CHAPTER FIVE

Basic Epidemiological Concepts in a Spatial Context

Elizabeth Eli Holmes

INTRODUCTION

It is obvious to any observer of epidemics that the spread of disease is unavoidably spatial. Disease moves from individual to individual following the network of contacts between individuals within a population. For many host-pathogen systems, the pattern of spread is a combination of local transmission out from a focus of infection and long-distance transmission, which establishes new foci. Yet most classical epidemiological theory glosses over the spatial dimension of disease transmission and instead assumes that every individual is equally likely to contact every other. A key question is to what extent do we lose insight or are quantitatively misled by modeling the intrinsically spatial process of disease spread with nonspatial theory. Obviously, spatial models are necessary to address spatial questions, such as the velocity at which disease spreads over a landscape or the spatial pattern of disease prevalence (see Murray 1990 and Cliff et al. 1981 for reviews). However, many of the most basic and important epidemiological questions are not spatial: Will a pathogen cause an epidemic? Can it invade and persist in a population? What fraction of the population will be infected? When will epidemics occur and reoccur?

Early epidemiological theory addressed these sorts of questions using simple nonspatial models of communicable disease such as the Kermack-McKendrick model (Kermack and

McKendrick 1927):

$$\frac{dS}{dt} = -\beta SI - bS + b$$

$$\frac{dI}{dt} = \beta SI - \mu I - bI$$

$$\frac{dR}{dt} = \mu I - bR. \qquad (5.1)$$

These models divide the population into susceptible, infected, and resistant (i.e., immune) classes, S, I, and R. All individuals are equally likely to contact every other individual in the population, and there are no differences among individuals except for their infection status. Individuals recover and become immune at rate μ and die at rate b. The population is assumed to be stable so that all individuals that die are replaced by births of new susceptible individuals. Susceptible individuals become infected at a rate βI that is simply proportional to the fraction infected.

From such simple nonspatial theory come some of our most fundamental principles about the dynamics of disease within populations, and although modern epidemiological models are far more realistic and introduce a variety of complexities (see Anderson and May 1991 for a review), these basic concepts are still used to understand and think about real disease systems. The purpose of this chapter is to discuss how fundamental epidemiological concepts change, qualitatively or quantitatively, when we move from a nonspatial to a spatial model of communicable disease.

Basic Epidemiological Concepts

1. The reproductive rate of a disease. The basic reproductive rate of a disease, R_0, is the expected number of new infections caused by one infected individual in a sea of susceptible individuals (MacDonald 1957).

2. Deterministic threshold theorem. Kermack and McKendrick (1927) showed that in a simple model with no influx of susceptible individuals, there exists a threshold density of susceptibles below which an epidemic cannot occur.

3. Stochastic threshold theorem. The Kermack-McKendrick model assumes that the population is large enough to be considered infinite and that densities can be considered continuous variables. Bailey (1975) studied stochastic versions of the Kermack-McKendrick model to explore the dynamics of disease in finite communities and showed that disease cannot become established in a population unless the size of the community is above the stochastic threshold.

4. Threshold for a disease to become endemic. The previous two points consider the dynamics of an epidemic. In this case, natural mortality, b, is negligible. If we consider instead the reintroduction of susceptible individuals due to births, then disease may persist within the population and become endemic. Analogous to the epidemic threshold, there exists an endemic threshold for the disease to establish and persist.

5. Equilibrium levels of disease. Kermack and McKendrick (1932) showed that Equation 5.1 predicts stable equilibrium levels of susceptible, infected, and resistant individuals. These levels are a function of the rates μ, β, and b.

6. Periodicity of epidemics. The long-term records of infectious diseases provide some striking examples of cyclic predator-prey dynamics. The Kermack-McKendrick model offers an explanation for these cycles as an intrinsic product of the interaction between hosts and pathogens (Soper 1929; Anderson and May 1983).

SPATIAL VERSUS NONSPATIAL

How does the spatial dimension change these basic concepts? To tackle this question, I discuss what two types of spatial models predict for concepts (1) through (6) above. The focus of the discussion will be on a cellular automata version of the Kermack-McKendrick model. This model is individual based and explicitly captures the notion of a spatial network of contacts between individuals. However, discussion of only one

type of model leaves one blind to the degree to which the results depend on a particular model structure, and much insight into results from one type of model is gained by comparing results with other types of models. Thus, in the concluding sections, the results for the cellular automata model are compared to a partial differential equation model that also incorporates local transmission.

Kermack-McKendrick Model

Before introducing the cellular automata model, let me briefly discuss the classical Kermack-McKendrick model (Eq. 5.1). The key feature of this model is the idea that disease transmission is described by the βIS term, which is known as the mass-action assumption. In this case, the rate at which susceptibles become infected is directly proportional to the fraction infected; doubling the number infected doubles the rate of disease transmission. For this model, the reproductive rate of the disease is

$$R_0 = \frac{\beta}{\mu + b}.$$

R_0 is analogous to the intrinsic rate of increase, r, in simple population growth models. R_0 is not generally observed because it is the maximum possible growth rate, which occurs at the very beginning of an epidemic when the population is 100% susceptible. See Bailey (1975) for a modern discussion of this model and Serfling (1952) for a historical review.

Cellular Automata Model of an Infectious Disease

In the basic lattice model, sites are distributed on a square lattice on which each site has a set of physically neighboring sites. Each site represents an individual that can have one of three states: susceptible, infected, or resistant. Disease transmission is modeled as a probabilistic process. Each infected site has an equal and independent probability, q, of infecting a susceptible neighbor. Thus the probability that a susceptible

site becomes infected is

1 − (the probability of not being infected)

$$= 1 - (1 - q)^{\text{number of infected neighbors}}. \quad (5.2)$$

At each time step, a site changes state based on the probabilities:

$S \to I$	$1 - (1 - q)^{\text{number of infected neighbors}}$
$I \to R$	μ
sites die and are reborn susceptible	b.

In the simulations discussed here, three different neighborhoods are compared: the four directly adjacent sites, the eight nearest neighbors, and the twenty-four nearest neighbors. Sites are updated synchronously, meaning time is discrete. Discrete time can affect the dynamics of cellular automata (Ingerson and Buvel 1984; Nowak, Bonhoefer, and May 1994), and to minimize these effects, the simulations were run with small transition probabilities to approximate continuous time.

In this discussion, the model with local transmission is often compared to the analogous model with global transmission. In the global model, the "neighborhood" is the entire population, and every infected site is equally likely to infect any susceptible site in the entire population. The rate of disease transmission is: (the proportion susceptible) × (the probability that a susceptible site becomes infected), that is,

$$S\left[1 - (1 - q_g)^{IN}\right] \approx (q_g N)IS$$

for q_g small. The parameter q_g is the probability that an infected individual infects a susceptible neighbor for the global model. Since transmission can be described as a linear function of IS for q_g small, the cellular automata model with global transmission is the analogue to the Kermack-McKendrick model with $\beta = q_g N$.

The definition of the disease reproductive rate in the cellular automata model is analogous to the definition for the Kermack-McKendrick model. It is the mean infectious period times the rate at which new infections are caused by an infected site that is surrounded by susceptible sites. Thus,

$$R_0 = \frac{qN_n}{\mu + b}$$

where N_n is the number of neighbors (4, 8, 24, or N). When a simulation of the model with local transmission was compared to one with global transmission, R_0 was kept identical in the simulations, but the new infections were distributed either locally or globally. Specifically, and to make the R_0's equivalent,

$$q_n = q_g \frac{\text{population size}}{\text{number of neighbors}} \tag{5.3}$$

where q_n and q_g are the transmission probabilities for the local and global models.

EPIDEMIOLOGICAL PRINCIPLES IN A SPATIAL CONTEXT

Epidemic Threshold

The Kermack-McKendrick threshold for an epidemic to occur in a population is $R_0 > 1/S$. The Kermack-McKendrick threshold is the motivation behind the concept of "herd immunity." This idea says that to protect a population from a disease, it is not necessary to vaccinate the entire population; it is enough to vaccinate a proportion $(1 - 1/R_0)$. One of the basic results from the cellular automata model with local transmission is that the Kermack-McKendrick threshold is overly conservative. In the cellular automata model with local transmission, each infected individual interacts with a relatively small neighborhood. From the infected individual's perspective, its "world" quickly fills with other infected individuals,

and the rate at which it causes new infections rapidly declines. To overcome this severe depletion of the local susceptible pool, the disease must have a higher R_0 in order to cause an epidemic (e.g., $R_0 = 1.29$ for four neighbors and $R_0 = 1.13$ for eight neighbors). This means that when one takes into account localized transmission, one needs to vaccinate a smaller proportion of the population. As an aside, the stochastic nature of transmission is often much more important than the local nature of transmission if only a few infected individuals are introduced into a susceptible population. In a stochastic model, there is some probability that the disease will go extinct by chance alone even though $R_0 > 1$. In a stochastic lattice model, this probability of chance extinction is $(1/R_0)^a$ where a is the initial number infected (Whittle 1955; Kendall 1965). When $a = 1$ and R_0 is small, this probability is quite high.

The traditional Kermack-McKendrick threshold described above is the threshold for the disease to increase when it enters a population. However, spatial models of disease introduce a new type of the epidemic threshold, which is the threshold reproductive rate for a pandemic. A pandemic is an epidemic that affects the entire population. The Black Plague in the 1300s and 1500s, the flu epidemic of the early 1920s, and the smallpox epidemics that decimated the Americas are real-world examples of pandemics that spread across entire continents. The cellular automata model predicts that there is a threshold R_0 for a pandemic to occur. At the reproductive rates near the threshold for a small innoculum of infection to increase in a population, the disease will begin to spread from the initial site of infection, but then die out. With a higher basic reproductive rate, the disease will spread farther and farther until, with a high enough R_0, it is able to spread throughout a very large population (Figure 5.1). The pandemic threshold for the cellular automata model with transmission to the four nearest neighbors is $4 < R_0 < 6$ (Kuulasmaa 1982). If R_0 is below this, the disease causes a small, spatially limited epidemic. If it is above, it can spread throughout even an infinite population. The probability of a pandemic also increases as R_0 increases above the pandemic threshold. From analogy with similar models (Ball 1983; Bramson, Durrett, and Swindle 1989) the

pandemic threshold depends on the size of the contact neighborhood and should tend to $R_0 = 1.0$ as the neighborhood size increases.

Threshold Community Size

The stochastic threshold theorem states that there is a threshold community size for a pandemic to occur in a population. This idea stems from work on a model equivalent to the cellular automata model with global transmission. In this model, smaller population size leads to an increased probability that the incipient epidemic will go extinct in the early stages merely by chance. If we assume that there are no births and no loss of immunity, the threshold population size is $N_t = \mu/\beta$ (Bailey 1975). Below this, the epidemic dies out with certainty, and above the threshold, pandemics occur with probability $1 - N/N_t$ where N is population size.

There is no real equivalent of this idea in the cellular automata model with local transmission. From the infected individual's perspective, the world is the same whether it is in an infinite population or in a population barely larger than its neighborhood size. Instead of a threshold community size, however, there exists a threshold neighborhood size below which the disease dies out. The existence of this threshold can be surmised by noting that the pandemic threshold is $4 < R_0 < 6$ when spread is to the nearest four neighbors, and that as the neighborhood size increases, the threshold approaches unity (see the previous section). In simulated epidemics that are started from one infected individual (Figure 5.1), the spatial size of the epidemic increases dramatically as the neighborhood size increases (note that R_0 is held constant between simulations).

FIGURE 5.1. Epidemics with increasing neighborhood size. The patterns show the distribution of susceptible (black regions) and immune individuals (gray regions) at the end of an epidemic. Each epidemic was started with one infected individual at the center of the grid. In each case, $R_0 = qN_x/\mu = 2.67$. The contact neighborhood is shown under each simulation (○ = neighbor of the x individual). $\mu = 0.15$, $b = 0$.

Equilibrium Levels of Disease

The previous discussion concerned growth and spread of a disease within a population. A disease can also persist at some basal level within a population, in which case it is said to be endemic. The Kermack-McKendrick model and the cellular automata model with global transmission predict that if a disease is able to cause an epidemic, then given sufficient time, the disease will become endemic at the equilibrium levels of

$$\hat{S} = \frac{\mu + b}{\beta} = \frac{1}{R_0}$$

$$\hat{I} = \frac{b}{\beta}(R_0 - 1)$$

$$\hat{R} = \frac{\mu}{\beta}(R_0 - 1). \tag{5.4}$$

Simulations of the cellular automata model with global versus local disease transmission show that, in general, localized transmission reduces the equilibrium level of infection within the population. Correspondingly, the equilibrium levels of susceptibility are higher. The lower equilibrium level of infection can be explained by the rapid depletion of the local susceptible pool which effectively reduces the transmission of the disease and creates isolated pockets of susceptibility (Figure 5.2). In a simulation with global transmission, infection is uniformly distributed, and such pockets of susceptibility do not occur.

The degree to which localized transmission decreases the infection levels, however, depends greatly on the reproductive rate of the disease and the length of the infectious period relative to an individual's lifespan. In Figure 5.3, the absolute difference in the susceptible equilibrium fraction observed in simulations of the local versus global model is shown as a function of R_0 and the ratio of lifespan to infectious period (μ/b). When the disease is chronic (when the infectious period is long relative to lifespan), the model with local transmis-

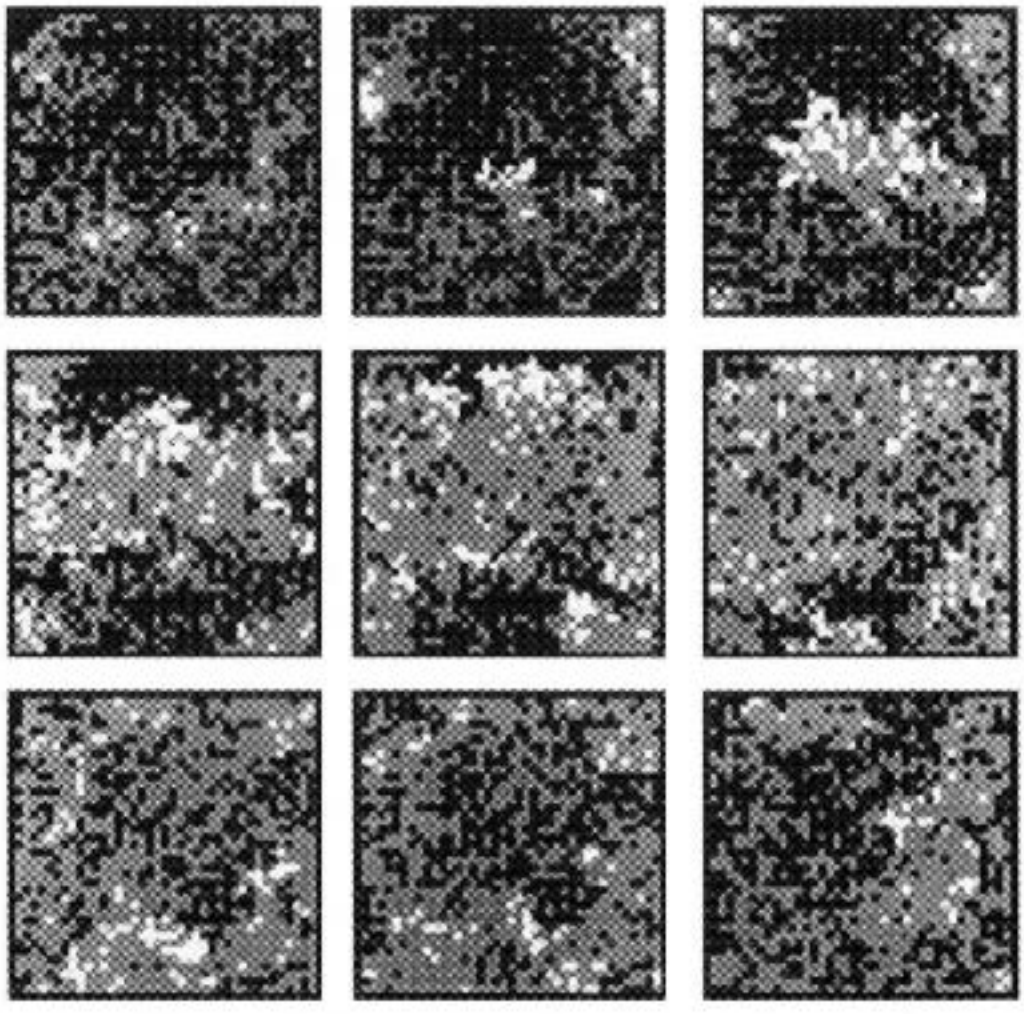


FIGURE 5.2. Endemic disease patterns. Sequential snapshots are at forty-cycle intervals. Black is susceptible; white is infected; gray is immune. The sequence begins one thousand cycles after an epidemic started with 50% infected and 50% susceptible in a random pattern. $R_0 = 4$, $\mu = 0.1$, $b = 0.01$ on a 40×40 grid with eight neighbors.

sion predicts the equilibrium levels of susceptibility are similar to those predicted by the simple nonspatial model. Figure 5.3 shows results when the neighborhood is the four nearest neighbors. The differences between the local and global models, however, shrink rapidly for larger neighborhood sizes. Localized transmission most strikingly reduces the level of infection in a population when the disease is acute, the disease reproductive rate is low, and the contact neighborhood is small.

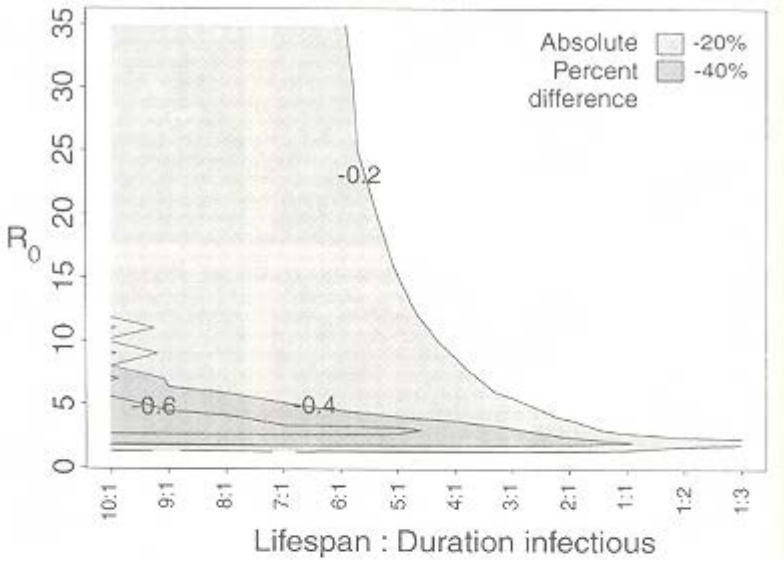


FIGURE 5.3. Equilibrium level of susceptibility in the cellular automata simulations with global versus local transmission (four neighbors). The shaded regions show the parameter space in which local transmission leads to a significantly higher equilibrium fraction of susceptible individuals. The absolute difference is calculated as the (level susceptible in the global model) − (level susceptible in the local model).

The Threshold for Endemicity

Localization transmission not only decreases the prevalence of disease in the population, but it also introduces a new type of endemic threshold. As discussed above, in a model with global transmission, the threshold for the disease to become endemic is $R_0 > 1.0$. It is the same regardless of the lifespan or the duration of the infectious period. When transmission is localized, in contrast, the ability of the disease to persist depends greatly on the ratio of the lifespan to the infectious period. If the infectious period is too short, the disease cannot persist even though R_0 is high. For example, in simulations with transmission to the nearest four neighbors on a 400×400 grid, the disease rapidly went extinct when the ratio of the lifespan to the infectious period was greater than four. For the

same parameters in a simulation with global transmission, the disease persisted and stabilized at a 10% infection level. In these simulations, disease extinction could be due to chance extinction when the number of infected sites is small. The probability of stochastic extinction by time t is approximately (Renshaw 1991)

$$\left(\frac{\mu t}{1+\mu t}\right)^{\hat{I}N}$$

where $\hat{I}$ is the equilibrium fraction infected. However, the rapid extinction at large grid sizes suggests that a threshold is being crossed in the simulations.

Analytical results from work on contact processes also imply that a new threshold exists. Loosely, one can argue as follows. If the rate at which individuals die is high compared to the rate at which they lose infection ($b \gg \mu$), then we have approximately a simple contact process in which

$S \rightarrow I$	from contact with neighboring infecteds
$I \rightarrow S$	at a constant rate, b, due to death of an infected and birth of a susceptible.

This process has been studied extensively, and the disease becomes endemic if $1.33 < R_0 < 4$ (Durrett and Levin 1994a,b). Numerically, the critical R_0 is $R_c \approx 1.64$ (Durrett 1991). If there are no births ($b = 0$), we have the simple epidemic with no recovery, and as discussed previously the disease dies out on an infinite grid if $R_0 < 4$. With this information, one can conjecture how the critical R_0 for persistence varies with the ratio of the lifespan to the infectious period, μ/b (Figure 5.4). Specifically, the critical R_0 increases as the infectious period shrinks relative to the lifespan. Thus as the lifespan increases, one can pass from a region where the disease can persist into a region where it cannot. Note that such a threshold does not exists when transmission is global; as the lifespan increases, $\hat{I}$ decreases and the probability of stochastic extinction increases, but a threshold is not crossed.

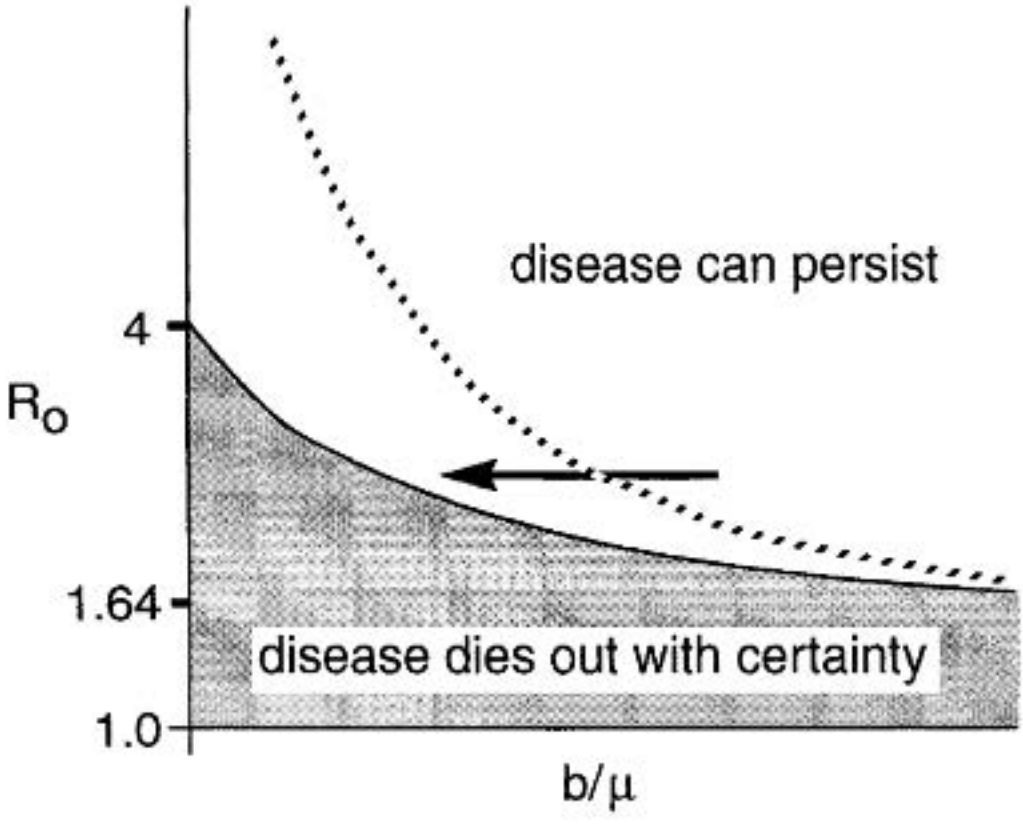


FIGURE 5.4. The critical reproductive rate for endemicity. When the ratio of the infectious period to the lifespan, b/μ, is very large, $R_c \approx 1.64$. When $b/\mu = 0$, R_c is at least greater than four. As b/μ is decreased (keeping R_0 constant), one can pass from a region where the disease can persist through a threshold b/μ to a region where the disease cannot persist. Note that the shaded area is the minimum region where the disease will die out because the critical R_0 at $b/\mu = 0$ is a lower bound. The actual line for the critical R_0 (designated by the dashed line) will lie somewhere above the shaded area.

The Periodicity of Epidemics

The seemingly regular recurrence of epidemics is certainly one of the most striking features from the historical records of directly communicable disease. For example, in large, non-immunized populations, measles epidemics tend to reoccur every one to two years; chickenpox and mumps, every two to four years; diphtheria epidemics, every four to six years; and smallpox, every five years (Anderson and May 1985a). The Kermack-McKendrick model presents a simple mechanism to

explain this periodicity. Analogous to the predator-prey cycles that occur in Lotka-Volterra models, cycles of disease reflect the tendency of the disease to decimate the host and then to decline until susceptibility increases again in the host population. The Kermack-McKendrick model predicts that epidemics (Soper 1929; Anderson and May 1992) reoccur every T years where

$$T \approx 2\pi \sqrt{\frac{L(D_i - D_l)}{R_0 - 1}} . \tag{5.5}$$

L is the mean lifespan $= 1/b$, D_i is the infectious period $= 1/\mu$, and D_l is the latent period. The time between epidemics is determined by how quickly susceptibility returns, reflected by L, and how quickly the disease increases in the population, reflected by $(R_0 - 1)/(D_i + D_l)$. Simulations show that the cellular automata model with global transmission produces epidemic cycles that are analogous to those in the Kermack-McKendrick model. The cycles in both models are damped, meaning their amplitude decreases with time; however, a variety of factors, such as seasonality in transmission rates, demographic stochasticity, age structure, and time lags can cause the cycles to persist (Anderson and May 1992).

To study the effect of local transmission on epidemic cycles, I simulated the cellular automata model with local transmission using neighborhoods of eight and twenty-four (400×400 grid). Damped cycles of epidemics appear in these simulations. As in the models with global transmission, the time between epidemics is negatively related to R_0 and positively related to lifespan. However, in the local transmission model, the time between epidemics is up to three times longer (Figure 5.5).

What causes these differences? Some insight can be gained by considering a plot of new infections per time step versus the product of the fraction infected times the fraction susceptible, IS. In a model with mass-action mixing, for example, the Kermack-McKendrick model or the cellular automata model with global transmission, this plot is linear because the number of new infections per time step is βIS (or $qNIS$ in the cellular

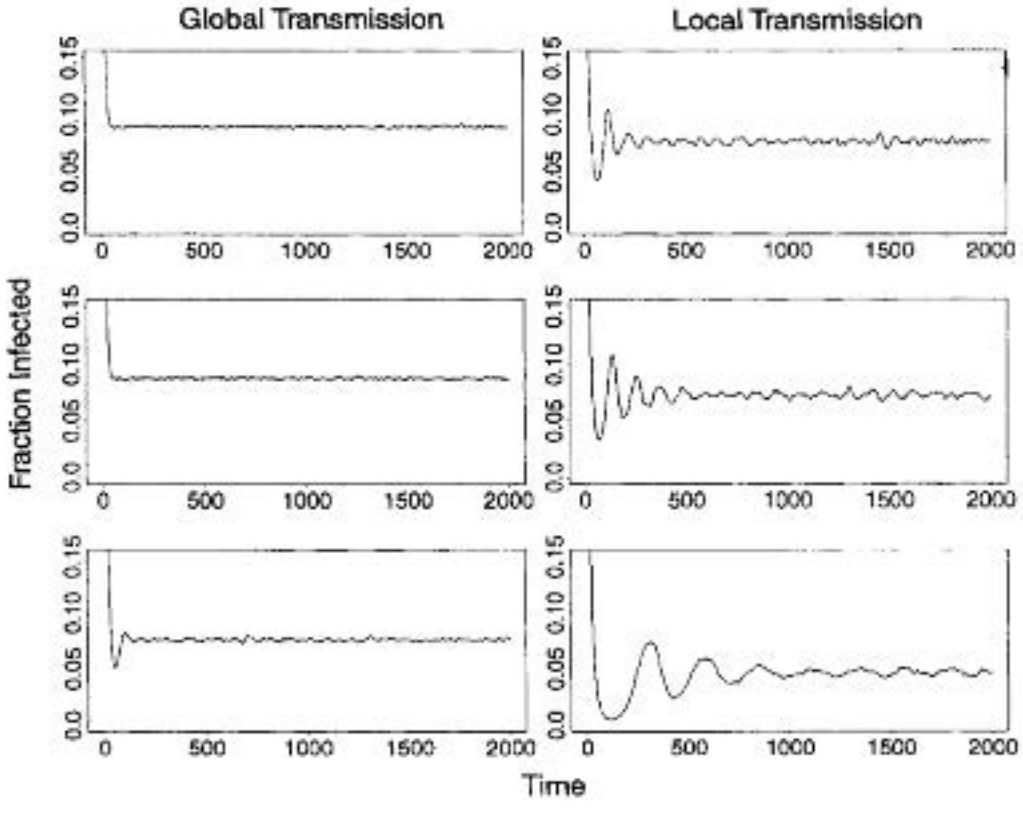


FIGURE 5.5. Epidemic cycles in the global and neighborhood transmission models. Simulations were started with 50% infected and 50% susceptible with infection randomly distributed on the lattice. Top, middle, and bottom rows are $R_0 = 35$, 20, and 5, respectively. $\mu = 0.1$, $b = 0.01$.

automata model). In the cellular automata model with local transmission, disease transmission is not via mass action, and this plot can be highly nonlinear. However, with time, the spatial pattern of infected, immune, and susceptible sites settles into an equilibrium, and at this equilibrium, the number of new infections per time step is approximately a linear function of IS. The slope of this line gives an estimate of the effective transmission rate, β_e which can be much less than the intrinsic transmission rate, $q_n N_n$. All this means is that near equilibrium the cellular automata model with local transmission can be approximated by a simple mass-action model with a new transmission parameter β_e. Indeed, the observed equilibrium fractions of infected, susceptibles, and immune in the local model are those predicted from the global model using β_e, b, and μ (Eq. 5.4). This lower effective transmission rate largely explains

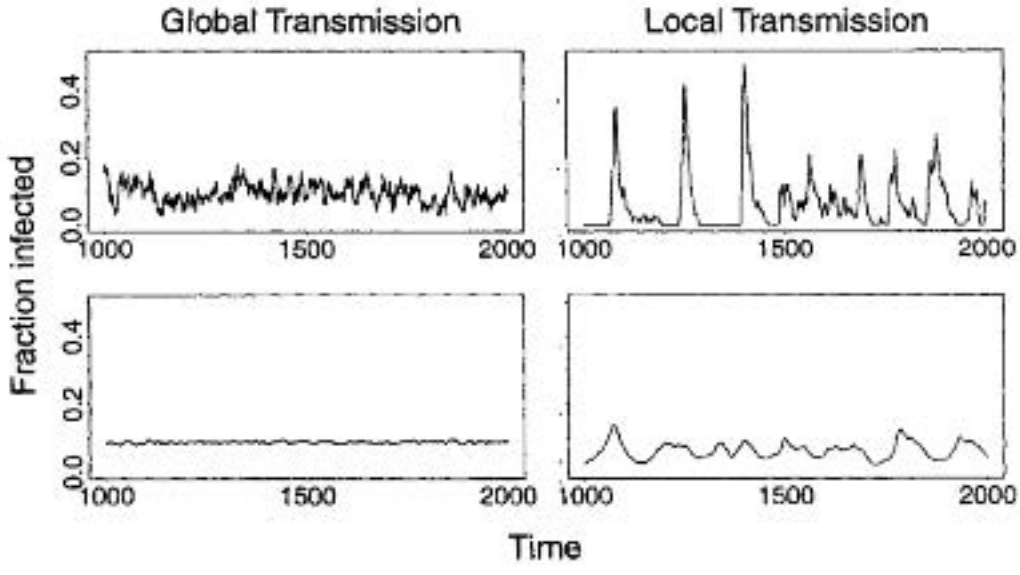


FIGURE 5.6. The infection within a subregion of the population versus within the total population. The top graphs show the fraction infected within a 10×10 subsection of the total grid. The bottom graphs show the fraction infected within the entire grid. On the left is the cellular automata model with global transmission, and on the right is the cellular automata model with a neighborhood of eight. $\mu = 0.1$, $R_0 = 25$, $b = 0.01$, 100×100 grid.

the lengthened periods of epidemic cycles. For example, replacing β by the β_e in Equation 5.5 gives a good prediction for the observed time between cycles in the local model (Figure 5.5).

At the level of the whole population, infection levels cycle smoothly. At the local level, the picture is very different. In the local transmission model, foci of infection move around in space (Figure 5.2). In any one small region, a cycle of epidemic occurs: Immunity increases, and the disease dies out, immune individuals die and are replaced by susceptible individuals, the disease reenters, and a new epidemic occurs. Cycles of epidemics sweep through the local population at irregular intervals, and the interepidemic period is a roughly negative function of R_0 (Figure 5.6). However, these cycles do not necessarily correspond to the cycles at the population level, and unlike the epidemic cycles at the population level, they do not damp out with time. Such stochastic local epidemics have been used to explain the irregularly occurring epidemics of measles in collections of small communities (Bartlett 1955).

The Character of Epidemics

The classic bell-shaped form of an epidemic is typical of diseases such as smallpox, measles, chickenpox, and rubella that have a high reproductive rate and short infectious period relative to the immune period (large μ/b). To look at the effect of local transmission on the size and duration of an epidemic, I compared epidemics in the cellular automata model using neighborhoods of four, eight, or twenty-four with those in the cellular automata model with global transmission. The reproductive rate of the disease was identical between comparisons. The models were compared over a range of R_0 and initial fraction infected. The epidemic started with one infected individual in the center of the grid, and there were no births or deaths ($b = 0$).

In general, an epidemic in the cellular automata model with local transmission sputtered slowly along compared to an epidemic in the model with global transmission (Figure 5.7), unless R_0 was well below the pandemic threshold. In this case, the epidemic in the neighborhood model died quickly. The maximum difference between epidemic durations occurred when the disease reproductive rate was just above the pandemic threshold (Figure 5.8b). At this R_0, the disease began to spread slowly through the entire population. The size of the epidemic was also generally smaller in the model with local transmission; however, the difference depended greatly on R_0 and was at a maximum between the pandemic threshold and $R_0 = 1$ (Figure 5.8a). At $R_0 = 1$, disease dies out quickly in both models, and above the pandemic threshold, the disease may spread slowly but eventually affects the entire population.

COMPARISON WITH OTHER SPATIAL MODELS

Clearly localized disease transmission changes the impact of disease on a population. However, the results discussed above emerge from a cellular automata model. This model is stochastic, individuals come in discrete units (it is not possible to have half an individual), and space is broken up into discrete patches.

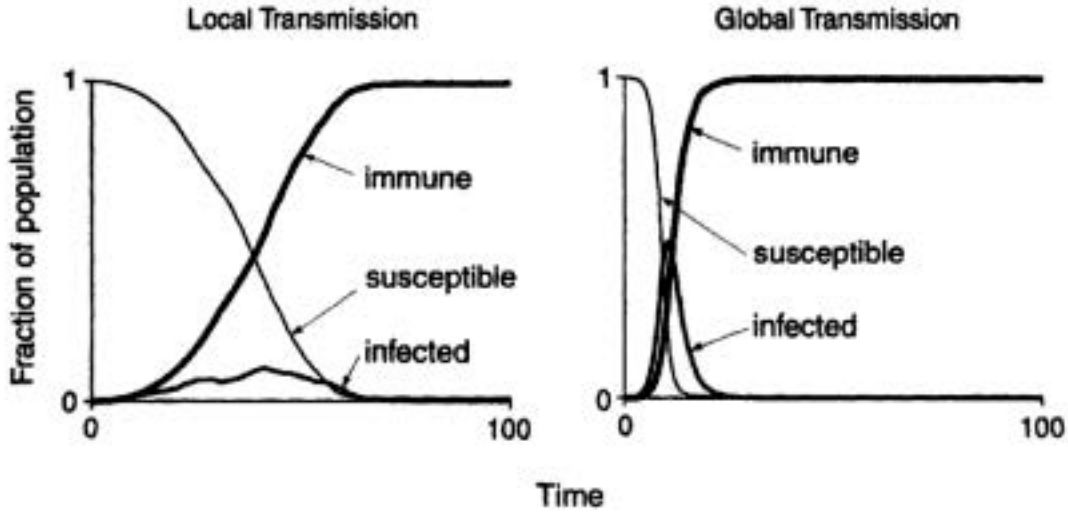


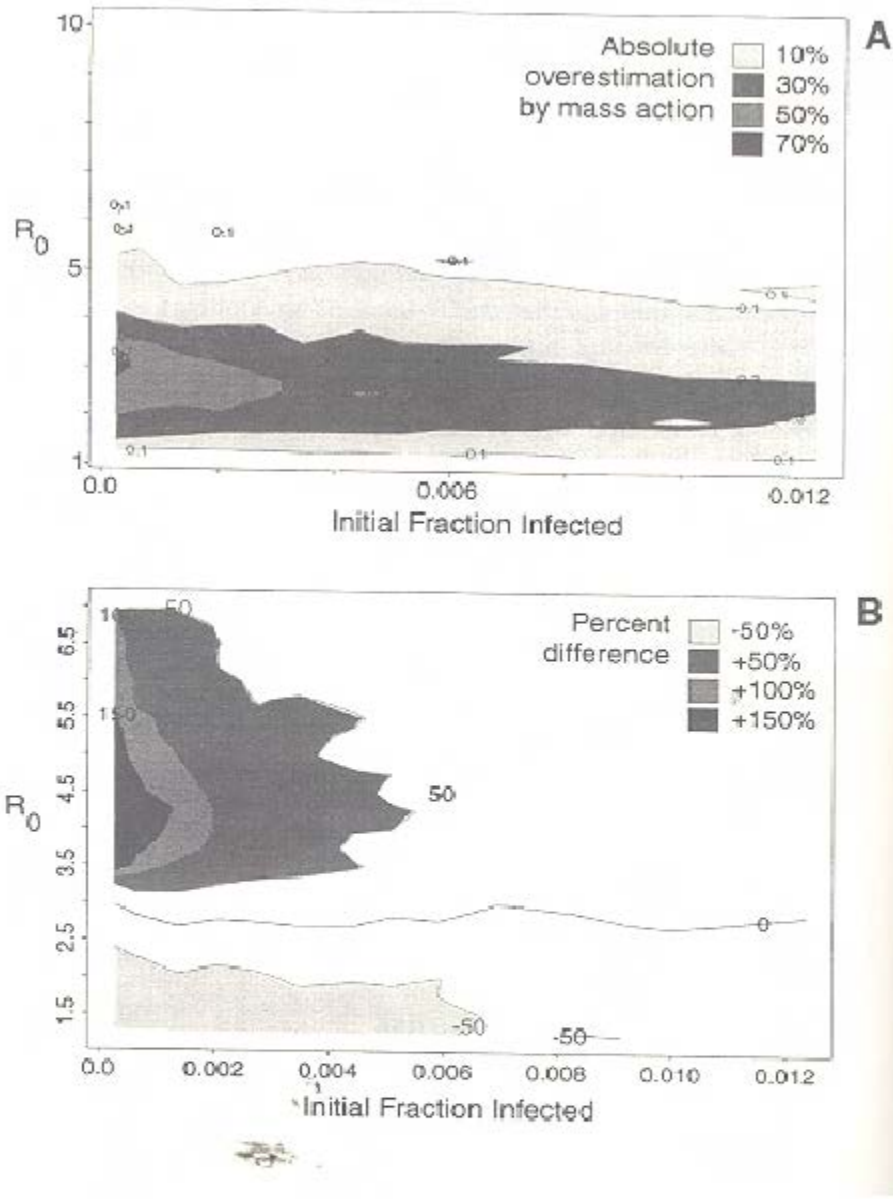
FIGURE 5.7. Epidemics in the cellular automata model with local versus global transmission.

To what extent are the results dependent on this cellular automata framework instead of localized transmission per se? Insight can be gained by looking at a partial differential equation model, the spatial Kermack-McKendrick model, which has localized transmission but a different basic framework. This model is deterministic, allows infinitesimal individuals, and evenly spreads individuals out on a plane rather than divides them among discrete patches.

Spatial Kermack-McKendrick Model

In the spatial Kermack-McKendrick model (see comments by D. G. Kendall in the paper by Bartlett 1957), susceptible individuals become infected by contact with surrounding infected individuals that are within some radius r. Within that radius, transmission occurs through mass-action mixing:

$$\begin{aligned}\frac{\partial S}{\partial t} &= -\beta S\tilde{I} \\ \frac{\partial I}{\partial t} &= \beta S\tilde{I} - \mu I \\ \frac{\partial R}{\partial t} &= \mu I. \end{aligned} \tag{5.6}$$



In this form, it is a partial differential equation model of an epidemic with no recovery; $\partial S/\partial t$ is the change in the proportion of susceptibles at location (x, y). The form of the model is quite similar to the Kermack-McKendrick model (Eq. 5.1) with b set to zero, but the contact rate is given by $\beta S\tilde{I}$ instead of βSI, where $\tilde{I}$ is a weighted function of the number of infected individuals within a radius, r, of a susceptible individual. The weighting is defined such that the total contact rate, $\beta\tilde{I}$, remains constant as the contact radius is reduced. In this way a susceptible individual is not less likely to become infected if they have a small contact radius; they simply contract the disease from a closer neighbor. The effect is to keep R_0 constant while studying the effect of shrinking the contact radius.

Epidemic, Endemic, and Neighborhood Thresholds

In the cellular automata model with local transmission, there are two types of epidemic threshold: the epidemic threshold for a small introduction of pathogen to increase locally before dying out, and the pandemic threshold for a pathogen to spread throughout the entire population. In the cellular automata model, both the epidemic and pandemic thresholds are greater than the simple epidemic threshold, $R_0 > 1$, from the nonspatial Kermack-McKendrick model. In contrast, it can be shown that both the epidemic and pandemic thresholds in the

FIGURE 5.8. Comparison of epidemic size and duration with global versus local transmission. (A) Shaded areas show the absolute difference in the size of the epidemics. Thus, an epidemic that infects 50% of the population in the global and only 20% in the local model gives an absolute difference of 30%. Local transmission causes the greatest decrease in epidemic size when R_0 is between 1.0 and the pandemic threshold. (B) The shaded areas show the parameter space where local transmission lengthens or shortens the duration of an epidemic. When R_0 is small, epidemics are small and die out quickly in the local model. As R_0 increases, the disease begins to spread throughout the population, and the epidemic duration is generally much longer than in the global model. The percentage difference is $100 \times$ (Neighborhood − global duration)/(global duration). Neighborhood equals the four nearest neighbors.

spatial Kermack-McKendrick model are the same, $R_0 > 1$, as for the nonspatial Kermack-McKendrick model. This contrasts greatly with the pandemic threshold ($4 < R_0 < 6$) in the cellular automata model with local transmission.

The difference between the pandemic thresholds in the cellular automata versus the spatial Kermack-McKendrick model is striking and emphasizes the implications of the different assumptions in the two models. In the spatial Kermack-McKendrick model, infinitesimal individuals exist; thus one can have infinitesimal infections produced in an infinitesimal time step, and these infinitesimal infections can then produce infection in the next time step. In the cellular automata model, individuals come in units of one, and transitions occur in discrete jumps. Also, in the spatial Kermack-McKendrick model, the contact rate per susceptible is always a linear function of the proportion infected (rate = $\beta \bar{I}$). In the lattice model with local transmission, the rate saturates (rate = $1 - [1 - q]^{IN}$).

The two models also differ in terms of a threshold neighborhood size. In the cellular automata model, there is a threshold neighborhood size below which the disease dies out and above which it can spread throughout the entire population. This threshold does not exist in the spatial Kermack-McKendrick model. As long as $R_0 > 1$, a pandemic can occur, regardless of the contact radius. The difference between the models emphasizes again the significance of the assumption that individuals come in discrete units.

Finally, the models differ with regard to the endemic threshold for persistence of disease in a population. In the nonspatial Kermack-McKendrick model, this is simply $R_0 > 1$; in the cellular automata model, the threshold is higher, and it depends on the ratio of the lifespan to the infectious period (Figure 5.4). The spatial Kermack-McKendrick model is again like the nonspatial model. The threshold is $R_0 > 1$ and is independent of the ratio of the lifespan to the infectious period. Furthermore, the equilibrium levels of susceptibility, infection, and immunity for the spatial Kermack-McKendrick model with recovery are given by the formulas from the nonspatial model (Eq. 5.4).

The Spatial Pattern of Disease

The cellular automata model produces fascinating, heterogeneous spatial patterns of disease that persist through time (Figure 5.2). The spatial Kermack-McKendrick model, in contrast, can produce traveling waves of infection, but at equilibrium it produces a uniform distribution of disease. For example, after introduction of the disease at one point in space, the pathogen spreads outward in a circular wave but then leaves uniform densities of susceptible, immune, and infected individuals in its wake. If there is no recovery, then the uniform densities will be $(1 - 1/R_0)$ susceptible and $1/R_0$ immune. If there is recovery, the densities will be those given in Equations 5.4.

The difference between the models stems from two basic differences in frameworks of the models: (1) in the cellular automata model, space is broken into discrete patches, and (2) in each patch, there are unstable dynamics. Point (2) comes about because neither susceptibility, immunity, nor infection are a stable endpoint for an individual. Instead individuals go through an inevitable cycle: susceptibility to infection to immunity and back to susceptibility. This combination of patches and unstable dynamics within a patch is the root of the interesting spatial patterns in a wide variety of cellular automata and coupled-lattice models.

To see how the patterns depend on unstable local dynamics, imagine instead that each patch in the cellular automata model represents a subpopulation and that in each subpopulation the dynamics are described by the Kermack-McKendrick model (Eq. 5.1). Like the cellular automata model, this is an epidemiological model in which there is a collection of discrete sites with local transmission between sites. However, now the local dynamics are stable because in an individual subpopulation the densities of susceptible, infected, and immune individuals will go toward the equilibrium values given in Equations 5.4. With patches and stable local dynamics, interesting spatial patterns of infection do not form. Instead, a uniform density at the equilibrium levels is produced (given by Eq. 5.4).

Are Cellular Automata Special?

The earlier discussion of the cellular automata model suggested that localized transmission changes the basic epidemic and endemic thresholds and creates new types of thresholds that are not seen in nonspatial models. At the end of this comparison between the cellular automata and partial differential equation models, however, it may seem that the changes in the fundamental epidemiological concepts have more to do with the specific framework of the cellular automata than with localized transmission per se. However, remember that the framework of the cellular automata is biologically intuitive. Individuals are discrete units. When the unit of a patch is an individual, then space is clearly divided into discrete patches. For the types of communicable diseases discussed here, disease dynamics are unstable at the level of the individual. The take-home message of this comparison between models is not that the results from the cellular automata model are somehow special, but that it is not localized transmission alone that is important. The combination of local transmission, discrete patches, and local unstable dynamics is key.

SUMMARY

One of the general principles to emerge from analyses of cellular automata models is that it is more difficult for a disease to persist when one considers the local nature of disease transmission instead of assuming that the population is like a well-mixed soup in which transmission is effectively global. In nonspatial models, disease persists if the reproductive rate is greater than unity. When transmission is highly localized, the reproductive rate must be much higher both for large epidemics to occur and for a disease to become endemic. Notably, this quantitative difference in the epidemic and endemic threshold does not occur in a partial differential equation model with local transmission. This points out that the combination of local transmission and the discrete individuals is

important when studying the effects of localized transmission. Second, the cellular automata model with local transmission introduces two new types of thresholds. The first is a threshold neighborhood size. In the lattice model, disease dies out quickly if the contact neighborhood is small and the reproductive rate is below some critical level. As the contact neighborhood is increased, the disease will at some point persist and be able to cause a pandemic. The second is a threshold ratio of lifespan to the duration of the infectious period. In the models with global transmission, the disease can persist even if the infectious period is extremely short. In the cellular automata model with local transmission, the disease goes extinct if the infectious period is too short. Finally, in cellular automata models, local transmission greatly lengthens time between epidemics. The interepidemic time can be three to four times longer than in the analogous model with global transmission. Interestingly, the difference can be explained in large part by a lower effective transmission rate in the model with local transmission.

While it is clear that local disease transmission alters many of these basic concepts from classical epidemiological theory, the magnitudes of the changes depend in large part on the parameters of the models. The effects of local transmission can be swamped by high interaction rates between individuals, moderate neighborhood size, or high turnover at the individual level. This begs the question of whether local disease transmission changes disease dynamics in the real world, and answering this question means estimating the model parameters for real diseases. However, the basic Kermack-McKendrick model and the cellular automata models are cartoons that incorporate a few basic aspects of real diseases. Because they are so simple, it is nontrivial to make meaningful associations between model parameters and data collected on real diseases. Here is a simple example of the types of parameter estimation problems that arise: Figure 5.3 shows the change in the impact of disease on a population if one assumes local versus global transmission. The plot axes are the disease reproductive rate and the ratio of the lifespan to the infectious period. Lifespan and infectious period could be estimated for real organisms and

their diseases, but in the model, organisms are assumed to die and lose infection at constant rate. If this is not the case for the real organism and disease (which will often be true), then lifespan and infectious period for the real organisms do not correspond to those parameters in the model. A more fruitful use of these simple models is to develop a general qualitative understanding of the behavior of complex spatial systems. Without such a foundation, the analysis of spatial models that include more biological realism becomes lost in a muddle of spatial and nonspatial effects.

Anderson and May. 1983. Vaccination against rubella and measles: quantitative investigations of different policies. Journal of Hygiene 90:259-325.

Anderson, RM and RM May. 1985a. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmers. Journal of Hygiene of Cambridge 94:365-436.

Anderson and May 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford Univ Press

Anderson and May 1992. Infectious Diseases of Humans: Dynamics and Control. Oxford Univ Press

Anderson, RM and RM May. 1992.

Bailey, NJT. 1975. The Mathematical Theory of Infectious Diseases and Their Application. C. Griffin.

Ball, F. 1983. The threshold behavior of epidemic models. Journal of Applied Probability 20:227-241.

Bartlett, MS. 1955. Stochastic Processes. Cambridge Univ Press.

Bartlett, MS. 1957. Measles periodicity and community size. Journal of the Royal Statistical Society A 120:48-70.

Bramson, M, R Durrett, and G Swindle. 1989. Statistical mechanics of crabgrass. Annals of Probability 17:444-481.

Cliff, A, et al. 1981. Spatial Diffusion. Cambridge University Press.

Durrett, R. 1991. The contact process, 1974-1989. Lectures in Applied Mathematics 27:1-18.

Durrett, R and SA Levin. 1994a. The importance of being discrete (and spatial). Theoretical Population Biology 46:363-394.

Durret, R and SA Levin. 1994b. Stochastic spatial models: a user's guide to ecological applications. Philosophical Transactions of the Royal Society of London B 343:329-350.

Ingerson, TE and RL Buvel. 1984. Structure and asynchronous cellular automata. Physica 10D:59-68.

Kermack William Ogilvy and McKendrick A. G. 1927. A contribution to the mathematical theory of epidemics. Proc. R. Society London A 115. http://doi.org/10.1098/rspa.1927.0118

Kendall, DG. 1965. Mathematical model of the spread of infection. In Mathematics and Computer Science in Biology and Medicine, 213-225. HM Stationery Office.

Kuulasmaa, K. 1982. The spatial general epidemic and locally dependent random graphs. Journal of Applied Probability 19:745-758.

MacDonald, G. 1957. The Epidemiology and Control of Malaria. Oxford Univ Press

Murray, JD. Mathematical Biology. Springer-Verlag

Novak, MA, S Bonhoeffer, and RM May. 1994. More spatial games. International Journal of Bifurcation and Chaos 4:33-56.

Serfling, RE. 1952. Historical review of epidemic theory. Human Biology 24:145-166.

Soper, MA. 1929. The interpretation of periodicity in disease prevalence. Journal of the Royal Statistical Society A 92:34-61.

Whittle, P. 1955. The outcome of a stochastic epidemic—a note on Bailey's paper. Biometrika 42:116-122.