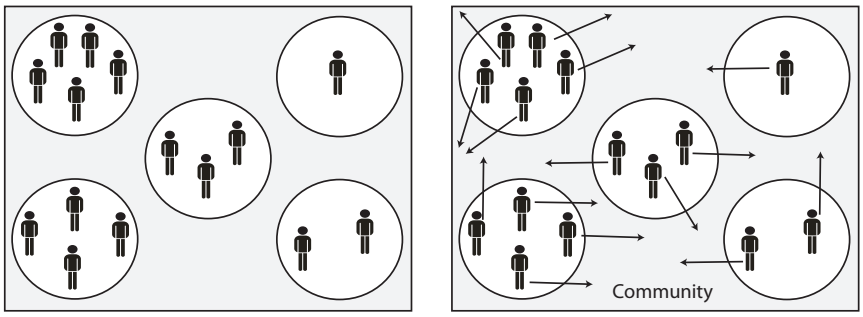

Fig. 4.1. Transmission units

4.2.3 Transmission units within populations

We have considered random mixing in small transmission units such as households. The transmission units can be assumed to be completely separate and independent of one another, as indicated in Figure **??**a. Under this assumption, an infected person in one transmission unit does not expose someone in another transmission unit. This is the assumption that underlies the simple chain binomial model discussed later. Alternatively, the individuals in the transmission units can be assumed to mix in the community at large as well and either expose each other to infection or be exposed to infection from some community source (Figure 4.1b). When we define this community structure, it allows that a susceptible individual can become infected if exposed to an infected person within the household as well as the possibility of being infected in the community at large during the course of an epidemic or over the duration of a study. The transmission units could be households, sexual partnerships, schools, workplaces, or day care centers, for example. These two differing assumptions underly the different approaches in Chapters 11 and 12.

More complex mixing models can be formulated where individuals mix in several transmission units as well as in the community at large. Figure 4.2 represents the mixing structure of a complex influenza model with households, daycare centers, schools, workplaces, neighborhoods, and communities. For example, schoolchildren mix at home, at school, the neighborhoods and the community at large. People are assumed to mix randomly within each structure. Network theory is used to study the contact patterns and social networks of actual populations and simulated populations formally (Morris and Kretzschmar, 1997; Koopman et al, 2000; Eubank et al 2004; Newman et al 2006; Meyers et al 2006).

4.2.4 Mutually exclusive subpopulations

Rather than small transmission units, we may think of a population as divided into large subgroups that mix with members of their own subgroups

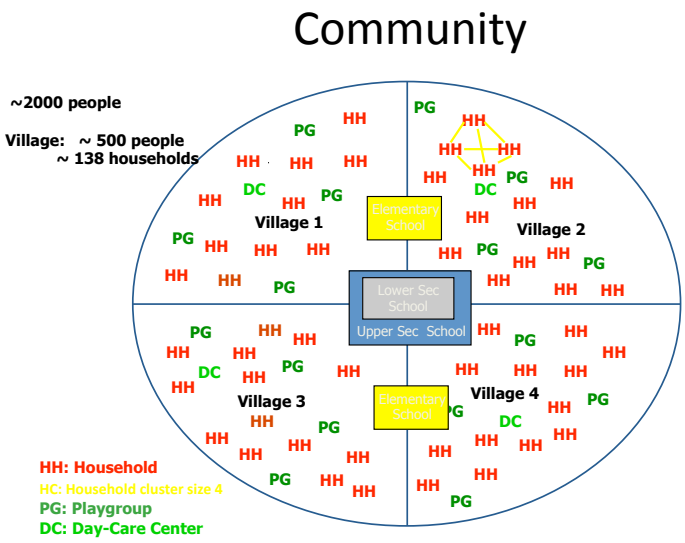

Fig. 4.2. Community structure of where individuals interact in more than one mixing group, including households within household clusters, neighborhoods, and the community, day care centers, play groups, and schools.

differently than with members of other subgroups. A common approach to modeling infectious diseases such as measles (McLean et al 1991) and chickenpox (Halloran et al 1994) is to divide the population into nonoverlapping age groups. In modeling sexually transmitted diseases, the population could be divided into groups with different activity levels (Hethcote and York 1984).

In a population composed of two mixing groups, group 1 and group 2, the contact pattern is described by a *mixing matrix* that has the same number of rows and columns as the number of mixing groups. The entries in the matrix represent the contact rates of individuals within and between the groups. The contact rate of individuals of group j with individuals of group i is denoted by c_{ij}. The mixing pattern of two groups is represented by the matrix

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} . \tag{4.1}$$

On the diagonal are the contact rates within groups, c_{11} and c_{22}. The entries c_{12} and c_{21} off the diagonal represent the contact rates between the groups corresponding to that row and column.

The average number of new infectives that one infective will be produce, R_0, (Chapter 1.3.3) will be higher in the group with the higher within-group contact rate, assuming that the transmission probability and infectious period are the same in both groups. If an epidemic occurs and there is contact between the two groups, the epidemic in the group with the higher contact rates will help drive the epidemic in the group with the lower rates. The group with the higher R_0 would then serve as a *core population* for transmission (Hethcote and York 1984) . The existence of a core group has consequences for intervention programs. It may be easy to reduce the average R_0 for the whole population below 1, while R_0 in the core population remains above 1, so that transmission will persist. In infectious diseases, the chain is only as weak as its strongest link.

Simple social contact data can be used to estimate age-specific transmission parameters for infectious respiratory spread agents (Wallinga et al 2006; Halloran 2006).

4.2.5 Population dynamics

Transmission models can be formulated as open populations with vital dynamics or as closed populations. There are two ways to enter and two ways to leave a population. Individuals can enter a population by being born into it or immigrating. Individuals can leave a population by dying or emigrating. Open populations may include just birth and death with no immigration or emigration. Open populations may also include just emigration, analogous to loss to follow-up. Open populations are analogous to dynamic cohorts. In a closed population, there are no births, immigration, deaths or emigration. The closed population is analogous to a closed cohort. Whether a transmission model is formulated with an open or closed population will depend on the circumstances and time frame of the study. Dynamic consequences of the assumptions are considered in Chapter 5.

4.3 Probability of discrete infection events

We consider the simple binomial model of transmission for discrete contacts and a simple model in continuous time.

4.3.1 Probability of infection in discrete time or contacts

The binomial model is often used to estimate the transmission probability as well as effects of covariates such as vaccination status. The basic idea of the binomial model is that exposure to infection occurs in discrete contacts, which can also be discrete time units of exposure. Generally it is assumed that each contact is independent of another. We have defined p as the transmission

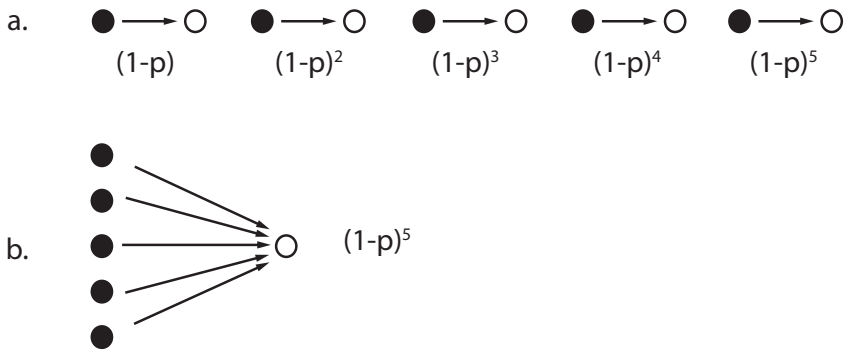

Fig. 4.3. a) The escape probability with five consecutive contacts. b) The escape probability with five simultaneous independent contacts, as in the Reed-Frost model. In both cases, the probability of infection is $1-(1-p)^5$.

probability during a contact between a susceptible person and an infectious person or other source of infection, such as an infectious mosquito. The quantity $q = 1-p$ is the probability that the susceptible person will not be infected during the contact, called the *escape probability*. For example, if the transmission probability for influenza is $p = 0.30$, then the escape probability for one contact is $q = 1-p = 0.70$. If a susceptible person makes n contacts with infectious people, then, assuming all contacts are equally infectious, the probability of escaping infection from all of the n contacts is $q^n = (1-p)^n$. The probability of being infected after n contacts with infectives is $1-q^n = 1-(1-p)^n$.

Suppose a person has five successive contacts with someone who has influenza (Figure 4.3a). What is the probability that the person will have become infected by the five contacts? In this example, $n = 5$. The calculation proceeds by first calculating the probability that the susceptible person will escape infection from all six contacts. Then this number is subtracted from one to get the probability that the person is infected at least once. If the probability of escaping infection from the first exposure is $q = 0.7$, then the probability of escaping infection from the second exposure is the probability of escaping the first one times the probability of escaping the second: $q{\cdot}q = 0.7{\cdot}0.7 = 0.49$. The probability of escaping infection from the third contact is similarly the probability of escaping infection from the first two contacts times the probability of escaping infection from the third, $q^2 \cdot q = 0.49 \cdot 0.7 = 0.34$. The probability of escaping infection from six successive contacts is $0.7^5 = 0.17$. The probability of becoming infected at least once is $1-(1-p)^n = 1-(0.7)^5 = 0.83$.

We have made an important assumption here. We assumed that each successive contact was not affected by any of the previous contacts. That is, the person did not develop immunity or become more susceptible as time went on. We also assumed that all of the contacts had the same risk of transmis-

sion. These assumptions may not be fulfilled. If so, the assumptions can easily be changed and a more complicated form of the binomial model developed. Becker (1989) discusses chain binomial models with random effects.

In a different problem, suppose a susceptible child attends school one day where five of the children simultaneously have influenza. What is the probability of becoming infected (Figure 4.3b)? Assume that the probability of becoming infected from one contact with one child with influenza is $p = 0.3$. Proceeding as before, the probability of escaping infection from one child is $q = 0.7$. Now we can calculate the probability of escaping infection from all six children, with $0.7^5 = 0.17$, so the probability of being infected on that day at school is $1 - q^5 = 0.83$.

Although the answers for the two examples are numerically the same, the biological assumptions in the two examples are different. In the example of influenza at school, each of the five *simultaneous* exposures to infection are the same, and that each additional child with influenza increases the probability of being infected independent of how many other infective children are present. The contacts and exposures to infection are assumed to operate the same as if they were successive and independent. The assumption of independence is commonly used in the binomial model, whether contacts are simultaneous or successive. For instance, this assumption is at the heart of the Reed-Frost model discussed below.

What if, however, biologically we think that once there is one infectious child in a classroom, then the room is saturated with infectious particles? Then adding more infectious children to the school will not increase the probability of becoming infected. We need to change our expression for the probability of becoming infected. If p is the probability of becoming infected from one infected person at school, then $q = 1 - p$ is again the escape probability from exposure to one infected. In contrast to the previous model, however, the probability of becoming infected from exposure to two or more infecteds at the same time is still p and the escape probability is still $q = 1 - p$. Under these biologic assumptions, the probability of becoming infected from one child with influenza on one day is $p = 0.3$, and the probability of becoming infected from simultaneous exposure to six children with influenza on one day is also $p = 0.3$. The Greenwood model (Greenwood, 1931) makes the assumption that the probability of infection on a given day does not change with increased number of infectives. The assumption is, however, seldom used in practice.

4.3.2 Other transmission models

Another way to model the probability of becoming infected is simply to multiply the number of contacts with infectives n times the transmission probability p, np. In the above influenza example, however, $np = 5 \cdot 0.3 = 1.5$. Since probabilities have to lie between 0 and 1, this approach obviously has limits. In particular, either n or p, or both need to be small. Another commonly used

expression for the probability of not becoming infected is e^{-np}, for the probability of becoming infected is $1 - e^{-np}$. In the influenza example above, then, the probability of not becoming infected is $e^{-5 \cdot 0.3} = e^{-1.5} = 0.22$ and for becoming infected is $1 - e^{-1.5} = 0.88$. Comparing this with the probability of being infected calculated from the binomial model, 0.83, we note that they are similar but not identical.

In the influenza example above, the transmission probability is high, and the product of np is large. If the transmission probability is much smaller or the contact rate is much smaller, or both, then the three methods for calculating the probability of becoming infected give similar answers. Suppose again that there are five infectious contacts in one day, but that the transmission probability of the infection in question is just $p = 0.001$. Then using the binomial model, the probability of becoming infected is $1 - (1 - p)^n = 1 - (.999)^5 = 0.00499$. Using the exponential expression, the probability of becoming infected is $1 - \exp(-5 \cdot 0.001) = 0.00499$, and based on the simple expression, $np = 5 \cdot 0.001 = 0.005$. There is little difference in the answers. In this example, the calculated np makes sense as the probability of becoming infected. The two simpler approaches are sometimes used as approximations for the binomial model. They are generally less time consuming to compute than the binomial model, which can be an issue in complex models. However, as we have just demonstrated, the approximation will not always be good. All three models require the same data for estimation of the parameters, namely the number of people who become infected, the number who do not, and the number of contacts made by each person up to when he or she becomes infected.

4.3.3 Probability of infection in continuous time

The above models assume discrete contacts or contacts within discrete units of time. Another approach to modeling the probability of becoming infected assumes that contacts occur in continuous time. The expression cp is the probability of being infected per unit time if all the contacts are with infectious persons, or c is the rate of infectious exposures and p is the transmission probability per exposure. Analogous to the discrete model, the expressions $\exp(-cp)$ and $1 - \exp(-cp)$ are the probabilities of escaping infection or becoming infected per unit time, respectively. If the exposure occurs over some time period Δt, then the probabilities of escape or of infection in the time interval Δt are $\exp(-cp\Delta t)$ and $1 - \exp(-cp\Delta t)$, respectively.

Another notation for the transmission rate per unit time of contact with an infective person is $\beta = cp$. Then the probabilities of escape or of infection in the time interval Δt are $\exp(-\beta\Delta t)$ and $1 - \exp(-\beta\Delta t)$, respectively. Unless data are available on the contact rate separate from the transmission probability, in this model the transmission rate will be estimated from data on the time interval of exposure and infection status of each person in the study.

4.3.4 Contacts with persons of unknown infection status

Sometimes contacts are made with persons or sources of unknown infection status. We denote the probability that an individual with whom a contact is made is infectious by P. Then the probability of being infected from a contact of unknown infection status is $\rho = pP$. The quantity ρ is not a transmission probability in the strict sense, but an infection probability. The probability of escaping infection from contact with someone of unknown infection status is $1 - \rho = 1 - pP$. Under the binomial model, the probability of becoming infected after n such contacts is $1 - (1 - pP)^n = 1 - (1 - \rho)^n$.

Suppose as in the influenza example above that $p = 0.3$ but that the contacts are with five individuals of unknown infection status. If the individuals are randomly chosen from a population where prevalence of influenza is $P = 0.4$, then the probability of being infected after five contacts is $1 - (1 - 0.3 \cdot 0.4)^5 = 0.47$.

An analogous expression can be developed for the continuous time model, as described in Chapter 2, since the hazard rate or incidence rate of infection as a function of the contact rate, the transmission probability, and the prevalence is $\lambda(t) = cpP$. The probability of escaping infection within some period of time Δt is $\exp(-cpP\Delta t)$, and of being infected is $1 - \exp(-cpP\Delta t)$. These examples demonstrate some of the options and subtleties inherent in different approaches to modeling the transmission process.

4.4 Chain Binomial Models

Chain binomial models are dynamic models developed from the simple binomial model by assuming that infection spreads from individual to individual in populations in discrete units of time, producing chains of infection governed by the binomial probability distribution. To use the model, one needs to know the number of susceptibles and number of infectives in each generation. The expected distribution of infections in a collection of populations after several units of time can be calculated from the chained, that is, sequential, application of the binomial model. The Reed-Frost and Greenwood models are examples of chain binomial models. As mentioned above, the Reed-Frost model assumes that exposure to two or more infectious people at the same time are independent exposures. The Greenwood model assumes that exposure to two or more infectious people at the same time is equivalent to exposure to one.

In the Reed-Frost model, the assumption is made that people pass through three states (Figure 4.4). They start out susceptible, denoted by S, then become infected and infectious, denoted by I, after which they recover with immunity, denoted by R. Models of this type of infection process are called SIR models for susceptible, infected, recovered. Sometimes the notation XYZ is used for the three states. This simple model assumes that there is no latent

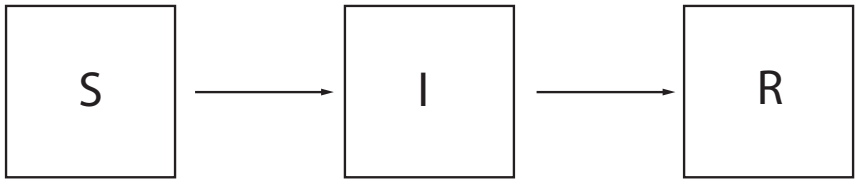

Fig. 4.4. Three states in the Reed-Frost chain binomial model. S, susceptible; I, infective; R, removed (immune).

period and that there are no asymptomatic infections. This model could be a simplified representation of influenza, measles, or chickenpox that ignores the latent period. Other examples include SIS models, in which people recover without immunity to become susceptible again, and SIRS models, in which people acquire immunity, but lose it again to become susceptible. An SEIR model allows people to pass through a latent period denoted by E. In the simple Reed-Frost model, one assumes that the population size is constant N. If there are the only three possible states, then each person in a population of N individuals is in one of these three states, where S_t is the number of susceptible people, I_t is the number of infectives, and R_t is the number of immune people at time t, where the subscript t denotes that the model is in discrete time. In contrast, in the continuous time differential equation models in the Chapter 5, the number of people in each state at the continuous time t is denoted by $S(t)$, $I(t)$, and $R(t)$.

As a simple example of the Reed-Frost chain binomial model, consider spread of infection in a transmission unit, such as a household, with three individuals, where one person is initially infected and the other two are initially susceptible (Table 4.1). The goal is to compute the probability of any of the possible chains. The model assumes that the initial infective is no longer infective after the first time unit. In the first time unit, one of three things can happen. One possibility is that neither of the two susceptibles become infected. A second possibility is that both of them become infected. A third possibility is that just one of them becomes infected. The probability that neither becomes infected is the probability that both escape infection, or q^2. In this case, the chain ends, so the probability of this chain is q^2. If both susceptibles become infected in the first time unit, the chain also ends. The probability of both becoming infected from the first exposure is p^2.

The probability that one person becomes infected from the first infected while the other does not is pq. This can happen two ways, so that the probability of just one of the susceptibles being infected from the initial infective person in the first time unit is $2pq$. If one susceptible is infected in the first time unit, then this person is the new infective who exposes the last remaining susceptible. Exposure of the last remaining susceptible can result in two possible outcomes. Either he becomes infected or he does not, with probabilities

Table 4.1. Chain binomial probabilities in the Reed-Frost model in households of size 3 with 1 initial infective and 2 susceptibles, $S_0 = 2, I_0 = 1$

Chain	Chain probability	at p=0.4	at p=0.7	Final number infected
$1 \longrightarrow 0$	q^2	0.360	0.090	1
$1 \longrightarrow 1 \longrightarrow 0$	$2pq^2$	0.288	0.126	2
$1 \longrightarrow 1 \longrightarrow 1$	$2p^2q$	0.192	0.294	3
$1 \longrightarrow 2$	p^2	0.160	0.490	3
Total	1	1.00	1.00	

p and q respectively. The *chained probabilities* are then $2pq \cdot p = 2p^2q$ and $2pq \cdot q = 2pq^2$, respectively.

In Table 4.1 the chain probabilities are calculated for two different values of p, $p = 0.4$ and $p = 0.7$. In 1000 groups of size three with one initial infective, at $p = 0.4$, 360 of the groups would be expected to have just one infected, 288 to have two infected, and $192 + 160 = 352$ to have three infected at the end. Similarly, at $p = 0.7$, 90 would be expected to have one infected, 126 to have two infected, and 784 to have three infected. Since there are two different chains by which all three people become infected, if we were not able to observe the actual chains, we would not know which path the chain had taken. That is, we may only have data on the number of people who get infected in each transmission unit or household. So we would have only final value data and observe the final size distribution.

The R_0 in the Reed-Frost model, assuming that the duration of infectiousness is one time unit, or $d = 1$, is $R_0 = pN$, or sometimes $R = p(N - 1)$, if there is one initial infective. More generally, $R = p(N - I_0)$, where I_0 is the number of initial infectives. In this example, if $p = 0.4$, then $R_0 = 0.4 \cdot 2 = 0.8$. If $p = 0.7$, then $R_0 = 0.7 \cdot 2 = 1.4$. In deterministic models, if $R_0 > 1$, the epidemic will always take off, and if $R_0 < 1$, the epidemic will never take off. An index that makes more sense in the probabilistic world of stochastic models is the probability that the epidemic will not take off.

Another index in stochastic models is the probability that an epidemic will not spread from the initially infected people, called the *probability of no spread*, denoted by P_{ns}. It can be calculated from the transmission probability p, or escape probability, $q = 1 - p$, the number of initially infected people in the population I_0, and the number of initially susceptible people S_0. The probability that a susceptible person escapes infection from all I_0 initial infectives is q^{I_0}. The probability that all S_0 of the initial susceptible people escape infection from all of the initial infectives is $P_{ns} = (q^{I_0})^{S_0}$. In the above example, with $p = 0.4$, the probability of no spread is $P_{ns} = (0.6^1)^2 = 0.36$. With $p = 0.7$, $P_{ns} = (0.3^1)^2 = 0.09$. The probability of no spread is the same as the probability that the infection chain ends with just the initial infectives. The terms *minor* and *major* epidemics distinguish situations in which there is

a little spread from the initial infectives from situations in which an epidemic gains momentum and is self-sustaining.

4.4.1 The Reed-Frost model

Based on the definition of the Reed-Frost model above, we write the transition probability of getting $I_{t+1} = i_{t+1}$ new infectives at time $t+1$, given $S_t = s_t$ and $I_t = i_t$ susceptibles and infectives one time period before as

$$\Pr(I_{t+1} = i_{t+1} | S_t = s_t, I_t = i_t) = \binom{s_t}{i_{t+1}} \left(1 - q^{i_t}\right)^{i_{t+1}} q^{i_t(s_t - i_{t+1})}, s_t \geq i_{t+1}. \quad (4.2)$$

Then, we can update the number of new susceptibles and recovered people, respectively, by the equations

$$S_{t+1} = S_t - I_{t+1}, \quad (4.3)$$

$$R_{t+1} = R_t + I_t = \sum_{r=0}^{t} I_r. \quad (4.4)$$

Since the population is closed, we have $S_t + I_t + R_t = N$ for all t. The epidemic process starts with $I_0 > 0$, and terminates at stopping time T, where

$$T = \inf_{t \geq 0} \{t : S_t I_t = 0\}. \quad (4.5)$$

Equations (4.2-4.4) form the classical Reed-Frost model. Formal mathematical treatment of the model involves formulation of the discrete, two-dimensional Markov chain $\{S_t, I_t\}_{t=0,1,...}$. I_t is the (binomial) random variable of interest, and S_t is updated using (4.3). The probability of a particular chain, $\{i_0, i_1, i_2, ..., i_T\}$, is given by the product of conditional binomial probabilities from (4.2) as

$$\begin{aligned} &\Pr(I_1 = i_1 \mid S_0 = s_0, I_0 = i_0) \Pr(I_2 = i_2 | S_1 = s_1, I_1 = i_1) \cdots \quad (4.6) \\ &\Pr(I_T = i_T | S_{T-1} = s_{T-1}, I_{T-1} = I_{T-1}) \\ &\quad = \prod_{t=0}^{T-1} \binom{s_t}{i_{t+1}} \left(1 - q^{i_t}\right)^{i_{t+1}} q^{i_t(s_t - i_{t+1})}. \end{aligned}$$

Table 4.2 shows the possible chains for a population of size 4 with one initial infective, *i.e.*, $S_0 = 3,\ I_0 = 1$.

In some cases, the distribution of the total number of cases, R_T, is the random variable of interest. We let J be the random variable for the total number of cases in addition to the initial cases, so that $R_T = J + I_0$. If we let $S_0 = k$ and $I_0 = i$, then the probability of interest is

$$\Pr(J = j | S_0 = k, I_0 = i) = m_{ijk}, \quad (4.7)$$

Table 4.2. Chain binomial probabilities in the Reed-Frost model in households of size 4 with 1 initial infective and three susceptibles, $S_0 = 3, I_0 = 1$

Chain	Chain probability	Final number infected
$i_0 \to i_1 \to i_2 \to ... \to i_T$		R_T
$1 \longrightarrow 0$	q^3	1
$1 \longrightarrow 1 \longrightarrow 0$	$3pq^4$	2
$1 \longrightarrow 1 \longrightarrow 1 \longrightarrow 0$	$6p^2q^4$	3
$1 \longrightarrow 2 \longrightarrow 0$	$3p^2q^3$	3
$1 \longrightarrow 1 \longrightarrow 1 \longrightarrow 1$	$6p^3q^3$	4
$1 \longrightarrow 1 \longrightarrow 2$	$3p^3q^2$	4
$1 \longrightarrow 2 \longrightarrow 1$	$3p^3q\,(1+q)$	4
$1 \longrightarrow 3$	p^3	4

where $\sum_{j=0}^{k} m_{ijk} = 1$. Then, based on probability arguments (*e.g.*, see Bailey 1975; Becker 1989), we have the recursive expression

$$m_{ijk} = \binom{k}{j} m_{ijj} q^{(i+j)(k-j)},\ j < k \tag{4.8}$$

$$m_{ikk} = 1 - \sum_{j=0}^{k-1} m_{ijk}. \tag{4.9}$$

Data are usually in the form of observed chains, $\{i_0, i_1, ..., i_r\}$, for one or more populations, or final sizes, R_T, for more than one population. With respect to the former data form, suppose that we have N populations and let $\{i_{k0}, i_{k1}, ..., i_{kr}\}$ be the observed chain for the k^{th} population. Then, from (4.6), the likelihood function for estimating $p = 1 - q$ is

$$\mathrm{L}(p) = \prod_{k=1}^{N} \prod_{t=0}^{r-1} \binom{s_{kt}}{i_{kt+1}} \left(1 - q^{i_{kt}}\right)^{i_{kt+1}} q^{i_{kt}(s_{kt} - i_{kt+1})}, \tag{4.10}$$

Whether data are available on observed chains or just the final size distribution, the simple Reed-Frost model assumes that transmission units are independent of one another as in Figure 4.1a. The initial infectives in the transmission unit somehow get infected, then the chain of infection unfolds within the transmission unit without any further introduction of infectives. Alternatively, one could assume that people, whether the initial infectives or the others in the transmission unit, are also exposed to infection outside the transmission unit in the community at large, as in Figure 4.1b, or in other mixing places. Longini and Koopman (1982) modified the Reed-Frost model for the case where there is a constant source of infection from outside the population that does not depend on the number of infected persons in the population. Analysis of data assuming transmission units in a community are

presented in Chapters 11 and 12. Becker (1989) gives details on different aspects of the Reed-Frost model and estimation of the parameters of interest from data. Bailey (1975) (Sec. 14.3) gives an example where (**??**) is used to estimate $\widehat{p} = 0.789 \pm 0.015$ (estimate ± 1 standard error) for the household spread of measles among children.

4.4.2 The Greenwood model

For the Greenwood model, the number of new infectives does not depend on the number of old infectives, but just on the presence of one or more infectives. Thus, the transition probability of getting $I_{t+1} = i_{t+1}$ new infectives at time $t+1$, given $S_t = s_t$ and $I_t = i_t$ susceptibles and infectives one time period before is

$$\Pr(I_{t+1} = i_{t+1} | S_t = s_t, I_t = i_t) = \left\{ \begin{array}{cc} \binom{s_t}{i_{t+1}} p^{i_{t+1}} q^{(s_t - i_{t+1})}, & s_t \geq i_{t+1} \text{ and } i_t > 0 \\ 0 & \text{otherwise} \end{array} \right\} \quad (4.11)$$

Analysis of this model is similar to that of the Reed-Frost model.

4.4.3 Stochastic realizations of the Reed-Frost model

Realizations of epidemics according to the Reed-Frost model in equations (4.2-4.4) can be simulated using a random number generator. At each time t, for each susceptible person exposed to I_t infectives, a random number between 0 and 1 is generated. If the random number is smaller than the infection probability $1 - q^{I_t}$, then the person becomes infected. If the random number lies between the infection probability and 1, then the person escapes infection in that time interval. The actually realized chain then depends on the series of random numbers that are generated, and varies from realization to realization. The probabilities in Tables 4.1 and 4.2 are the expected probabilities of particular chains if a large number of epidemics are simulated.

Tables 4.3 through 4.5 show realizations of stochastic epidemics in a population with 20 people. Table 4.3 show ten epidemics in populations of size 20 and $p = 0.05$. Ten epidemics were run with one initial infective, $I_0 = 1$, $S_0 = 19$, the other ten epidemics were run with three initial infectives, $I_0 = 3$, $S_0 = 17$. The underlying Reed-Frost model is identical for both types of run, just the initial conditions are different. The $\mathrm{R}_0 = 1.0$, without taking into account the initial infectives. Taking into account the number of initial susceptibles, the initial reproductive numbers are 0.95 and 0.85, respectively. With one initial infective, the probability of no spread is $P_{ns} = (0.05^1)^{19} = 0.377$, with three initial infectives, it is $P_{ns} = (0.05^3)^{17} = 0.073$. The number of initial infectives is important on how long the chain is, whether any further infections occur, and the average number of final infectives. The chains in the table demonstrate the randomness of the epidemics and how in nature,

Table 4.3. Ten stochastic epidemics with the Reed-Frost model, 20 people, $p = 0.05$

	1 initial infective, $I_0 = 1$		3 initial infectives, $I_0 = 3$	
Epidemic	Final infected	Chain	Final infected	Chain
1	1	$1 \to 0$	8	$3 \to 2 \to 2 \to 1 \to 0$
2	1	$1 \to 0$	8	$3 \to 2 \to 1 \to 2 \to 0$
3	8	$1 \to 3 \to 3 \to 1 \to 0$	14	$3 \to 4 \to 3 \to 3 \to 1 \to 0$
4	1	$1 \to 0$	4	$3 \to 1 \to 0$
5	1	$1 \to 0$	11	$3 \to 2 \to 1 \to 4 \to 1 \to 0$
6	2	$1 \to 1 \to 0$	4	$3 \to 1 \to 0$
7	1	$1 \to 0$	4	$3 \to 1 \to 0$
8	1	$1 \to 0$	14	$3 \to 3 \to 3 \to 2 \to 2 \to 1 \to 0$
9	4	$1 \to 1 \to 1 \to 1 \to 0$	6	$3 \to 2 \to 1 \to 0$
10	1	$1 \to 0$	10	$3 \to 1 \to 3 \to 2 \to 1 \to 0$

Table 4.4. Ten stochastic epidemics with the Reed-Frost model, 20 people, $p = 0.06$

	1 initial infective, $I_0 = 1$	
Epidemic	Final number infected	Chain
1	1	$1 \to 0$
2	1	$1 \to 0$
3	2	$1 \to 1 \to 0$
4	8	$1 \to 2 \to 4 \to 1 \to 0$
5	10	$1 \to 2 \to 4 \to 2 \to 1 \to 0$
6	2	$1 \to 1 \to 0$
7	12	$1 \to 1 \to 3 \to 3 \to 1 \to 2 \to 1 \to 0$
8	8	$1 \to 2 \to 3 \to 2 \to 0$
9	9	$1 \to 3 \to 2 \to 1 \to 1 \to 1 \to 0$
10	14	$1 \to 3 \to 3 \to 2 \to 2 \to 2 \to 1 \to 0$

given the same conditions, that many different outcomes can occur merely by chance.

In Table 4.4, the transmission probability is increased to 0.06, so that $R_0 = 1.2$. Taking into account the one initial infective, the reproductive number is 1.14, and the probability of no spread is $P_{ns} = (0.06^1)^{19} = 0.309$.

In Table 4.5, the transmission probability is increased to 0.06, so that $R_0 = 2.0$. Taking into account the one initial infective, the reproductive number is 1.9, and the probability of no spread is $P_{ns} = (0.1^1)^{19} = 0.135$. A clear bimodal distribution has emerged at this higher transmission probability. Two of the epidemics produce only one more infective, but in the other eight, a majority of the population becomes infected.

Table 4.5. Ten stochastic epidemics with the Reed-Frost model, 20 people, $p = 0.1$

	1 initial infective, $I_0 = 1$	
Epidemic	Final number infected	Chain
1	2	$1 \to 1 \to 0$
2	16	$1 \to 3 \to 2 \to 2 \to 3 \to 3 \to 2 \to 0$
3	17	$1 \to 1 \to 3 \to 5 \to 6 \to 1 \to 0$
4	17	$1 \to 1 \to 1 \to 2 \to 3 \to 3 \to 4 \to 1 \to 1 \to 0$
5	17	$1 \to 2 \to 4 \to 6 \to 4 \to 0$
6	16	$1 \to 1 \to 2 \to 2 \to 4 \to 5 \to 1 \to 0$
7	14	$1 \to 1 \to 2 \to 4 \to 6 \to 0$
8	19	$1 \to 3 \to 3 \to 6 \to 3 \to 3 \to 0$
9	17	$1 \to 1 \to 3 \to 4 \to 4 \to 3 \to 1 \to 0$
10	2	$1 \to 1 \to 0$

4.5 Stochastic simulation models

The simple Reed-Frost model is the basic building block of small- and large-scale stochastic simulation models of infectious disease spread and studies of interventions. Such models need to include 1) the natural history of the infection of interest, 2) the demographics of the relevant population, 3) the contact structure and assumptions about where and how transmission occurs, 4) models of the interventions and assumptions about how they will affect transmission, natural history, or the contact structure. Halloran et al (2002) and Longini et al (2007) examined vaccination strategies for smallpox. Several studies of interventions for pandemic influenza have made use of such simulation models (Longini et al 2004; Longini et al 2005; Germann et al 2006; Ferguson et al 2006; Halloran et al 2008). Here we present one example of a stochastic simulation model used to examine potential indirect, total, and overall effects of cholera vaccination.

4.5.1 Endemic cholera and vaccination

In the mid 1980's, a randomized vaccine trial with OCV in Matlab, Bangladesh, yielded 70% direct vaccine efficacy for up to two years (Clemens et al 1990; Durham et al 1998) . Information about Matlab, Bangladesh was used to construct a model of the population as it was in 1985, consisting of 183,826 subjects (Longini et al 2007). These subjects were mapped into families and families were distributed in baris, i.e., patrilineally related household clusters. In the model, baris are further clustered into sub-regions of about 6 square km in size considered to be the geographic cholera transmission areas. The model represents the number of contacts that a typical person makes with sources of potential cholera transmission in the course of a day. The age and bari size distributions of the population are based on data from Ali et al (2005). (See

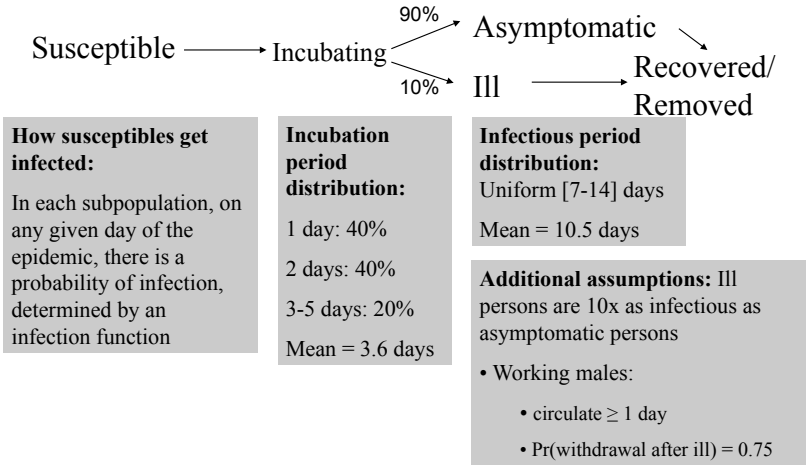

Fig. 4.5. Modeled natural history of cholera. Newly infected people pass through the incubating state (mean 3.6 days) and infectious state (mean 10.5 days) after which they recover with immunity or die. The probability distributions of the incubation and infectious periods are shown. It is assumed that 10% of infected people develop overt cholera symptoms and 90% are asymptomatic. Symptomatic people are assumed to be ten times as infectious as asymptomatics. Additionally, the model allows for 75% of symptomatic working males to withdraw to their sub-region. (Longini et al 2007)

Chapter 13.2.6). Women and children are assumed to come into contact with sources of infection in the sub-region where they live, while working males are assumed to make contact with infective sources in the sub-region where they live as well as where they work. The modeled natural history of cholera is described in Figure 4.5.

The model was calibrated to cholera illness incidence data from a large cholera vaccine trial in the Matlab field area of the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), that took place from 1985 - 1989, described in Chapter 13.2.6. Oral cholera vaccine or placebo (killed *E. coli*) was offered to children 2 – 15 years old and women greater than 15 years old. Matlab has cholera transmission year around, but it generally experiences a large cholera epidemic from September through December

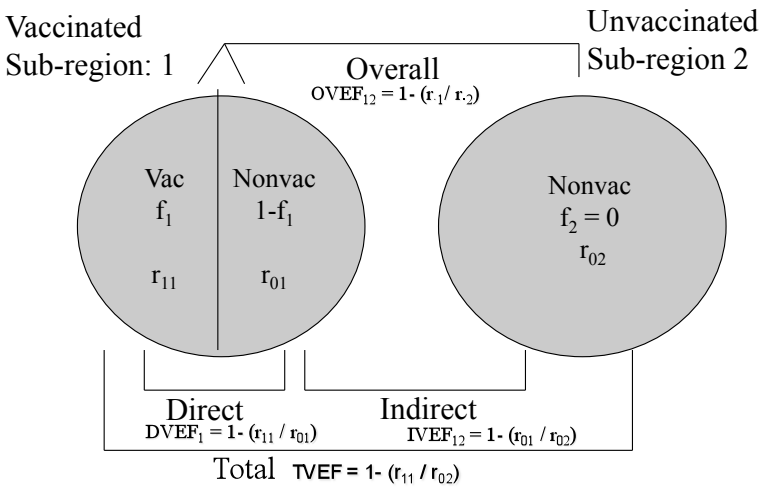

Fig. 4.6. Schematic of Effectiveness Comparisons for Two Sub-regions. Sub-region 1 has a fraction $f_1 > 0$ people vaccinated, while the comparison sub-region 2 has nobody vaccinated, i.e., $f_2 = 0$. (from Longini et al 2007)

and then a somewhat smaller epidemic from March through May every year (Longini, et al 2002).

Cholera risk was assessed in individuals residing in different sub-regions in the field trial area. Sub-regions are useful for cluster analysis because they are geographically discrete with local sources of water. An infection function was defined that gives each susceptible person's daily probability of infection from all possible sources of infection created by infected people excreting cholera vibrios into the environment or through more direct contact similar to that in the Reed-Frost model with environmental exposure outside the transmission units. The probability of infection is proportional to the number of vaccinated and unvaccinated people in the sub-region where contact is specified to occur. The model of vaccine effect assumed that immunity resulted in a proportional reduction in the probability of infection per contact with an infectious source, i.e., a leaky vaccine. Results were averaged over all the sub-regions within vaccination coverage strata.

As described in Chapter 2, the indirect, overall, and total vaccine effectiveness were based on the reduction in infection rates when comparing the appropriate groups within a sub-region with no vaccination to a compara-

Table 4.6. Vaccination coverage, average incidence rates, and direct effectiveness (calibration runs) (Longini et al 2007)

Vaccination Coverage (%)		Mean Cases/1,000 (95% CI)				Mean Direct Effectiveness (%)(95% CI)	
Population		Vaccinated		Placebo			
Target	Overall	Observed	Simulated	Observed	Simulated	Observed	Simulated
14	9	2.7 (1.9 to 3.5)	2.8 (0.5 to 6.1)	7.0 (6.5 to 7.5)	7.8 (1.9 to 14.8)	62	65 (52 to 77)
31	20	2.5 (2.0 to 3.0)	1.7 (0.3 to 3.8)	5.9 (5.4 to 6.4)	4.7 (0.9 to 10.2)	58	65 (55 to 76)
38	25	4.7 (1.2 to 2.0)	1.3 (0.2 to 3.4)	4.7 (4.2 to 5.2)	3.8 (0.8 to 8.6)	67	65 (54 to 77)
46	30	2.3 (1.9 to 2.7)	1.0 (0.1 to 2.5)	4.7 (4.2 to 5.2)	2.8 (0.5 to 6.8)	52	66 (54 to 79)
58	38	1.3 (1.0 to 1.6)	0.6 (0.1 to 1.8	1.5 (1.2 to 1.8)	1.8 (0.3 to 4.8)	14	66 (51 to 50)

ble sub-region with a fraction $f > 0$ of the population vaccinated (Figure 4.6). Let r_{ij} denote the cholera infection rate for people in sub-region j with vaccination status i, where $i = 0$ for unvaccinated and $i = 1$ for vaccinated. The indirect effect of vaccination is measured by comparing the infection rates between the unvaccinated in the two sub-regions. Thus, the indirect vaccine effectiveness, i.e., IVEF, when comparing sub-region 1 to 2 is $\text{IVEF}_{12} = 1 - (r_{01}/r_{02})$. The overall effect of vaccination is measured by comparing the average (over the vaccinated and unvaccinated groups) infection rates between the two sub-regions. Thus, the overall vaccine effectiveness, i.e., OVEF, is $\text{OVEF}_{12}12 = 1 - (r_{\cdot 1}/r_{\cdot 2})$, where the $\cdot$ indicates averaging over the vaccinated and unvaccinated. The total effect of vaccination is measured by comparing the infection rate in the vaccinated in sub-region 1 to the unvaccinated in sub-region 2. Thus, the total vaccine effectiveness, i.e., TVEF, is $\text{TVEF}_{12} = 1 - (r_{11}/r_{02})$. In general, these effectiveness measures could be computed across any gradient of coverage, $|f_1 - f_2|$, other than those with $f_2 = 0$.

The direct effectiveness compares the vaccinated to the unvaccinated within a sub-region. The direct effect of vaccination is measured by comparing the infection rates in the vaccinated and unvaccinated in the same sub-region. The direct vaccine effectiveness, i.e., DVEF, is $\text{DVEF}_1 = 1 - (r_{11}/r_{01}))$.

The simulation model was calibrated using cholera incidence data observed in the first year of the vaccine trial (Table 4.6) over a 180 day period to capture all the cholera transmission during the large annual cholera outbreak. This was done by varying the transmission probability, p such that the differences between the observed incidence rates and the simulated incidence rates in Table 4.6 were minimized. The estimated reproductive number was 5.0 with a standard deviation of 3.3. The vaccine coverage levels in the target population and the effective coverage in the entire population from the trial are summarized in Table 4.6 (See also Table 13.5). Vaccinated people receive an effective regimen of two doses. The observed cholera incidence rates among the unvaccinated ranged from a high of 7.0 cases/1,000 over 180 days for the sub-regions with the lowest coverage in the target population, centered at 14%, to 1.5 cases/1,000 for the highest coverage, centered at 58%. The ob-

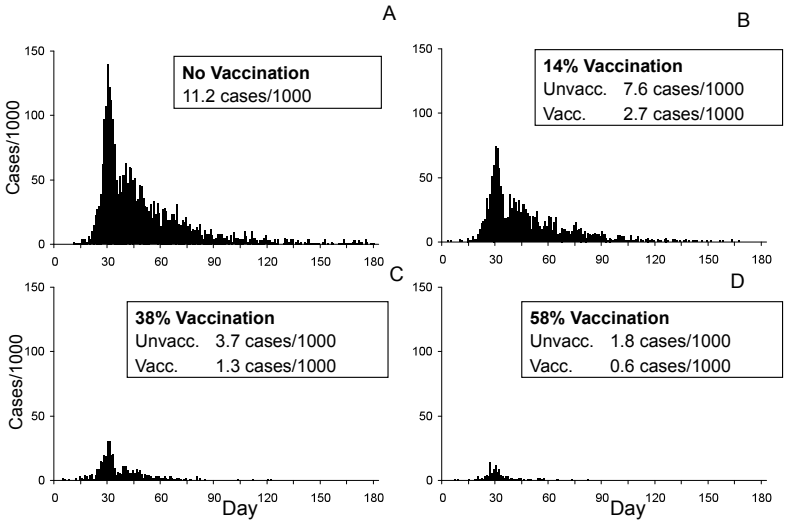

Fig. 4.7. Simulated number of cholera cases/1,000 over a 180 day period in the Matlab study population for a single stochastic realization A. No vaccination; B. 14% vaccination coverage of women and children; C. 38% vaccination coverage; D. 58% vaccination coverage. (from Longini et al 2007)

served cholera incidence rates among the vaccinated ranged from a high of 2.7 cases/1,000 for the sub-regions with the lowest coverage to 1.3 cases/1,000 for the highest coverage. Vaccine efficacy for susceptibility was set to VE_S= 0.7 (Clemens et al 1990; Durham et al 1998) and for infectiousness to VE_I= 0.5. The simulated incidence rates provided a good fit to the data based on a χ^2 goodness-of-fit test for frequency data ($p = 0.84$, 9 degrees of freedom). Figure 4.7 shows the number of cases over time comparing the unvaccinated to the vaccinated populations for different levels of coverage.

For effectiveness measures, a comparison was made between the intervention sub-regions to hypothetical sub-regions that receive no vaccine, i.e., $f = 0$. Table 4.7 shows the indirect, total and overall effectiveness estimated by the model for possible coverage levels when comparing coverages in the entire population, two years of age and older, ranging from 10% to 90% compared to no vaccination. For example, the average indirect effectiveness, comparing a population with a coverage of 30% to one with no vaccination is 70%. This indicates that on average, the cholera incidence among unvaccinated people in a population with 30% coverage would be reduced by 70% compared

Table 4.7. Average indirect, total,and overall effectiveness of vaccination, and cases prevented per 1,000 two-dose regimens (Longini et al 2007)

Vaccination	Mean effectiveness (%) (95% CI)			Mean cases prevented per
coverage (%)	Indirect	Total	Overall	1,000 dose regimens (95%)
10	30 (−39 to 83)	76 (47 to 95)	34 (−30 to 84)	40 (−34 to 97)
30	70 (31 to 93)	90 (76 to 98)	76 (44 to 95)	30 (17 to 36)
50	89 (72 to 98)	97 (91 to 99)	93 (82 to 99)	21 (19 to 23)
70	97 (91 to 99)	99 (97 to 100)	98 (95 to 100)	16 (15 to 17)
90	99 (98 to 100)	100 (99 to 100)	100 (99 to 100)	13 (12 to 14)

with a completely unvaccinated population. At this level of coverage, the total effectiveness of 90% indicates high protection for a vaccinated person in a population with 30% vaccination coverage, while the overall effectiveness of 76% indicates a good overall reduction in risk to the overall population. According to the model, around 40 cases of cholera are prevented per 1000 two-dose regimens of vaccine at low coverage to 13 cases at high coverage. At coverage levels of 50% and higher, all levels of effectiveness exceed 85%.

From Table 4.6, we see that the simulated direct effectiveness at all coverage levels is estimated from the simulations to be about 66%, while the vaccine efficacy for susceptibility, VE_S, is preset at 70%. This small underestimation is due to the fact that the model assumes the vaccine effect to be a 70% reduction in the risk of infection per contact with an infective source, i.e., a leaky effect, but the risk ratio estimator of vaccine effectiveness over the entire cholera epidemic is used. The risk ratio estimator allows comparison with the primary estimator used in the cholera vaccine trial in Matlab. The analysis implies that mass immunization with oral killed whole-cell based vaccines could possibly achieve a substantial reduction in cholera in an endemic setting, even with modest levels of vaccine coverage, due to the combination of direct and indirect vaccine protective effects.

4.5.2 Use in Study Design

In general, stochastic simulation models are useful for generating simulated data with variability so that methods of analysis can be used and compared. Stochastic computer simulations are especially useful in helping to design studies and to develop new methods of analysis (see for example, Halloran et al (2002), or Golm, et al, 1999 or Longini, et al, 1999). Deterministic models do not generate variability, but can be used to understand properties of the transmission system.

Problems

4.1. Reed-Frost Model
Compute the average number of people infected in the four examples of 10

epidemics in Tables 4.3 through4.5. Make histograms of the number of people infected in each set of 10 epidemics and compare the shapes of the histograms.

4.2. Final size distribution of the Reed-Frost model
(a) Compute the final size distribution from Equations 4.8 and 4.9 for households of size 4 with one initial infective when $p = 0.4$.
(b) Compare it to the final size distribution obtained using the chain probabilities in Table 4.2.