4

Concepts of Transmission and Dynamics

M. ELIZABETH HALLORAN

Transmission from one host to another is fundamental to the survival strategy of most infectious agents. Each microbe has its own life cycle, modes of transmission, population dynamics, evolutionary pressures, and molecular and immunologic interaction with its host. The transmission cycle may involve a particular insect or other vector, and consequently its ecology. Studies and interventions need to take the particular transmission, dynamics, and biology of each infectious agent into account.

Some underlying principles of transmission and dynamics, however, are common to many infectious diseases. These principles are captured in a wide variety of mathematical and statistical models. Since the human host population is the ecological niche for the infectious agent, some of the principles come from general theories of populations, evolution, and ecology (see Burnet and White 1972, McNeill 1976). Other principles have their origins in infectious disease epidemiology.

Many different questions motivate quantitative transmission models. A few examples follow.

- What is the probability that transmission will occur after a susceptible host is exposed to infection? How do transmission dynamics and interventions influence the evolution of a microbe? How do different models of transmission influence our thinking? How do different assumptions about human contact patterns influence the design and analysis of field studies?
- Under what conditions will an epidemic occur? Will an infectious agent become established in a population and either persist or die out? Will a microbe establish itself within a host and avoid immune surveillance and clearance?
- What interventions can prevent an epidemic or eliminate endemic transmission? What interventions will reduce transmission and by how much? What will the long-term effects of an intervention be in a population? What is the best intervention type and resource allocation strategy? What is the optimal timing? How do different subpopulations influence transmission of an infectious agent and choice of intervention strategies?

When the appropriate data are available, the models can be used to estimate quantities of interest to answer the above questions.

How we think about the transmission dynamics of an infectious agent within a host population influences how we design and interpret epidemiologic studies. It can influence our choice of interventions. Mixing structures, contact patterns, and subpopulations can affect both transmission dynamics and the results of studies. In this chapter, we consider some basic principles and simple models of transmission and population dynamics of infectious diseases. We focus on aspects of transmission and dynamics that have consequences for the design of studies and interpretation of results.

STATES OF INFECTION WITHIN A HOST

The natural history of infection within a host can be described with reference to either infectiousness or disease (Fig. 4–1). Both time lines begin with the successful infection of the susceptible host by the microbe. The natural history of infectiousness includes the *latent period*, the time interval from infection to becoming infectious, and the *infectious period*, during which time the host could infect another host or vector. Eventually the host becomes noninfectious, either by clearing the infection, possibly developing immunity, or by death. The host can also become noninfectious while still harboring the microbe. The host may become an infectious *carrier* if he or she recovers from disease but remains infectious (i.e., asymptomatically infected).

The natural history of disease in the infected host includes the *incubation period*, the time from infection to symptomatic disease, and the *symptomatic period*. The probability of developing symptomatic disease after becoming infected is the *pathogenicity* of the interaction of the microbe with the host. Eventually the host leaves the symptomatic state, either by recovering from the symptoms or by death. If the microbe has provoked an autoimmune response in the host, symptoms can continue even after the microbe is cleared. An *inapparent case* or *silent infection* is a successful infection that does not produce symptoms in the host. Inapparent cases can be infectious.

While the disease process and its associated time line are important to the infected person and to a physician, the dynamics of

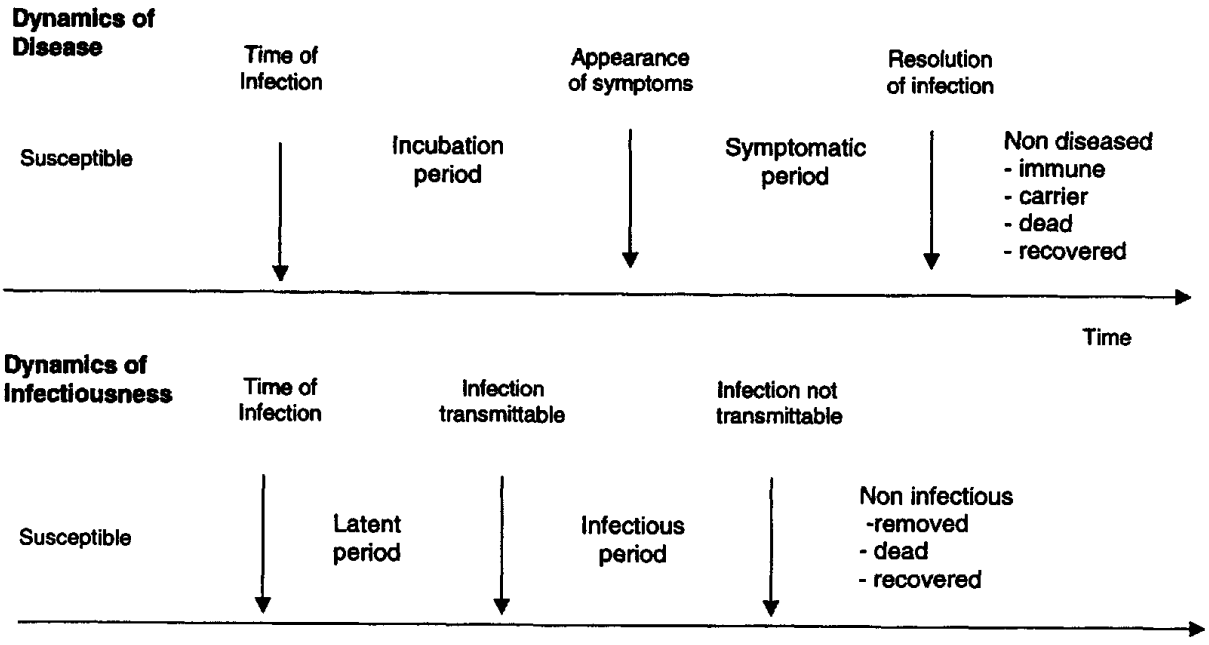


Figure 4–1. Natural history time lines for infection and disease.

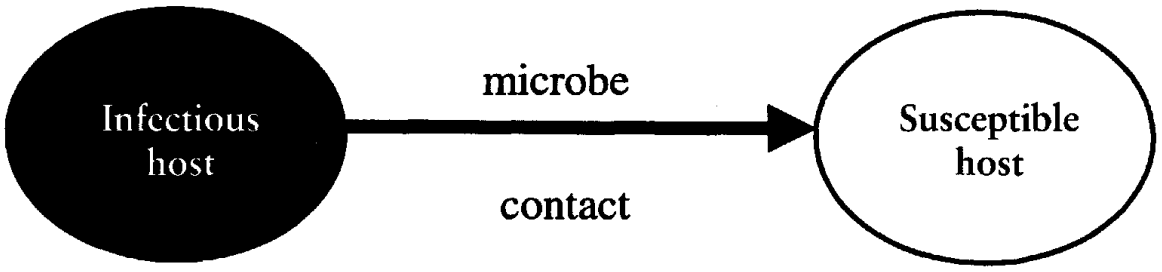
infectiousness are important for propagation of the microbe and for public health. The relation of the two time lines to one another is specific to each microbe and can have important implications for study design, modeling, and public health.

For example, Elveback and colleagues (1976) developed an influenza model that distinguished between illness and infection attack rates. The infected people become infectious, but only a fraction of them develop overt disease. In many studies of infectious agents, it is easier to use overt disease as the outcome, rather than infection, since infection may be difficult to ascertain. If many infections are inapparent, however, using overt disease would result in an underestimate of the level of exposure to infection in the population. Estimation of the incubation and latent periods can be difficult because the time of infection as well as the onset of infectiousness are often difficult to observe.

Human immunodeficiency virus (HIV) poses a particular problem for public health because the virus has a short latent period and a long incubation period. A person infected with HIV can infect many people before symptoms develop. *Plasmodium falciparum*, one of the organisms that causes human malaria, has an incubation period of about 14 days, but the infective stages do not appear until about 10 days after the first symptoms. Thus early treatment of symptoms with a drug that also kills or prevents infective stages could have an important effect on transmission. In chickenpox, the latent period is about two days shorter than the incubation period. Thus by the time symptoms appear, a child can infect many other children. Keeping children with symptomatic chickenpox out of school might not have a large effect on transmission. Gonorrhea infection in women is often asymptomatic, so women often go untreated. In men, the infection is often quite painful, leading them to seek treatment. Thus the duration of infectiousness tends to be shorter in men than in women for reasons related to the different time lines of disease in men and women.

TRANSMISSION MODELS

One measure of the success of an infectious agent is how effectively it is transmitted. The *transmission probability p* is the probability that, given a contact between an infective source and a susceptible host, successful transfer of the microbe will occur so that the susceptible host becomes infected (Fig. 4–2). The transmission probability is a key quantity both in epidemiology and in infectious disease models. There are many different ways of modeling the probability of becoming infected upon repeated exposure to infection. We consider the simple binomial model



Transmission probability depends on:

- **Type and definition of contact**
- **Microbe**
- **Infectious host**
- **Susceptible host**

Figure 4–2. Transmission from an infective to a susceptible host during contact.

of transmission for discrete contacts and, briefly, a simple model in continuous time.

Binomial Models of Probability of Infection

The binomial model of transmission can help answer several questions. What is the effect of an intervention in a population? How do we interpret our assumptions about how a risk factor or intervention affects the transmission probability? How well is one infectious agent transmitted compared to another? When the appropriate data are available from field studies, the binomial model is often used to estimate the transmission probability.

The basic idea of the binomial model is that exposure to infection occurs in discrete contacts and that each contact is independent of another. We define p as the transmission probability during a contact between a susceptible person and an infectious person or other source of infection, such as an infectious mosquito. Then the probability that the susceptible person will not be infected during the contact is $q = 1 - p$. The quantity q is called the *escape probability*. For example, if the transmission probability for herpes simplex 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 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 six successive sexual contacts with someone who has genital herpes (Fig. 4–3A). What is ithe probability that the person will have become infected after six contacts? In this example, $n = 6$. 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: $q \times q = 0.7 \times 0.7 = 0.49$. 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: $q^2 \times q = 0.49 \times 0.7 = 0.34$. The probability of escaping infection from six successive contacts is $0.7^6 = 0.12$. The probability of becoming infected at least once is $1 - (1 - p)^n = 1 - (0.7)^6 = 0.88$.

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 transmission. These assumptions may not be fulfilled. If so, the assumptions can easily be changed and a more complicated form of the binomial model developed.

In a different problem, suppose a susceptible child attends school one day where six of the children simultaneously have influenza. What is the probability of becoming infected (Fig. 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 being infected from all six children, with a $0.7^6 = 0.12$, so the probability of being infected on that day at school is $1 - q^6 = 0.88$.

Although the answers for the two examples are numerically the same, in the second example we made a different biological assumption than in the first. In the example of influenza at school, we assumed that each of the six *simultaneous* exposures to infection is 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 si-

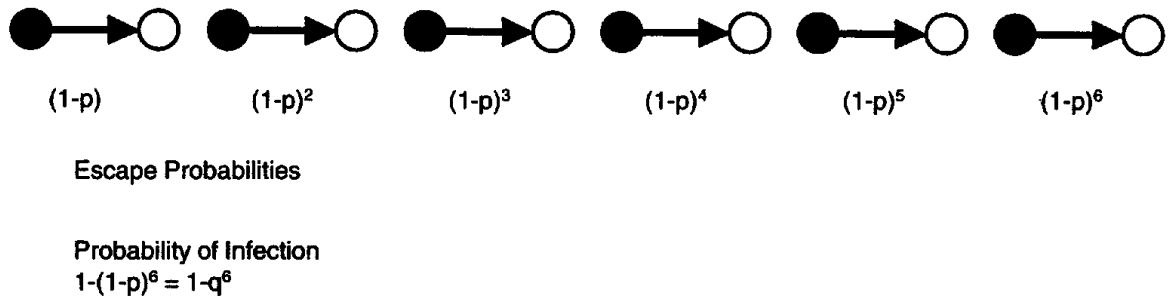


Figure 4–3A. The probability of infection with six consecutive contacts.

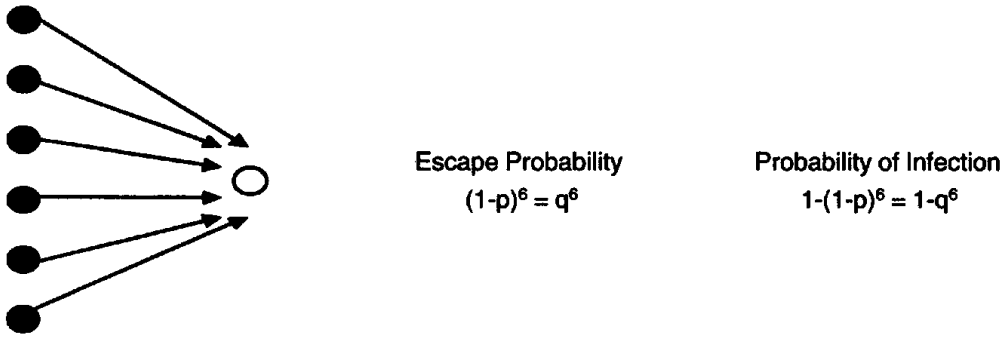


Figure 4–3B. The probability of infection with six simultaneous contacts.

multaneous or successive. For instance, this assumption is at the heart of the Reed-Frost model discussed later.

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 infectives 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 seldom used in practice, however. We could make assumptions between the two extremes, but there are generally not enough data to support using more complex models.

As discussed in Chapter 5, Overview of Study Design, the binomial model is useful in estimating the transmission probability if data are available on the number of potentially infectious contacts that susceptibles in a study population make as well as the number of susceptibles who become infected.

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 previous example of herpes, however, $np = 6 \times 0.3 = 1.8$. 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}. So, the probability of becoming infected is $1 - e^{-np}$. In the herpes example above, then, the probability of not becoming infected is $e^{-6\times0.3} = e^{-1.8} = 0.17$ and for becoming infected is $1 - e^{-1.8} = 0.83$. Comparing this with the probability of being infected calculated from the binomial model, 0.88, we note that they are similar but not identical.

In the herpes example, 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 six 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 $1 - (1 - p)^n = 1 - (.999)^6 = 0.00599$. Using the exponential expression, the probability of becoming infected is $1 - \exp(-6 \times 0.001) = 0.00598$, and based on the simple expression, $np = 6 \times 0.001 = 0.006$. There is a 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. If used for estimation, all three models require the same data. In general, it is good to use the binomial model if feasible.

Continuous Models for Probabilitiy of Infection

The binomial model assumes discrete contacts or discrete units of time. Another approach to modeling the probability of becoming infected assumes that contacts occur in continuous time. This approach is usually based on the contact rate per unit time, which we denote by c. Thus cp is the probability of being infected per unit time if all the contacts are with infectious persons. 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 are $\exp(-cp\Delta t)$ and $1 - \exp(-cp\Delta t)$, respectively. The data needed for using this approach to estimate the transmission probability are the contact rate with infectives per unit time, the time interval, and the infection status of each person in the study.

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 genital herpes example above that $p = 0.3$ but that the contacts are with six individuals of unknown infection status. If the individuals are randomly chosen from a population where the prevalence of genital herpes is $P = 0.4$, then the probability of being infected after six contacts is $1 - (1 - 0.3 \times 0.4)^6 = 0.54$.

An analogous expression for the infection probability can be developed for the continuous time model. The probability of being infected per unit time is the incidence rate or hazard rate of infection. An expression for the incidence rate, I, as a function of the contact rate, the transmission probability, and the prevalence is $I = cpP$. This expression

for the incidence rate as a function of prevalence is a fundamental relation of dependent happenings (Ross 1916) in infectious diseases, discussed in more detail in Chapter 5, Overview of Study Design. The probability of escaping infection within some period of time Δt is $\exp(-cpP\Delta t)$, and of being infected is $1 - \exp(-cpP\Delta t)$. At the population level, the probability of becoming infected in some period of time is closely related to the incidence proportion.

These examples show some of the options and subtleties inherent in different approaches to modeling the transmission process.

Modeling Risk Factors for the Transmission Probability

How do risk factors or interventions play a role in the probability of becoming infected during a contact between an infectious person and a susceptible person? How does the choice of models affect our answer? Suppose we are doing a study of a vaginal foam to prevent genital herpes transmission. We believe that the vaginal foam reduces the probability of transmission per sex act by 80%. We might formulate our binomial model so that foam reduces the transmission probability, p, by 80% in everyone who uses it and at every sex act with an infective. Then the transmission probability in people using foam, p_{foam}, would be 20% of that in people not using it, so that $p_{\text{foam}} = 0.20p$. Since the factor 0.20 multiplies the baseline p, we are assuming that the foam has a *multiplicative effect* on the transmission probability. The protection is not complete, since the people using foam still have a transmission probability of $0.20p$. Thus, a multiplicative effect is sometimes called *leaky*, because it denotes only partial protection, allowing microbes to get through the defense. Note that we have also assumed that the effect is the same in everyone and for every contact.

Suppose we want to evaluate the effect of using vaginal foam in a study population of 2000 sexual partnerships, where one partner is infected in each partnership. Half of the partnerships use foam, the other half do not. We decide to use the incidence proportion ratio at the end of the study to estimate the relative risk of infection with and without foam. The first study we conduct is one month long and each partnership has exactly five contacts during that time. If $p = 0.25$, then $p_{\text{foam}} = 0.20 \times 0.25 = 0.05$. What is the expected incidence proportion at the end of one month?

In the group not using vaginal foam, the probability of becoming infected is $1 - (1 - p)^5 = 1 - 0.75^5 = 0.76$, so the expected number of infections in that group is 1000 people $\times$ 0.76 = 760. In the group using foam, the probability of becoming infected is $1 - (1 - 0.05)^5 = 1 - 0.95^5 = 0.23$, so the expected number of infections in that group is 1000 people $\times$ 0.23 = 230. The incidence proportion ratio we would expect to see at the end of one month is (230/1000)/(760/1000) = 0.30. The incidence proportion ratio, 0.30, is not equal to the multiplicative effect of the foam on the transmission probability, 0.20. The efficacy of the vaginal foam based on the incidence proportion ratio would be estimated to be $1 - 0.30 = 0.70$, not $1 - 0.20 = 0.80$, the efficacy per single contact.

What happens to the incidence proportion ratio if we continue the study for two months? Suppose that after two months, each partnership has had exactly ten sexual contacts. Now the expected number of infections in the control group is $(1 - .75^{10}) \times 1000 = 943$, while in the group using vaginal foam, it is $(1 - 0.95^{10}) \times 1000 = 401$. We expect to see an incidence proportion ratio after two months of (401/1000)/(943/1000) = 0.43. The incidence proportion ratio has increased from 0.30 to 0.43. The efficacy appears to be 0.57, not 0.70, as it would after one month, or 0.80, for a single contact. The intervention seems less efficacious after two months even though the effect of the vaginal foam on the transmission probability has not waned.

As the number of exposures in the two groups increase, the incidence proportion ratio will continue to increase toward one as it did from one month (five contacts) to two

months (ten contacts total), and efficacy will appear to decrease. Eventually everyone in both groups will become infected under the multiplicative assumption if they are exposed often enough. This illustrates the meaning of a multiplicative or leaky model at the transmission probability level. In principle, people can still become infected if exposed often enough. A different model might assume that vaginal foam protected 80% of the users completely, while 20% not at all. In this situation, at least 80% of the 1000 people using foam in the study would never become infected. This illustrates the difference between assuming a multiplicative model where the effect is the same for everyone and assuming a heterogeneous distribution of protection. Smith and colleagues (1984) and Halloran and associates (1991, 1992) provide further discussion of this point.

Suppose we use the model in continuous time developed earlier, and we assume that the protective effect has the same multiplicative effect on the transmission probability. In this partner study the contacts are all potentially infectious, $P = 1$, thus $I(t)_{\text{foam}} = 0.20cp$. Using this expression, the incidence rate ratio will be 0.20, giving the same answer as the multiplicative effect on the transmission probability. The expected incidence proportions in the two groups are not the same as those obtained using the binomial model. For the nonfoam group, the probability of being infected after one month is $1 - \exp(-5 \times 0.25) = 0.713$ and for the foam group it is $1 - \exp(-5 \times 0.05) = 0.221$, so the expected number of infections is 221 in the group using vaginal foam and 713 in the nonfoam group. The number of expected infections is different than calculated above from the discrete binomial model. The incidence proportion ratio is (221/1000)/(713/1000) = 0.31 after one month, which is similar though not identical to that calculated above.

In summary, there are three important points: (1) The binomial model of infection is widely used in practice. (2) There are differences between assuming that a contact process occurs discretely or that it occurs in continuous time. It is not possible to say that one approach is better than the other. They are simply different, and sometimes produce different answers. (3) The effect of a risk factor at the level of the transmission probability might be different from the apparent effect that will be estimated if using the incidence rate ratio or the incidence proportion ratio. Care should be taken to be precise in interpreting estimated relative risks.

BASIC REPRODUCTIVE NUMBER

Another key quantity in infectious diseases is the basic reproductive number, R_0, pronounced "are-zero" or "are-naught." The concept comes from general population theory. Understanding R_0 is important for public health applications and for describing the population biology of a parasite in a population of hosts. For small microbes such as viruses and bacteria, also called *microparasitic diseases* in the population biology literature, R_0 is defined as the expected number of new infectious hosts that one infectious host will produce during his or her infectious period in a large population that is completely susceptible. R_0 does not include the new cases produced by the secondary cases, or cases further down the chain. It also does not include secondary cases who do not become infectious.

For example, if $R_0 = 6$ for mumps in a human population, then one infectious person in that population would be expected to produce six new secondary infectious cases if the population were completely susceptible. If the infectious person produced three additional cases who were not infectious, R_0 would still be 6.

For microparasitic infections, R_0 is the product of the contact rate c, the duration of infectiousness d, and the transmission probability per contact with the infectious person, p. The average number of contacts made by an infectious person is the product of the contact rate and the duration of infectiousness—cd. The number of new infections produced by one infective during his or her infectious period is the product of the number of contacts in that time interval and the transmission probability per contact:

$$R_0 = \begin{matrix}\text{number of} \\ \text{contacts} \\ \text{per unit} \\ \text{time}\end{matrix} \times \begin{matrix}\text{transmission} \\ \text{probability} \\ \text{per contact}\end{matrix} \times \begin{matrix}\text{duration of} \\ \text{infectiousness}\end{matrix} = cpd.$$

As presented here, the expression assumes that everyone who gets infected becomes infectious. A term could be included for the probability of becoming infectious after infection. The simplest assumption is that the recovery rate, r, is constant. Then the duration of infectiousness equals the reciprocal of the rate of recovery from the infectiousness, that is $d = 1/r$. Another expression for R_0 is then $R_0 = cp/r$.

R_0 summarizes many important aspects of an infectious agent in a host population in one quantity. It allows comparison of seemingly disparate diseases from the viewpoint of population biology. A value of R_0 is not specific to a microbe, but to a microbe population within a particular host population at a particular time. Contact rates relevant for respiratory transmission will be lower in rural areas than in more densely populated urban areas. So, for example, we expect the R_0 of mumps to be lower in rural than in urban areas. The R_0 of malaria may be low during the season of low mosquito density but high during the season in which mosquitoes are plentiful. The R_0 of HIV in a sexually active population of single people might be much higher than it is in a population of fairly monogamous married couples.

R_0 is dimensionless. It represents the number of new infectious cases per index infectious case (i.e., referent or original case). Without further information about the magnitude of the quantities composing R_0 we cannot conclude much about the time frame of an epidemic, the transmissibility of the microbe, or the contact rate. R_0 is about 2 to 3 for influenza in some populations and also about 2 to 3 for HIV in some populations. Influenza has a relatively high transmission probability and short duration of infectiousness. The influenza virus spreads on a different time scale than HIV, which has a low transmission probability and longer duration of infectiousness. If we were to know only that $R_0 = 3$ for both, then we would know that they both could easiliy produce epidemics, but we would not be able to draw conclusions about the relative time frames of the two. For that, we require further information.

The R_0 for indirectly transmitted diseases depends on the product of the two components of transmission. Indirectly transmitted diseases are those in which an infectious agent is transmitted between two different host populations. An example is the vector-borne disease malaria, which is transmitted from humans to mosquitoes and back to humans.

The definition of R_0 assumes that all contacts are with susceptibles. In real populations, however, people are often immune to a parasite. Under these circumstances, the expected number of new cases produced by an infectious person is less than R_0 and is called the *effective reproductive number*, denoted by R. If x is the proportion of a randomly mixing, homogeneous population that is susceptible, R is the product of R_0 times the proportion x of the contacts made with susceptibles:

$$R = R_0 x. \qquad (1)$$

Suppose that $R_0 = 3$ for influenza in a population and that one-half of the population is immune. Then the effective reproductive number for influenza is $R = 3 \times 0.5 = 1.5$. A case of influenza would produce on average only 1.5 new secondary cases rather than 3 in this population.

R_0 and Public Health

Under what conditions will an epidemic occur? In general, for an epidemic to occur in a susceptible population, R_0 must be greater than one. If R_0 is less than one, an average case will not reproduce itself, so a microbe will not spread. Since R_0 is an average, a particular infectious person could produce more than one infective case, even when R_0 is less than 1, so there may be a small cluster of cases. We would not, however, expect a self-sustaining outbreak.

When a microbe has established itself and is endemic so that, over time, the average in-

cidence does not change, then each infectious case must be producing on average one infectious case, that is, replacing itself. Otherwise the average incidence would either be increasing or decreasing. Thus at equilibrium, on average, $R = 1$.

How might we reduce or eliminate an infectious agent from a host population? If we want to reduce transmission so that the microbe will die out, then we must keep the average number of secondary cases produced by one infectious case below 1, R is less than 1. Suppose that $R_0 = 3$ for genital herpes in a population. To prevent an epidemic, we would need to decrease the contact rate by more than a factor of three. Alternatively, if vaginal foam reduced the transmission probability by 80%, then R_0 would be reduced to $0.2 \times 3 = 0.6$ if everybody used it. Thus, an epidemic might effectively be prevented either by reducing the contact rate or by use of an effective vaginal foam. Suppose that without treatment, an average case of tuberculosis is infectious for one year. If an average case produces five other cases, then $R_0 = 5$. By using active case detection, it might be possible to find cases in the first month of being infectious and treat with antibiotics. If the treated cases become noninfectious within two weeks after beginning treatment, then they would be infectious on average for only six weeks rather than 52 weeks. The R_0 would be reduced to about $(6/52) \times 5 = 0.6$.

What fraction, f, of the population do we need to vaccinate to produce enough immune people so that the infective people will not each be able to infect on average one other person? If the fraction of susceptibles is low enough, the probability that an infective host has contact with a susceptible host before recovering will be very low. The microbe will not be able to persist. Suppose that a vaccine confers complete and lifelong immunity in everyone who is immunized. If f is the fraction vaccinated before the age of first infection, then $1 - f$ would be the maximum fraction of the population that is susceptible, not taking into account additional immune people who have already had the disease. Substituting $1 - f$ for x in expression (1) for R, in principle, we need to vaccinate a fraction f such that

$$R = R_0(1 - f) < 1,$$

to eliminate transmission. The fraction that needs to be immunized to eliminate transmission is

$$f > 1 - (1/R_0).$$

Assume that $R_0 = 3$ for influenza in a population. Under the assumption of random mixing, the fraction that needs to be immunized before the age of first infection is $f = 1 - (1/R_0) = 1 - (1/3) = 0.67$. A higher R_0 requires immunization of a higher fraction to eliminate transmission (Fig. 4–4).

In the preceding example, we assumed that vaccination conferred complete protection. However, an intervention might provide only partial protection, such as the example of using vaginal foam that we presented. In that example, protection was just 80%. A vaccine might provide only partial protection and be just 90% or even 50% efficacious. If in the influenza example the transmission probability per contact in the vaccinated people is reduced by 90%, then the probability of infection in the vaccinated is just the factor $\theta = 0.10$ of that in the unvaccinated. If $R_0 = 3.0$ and everyone is vaccinated, then $R = 0.10 \times R_0 = 0.30$. The vaccine might be successful in preventing the spread of influenza. If the protective efficacy is just 0.50, however, then $\theta = 0.50$. Even if everyone is vaccinated, $R = 0.50 \times 3.0 = 1.50$. Since R is greater than 1, we would not expect to eliminate transmission with this vaccine.

In the preceding example, we assumed that the intervention had the same effect on everyone. An intervention might reduce the transmissin probability, the contact rate, or the duration of infection the same in everyone. As mentioned earlier, though, a risk factor or intervention may have different effects in different people. A vaccine might completely protect some people but fail completely in others. Then the expression for R takes into account the heterogeneities.

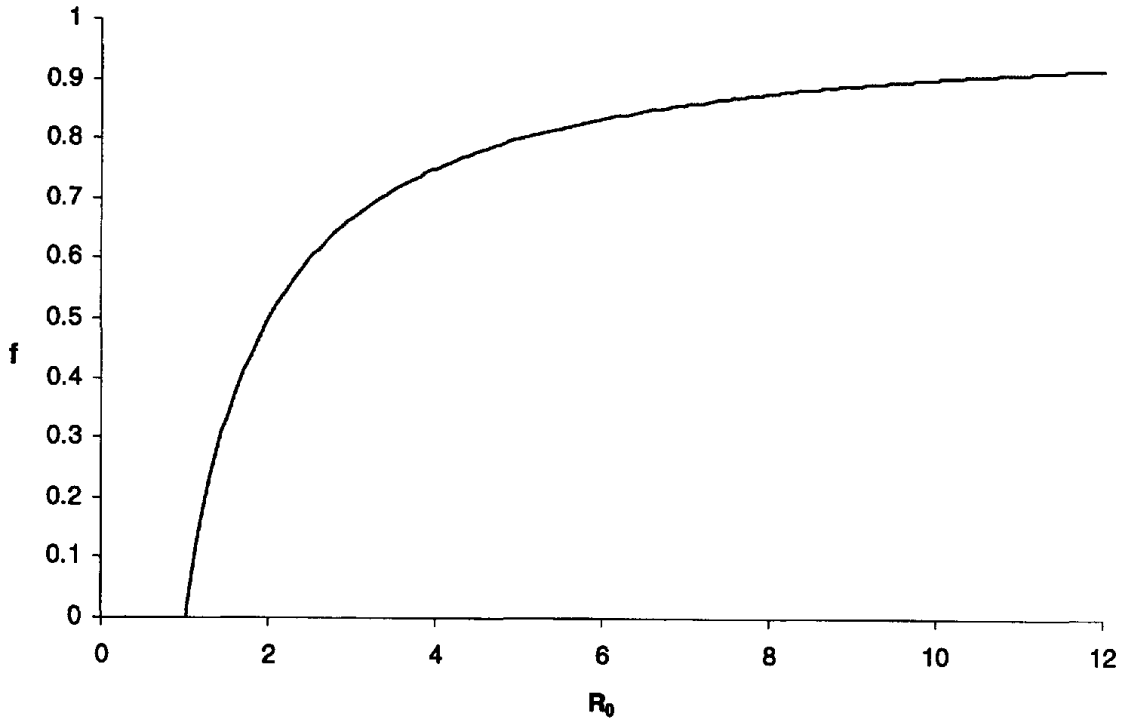


Figure 4–4. The fraction, f, of a population needed to be vaccinated with a completely protective vaccine to eliminate transmission as a function of R_0, the basic reproductive number.

In a simple example, suppose that a vaccine completely protects a proportion, h, of those who are vaccinated, while it fails in the remainder, $1 - h$, of the individuals who receive it. Suppose that, again, a fraction, f, of the population is vaccinated. Then the fraction of the population protected by immunization is hf, and $R = R_0(1 - hf)$. Then the fraction of the population that needs to be immunized to eliminate transmission is

$$f > \frac{1 - (1/R_0)}{h}.$$

Assume as in the preceding influenza example that $R_0 = 3$. Assume that the vaccination fails completely in the fraction $1 - h = 0.15$ while conferring complete and long-lasting protection in the other fraction $h = 0.85$. The fraction, f, that must be vaccinated to eliminate transmission increases to

$$f = \frac{1 - (1/R_0)}{h} = \frac{0.67}{0.85} = 0.79.$$

If the vaccine fails in 40% of the vaccinated people, then the fraction that must be vaccinated is $0.67/0.60 > 1.0$. With such a vaccine at the high failure rate, elimination of transmission would not be possible even if everyone were vaccinated.

HERD IMMUNITY

Herd immunity describes the collective immunological status of a population of hosts, as opposed to an individual organism, with respect to a given microbe (Anderson and May 1982). Herd immunity of a population can be high if many people have been immunized or have recovered from infection with immunity, or be low if most people are susceptible. If x is the proportion susceptible in expression (1), then $(1 - x)$, the proportion immune, gives some measure of the herd immunity. For any given microbe and host population, as herd immunity increases, R will decrease.

Seroprevalence of protective antibodies against an infectious agent is a measure of

herd immunity. In Figure 4–5, the age-specific seroprevalences, that is, proportions of people with anti-hepatitis A virus (HAV) IgG and anti-hepatitis E virus (HEV) IgG in a collection of communities in Vietnam (Hau et al. 1999) are plotted. Seroprevalence of anti-HAV IgG rises very quickly with age, essentially reaching 1.00. The seroprevalence of anti-HEV IgG, on the other hand, is very low. The area under the histograms, adjusted for the varying sizes of the age groups, can be regarded as the level of herd immunity. The herd immunity for HAV is high and that for HEV is low. On average, 97% versus 16% of the people have antibodies against the two diseases. There is concern that the population is susceptible to an outbreak of HEV. Fine (1993) reviews herd immunity.

COMPARING INTERVENTIONS

We can also use the basic reproductive number to help choose among intervention strategies. Which intervention strategy has the largest effect on R_0? Given how much each intervention costs, which has the greatest effect for the amount of money spent? How does our choice of model for R_0 affect our conclusions?

Historically, the concept of R_0 played an important role in the choice of malaria intervention campaigns. In malaria, several interventions can be used against the vector, anopheline mosquitoes. Mosquitoes lay their eggs in water, and the hatched larvae need to breathe at the water surface. Thus either draining breeding pools or putting oil on the water surface will reduce the number of larvae that grow to adult mosquitoes. Some vector mosquitoes will bite nonhuman animals as well as humans. By increasing the number of nonhuman animals available for mosquitoes to bite, the biting rate (i.e., contact rate) on humans will be decreased. Often malaria mosquito vectors tend to bite people indoors, then rest on the walls while they excrete some of the blood fluid. This behavior makes spraying walls with insecticides a useful intervention. How can we derive an expression for R_0 for malaria, then use it to compare intervention strategies? What different aspects of the transmission cycle go into the expression?

Malaria is an indirectly transmitted dis-

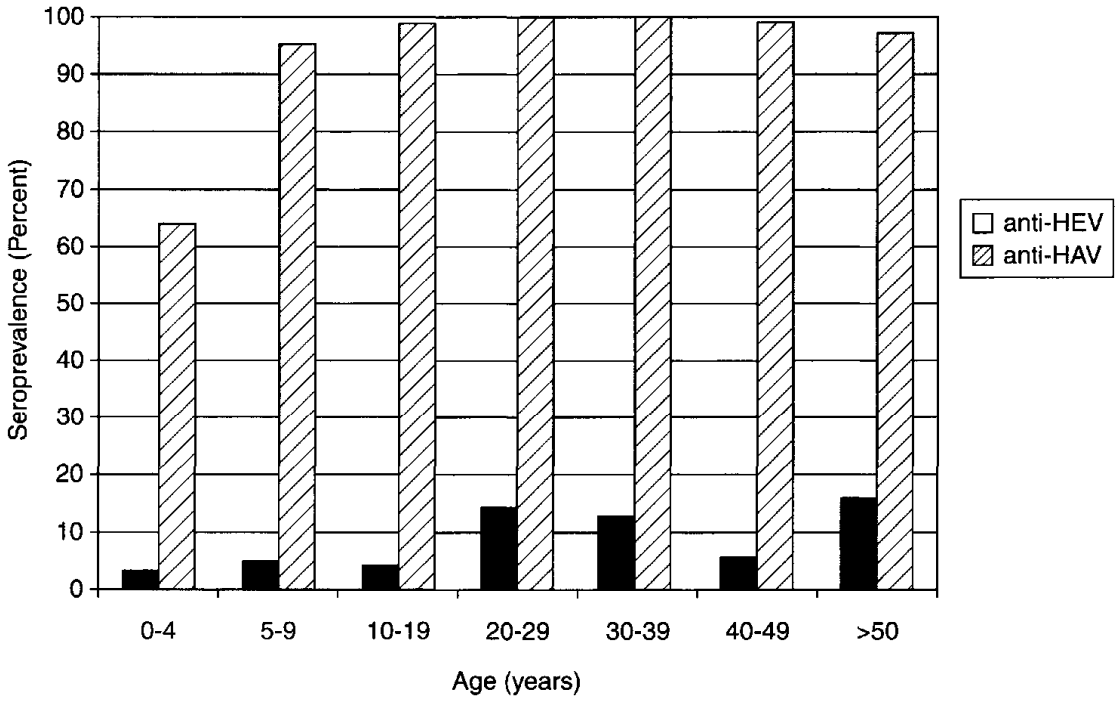


Figure 4–5. Age-specific prevalences of anti-HEV and anti-HAV immunoglobulin G in Vietnam. Source: Hau et al. 1999

ease in that it is transmitted from human to human via female anopheline mosquitoes. We can also say it is transmitted from mosquito to mosquito via the human. Thus, the R_0 expression is composed of two parts, the part from mosquito to human and the part from human to mosquito. We need the contact rates from mosquitoes to humans and from humans to mosquitoes, the two transmission probabilities, and the duration of infectiousness in mosquitoes and humans (Fig. 4–6; Table 4–1).

The expression for R_0 depends on the model we choose for a disease, that is, what components of the life cycle that we include. We consider here two simple models based on the early Ross (1911) and Ross-Macdonald (Macdonald 1957) models (see Aron and May 1982). We assume the humans become infected and infectious at some rate depending on the mosquito biting rate and transmission probability, then recover at some rate, r, without developing immunity. We do not include birth or death of humans in our expression. The duration of infectiousness of humans is $1/r$.

Mosquitoes become infected by biting infective humans. The mortality rate, μ, of mosquitoes is assumed to be independent of whether they are infected. Mosquitoes do not recover from malaria, so the duration of infectiousness is the reciprocal of the death rate, $1/\mu$. The factor, b, is the transmission probability to humans per bite by an infectious mosquito, and c is the transmission probability to mosquito from an infective human.

To get expressions for the two contact rates, we define the quantity, a, as the number of bites per unit time on humans by a single female mosquito, or the contact rate for female mosquitoes with humans per unit time. The quantity a is a composite of the rate at which mosquitoes take blood meals and the proportion of those blood meals that are taken on humans. We assume that there is some constant number, M, of female mosquitoes and a constant number, N, of humans, so that the number of female mosquitoes per human host is $m = M/N$. The factor, $ma = aM/N$, is the rate of bites received by one human per unit time.

The component ac/r represents the part of R_0 of mosquitoes being infected by humans. The component mab/μ represents the part of R_0 of humans becoming infected by mosquitoes. Both components have the form of contact rate times transmission probability

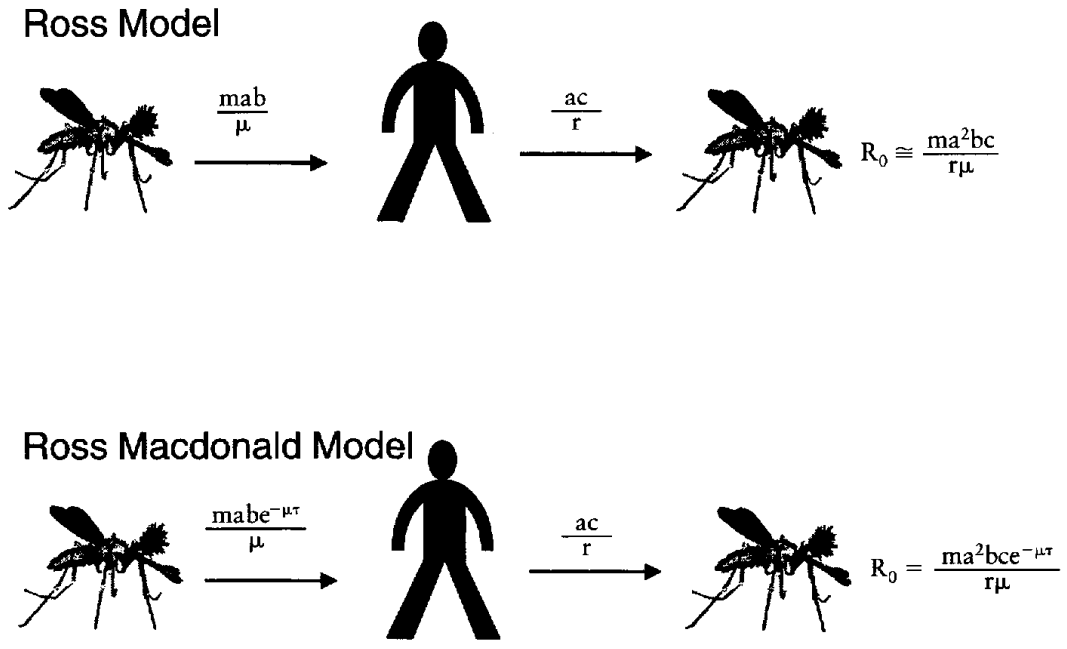


Figure 4–6. R_0 expression for two different malaria models.
Source: Mosquito image used with permission from the American Museum of Natural History.

Table 4–1 Quantities for the R_0 for Malaria

Term	Meaning
N	the size of the human population
M	the size of the female mosquito population
m	= M/N, the number of female mosquitoes per human host
a	the rate of biting on humans by a single mosquito (number of bites per unit time)
b	the transmission probability from an infective mosquito to a human
c	the transmission probability from an infective human to a mosquito
r	the recovery rate for humans
μ	the mortality rate for mosquitoes
τ	the latent period of the malaria parasite in the mosquito

times duration of infection. The basic reproductive number for the simple Ross model is

$$R_0 = \frac{ac}{r} \times \frac{mab}{\mu} = \frac{ma^2bc}{r\mu}.$$

As an example, suppose that the transmission probabilities $b = c = 0.1$, $M = 1{,}000{,}000$ mosquitoes, $N = 1000$ humans, $a = 0.1$ bites on human per day, $\mu = 0.1$ per day, and $r = 0.004$ per day. Then $R_0 = 250$, very high indeed.

This original simple model by Ross does not include the latent period (or external incubation period) in the mosquito. The latent period of the malaria parasite in mosquitoes is an important component, because it is on the order of the life expectancy of the mosquito. Thus a large proportion of the mosquitoes who become infected never become infective before they die. In the late 1940s, George Macdonald (1957) added the latent period of malaria in the mosquito to the model. If τ (pronounced "tau") is the latent period of the malaria parasite in the mosquito, the probability of a mosquito surviving the latent period to become infectious without dying is $e^{-\mu\tau}$. The expression for R_0 for the expanded Ross-Macdonald model is

$$R_0 = \frac{ac}{r} \times \frac{mabe^{-\mu\tau}}{\mu} = \frac{ma^2\, bce^{-\mu\tau}}{r\mu}.$$

Suppose that the latent period $\tau = 10$ days and that the other quantities have the same values as above. Then using the Ross-Macdonald model, $R_0 = 92.0$, considerably lower than the R_0 calculated above. Thus the expression for R_0 and the underlying model can influence the value obtained for R_0.

The Ross-Macdonald R_0 played an important role in the decision of the World Health Organization to launch the malaria eradication campaign in the 1950s. This eradication campaign was based on spraying the insecticide DDT on the insides of houses to kill the mosquitoes resting after taking a blood meal, drastically increasing the mortality rate of mosquitoes, μ. Interventions up until that time had been aimed at decreasing the number of female mosquitoes, M, thus m, the number of female mosquitoes per human, by draining breeding pools or spraying water with oil. Another intervention was to add animals to the environment, so that the mosquitoes might bite the animals instead of the humans. This would reduce a, the rate of biting on humans by a single mosquito.

By seeing how these three quantities, μ, M, and a, enter into R_0, it is possible to get an idea which type of intervention would have the strongest effect. By inspection, M in ma, enters linearly into R_0. Because the biting rate of the mosquito is included twice, the squared power of a enters R_0. The life expectancy of the mosquito also enters twice into R_0, once linearly, and then exponentially. If the number of mosquitoes is reduced by one-half, R_0 will just decrease by a factor of two. If a is reduced by one-half, R_0 will decrease by a factor of four. If μ is increased by

a factor of two, the decrease in R_0 will be by a factor greater than four.

Continuing the preceding example, suppose interventions change the parameters M, a, or μ by factors of 2. If the abundance of mosquitoes is reduced to M = 500,000, then $R_0 = 46.3$. If the biting rate is reduced to $a = 0.05$, then $R_0 = 23.2$. If the mortality rate of the mosquitoes increases by a factor of 2 to $\mu = 0.2$, then $R_0 = 16.9$. Thus increasing the mortality rate, thereby reducing the life expectancy of the mosquito, has the strongest effect on R_0. The goal of the eradication program was to decrease $R_0 < 1$. In this example, to reduce R_0 so that it is less than 1, then μ must be increased to somewhat more than 0.4, for a life expectancy of about 2.5 days. Thus mortality of the mosquito needs to be increased by a factor of 4. The abundance of mosquitoes would have to be reduced by a factor of at least 92, however, to reduce R_0 below 1. We leave it as an exercise for the reader to calculate the changes in R_0 if the same interventions were used under the simple Ross model rather than the Ross-Macdonald model.

The prior calculations show how the Ross-Macdonald model could have a strong influence in embarking on the eradication campaign. Other factors, such as the beginning of the appearance of insecticide resistance, also put pressure on the campaign. The eradication campaign was abandoned in the late 1960s. After that time the goal was to achieve a new host-parasite balance. More complex models of malaria were developed that included immunity and superinfection (Dietz et al. 1974, Struchiner et al. 1989, Halloran et al. 1989). These models allow modeling of the effect of vaccination.

Factors contributing to malaria epidemics can also be understood using the Ross-Macdonald expression. The mortality rate of mosquitoes depends heavily on the weather. In particular, mosquitoes live longer in higher humidity. Also, the latent period, or extrinsic cycle, of the malaria parasite within the mosquito depends on the temperature. Thus increased humidity would reduce μ, increasing R_0. Similarly, high temperatures would reduce τ, also increasing R_0.

In this section we have shown that the computed value of R_0 depends on what is included in the model. Also, the effect of interventions can be compared using R_0, but conclusions will also depend on what is included in the model.

EVOLUTIONARY USES OF R_0

R_0 can be used to quantify evolutionary concepts. *Virulence* is a measure of the speed with which an organism kills an infected host. We denote the disease-dependent death rate, or virulence, by α. If r is the recovery rate from infectiousness, and α the virulence, then the duration of infectiousness is $d = 1/(r + \alpha)$ and $R_0 = cp/(r + \alpha)$. Since R_0 is a function of the time spent in the infective state, R_0 could decrease as virulence increases. If the microbe is highly virulent so that it kills its host quickly, then R_0 could be less than 1, and the microbe will die out. For example, suppose that the microbe does not kill the host and that the host usually recovers from infectiousness in about $d = 10$ days. Then $r = 0.1$ per day. If $R_0 = 3.0$ for this disease, then $cp = rR_0 = 0.1 \times 3.0 = 0.3$. If instead the microbe kills the host on average in a little over three days when the host does not recover first, then $\alpha = 0.3$ per day, and $R_0 = 0.3/(0.1 + 0.3) = 0.75$. In this case $R_0 < 1$, so the microbe will not be successful. If, on the other hand, the microbe kills the host only after about 10 days on average when the host does not recover first, then $R_0 = 0.3/(0.1 + 0.1) = 1.5$. In this case, $R_0 > 1$. Viewed in this way, there is evolutionary pressure on microbes to become less virulent and to develop a more benign relation to the host.

In some diseases, hosts become more infectious when they become sicker, so the transmission probability increases at the same time that virulence increases. Thus, R_0 could increase as virulence increases, putting evolutionary pressure on the agent to increase virulence. The balance depends on the particular microbe. In the above example, suppose that even though the microbe kills the host on average after about three days, the transmission probability, p, also increases

by a factor of two. Then $R_0 = (c \times 2p)/(r + \alpha) = (2 \times 0.3)/(0.1 + 0.3) = 1.5$. In this case, R_0 is greater than 1, so we would expect the microbe to be successful. The increased virulence was offset by the increased transmission probability to keep $R_0 > 1$.

The *case fatality rate* is the probability of dying from a disease before recovering or dying of something else. In the notation used here, the case fatality rate is $\alpha/(r + \alpha)$ (ignores other death causes). If virulence is $\alpha = 0.3$ per day, and the recovery rate is $r = 0.1$ per day, then the expected case fatality rate is $0.3/(0.1 + 0.3) = 0.75$. This means that 75% of the people die before recovering. As virulence increases, the case fatality rate increases.

R_0 IN MACROPARASITIC DISEASES

The concept of R_0 comes from general population theory and refers to the expected number of reproducing offspring that one reproducing member of the population will produce in the absence of overcrowding. With larger parasites such as worms, called *macroparasites*, we define R_0 to be the expected number of mature female offspring that one female will produce in her lifetime. This contrasts with the definition of R_0 for microparasites, or microbes, which refers to the number of new infectious *hosts* produced by one infectious host.

For example, the disease schistosomiasis is caused by large, sexually reproducing worms called schistosomas that can live for over two decades within a human host. If a female schistosoma worm has an $R_0 = 2$ in a population of human hosts and an intermediate host population of snails, then the average female schistosoma produces two mature female worms. Most of the thousands of eggs produced by the adult female do not survive passage through the environment and the intermediate snail hosts to establish themselves in another human host. The two new successful worms could be in one new human host, or in two different hosts. The $R_0 = 2$ refers to the number of worms, not to the number of hosts. There are some further complexities in calculating thresholds, because there must be at least one male worm in the human host for the female to reproduce.

What is important in designing interventions against macroparasitic diseases? The total number of parasites in a host is often more important than whether a host is infected, because the level of morbidity of a host can be associated with the number of parasites the host carries, or *parasite burden*. Some hosts can have very heavy infection, that is many worms, while others have very light infection. Chemotherapy that targets people with heavy parasite loads could have a greater effect on transmission and morbidity than untargeted therapy.

CAVEATS

The previous sections demonstrate that R_0 is a conceptually useful measure that provides a summary of several aspects of an infectious disease. However, the simple relations described earlier usually do not hold. Heterogeneities in the contact rates, transmission probabilities, and infectious periods produce different R_0s in different subgroups. If members of a group who live near each other are not immunized, then it is possible for transmission to occur in that group, even when transmission has been eliminated in other segments of the population. The contact rate can increase locally if people move into crowded conditions, such as into college dormitories, military barracks, or refugee camps. Especially when transmission is tenuous or near elimination, heterogeneities can play an important role in determining whether a microbe can persist in a population. Anderson and May (1991) present an extensive overview of R_0. R_0 is a relatively static concept. Further understanding of infectious diseases in populations requires study of transmission dynamics.

DYNAMICS OF INFECTION IN A POPULATION

Under some circumstances an infectious agent will invade and establish itself in a susceptible host population, with an ensuing

epidemic, then die out again. Some infectious agents will invade, however, and after an initial epidemic, *persist*. They become *endemic*, with either fairly stable, possibly seasonal transmission, or other epidemic patterns. In addition to the consderations of R_0 described above, under what conditions might persistence or dying out happen?

CONTACT PROCESS AND RANDOM MIXING

To describe the spread of an infectious agent in a human population, we need to describe how the human hosts and any vectors contact each other in some way that the infectious agent can be spread. There are different ways to think about how individuals in populations make contact. One is that people behave like gas molecules with the rate of contact being determined by density. If people were pressed more closely together, as in an urban environment, they would contact each other more often than if they were less densely distributed, as in a rural environment. Hence, for diseases such as measles, influenza, or mumps that spread by airborne or droplet transmission, population density plays a role in determining the value of R_0. Alternatively, contacts can be selected, such as in sexual contacts or injection of intravenous drugs. In this case, R_0 is determined more by social behavior. In many cases, both density and social choice will play a role in determining contact rates and mixing patterns.

Regardless of how contacts arise, the simplest assumption about the contact pattern in a population is that of *random mixing*. Figure 4–7 schematically represents random mixing, with the figures being evenly distributed in the space. Under the assumption of random mixing, every person has an equal chance of making contact with each other person. Consequently, every person also has an equal chance of being exposed to infection because every person is equally likely to make contact with an infectious person. The assumption of equal exposure to infection of people in the comparison groups, and whether it is valid, is important in many studies of interventions and risk factors affecting susceptibility. As in the prior discussion, we denote by c the constant contact rate that does not change over time in a randomly mixing population.

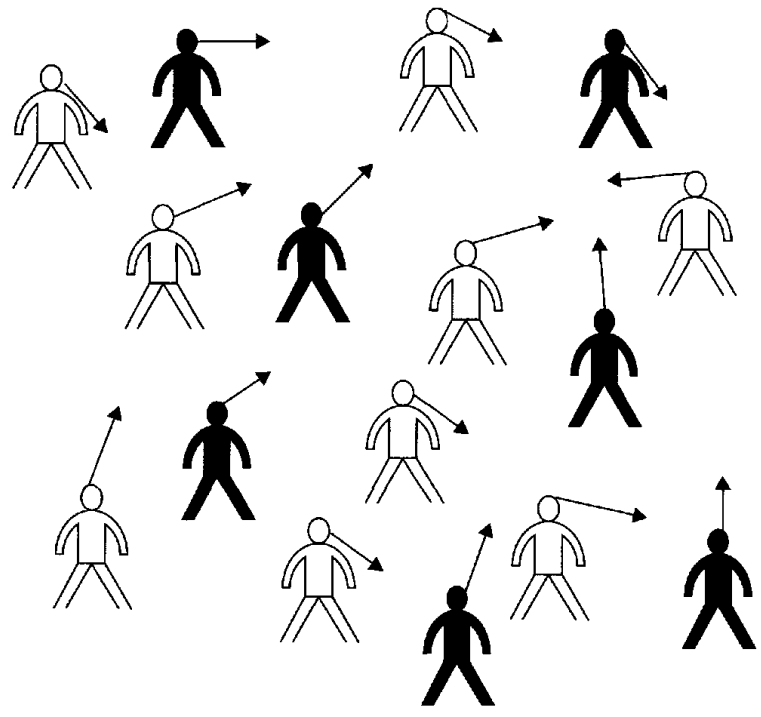


Figure 4–7. Random mixing. Solid figures denote infective people. Open figures denote susceptible people.

STATES OF THE HOST POPULATION

Suppose we choose to model an infectious disease by allowing people in the human population to pass through three different states (Fig. 4–8A,B). They start out susceptible, denoted by X, then become infected and infectious, denoted by Y, after which they recover with immunity, denoted by Z. Models of this type of infection process are called SIR models for susceptible, infected, recovered or removed. 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. We use the notation XYZ here, rather than SIR, because we use I for incidence rate and R for incidence proportion. If these are the only three states possible, then each person in a population of N individuals is in one of these three states, where $X(t)$ is the number of susceptible people at time t, $Y(t)$ is the number of infectives, and $Z(t)$ is the number of immunes. This simple model ignores the latent and incubation periods, and assumes that infection, disease, and infectiousness occur simultaneously. This model could be a simplified representation of influenza, measles, or chickenpox.

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. In a closed population, there are no births, immigration, deaths, or emigration. We first consider a closed population of N initially susceptible people who are assumed to be mixing randomly with contact rate c (Fig. 4–8A). The population is analogous to a closed cohort. Initially, at time $t = 0$, everyone in the population is in the susceptible state X.

DYNAMICS OF AN EPIDEMIC

Suppose that a microbe such as an influenza virus is introduced into a closed population, so that one person enters the infectious state, Y (Fig. 4–9). Alternatively, an infectious person or several infectious persons besides the N initially susceptible might enter the population. Here we consider a simple deterministic, mass mixing model of the spread of infection. A deterministic, non-chaotic, model always gives the same answer and usually solves equations for populations rather than for discrete individuals.

The infection spreads from the first infective to the average number R_0 of susceptibles. If people recover at the rate r, then they are infectious on average for the time

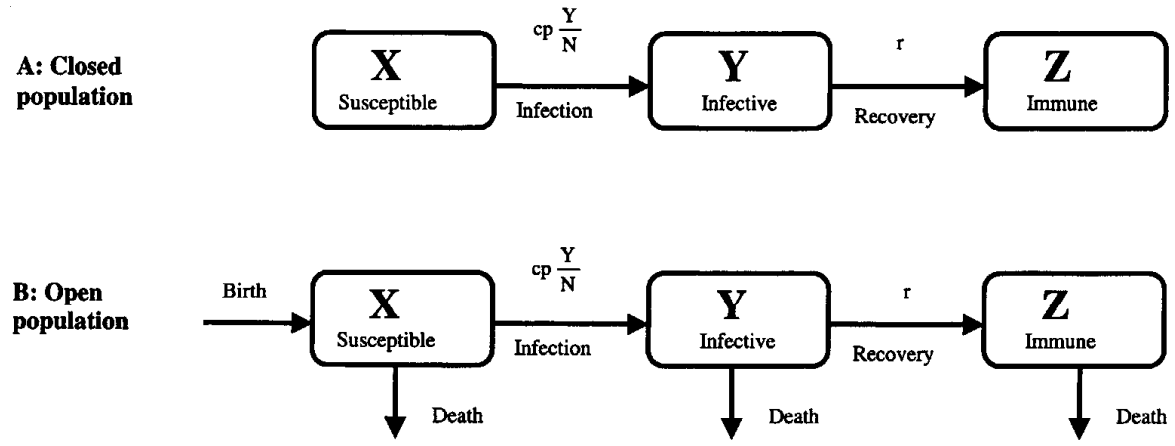


Figure 4–8A,B. Transmission model for an infectious disease in a host population. The three compartments represent susceptible (X), infective (Y), and immune (Z) hosts at time t. The total host population is of size $N = X + Y + Z$. Susceptible hosts become infected at an incidence rate (force of infection) of cpY/N, where c is the contact rate, p is the transmission probability, and Y/N is the prevalence of infective hosts at time t. The rate of recovery is r. Arrows represent transitions in and out of compartments.

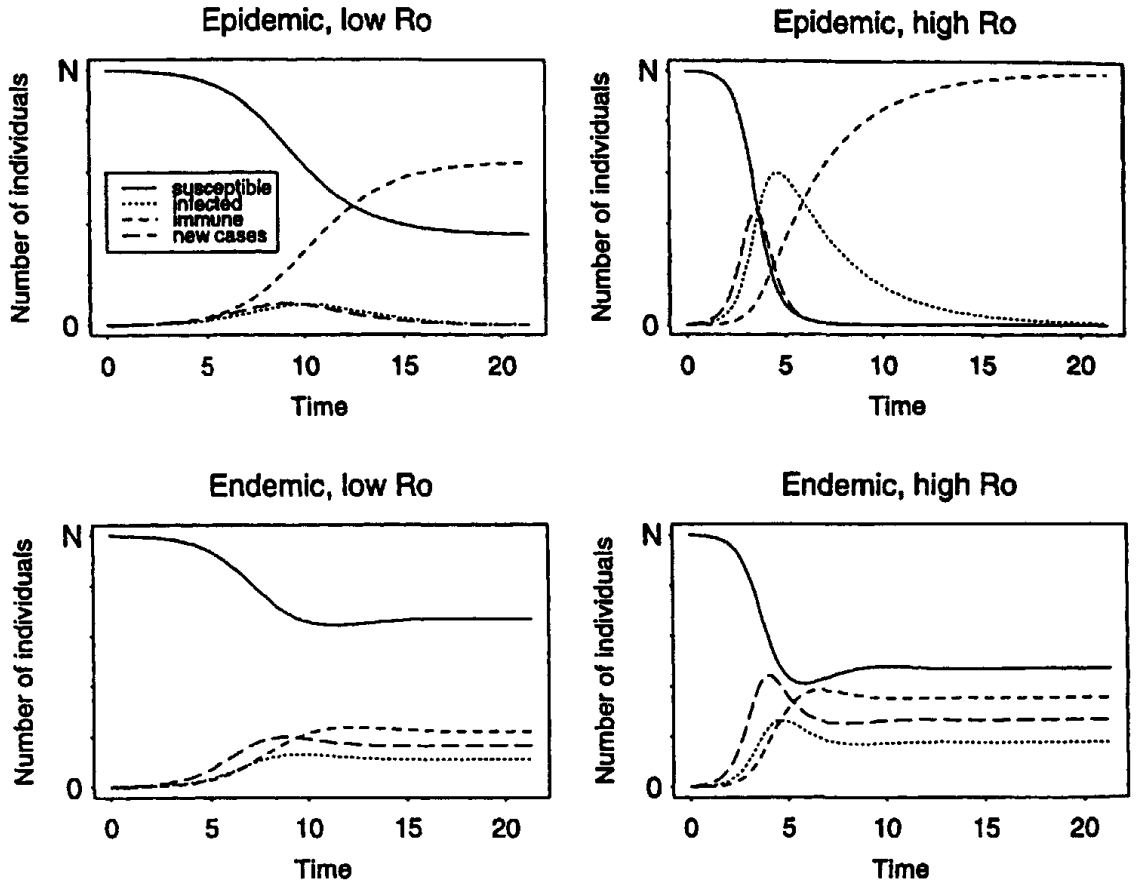


Figure 4–9. Comparison of the spread of an infectious disease in a closed or open population. The infectious agent is introduced into a population of N susceptibles. Susceptible people become infected and infectious, then develop immunity. **Top left:** Epidemic in a closed population, low R_0. The epidemic dies out before all susceptibles become infected. **Top right:** Epidemic in a closed population, higher R_0. Everyone becomes infected during the epidemic. There are no infectives left as the epidemic dies out. **Bottom left:** Epidemic followed by endemic persistence in an open population, low R_0. The infectious agent does not die out due to the supply of new susceptibles. Prevalence of susceptibles, infectives, and immune people is in dynamic equilibrium. The number of new incident cases is steady. **Bottom right:** Epidemic followed by endemic persistence in an open population, high R_0.
Source: Halloran, 1998.

period $d = 1/r$. If $R_0 > 1$, the epidemic is expected to spread. The first infective eventually recovers with immunity into state Z, while the infection spreads from those he or she infected to more susceptibles. In Figure 4–9, the number of infectives, $Y(t)$, initially increases. As the epidemic spreads, the number of susceptibles, $X(t)$, decreases, while the number of people with immunity, $Z(t)$, begins to increase. Incidence and prevalence of infection will increase until the number of susceptibles available becomes a limiting factor. Then the number of new cases and the prevalence begin to decrease until the microbe dies out and no people are left in the infective compartment, $Y(t)$. A microbe in a closed population where people recover with long-lasting immunity will inevitably die out, because the key to *persistence* in a host population is a continuous supply of susceptibles. The susceptibles can be produced either by births or immigration into the population, or by recovery without immunity or by waning of immunity after it is acquired. In this example of a closed population, however, no new susceptibles are produced.

The dynamics of the epidemic are described by three differential or difference

equations that express the rate of change of the number of people in each of the three states. The rate at time t at which people leave the susceptible compartment X and become infected is simply the incidence rate, $I(t)$, or similarly the hazard rate or force of infection. The prevalence of infectives at time t, $P(t)$, is the number of infectious people, $Y(t)$, divided by the size of the population N, or $Y(t)/N$. The expression for the incidence rate as a function of prevalence in the epidemic is

$$I(t) = cpP(t) = cp\,\frac{Y(t)}{N}.$$

This is the dependent happening expression discussed in Chapter 5. The change in the number of susceptibles, the population-at-risk to become infected, $\Delta X(t)$, per small interval of time, Δt, at time t equals the incidence rate, $I(t)$, times the size of the population-at-risk, $X(t)$. The change in the number of infectives, $\Delta Y(t)$, is the difference between the number of new infections and the number of infectives developing immunity. The number of infectives developing immunity in the time interval Δt is the change in the number of immunes, $\Delta Z(t)$. The three difference equations for the epidemic model are then

change in susceptibles:

$$\frac{\Delta X(t)}{\Delta t} = -I(t)X(t) = -cp\frac{Y(t)}{N}X(t),$$

change in infectives:

$$\frac{\Delta Y(t)}{\Delta t} = cp\frac{Y(t)}{N}X(t) - rY(t),$$

change in immunes:

$$\frac{\Delta Z(t)}{\Delta t} = rY(t).$$

More commonly, differential equations are used, but we avoid the notation here.

We can associate aspects of the epidemic process with common epidemiologic measures. An estimate of the incidence rate, $I(t)$, estimates $cpY(t)/N$. A cross-sectional study to estimate prevalence, $P(t)$, of current infection would yield an estimate of $Y(t)/N$. The number of new infections in an interval of time estimates $[cpY(t)/N]X(t)\,\Delta t$, the incidence rate times the number at risk for the event times the time interval. The epidemic process of a disease producing long-lasting immunity in a closed population is always either increasing or decreasing. An important consequence for conducting studies in epidemics in closed populations is that there is no stationary state of the disease process. Thus epidemiologic methods, study designs, or analytic methods that assume stationarity of the disease process are not applicable under epidemic conditions.

The epidemic process also depends on the population biology. Since R_0 is the product of the contact rate, the transmission probability, and the infectious period, in this model, $R_0 = cp/r$. The expected number of new cases per infective host decreases from R_0 to $R = R_0x$, where $x = X(t)/N$, the proportion still susceptible at time t. The epidemic peaks and begins to decrease when R is less than 1, so that $X(t)/N$ is less than $1/R_0$, that is, when the proportion of the population still susceptible becomes less than the reciprocal of the basic reproductive number. Not all the susceptibles need to become infected before the microbe dies out. The greater R_0, the fewer susceptibles will be left when the epidemic peaks and the fewer susceptibles will be left at the end of the epidemic (Fig. 4–10). Thus the incidence proportion, or attack rate, after an epidemic provides information on R_0. If an intervention reduced some aspect of R_0, then the intervention would result in the epidemic peaking when a greater proportion of the population was still susceptible, and fewer people would become infected before the epidemic died out.

TRANSMISSION IN AN OPEN POPULATION

An open population can have people entering, leaving, or both. In an open population, the susceptibles form a dynamic cohort with

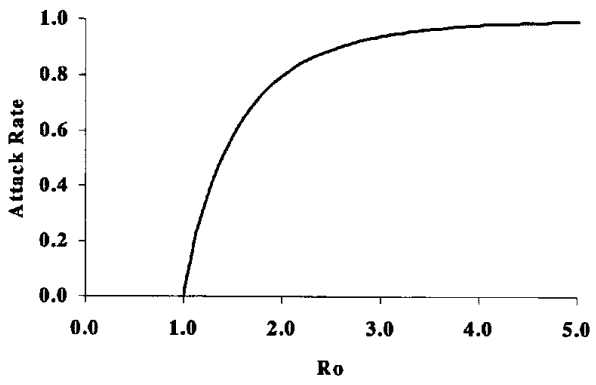


Figure 4–10. The attack rate as a function of the basic reproductive number, R_0.

the population-at-risk changing over time (see Fig. 4–8B). In an open population, if the replenishment of susceptibles is fast enough compared with the dynamics of the microbe, then the microbe will not necessarily die out. The microbe can invade the population, establish itself, persist, and become endemic. It is possible, however, that the microbe will die out if the replenishment of susceptibles is not fast enough in comparison to the spread of immunity to the microbe. Microbes can persist by hopping from one population to another, then returning to one where the susceptibles have had time to replenish themselves.

When a disease is first introduced into a population, there will be a period when the dynamics are not stationary. As stated in the section, Dynamics of an Epidemic, epidemiologic methods that assume stationarity of the disease process cannot be used during the epidemic phase. If the infectious agent has achieved a dynamic equilibrium, however, then some relations might be applicable. An open population with a dynamic cohort at risk for infection is amenable to many of the study designs regularly used in dynamic cohorts. In choosing study designs and methods of analysis, we need to consider whether the dynamics of transmission are at equilibrium or changing over time.

WITHIN HOST DYNAMICS

The dynamics of the infectious agent *within* a host also can be described by dynamic models (Antia et al. 1996, 1998, Pilyugin et al. 1997). These models describe the interaction of the microbe with the immune cells or antibodies that might attack it, and its target cells within the host. Similar concepts from population theory are used to model the within-host dynamics of the infectious agent and to model the infectious agent circulating in the human population. For example, R_0 for a virus within a host describes the number of new viral particles successfully produced by one virus particle. The various immune compartments such as T cells, B cells, and memory cells can be included in the dynamic models.

Chain Binomial Models

The deterministic, mass action, dynamic models described above are useful for exploring scientific and biologic questions. However, they have not been used much for estimating quantities of interest. *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 (i.e., individual "links" of the chain) of time, producing infection chains governed by the binomial (i.e., dichotomous outcomes such as yes/no or newly infected/not newly infected) probability distribution. 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 before, 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 infectious person.

As a simple example of the Reed-Frost chain binomial model, consider spread of infection in a group of three individuals, where one person is initially infected and the other two are initially susceptible (Table 4–2). We assume that the initial infective is no longer infective after the first time unit. In the first

Table 4–2 Chain Binomial Reed-Frost Model in Groups of Size 3 with 1 Initial Infective and 2 Susceptibles

Chain	Chain Probability	at p = 0.4	at p = 0.7	Total Infected
$1 \rightarrow 0$	q^2	0.360	0.090	1
$1 \rightarrow 1 \rightarrow 0$	$2pq^2$	0.288	0.126	2
$1 \rightarrow 1 \rightarrow 1$	$2p^2q$	0.192	0.294	3
$1 \rightarrow 2$	p^2	0.160	0.490	3
Total	1	1.00	1.00	

time unit, one of three things can happen: neither of the susceptibles will become infected; both of them will become infected; or just one of them will become 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 while the other does not is pq. Since this can happen two ways, then the probability of just one being infected in the first time unit is $2pq$. If one of the susceptibles becomes infected in the first time unit, then this person is the new infective who can expose the last remaining susceptible. Exposure of the last remaining susceptible can result in two possible outcomes. Either the susceptible becomes infected or does not, with probabilities p and q, respectively. The *chained probabilities* then are $2pq \times p = 2p^2q$ and $2pq \times q = 2pq^2$, respectively.

In Table 4–2 the chain probabilities are calculated for two different values of p, $p = 0.4$ and $p = 0.7$. If we were to have 1000 groups of size three with one initial infective, at $p = 0.4$ we would expect 360 groups to have just one infected, 288 to have two infected, and 192 + 160 = 352 to have three infected at the end. Similarly, at $p = 0.7$, we would expect 90, 126, and 784, respectively. Note that 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. In this case, we would have only final value data, that is, data on how many were infected in each household at the end. This is also called the final size distribution.

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

The probability that an epidemic will not spread from the initially infected people is 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 Y_0, and the number of initially susceptible people X_0.

The probability that a susceptible person escapes infection from all Y_0 initial infectives is q^{Y_0}. The probability that all X_0 of the susceptible people escape infection from all of the infectives is $P_{ns} = (q^{Y_0})^{X_0}$. 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 terms *minor* and *major* epidemics distinguish situations in which there is little spread from the initial infectives from situations in which an epidemic gains momentum and is self-sustaining.

Chain binomial models can be used to estimate the transmission probability from data gathered on each generation of infection or from the final distribution of infectives within a collection of households or

other small transmission units after an epidemic has occurred. Abbey (1952), Bailey (1957), and Becker (1989) discuss chain binomial models. An important assumption of the simple version of the Reed-Frost model is that the households or mixing groups are each independent of one another. Below we present an extension of the model that allows for interaction outside of the households within the community.

Stochastic Models

Stochastic models, which incorporate elements of chance, are commonly used in infectious disease modeling (Chiang 1980). For example, the Reed-Frost model can be simulated using a random number generator at each step for each person to decide whether an exposed person becomes infected. In contrast, in deterministic (i.e., nonstochastic), mass action models, fractions of a population are assigned to a particular state at any given time. 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, 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.

In a staged Markov model, individuals rather than populations move through states. Whether a person moves to the next state in a given time unit has a certain probability and is random. The Markov property means that the probability of moving to the next state is independent of the time already spent in the current state. The relation between the stochastic formulation of epidemic models and the deterministic formulation has been studied in detail. In many situations, the deterministic model gives an average of the behavior of the stochastic model. However, more situations lead to extinction of infection in stochastic models than in deterministic models.

Complex Dynamic Models and Simulation

Many questions of interest require more complex models than we can present here. What are the age-related changes in infection and disease? What is the advantage of using a targeted versus an untargeted strategy? Will natural immunity wane if transmission is too low? If many people are vaccinated, the incidence of infection will decrease, so that the average age of infection in the susceptibles will increase. Some diseases, such as mumps, chickenpox, and rubella, are more serious if acquired at older ages. Thus the number of total cases could decrease owing to a vaccination program at the same time that the number of serious cases would increase. For example, rubella is a mild disease in children, but it can result in congenital defects if a pregnant woman becomes infected. If many, but not all, young people are vaccinated, then transmission will be reduced. The people who were not vaccinated will acquire rubella at a later age than if no one was vaccinated (Knox 1980, Ukkonen and von Bonsdorff 1988). Thus it is possible that the number of babies born with congenital defects could increase, even though fewer people contract rubella.

A similar concern about introducing varicella vaccination in the United States was raised. The question was whether vaccination, especially if the fraction vaccinated was not high, could increase the number of primary chickenpox cases in older age groups who have more severe morbidity. Halloran and colleagues (1994) studied several different scenarios and found that likely vaccination would not result in more severe cases. Models including age (Schenzle 1984) and mixing structures are required to study complex questions such as this one. The general rule is that a model has to contain the characteristics related to the question you are asking or you cannot get an answer. Anderson and May (1991) provide an extensive overview of deterministic models to study dynamics of infectious disease and interventions.

Several caveats should be kept in mind in considering the results of complex models and computer simulations. Regardless of

how complex the model is, it is always a simplification of reality. Someone made choices in choosing what would be included in the model. These choices affect the results produced by the model. Models are excellent at forcing us to make both our assumptions and our ignorance explicit. Often, too few data are available to estimate the parameters, and the results usually underestimate the uncertainty of the knowledge. Regardless of these caveats, models are very useful in sharpening our thinking and especially in gaining qualitative understanding of complex processes.

USING DYNAMIC CONCEPTS TO INTERPRET STUDIES

We now illustrate how understanding the transmission dynamics of an infection can help interpret the results of a study. Two hypothetical investigators who conducted separate studies of gonorrhea in a heterosexual population of men and women come to different conclusions. The subscript m and f denote men and women, respectively. The first investigator conducted a study in clinics using a sound sampling scheme with good ascertainment. The results showed that the incidence rate and number of new clinical cases of gonorrhea are higher in men than women, $I_m > I_f$. The investigator concluded that gonorrhea is a greater problem in men than women. The second investigator conducted a population-based study that was also well designed, and found that the prevalence of gonorrhea infection is higher in women than in men, $P_f > P_m$. She concluded that the problem is greater in women. How can transmission concepts help us think about this paradox (Fig. 4–11)?

Assume that gonorrhea transmission has been fairly constant over a period of time in this population. Women can be infected with gonorrhea for a long time before they develop symptoms, whereas men develop symptoms quickly and go for treatment. Thus the infectious period in men is shorter than in women, d_m is less than d_f. Generally the transmission probability from females to males is lower than that from males to females. However, to make this point as simply as possible, we assume here that they are equal, so that $p_{fm} = p_{mf} = p$. Assume that the population has an equal number of men and women, $N_m = N_f = N$, that the rate of new partners (contact rate) is the same in both, $c_m = c_f = c$, and that men and women mix randomly with the opposite sex.

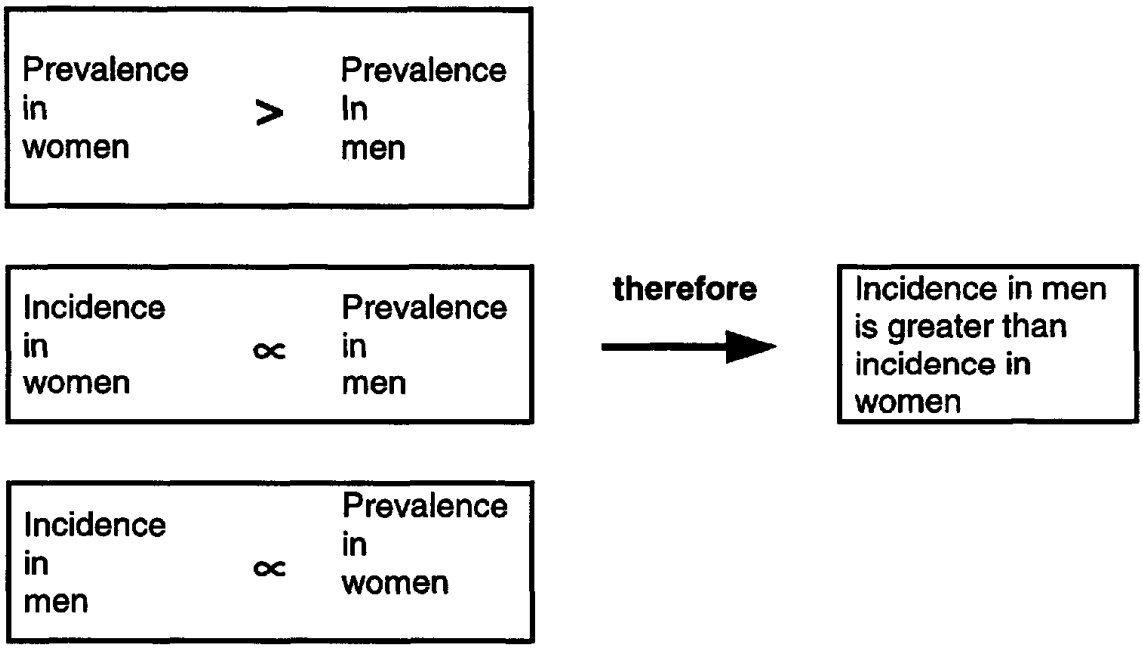


Figure 4–11. Relation of incidence rate and prevalence of gonorrhea in men and women.

Prevalence of infection in women is higher than in men, largely because the duration is longer, so there is a greater number of susceptible men than women who are at risk to become new cases $(1 - P_m) > (1 - P_f)$. Susceptible men make the same number of contacts per unit time and have the same transmission probability as the women, but their contact pool, the women, has a higher prevalence of infection, so the incidence rate is higher in the men, $I_m = cpP_f > cpP_m = I_f$. The combined effect in the men of a higher incidence rate and a greater proportion of susceptibles results in a higher number of new cases in men than in women. If we were to conduct a study in a clinic based on incidence rate or number of new cases, we would conclude that the problem is more serious in men. If we were to conduct a prevalence study in the population, we might think the problem is more pronounced in women. The males and females are related through the dynamic transmission process, and the paradox is resolved.

If we can reduce the population prevalence in women, for example, by shortening the duration of infection by early detection and treatment (Thomas et al. 1998, 1999), it would reduce the incidence rate in men, and consequently the prevalence in men. This, in turn, will reduce the incidence rate in women, and consequently contribute further to decreasing the prevalence in women. The dependence of events in infectious diseases results in interventions having greater overall effects than would be expected from just the direct effects in the individuals receiving the intervention. We leave it as an exercise for the reader to develop an expression for the basic reproductive number for this situation. What would happen to the rate of new partnerships (contact rate) c_f and c_m if there were twice as many men as women? Consider what would happen if the transmission probability from men to women, p_{mf}, were twice as high as that from women to men, p_{fm}.

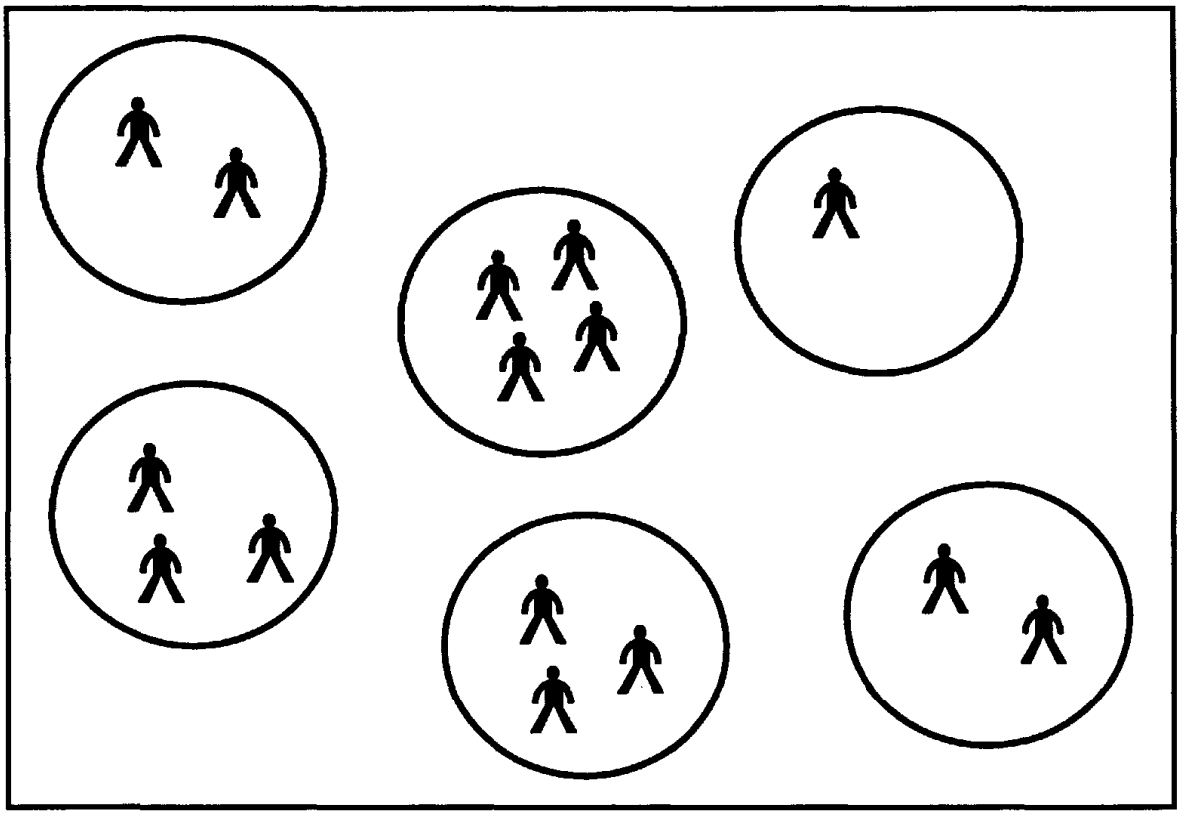
NONRANDOM MIXING

We conclude this chapter by considering some more ideas on contact processes and patterns within populations. Contact patterns play a central role in determining transmission and exposure to infection. Most populations do not mix randomly but are composed of different types of small transmission units or subpopulations that mix with their own members differently than with other subpopulations. The groups could be sexual behavior groups, different age groups within a school, or households in a community. Individuals may belong to several different mixing groups, including families, schools, and neighborhoods. Our scientific questions and the purpose of our investigation will determine in large part how we choose to think about the structure of the population and how the individuals and groups within it mix. Are we modeling the long-term effects of intervention? Do we want a model of transmission that allows us to estimate meaningful parameters from the data we collect? Are we interested in understanding social networks?

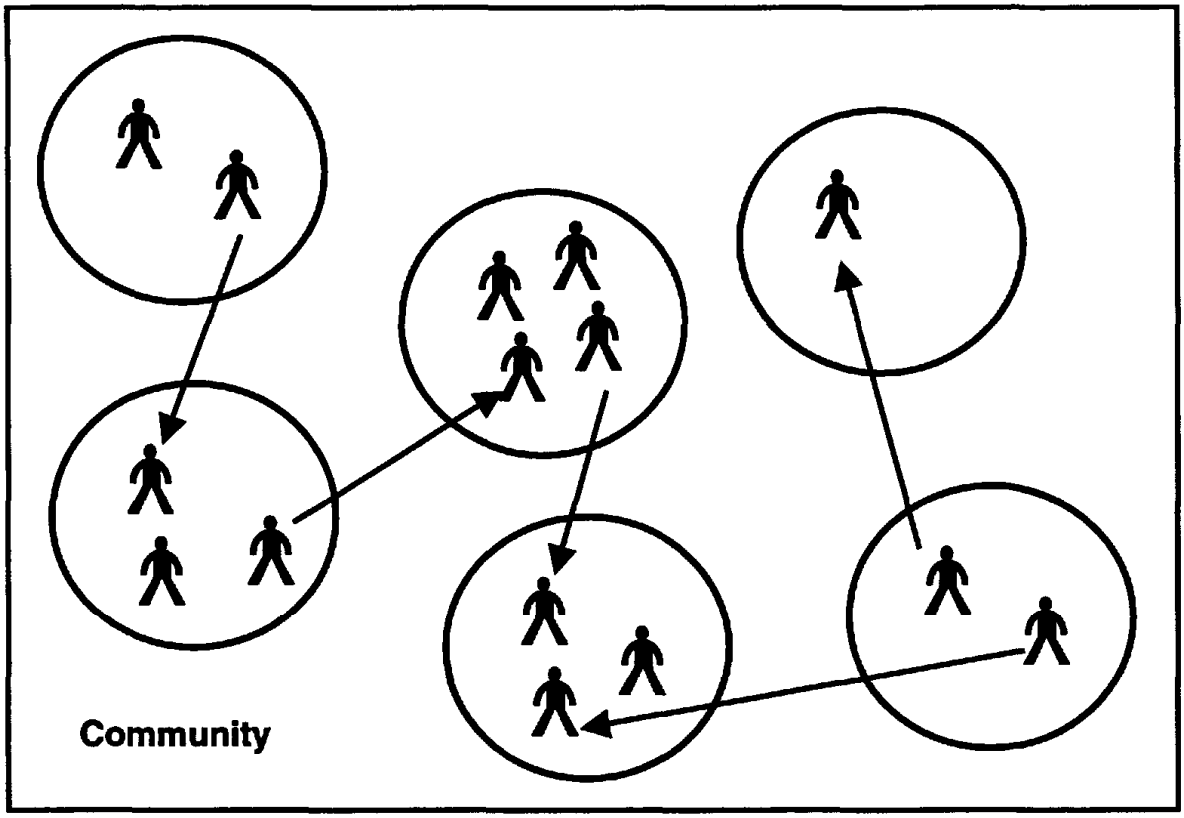
Transmission Units within Populations

So far we have considered two possible mixing patterns. One was random mixing in a large population. The other, in the context of the chain binomial model, was a collection of small populations that mixed randomly within themselves but did not interact with one another (Fig. 4–12A). Now we consider a combination of the two with a larger population being composed of smaller transmission units. Individuals mix with the others in their own transmission units in one way and with members of the community who belong to other small transmission units in a different way. The transmission units could be households, sexual partnerships, schools, workplaces, or day care centers, for example (Fig. 4–12B,C,D).

In the simplest case, an individual belongs to one small transmission unit and interacts randomly with the others in the population. For example, a person may be in a steady but nonmonogamous sexual partnership and so have contact with the partner but also sexual contact with other individuals in the community at large. As another example, a person may have contact in a family

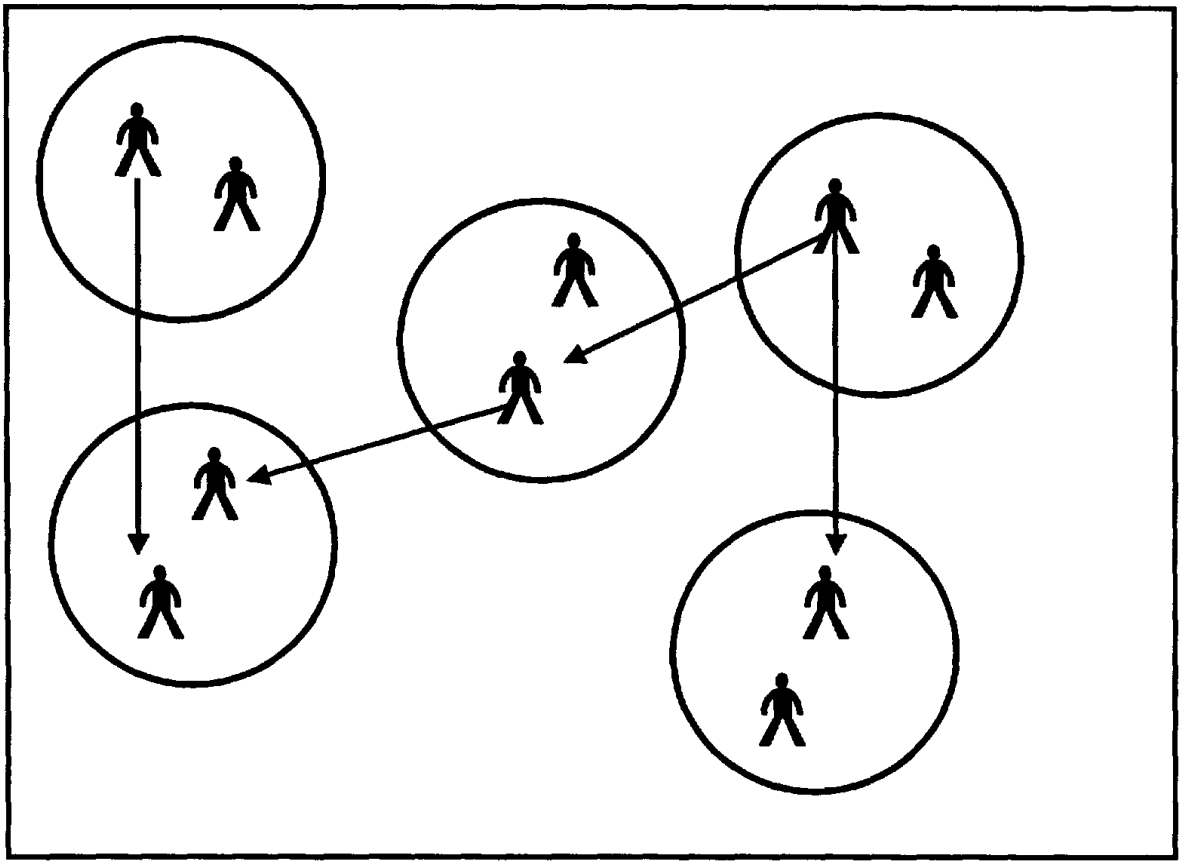


A

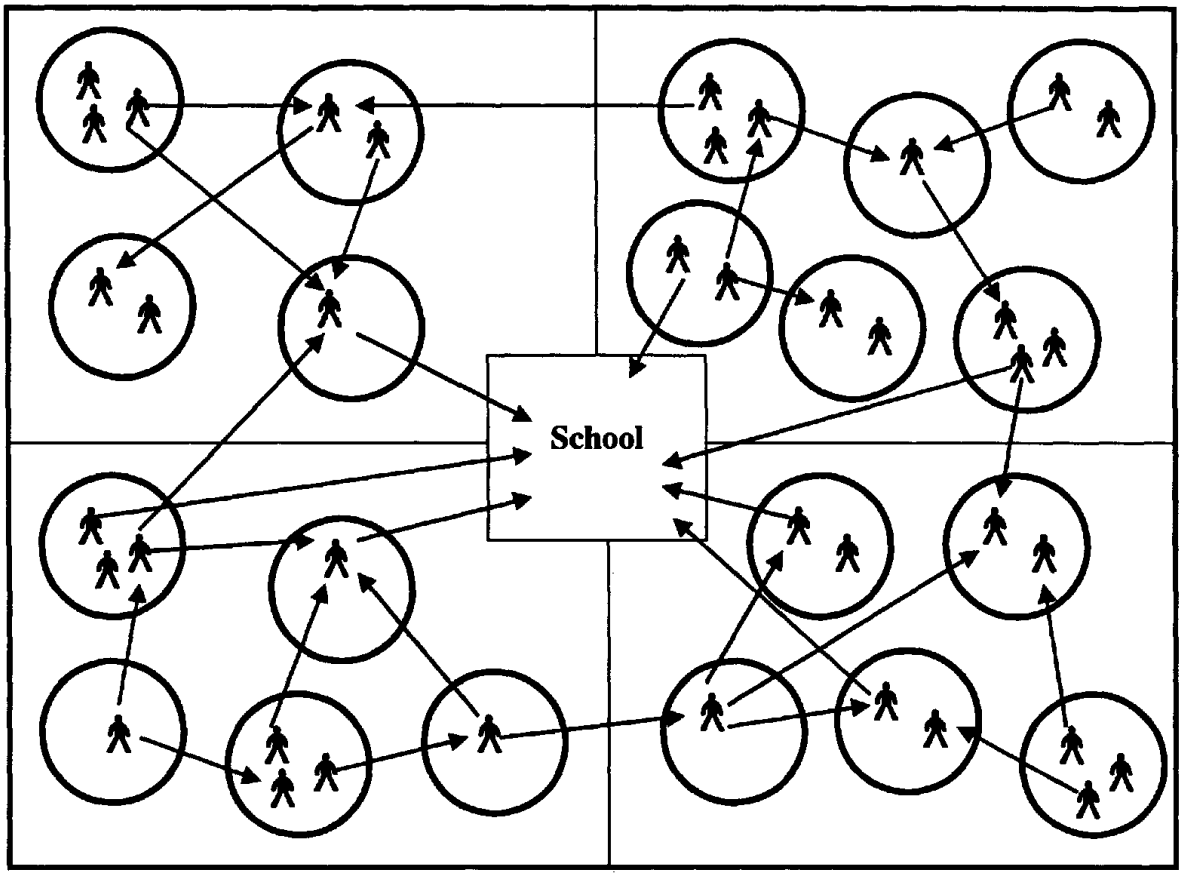


B

Figure 4–12A and B. A: Independent transmission units. B: Transmission units within a community.



C



D

Figure 4–12C and D. C: Nonmonogamous sexual partnerships with contacts in the community. D: Contact patterns in a community of households in four neighborhoods with one school.

household, but also within the community at large. When we define the community structure in this way, 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. Longini and Koopman (1982) formulated a model that contains both the probability of being infected within a household and within the community at large. This model can be used for studying transmission in sexual partnerships by assuming that the households are all of size two. We can allow for the fact that some people do not have steady sexual partners by allowing for singles (households of size one) as well as partnerships. This is the basis of the model for the augmented study design discussed in Chapter 5.

Subpopulations

Rather than small transmission units, we may think of a population as divided into large subgroups that mix more with their own members than with other groups. For example, we may observe two large communities that have little interchange between them. Alternatively, we may divide a population into two differently sexually active subgroups that have some contact with each other.

In a population composed of two mixing groups, group 1 and group 2 (Fig. 4–13), 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 ($i,j = 1,2$) is denoted by c_{ij}. The mixing pattern of two groups is represented by the matrix,

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

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

R_0 will be higher in the group with the higher 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 serve as a *core population* for transmission. Thomas and Tucker (1995) have reviewed this and other concepts of core groups for sexually transmitted diseases. 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.

Hethcote and York (1984) examined different strategies for reducing gonorrhea, taking into account sex workers who acted as a core group and their contacts within the general population. They found that an intervention program generally needs to be targeted at the subpopulation with the higher R_0, in this case, the core population of sex

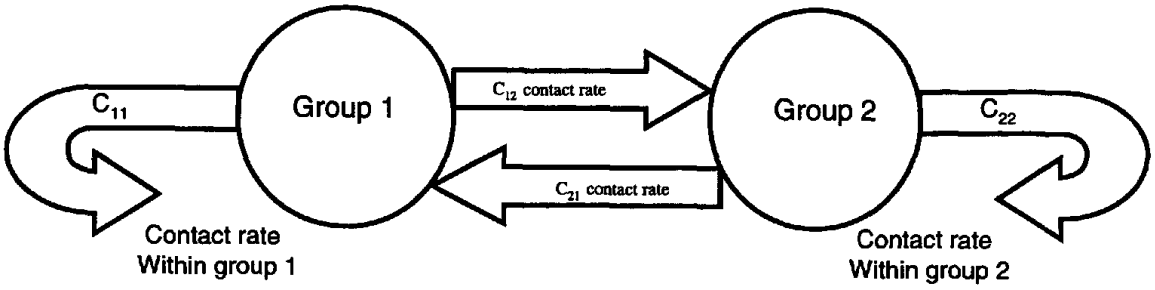


Figure 4–13. Mixing pattern of two groups in a population.

workers, to have most effect. In general, when planning interventions in situations with heterogeneous transmission or levels of infection, targeting therapy or prevention to the groups with the highest transmission or levels of infection is often most effective in reducing infection in the population at large. In the section, R_0 in Macroparasitic Diseases, we mentioned that often a small proportion of the population has a very heavy parasite burden. This is another example in which targeting the therapy enhances the effectiveness of the intervention.

Another approach to describing contact patterns is *social networks* of people (Morris and Kretzschmar 1997, Koopman et al. 2000). In this approach contacts are made through pair-formation and dissociation processes. The approach also allows that several people can be in contact with each other during an interval of time, and then dissociate. The transmission patterns produced by this *concurrency* of contacts and the resulting networks can be compared to transmission in which all contacts are sequential. In general, epidemic spread is more rapid if several people can make contact simultaneously.

Like much else in infectious diseases, contact patterns are often difficult to determine and usually are not measured. When conducting studies in infectious diseases where transmission plays a role, it is important to formulate explicitly the underlying assumptions that are being made with respect to contact patterns and exposure to infection. Since groups with different contact rates and mixing patterns could have different exposure to infection, consideration of the contact patterns could be important for interpreting measures of effect. Failure to take into account unequal exposure to infection in the groups being compared can produce biased estimates of effect. In Chapter 5, we demonstrate how differential contact rates can affect estimates of effect.

SUMMARY

Several principles of transmission and dynamics are common to many infectious diseases. These include the transmission probability, the basic reproductive number, conditions for an epidemic, and the role of contact and mixing patterns. The transmission probability is a measure of the ability of a microbe to spread from an infected to a susceptible host during a contact. The binomial model of transmission is widely used to quantify transmission concepts and to estimate the transmission probability. The basic reproductive number, R_0, describes the potential of a microbe to spread in a population. Dynamic models are used to understand the spread of infection and the role of interventions over time. Assumptions about mixing and contact patterns influence the interpretation of epidemiologic studies in infectious diseases.

Dr. Halloran was partially supported in writing this chapter by NIH grants R01-AI32042 and R01-AI40846. The mosquito image was used with permission from the American Museum of Natural History.

REFERENCES

Abbey H. An examination of the Reed-Frost theory of epidemics. Hum. Biol. 24: 201–233, 1952.

Anderson RM, and May RM, eds. Population Biology of Infectious Diseases. Berlin: Springer-Verlag, 1982.

Anderson RM, and May RM. Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press, 1991.

Antia R, Koella JC, and Perrot V. Models of the within-host dynamics of persistent mycobacterial infections. Proc. R. Soc. Lond. B. 263: 257–263, 1996.

Antia R, Pilyugin SS, and Ahmed R. Models of immune memory: on the role of cross-reactive stimulation, competition, and homeostasis in maintaining immune memory. Proc. Natl. Acad. Sci. USA. 95: 14926–31, 1998.

Aron JL, and May RM. The population dynamics of malaria. In: Anderson RM, ed. Population Dynamics of Infectious Diseases: Theory and Application. London: Chapman and Hall, 1982: 139–179.

Bailey NTJ. The Mathematical Theory of Epidemics. London: Griffin, 1957.

Becker, NG. Analysis of Infectious Disease Data. London: Chapman and Hall, 1989.

Burnet M, and White DO. Natural History of Infectious Disease. 4th ed. Cambridge: Cambridge University Press, 1972.

Chiang CL. An Introduction to Stochastic Processes and Their Applications. Huntington, NY: Robert E. Krieger, 1980.

Dietz K, Molineaux L, and Thomas A. A malaria model tested in the African savannah. Bull. WHO. 50: 347–357, 1974.

Elveback LR, Fox JP, Ackerman E, et al. An influenza simulation model for immunization studies. Am. J. Epidemiol. 103: 152–165, 1976.

Fine P. Herd immunity. Epidemiol. Rev. 15: 265–302, 1993.

Golm GT, Halloran ME, and Longini IM. Semiparametric methods for multiple exposure mismeasurement and a bivariate outcome in HIV vaccine trials. Biometrics. 55: 94–101, 1999.

Greenwood M. On the statistical measure of infectiousness. J. Hyg. Camb. 31: 336–351, 1931.

Halloran ME, Struchiner CJ, and Spielman A. Modeling malaria vaccines II: population effects of stage-specific malaria vaccines dependent on natural boosting. Math. Biosci. 94: 115–149, 1989.

Halloran ME, Haber MJ, Longini IM, and Struchiner CJ. Direct and indirect effects in vaccine field efficacy and effectiveness. Am. J. Epidemiol. 133: 323–331, 1991.

Halloran ME, Haber MJ, and Longini IM. Interpretation and estimation of vaccine efficacy under heterogeneity. Am. J. Epidemiol. 136: 328–343, 1992.

Halloran ME, Cochi S, Lieu T, Wharton M, and Fehrs LJ. Theoretical epidemiologic and morbidity effects of routine immunization of preschool children with live-virus varicella vaccine in the U.S. Am. J. Epidemiol. 140: 81–104, 1994.

Hau CH, Hien TT, Tien NTK, et al. Prevalence of enteric hepatitis A and E viruses in the Mekong River Delta region of Vietnam. Am. J. Trop. Med. Hyg. 60: 277–280, 1999.

Hethcote HW, and Yorke JA. Gonorrhea transmission dynamics and control. Lecture Notes in Mathematics 56. Berlin: Springer-Verlag, 1984.

Knox EG. Strategy for rubella vaccination. Int. J. Epidemiol. 9: 13–23, 1980.

Koopman JS, Chick SE, Riolo CS, et al. Modeling infection transmission through networks in geographic and social space using the GERMS framework. J. STD. (in press) 2000.

Longini IM, and Koopman JS. Household and community transmission parameters from final distributions of infections in households. Biometrics. 38(1): 115–126, 1982.

Longini IM, Hudgens MG, Halloran ME, and Sagatelian K. A Markov model for measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines. Stat. Med. 18: 53–68, 1999.

Macdonald G. The Epidemiology and Control of Malaria. London: Oxford University Press, 1957.

McNeill WH. Plagues and Peoples. New York: Doubleday, 1976.

Morris M, and Kretzschmar M. Concurrent partnerships and the spread of HIV. AIDS. 11: 641–648, 1997.

Pilyugin S, Mittler J, and Antia R. Modeling T-cell proliferation: an investiation of the consequences of the Hayflick Limit. J. Theor. Biol. 186: 117–129, 1997.

Ross R. The Prevention of Malaria. 2nd ed. London: John Murray, 1911.

Ross R. An application of the theory of probabilities to the study of a priori pathometry, Part 1. Proc. R. Soc. Series A. 92: 204–230, 1916.

Schenzle D. An aged-structured model of pre- and post-vaccination measles transmission. IMA J. Math. Appl. Med. Biol. 1: 169–191, 1984.

Smith PG, Rodrigues LC, and Fine PEM. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. Int. J. Epidemiol. 13(1): 87–93, 1984.

Struchiner CJ, Halloran ME, and Spielman A. Modeling malaria vaccines I: new uses for old ideas. Math. Biosci. 94: 87–113, 1989.

Thomas JC, and Tucker M. The development and use of the concept of a sexually transmitted disease core. J. Infect. Dis. 174 (Suppl): S134–43, 1995.

Thomas JC, Eng E, Clark M, Robinson J, and Blumenthal C. Lay health advisors: sexually transmitted disease prevention through community involvement. Am. J. Public Health. 88: 1252–1253, 1998.

Thomas JC, Lansky A, Weiner DH, Earp JA, and Schoenbach VJ. Behaviors facilitating sexual transmission of HIV and STDs in a rural community. AIDS and Behavior. 3: 257–268, 1999.

Ukkonen P, and von Bonsdorff C-H. Rubella immunity and morbidity: effects of vaccination in Finland. Scand. J. Infect. Dis. 20: 255–259, 1988.