Epidemics in Nonrandomly Mixing Populations: A Simulation

LISA SATTENSPIEL
*Department of Anthropology, University of Michigan,
Ann Arbor, Michigan 48109*

ABSTRACT   Two stochastic, discrete-time simulation models for the spread
of an epidemic through a population are presented. The models explore the
effects of nonrandom mixing within the population and are based on an SIR
epidemic model without vital statistics. They consider a population of preschool
children, some of whom attend child care facilities. Disease transmission oc-
curs both within the home neighborhood and at the child care facility used, if
any. The two models differ in population size used, population density, the
proportions of children using different kinds of care, and the functions used for
calculating the probability of disease transmission. Results are presented for
seven different variables—length of the epidemic in weeks, number of cases,
number of cases in each kind of care (two day care centers, private homes, and
children staying at home), and the number of private home providers affected
by the epidemic. In addition, the distribution of total epidemic size and the
progress of an epidemic are estimated from 25 epidemic trials. The effects of
the location of homes of initial cases, the type of care used by initial cases, and
the density of the population are discussed. Results from the simulation con-
firmed the importance of type of care on the risk for disease transmission.
Results from all runs of the simulation showed that children who attended a
day care center were most likely to become infected, children who went to a
private home were intermediate, and children who did not use any day care
facility were at the lowest risk. The size and length of the epidemics were
related to the presence of the disease in day care centers, regardless of the
location of the initial case, and the time at which the disease entered the
center(s). The simulations also showed that the geographical distribution of
the homes of children attending a particular center was a critical feature
involved in the production of epidemics. The center with more widely distrib-
uted homes of students was less likely to experience a major epidemic than the
center with clustering of student's homes within a neighborhood. This indi-
cates that it is not simply attendance at a day care center that is critical for
disease spread, but that the nature of the population of children attending a
center is also of critical importance in the actual risk for disease spread within
the center. These results are discussed with reference to the spread of hepatitis
A among day care centers in Albuquerque, New Mexico.

In recent years there has been a growing
interest in mathematical models for the
spread of a disease through a population.
Most of the models are variations upon a
single basic theme. The population is divided
into a few classes of individuals according to
disease status. Mixing of individuals occurs
at random with transmission of the disease
being dependent upon the probability of an
infective individual's coming into contact
with one who is susceptible to the disease.

In the simplest model (called the SI model),
the population is divided into only two
classes, susceptibles and infectives. The pop-

ulation size is assumed to be constant, and it is also assumed that there is no migration and that there are no births or deaths in the population. The probability that a susceptible individual will become infected in a given period of time is proportional to the amount of contact between susceptibles and infectives. Once an individual is infected, there is no change in status, i.e., there is no recovery.

A complete solution to this model is given in Bailey (1975). Because an individual remains infective for all time once he becomes infected, the ultimate result of this model is that all individuals in the population become infected. This is not a realistic result. Consequently several modifications have been made to this model. These modifications include allowing recovery with permanent immunity (SIR model), incorporating vital statistics (births, deaths, and migration), assuming a discrete-time process with a binomial distribution for new infections rather than a continuous-time process (chain binomial model), considering the effects of the age structure of the population, considering the effects of spatial clustering, and several other topics.

A major assumption of almost all of the models, including most of the modified versions of the simple model, is that the individuals within the population mix randomly. This is not a general characteristic of most human populations. Usually a large population is divided into several smaller groups, within which mixing is random, but between which there is a limited amount of interaction. This subdivision of the population can be the result of geographic separation of the groups or it can result from structuring of the population along social networks. The patterns of movement of individuals among these groups can profoundly affect the course of the disease through the large population.

Sattenspiel (1984, 1987) developed a mathematical model to describe the spread of a disease through a subdivided population. This model shares a characteristic with all but the simplest of mathematical models for disease spread—that analytical solutions are not feasible. Qualitative methods such as stability analysis may be used, but these usually provide information only about the ultimate effects of an epidemic given long periods of time and near the equilibrium points of the model. The behavior of the disease during the early and middle stages of an epidemic is not easily determined unless an analytic solution of the model is found. Simulation is one method that can be used to study the progress of a disease throughout the course of an epidemic.

Simulation models which involve the spatial structure of populations can be divided into two types. The first type considers a single population within which there is diffusion of infectives. Typically these models consider the population to be distributed at the points of a lattice. All points contain individuals and no points contain more than one individual. Transmission occurs between neighbors; usually the eight nearest neighbors are chosen so that the difference between the probability of transmission to an individual on the boundary of the lattice and the probability of transmission to an individual in the interior of the lattice is minimized. The initial infective is assumed to be at the center of the lattice, and diffusion proceeds outward toward the boundaries.

Bailey (1967) used a chain binomial model to simulate the spread of an epidemic with and without recovery over a lattice. He studied the effects on the epidemic patterns of varying the probability that the disease is transmitted from an infective to a susceptible. Bartlett (1961) used an epidemic model with recovery and births (SIR model with vital statistics) to simulate the spread of measles over a lattice. He studied the time to extinction of the disease and whether there were recurrent outbreaks, as well as the effects of seasonal differences in the probability of infection.

The second type of simulation model for the spread of disease in a spatially structured population considers populations which are divided into several different groups. These groups can be separated on the basis of geographic distance or they can be separated because of social structure. In either case the probability of transmission is influenced by whether two individuals are in the same group or in different groups.

Baroyan et al. (1969, 1971) simulated the spread of influenza among cities in the U.S.S.R. They used an SIR model with movement between cities. Because of the difficulties of obtaining actual movement data they estimated the probability of movement by using a function that was proportional to the product of the sizes of the populations between which movement occurred. The risk of infection of a susceptible was determined by adding the probability of infection by an in-

fective within the same city to the probability of infection by an infective from another city who visited the city of the susceptible individual. The results from the simulation were compared to actual values for influenza epidemics in the U.S.S.R. and appeared to provide good estimates of the spread of the disease through the country.

Cliff et al. (1981) studied the spread of measles over the island of Iceland. In addition to statistical analysis of the actual data from Iceland, they explored a variety of models for the spread of disease, including four models based on time-series analysis and two models based on diffusion. The diffusion models were analyzed for single populations as well as for multiple groups. Results from the various models were then compared to the real data and the usefulness of specific models and of models in general was evaluated.

Ewy et al. (1972) developed a simulation model which included the effects of the social structure on the spread of disease. Factors which were incorporated into the model include family structure, preschool play groups, attendance at schools for older children, age of individuals, and various other social factors. The probabilities of disease transmission were a function of the amount of contact between a susceptible and infectives and were increased if the individuals were members of the same social group. This model was not applied to any specific disease, but it did consider the effects of two interacting infectious agents in addition to a situation with one infectious agent. A similar model was used by Elveback et al. (1971) to study the spread of enteric diseases (which includes hepatitis A) among a heterogeneous community. This model was also primarily concerned with the effects of two interacting agents (specifically live-type polio vaccine and Coxsackie virus). These two models and several other simulation models for the spread of disease among human populations are discussed more fully in Ackerman et al. (1984).

The simulation model described in this paper is similar to the models developed by Elveback et al. (1971) and Ewy et al. (1972) and considers the spread of a disease in a population where the probability of disease transmission between individuals is a function both of the distance between individuals and of whether two individuals attend the same social facility. Specifically, the model considers the spread of a disease among pre-

school children, some of whom attend day care centers. The model is of the second type described above, where the population is divided into several discrete groups. A significant difference between this model and the model of Elveback et al. (1971) is that this model attempts to look at the specific effects on the disease spread of the structuring of the population. The Elveback model incorporates the population structure because of a recognized need to have a greater degree of complexity in order to answer questions about the interaction between two or more infectious agents, but is less concerned with the actual effects of the structuring itself.

The model presented here is applied to a specific disease, hepatitis A, that occurs with relatively high frequency in day care centers. Hepatitis A is a viral disease which produces mild diarrhea and which may produce jaundice and more severe symptoms. Like polio it is an enteric disease with a fecal-oral mode of transmission. In recent years day care centers have become major foci for the transmission of several such diseases. Around 30% of all cases of hepatitis A in some parts of the country are associated with day care centers (Hadler et al., 1980). This increased risk associated with day care centers is thought to arise from two main causes, an increase in the use of centers, especially by diapered children, and the fact that hepatitis A is usually asymptomatic among such children. The former factor provides a necessary degree of contact between susceptible and infective individuals, while the latter factor guarantees that control of the disease within day care centers will be very difficult. Because a large proportion of the infected young children do not show symptoms, they tend not to be isolated from other individuals. The disease is not detected within a center until a staff member, parent, or older sibling becomes infected. By that time, a large number of individuals are likely to have been exposed to the disease. Knowledge of the mechanisms for disease spread within and between day care centers is necessary for the development of effective control measures.

MATERIALS AND METHODS

The model used for the simulation in this study was based on an epidemic model with three classes of individuals; susceptibles, infectives, and removed, and with no vital statistics and no movement of individuals into or out of the population (SIR model). In addi-

TABLE 1. A comparison of the two simulation models

Model 1	Model 2
200 children	500 children
100 × 100 grid	500 × 500 grid
population density of 1 child/50 sq. units	population density of 1 child/500 sq. units
2 day care centers	2 day care centers
20 private homes	40 private homes
38% no care, 38% private home, 24% day care center	53% no care, 36% private home, 11% day care center
Maximum probability of transmission due to geographic proximity of 0.5	Maximum probability of transmission due to geographic proximity of 0.5
Maximum probability of transmission in day care center of 1	Maximum probability of transmission in day care center of 0.6
Maximum probability of transmission in private home of 1	Maximum probability of transmission in private home of 0.4

tion to these restrictions, several other assumptions were incorporated into the model. These include the following: (1) the disease has no latent period, (2) recovery confers permanent immunity, (3) the disease begins in one individual only, (4) individuals do not change home location or day care facility at any time during the simulation, and (5) no two individuals live at the same location. This model only considers children of preschool age, no age structure is included, and no other family members or non-preschool contacts are included. The probability that the disease is transmitted is a function only of the distance between individuals and of the attendance at a day care facility. Both care in a day care center and care in a private home were included in the model.

Two slightly different models were used in the study. These models will be designated as Model 1 and Model 2. The major difference between the models is in the size of the population considered. Model 1 used a population of 200 children; Model 2 increased the population size to 500 children. This change and other differences between the models will be discussed in greater detail below and are summarized in Table 1.

Children in both models were assigned home geographic coordinates randomly with reference to a square grid. These homes are not evenly spaced and so do not constitute a lattice as described by Bailey (1967) and Bartlett (1961). The children were also assigned to a day care center, to a private home, or to no care. For both models, private homes ranged in size from 1 to 7 children and day care centers ranged in size from 23 to 29 children.

The proportions of children in the different kinds of day care facilities were not the same in the two models. Licensing laws in New

Mexico prescribe a maximum of six children per home for private home care. Day care centers range in size from about 15–20 children to over 200 children. Since a major focus of the simulation was on the spread of disease within and between centers it was necessary to include two centers. However, with a population size of only 200 or 500 children for the simulation it was necessary to consider an expected center size of no greater than 20 children. This means an expected value of 40 children attending day care centers. In Model 1, as a result of the random assignment of children to child care facility, 24% of the population of 200 were assigned to a day care center, and the remaining children were divided about equally into two groups; those attending private homes and those who stayed at home. These proportions are quite different from those actually seen in the United States, but with a sample size of 200, a minimum center size of 15, and at least two centers required, it is not possible to approximate the actual proportions. A major reason for increasing the sample size to 500 children was to allow the necessary minimum of children in day care centers while still being able to approximate the observed proportions.

The proportions of children using particular types of care in Model 2 were derived from published values of the percentages of children in different situations. The data on use of child care facilities have been published only for working women [Lueck et al., 1982]. However, many women who do not work still choose to send their children to some form of child care. It was therefore necessary to estimate the percentage of all women who are in the labor force and to estimate the probability that the children of a nonworking parent would attend day care.

These values were determined from information in Lueck et al. (1982). From these data it was possible to derive the relative proportions of children in various kinds of care. These were determined to be 10% in day care centers, 30% in private homes, and 60% with no outside care. A sample size of 500 children was sufficient to guarantee that there would be two day care centers with approximately 20 children each, but that the total number of children attending day care centers would be only about 10% of the total number of children.

Assignment to a particular type of care for any one child was made randomly with the constraints on proportions of children in each type of care described above. The assignment of day care center children to a particular center was also made randomly. Private-home children were usually assigned to the available home closest to their place of residence. However, although private homes are usually chosen from within the same neighborhood as the child's home, a percentage of children attend a more distant home, usually because the home is nearer to the parent's workplace, because it is the home of a relative or friend, or because the provider has special qualities the parents want for the child. To take this into account 20% of the possible numbers assigning a child to a private home designated that a distant home should be used.

Most states allow a maximum of six children to be taken care of in one private home in order for the home to be licensed. Although many homes are not licensed and therefore do not have to comply with the laws, it is relatively rare for a provider to take care of more than six full-time children. On the other hand, it is relatively common for a provider to take care of fewer than six children. To guarantee that there would be a range of home sizes in the study, a maximum of six children per home and an average of four children per home were assumed.

The probability of transmission of the disease from an infective child to a susceptible child depended on two factors, the geographic distance between the homes of the children and whether they attended the same day care facility. These two factors are independent of each other, so that the probability of transmission is the sum of the contributions from each factor.

The probability of transmission as a result of geographic proximity was determined by a function that was inversely proportional to the square of the distance between children. The probability of transmission due to day care attendance was automatically set to zero for a child who attended no day care facility and was a function of the proportion of infective children in the facility attended for all other children.

The simulation of disease spread begins with one case of the disease. Six (Model 1) or seven (Model 2) different starting points were used. The starting points were chosen from throughout the entire geographic area and included all three day care situations (day care center care, private home care, and staying at home). At the beginning of each cycle (1 cycle = 1 week), the probabilities of transmission for all susceptibles were computed. This was done by looking at all pairs of individuals in the population. Each time there was a susceptible-infective pair, an increment to the transmission probability for the susceptible was calculated, taking into account the distance between the homes of the two individuals and whether they attended the same child care facility. The total probability of transmission for a given susceptible was the sum of each of these increments. This probability was then compared to a uniformly distributed random number between 0 and 1 generated by the program URAND (Forsythe et al., 1977). If the random number was less than or equal to the calculated probability of transmission and if the random number was not equal to zero, then transmission occurred and the susceptible became infective. Infective individuals remained infective for 3 weeks. At the end of this time, the infective individuals recovered and could neither transmit the disease nor become reinfected. The simulation continued for 52 cycles or until there were no more infectives, whichever came first.

RESULTS
Model 1

The simulation was repeated twenty-five times for each of the six starting points in Model 1. Eight items were sufficient to characterize each run. These included the length of the outbreak in weeks, the total number of cases, the number of cases in day care center 1, the number of cases in day care center 2, the number of cases among children who did not attend day care, the number of cases among children who attended a private home, the number of private homes in which there were cases, and the actual homes which

TABLE 2. Means and ranges of simulation results, Model 1 [1]

Variable	Child number					
	132	33	19	122	103	90
Means						
Type of care used	ph	nc	dc2	nc	dc1	ph
Weeks per epidemic	9.24	3.24	12.24	5.60	19.32	9.92
Number of cases	10.80	1.12	17.64	3.96	51.56	17.04
Cases in dc1	1.84	0	4.60	0.92	23.00	6.44
Cases in dc2	1.16	0	4.44	0.08	5.36	1.72
Cases in ph	4.92	0	6.04	0.52	15.28	5.76
Number of ph providers	1.76	0	2.00	0.32	5.52	2.20
Cases in no care	2.88	1.12	2.56	2.44	7.92	3.12
Ranges						
Weeks per epidemic	3–38	3–6	3–31	3–21	8–37	3–27
Number of cases	1–70	1–2	1–66	1–44	29–86	1–64
Cases in dc1	0,23	0	0,23	0,23	23	0,23
Cases in dc2	0–17	0	1–13	0,2	0–15	0–12
Cases in ph	1–21	0	0–22	0–8	1–36	1–18
Number of ph providers	1–8	0	0–7	0–4	1–11	1–7
Cases in no care	0–9	1–2	0–14	1–11	4–12	0–14

[1] dc = day care center; nc = no care; ph = private home.

(a)
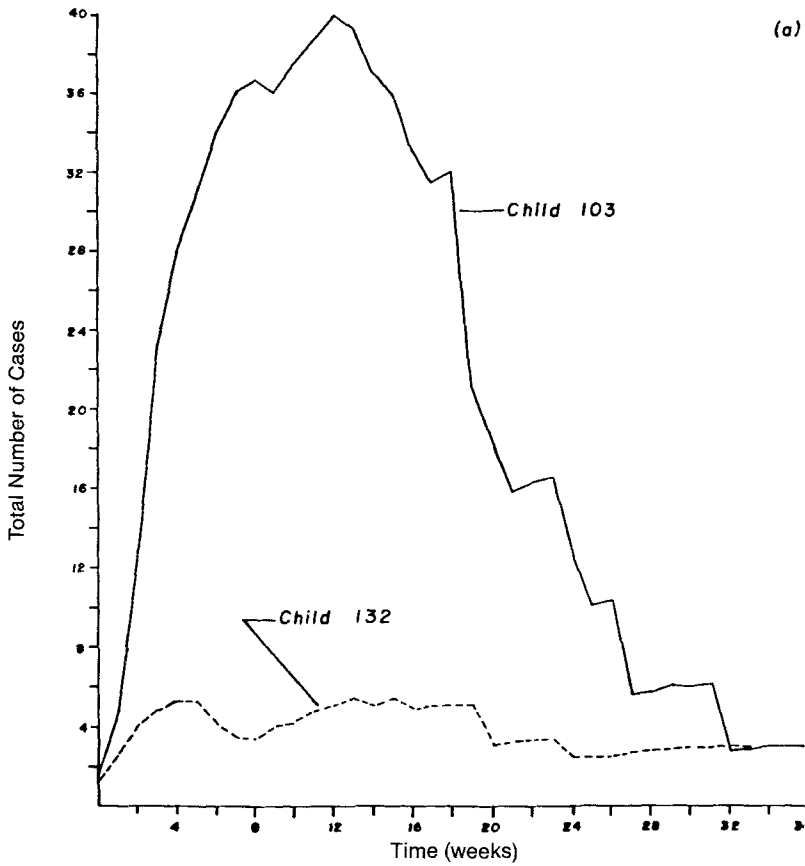
were affected. The means and ranges for the first seven items were calculated using the data from all of the twenty-five runs. These values are presented in Table 2.

The mean length of an epidemic appears to be related to the type of care experienced by the initial case. The shortest epidemics occurred when the initial infective was a child who did not use day care. The longest epidemics resulted when the initial infective attended a day care center. When the initial case attended a private home the epidemic was of intermediate length. This result occurs because of the existence of larger numbers of contacts for children in day care centers, which would allow the disease to

maintain itself for longer periods of time. The children who did not use day care would have contact with neighbors only, and the number of neighbors was ordinarily quite small.

The total number of cases is not as closely related to the type of care used. When the initial case was one of the two children who did not use day care, there were significantly smaller outbreaks than when the initial case was one of the other four children, but the children in private homes and the child in day care center 2 all caused outbreaks of about the same size (Table 2). When the initial case attended day care center 1, the outbreaks affected, on average, 25% of the entire
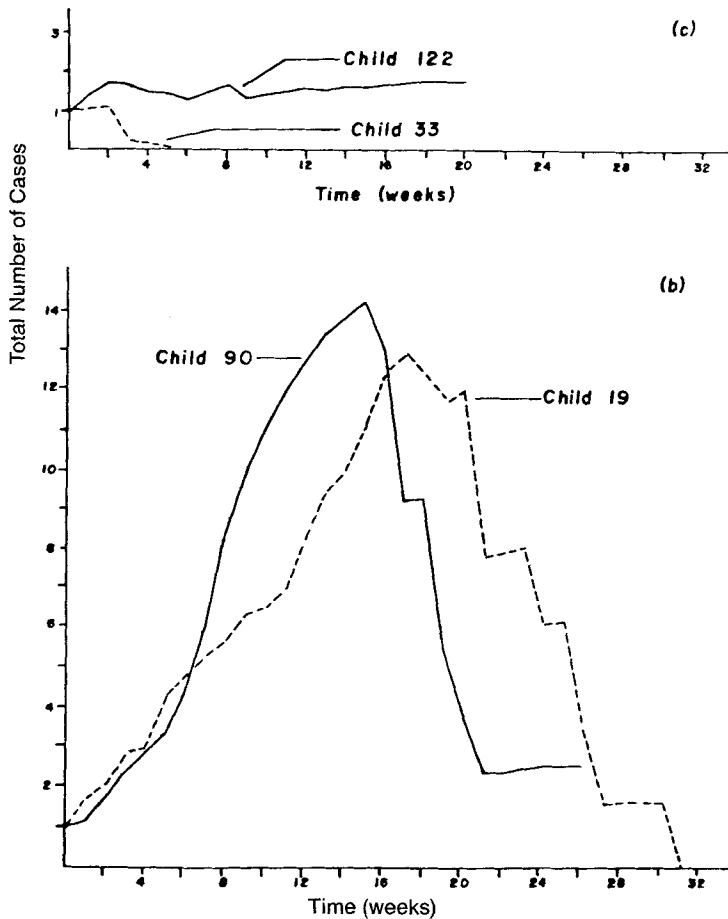

Fig. 1. Total number of cases of the disease per week, averaged over all 25 runs of the simulation, Model 1. a: Starts of epidemic at child 103 (day care center 1) and child 132 (private home). b: Starts of epidemic at child 19 (day care center 2) and child 90 (private home). c: Starts of epidemic at child 33 (no care) and child 122 (no care).

population and the mean number of cases was 3–4 times as high as outbreaks caused by the other children using day care facilities.

The reasons for this are not clear but they appear to be related to the geographic distribution of children attending day care center 1. Every time the disease was introduced into day care center 1, no matter what the starting point, all the children in the center became ill. On the other hand, an outbreak in day care center 2 never reached every child.

The distribution of homes of children attending center 2 is much more uniform than that for center 1. There are several clusters of children attending center 1 where the probability of transmission will be influenced by both day care transmission and transmission between neighbors. This means that the disease will spread much more rapidly within the center, and therefore it is much more likely that the critical value will be reached, above which the probability of transmission approaches one. Each of the children attending the center has contacts within his neighborhood, so that the disease is likely to spread to other children, some of whom attend other facilities. When the initial case was in center 1 (child 103 in Table 2), every run resulted in a major epidemic, so the average number of cases was very large.

In addition to the means and ranges of the above variables, the progress of an "average" epidemic for a particular starting point was studied. This was done in two ways; by looking at the total number of cases at each week during an outbreak, and by looking at the total number of new cases at each week. The values for a specific week were averaged over all twenty-five runs. A value of zero was assumed for every week after the extinction of an outbreak. The resulting graphs represent the average effects of an epidemic from a given starting point, expressed in terms of the total number of cases each week, calculated from the beginning of the epidemic (Fig. 1a–c), and also expressed in terms of the average number of new cases per week throughout the outbreak (Fig. 2a–c). These graphs provide a means of comparing the dependence of the form of the epidemic on the particular starting point used.

The number of runs of the simulation which involved day care center 1 and the time at which a case of the disease entered the center were the most important factors influencing the patterns of disease spread. Transmission within center 1 was always

rapid, which resulted in a sharp rise in the total number of cases shortly after introduction of the disease into the center. Early introduction of the disease into day care center 1 resulted in an average epidemic which increased in number of cases rapidly, but declined slowly (Fig. 1a). The curve of the average number of new cases per week shows a rapid early growth rate for the epidemic, which then decreases throughout the length of the epidemic (Fig. 2a).

Later introduction of the disease into day care center 1 has the opposite effect on the total number of cases. In this situation there is a gradual increase in the number of cases followed by a more rapid decline (Fig. 1b). The growth rate of the epidemic, as indicated by the average number of new cases, is relatively constant initially, then increases rapidly, and is followed by a more gradual decline (Fig. 2b).

When the disease did not enter day care center 1, the average total number of cases was very small and relatively constant throughout the course of an epidemic (Fig. 1c). The average growth rate for epidemics beginning with child 122 increased midway through the "average" epidemic because of one run of the simulation in which the disease entered day care center 1 (Fig. 2c). If this one run is excluded from the calculations of the average number of new cases, the average number of new cases for both epidemics beginning with children who do not attend day care decreases gradually throughout the course of the epidemic.

Model 2

In Model 2, a total of 25 runs were made for each of seven starting points. An additional child who did not use day care was included because the first two chosen had no neighbors to whom the disease could spread. This third child was chosen from the same general geographic area as one of the other two. There was one neighbor within the critical distance to this child, but the distance was so far that even in this case transmission never occurred. Results will be discussed, therefore, for only the remaining four starting points.

The means and ranges of each item characterizing these epidemics are given in Table 3. For every one of the seven variables the two children attending private homes were very similar and the two children attending day care centers were very similar. The type
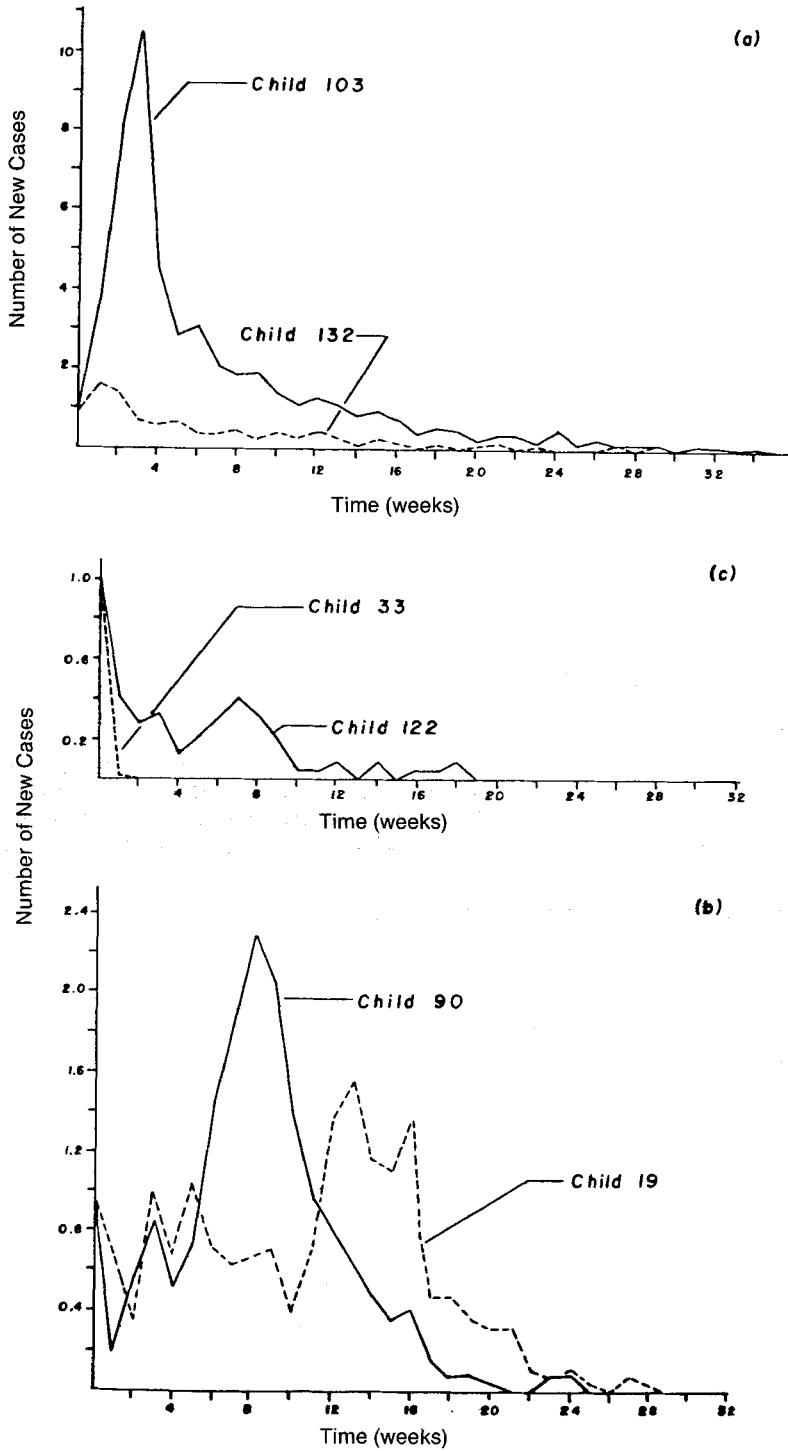
Fig. 2. Number of new cases of the disease each week, averaged over all 25 runs of the simulation, Model 1. a: Starts of epidemic at child 103 (day care center 1) and child 132 (private home. b: Starts of epidemic at child 19 (day care center 2) and child 90 (private home). c: Starts of epidemic at child 33 (no care) and child 122 (no care).

TABLE 3. Means and ranges of simulation results, Model 2[1]

Variable	Child number						
	111	149	241	315	100	132	198
Means							
Type of care used	dc1	ph	ph	dc2	nc	nc	nc
Weeks per epidemic	10.52	4.96	5.56	12.96	3.00	3.00	3.00
Number of cases	34.96	2.48	2.48	30.68	1.00	1.00	1.00
Cases in dc1	28.68	0	0	4.64	0	0	0
Cases in dc2	4.69	0	0	23.40	0	0	0
Cases in ph	1.08	2.48	2.28	1.96	0	0	0
Number of ph providers	0.52	1.00	1.00	0.76	0	0	0
Cases in no care	0.52	0	0.20	0.72	1.00	1.00	1.00
Ranges							
Weeks per epidemic	7–16	3–10	3–16	9–19	3	3	3
Number of cases	28–62	1–6	1–8	23–62	1	1	1
Cases in dc1	28–29	0	0	0,29	0	0	0
Cases in dc2	0–24	0	0	22–24	0	0	0
Cases in ph	0–8	1–6	1–6	0–9	0	0	0
Number of ph providers	0–2	1	1	0–3	0	0	0
Cases in no care	0–3	0	0–2	0–3	1	1	1

[1]dc = day care center; nc = no care; ph = private home.

of care attended was overwhelmingly the most important factor influencing the type of outbreak seen.

The curves for the average total number of cases and for the average number of new cases also give evidence of the importance of type of care in determining the risk of disease spread. When the initial infective attended a private home, the total number of cases gradually increased a small amount and then decreased rapidly until the disease died out. The growth rate decreased rapidly at first and then increased temporarily when there was a small outbreak within the private home attended (Fig. 3a, 4a).

When the initial infective attended a day care center, the total number of cases increased rapidly and was followed by a rapid decrease. The curves for the average number of new cases indicate rapid initial growth followed by a gradual decrease in the growth rate until the epidemic dies out (Fig. 3b, 4b).

This dependence of the epidemic spread on the type of care used by the initial case is due to the lower population density in Model 2 than in Model 1. The effect of the lower density was to make neighborhood transmission much more difficult. The three children who did not attend any form of care were never able to transmit the disease. Of the 50 runs involving an initial case who attended a private home, there was no transmission to a day care center, and only 5 runs involved

transmission to a neighbor who did not use day care facilities. The cases that did occur, with the exception of the few cases among children who stayed at home, were all in children who attended the same private home as the initial case.

When the initial case was in a day care center, because there were 4–5 times as many contacts, the probability that at least one child in the center had a close neighbor was quite high. Also, because each infected child was infective for 3 weeks, the probability that the disease would reach all or almost all of the children in the center was nearly one; and therefore, the probability that the disease spread to children outside the center by neighborhood transmission was high. However, because of the lack of geographic spread, the number of children outside the initial center who contracted the disease was usually small. The largest outbreaks occurred when the disease spread from one day care center to the other, and these outbreaks did not include very many children from outside the day care center population.

With the low population density of this model, it is only in day care centers that there are a sufficient number of susceptibles clustered together so that the disease can reach epidemic proportions. Private homes did not provide the critical number of susceptibles, and in no case were there sufficient numbers of neighbors to allow for the devel-
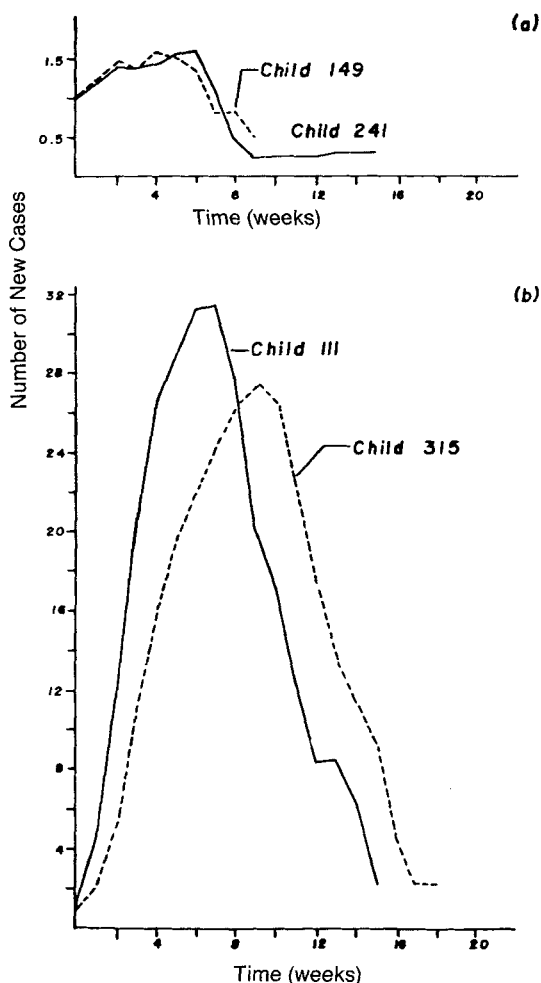
Fig. 3. Total number of cases of the disease per week, averaged over all 25 runs of the simulation, Model 2. a: Starts of epidemic at child 149 (private home) and child 241 (private home). b: Starts of epidemic at child 111 (day care center 1) and child 315 (day care center 2).
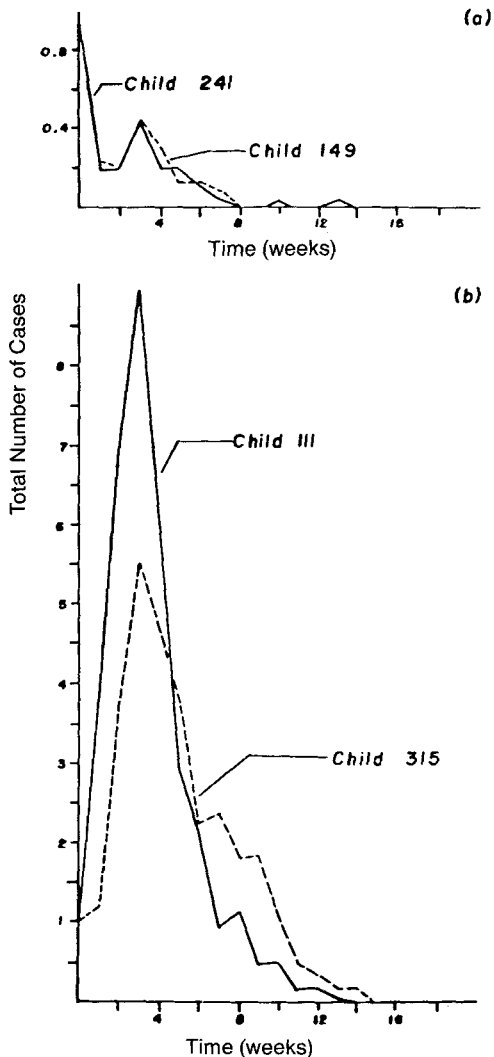

Fig. 4. Number of new cases of the disease each week, averaged over all 25 runs of the simulation, Model 2. a: Starts of epidemic at child 149 (private home) and child 241 (private home). b: Starts of epidemic at child 111 (day care center 1) and child 315 (day care center 2).

opment of an epidemic if the initial case did not attend a day care center. This threshold behavior for the existence of epidemic outbreaks is a well-known result both for theoretical models (originally described by Kermack and McKendrick, 1927) and for the behavior of epidemics in actual populations (for example, Black [1966] describes the effects of population size on the presence of measles in island populations in the Pacific Ocean).

The distinction between the two day care centers that was observed in Model 1 was not observed in Model 2. In every run in which the disease entered a day care center (either one) there were at most three unaffected children attending the center by the end of the epidemic. However, there is no chance of a contribution from the neighborhood for children in either day care center, since there are no children attending the centers who live close enough to another child from the

same center for neighborhood transmission to occur. Because of this, the epidemic patterns to be expected from initial cases in each center would be similar. In a real population, however, some of the centers are likely to have a number of children living in the same neighborhoods, so that the behaviors seen in model 1 are not unexpected even with the lower population densities present in actual communities.

DISCUSSION

The results of these simulations confirm the observation that children who attend day care centers are at much higher risk for many diseases than children who attend either private homes or who do not attend any form of day care. However, the simulations also indicate several other aspects of disease behavior that may not be so obvious. They clearly showed that day care centers not only increase the risk to individual children, but they increase the severity and duration of community-wide epidemics. By providing a much larger population of susceptible individuals who have come together from many different regions within a community and who interact over long periods of time, day care centers both help to sustain epidemics for a longer period of time and help in the spread of the disease throughout a community.

The Model 1 simulations also showed dramatically that slight variations in the degree of geographic clustering of children attending a particular center cause a large increase in the size and severity of epidemics, both within that particular center and throughout the community. This indicates a need to look at contact among children outside the centers themselves in determining relative risks of various centers within a community. Those centers in which there are a number of children from the same neighborhoods or with other outside social contacts are likely to be foci for more severe epidemics than centers without such clustering, and therefore relatively more attention should be paid to preventive and control measures in these centers.

The major differences in the structure of Model 1 and Model 2 are different sample sizes, different population densities, and changes in day care transmission probabilities. In Model 1, the function for day care transmission was such that the maximum probability of transmission in any one week

was one whenever a critical number of children were infective. Because of this, together with the geographic clustering of children attending center 1, whenever the disease entered center 1, all children rapidly became infected. This is not a realistic result. With any disease there are always some individuals who do not become infected, either because they are already immune or because they are more resistant or because they do not come into enough contact with an infected person.

For a disease such as hepatitis A, which is a disease with fecal-oral transmission, and which can be a problem in day care centers accepting diapered children, it has been estimated that about 50% of the children in a center may actually become infected. Hadler et al. (1980) performed serologic testing of children in two day care centers in Phoenix, Arizona, which had been involved in an outbreak of hepatitis A. They found that 53% of the children at one center and 44% of the children at another center possessed antibodies to the disease. Only 1.5% in the first center and 8.2% in the second center had symptomatic cases of the disease. Although serologic testing was not done in New Mexico, the percentages of clinical cases resulting from outbreaks in ten centers in Albuquerque in 1979 are similar to those found by Hadler et al. The range was about 1%–8% and the mean was about 4%. It is therefore likely that about half of the children in the Albuquerque centers were also infected during the course of an epidemic.

In order to retard the spread of the disease within a center, in Model 2 the function for day care transmission was changed so that the maximum probability of transmission in any 1 week was only 0.6 for day care centers and 0.4 for private homes. However, because an infected child could transmit the disease for 3 weeks the effect of lowering this probability was minimal. In all cases where the disease was introduced into a day care center in Model 2, most or all of the children in the center contracted the disease. In future models this probability would need to be lowered even further to guarantee that only about half of the children would become infected during the course of an outbreak.

The use of a sample size of 500 in Model 2 guaranteed that the proportions of children in different types of care would approach the actual values. This resulted in a much higher proportion of children attending private

homes and of children who used no form of care than in Model 1. The probability of transmission among these children was much smaller than for those in day care centers, primarily because the risk of exposure was much less. The effect of this would be to decrease the size and the length of epidemics. Comparisons of Tables 2 and 3 show that this is the case.

However, even with a sample size of 500, there could only be two very small day care centers included in the simulations. Communities which tend to have problems with day-care-related diseases have a number of centers ranging up to several hundred children in size. Hadler et al. (1982) found that in Maricopa County, Arizona, from August 1977 to August 1979, only 3% of the day care hepatitis outbreaks occurred in centers that enrolled <20 children, while 53% occurred in centers that enrolled >50 children. This provides further evidence that the maximum probabilities of transmission used in the simulations are much higher than those actually present.

The effects of the difference in population density between Model 1 and Model 2 have already been discussed. The patterns of epidemic spread seen in Model 2, while less interesting, are more realistic than the patterns seen in Model 1. The number of neighborhood contacts of most preschoolers (and especially infants and toddlers) is usually fairly low. The population density of a real city, in terms of preschool children, is probably nearer to the density used for Model 2 than for Model 1. When a population is limited to preschool children only, as was the case for both models, the likelihood of neighborhood transmission is small, so that epidemics involving transmission directly between children are likely to involve only day care centers, with a sprinkling of children who attend private homes and children who use no form of care.

Neither simulation included household contacts of children in day care, nor did they consider other sources for cases of the disease. However, data from Bernalillo County, New Mexico, show that from January 1980 through October 1982 only 21% of all cases of hepatitis A were documented as day care related. Thirty-four percent of the cases were in children, but only 30% of the cases in children were day care related (Bernalillo County Health Department, unpublished). It is clear that although day care facilities are

at high risk for hepatitis A, there are many other ways in which the disease is transmitted. The source of infection for children who come down with hepatitis A appears to be, in many cases, household contacts rather than other children of the same age. Therefore, the simulations only consider one of the major risk factors for hepatitis spread in a real population and should not be generalized to describe the entire range of behavior of the disease.

The exclusion of household contacts from the simulation is a major drawback. Two-thirds of all day-care-related cases do not occur in children attending the centers. Parents and other relatives and friends usually have a much larger number of contacts than preschool children. It is possible that if chains of infection were traced a much higher proportion of all cases of hepatitis A would be linked to transmission within a day care center, by an indirect route from household contact to friend to friend, etc. The influence of household contacts on the spread of hepatitis A would be an important factor to include in future models.

Another assumption of the simulation models was that there was no latent period for the disease. However, hepatitis A has a total incubation period averaging 28–30 days, with maximum infectivity during the latter half of the incubation period and the first week following onset of symptoms. This means that the disease has a latent period of about 1–2 weeks (Benenson, 1981). This could be incorporated into the simulation by adding an additional state after infection with the disease but before an individual becomes infectious. The length of the actual infectious period would remain at three weeks, but the onset of this period would be delayed. It is likely that this would change the shape of the epidemic curve, making epidemics longer and less severe at any one time, but it is not likely that the relative risks to individuals and the sizes of the epidemics would change significantly. Longini (1986) and Longini and Koopman (1982) have shown that for a discrete time epidemic model with permanent immunity, such as the models for these simulations, the final attack rates for a model with a latent period are identical with the final attack rates for a model without a latent period.

It would also be important to incorporate age of the children into the simulation. It has been shown (Hadler et al., 1982) that the risk

of outbreaks of hepatitis A is highly corre-
lated with the number of diapered children
in a center. If only symptomatic cases are
taken into account, the number of cases does
not vary much among preschool children.
However, if asymptomatic cases are in-
cluded, the risk is much greater for children
from 1 to 2 years of age than for other ages
(Hadler et al., 1980). This differential risk is
an important factor to take into account.

Even though the simulations described
above are simplified models, some important
results can be derived from them. The results
from both models, and especially from Model
2, show the importance of day care centers as
a focus for disease spread. The most critical
factor involved is size of the centers, as evi-
denced by the difference between private
home care and care in a day care center. This
factor affects the number of available con-
tacts, so that children who attend day care
centers are at much higher risk for contract-
ing a case of the disease than children who
attend private homes and children who do
not attend any sort of child care facility. This
is a well-known fact and serves primarily to
indicate that the models are reasonably re-
alistic, in spite of the simplifying assump-
tions made.

A more important, and less obvious, result
is the effect of the number of children from
the same neighborhood attending the same
day care center. The results from Model 1
indicate that the probability of a major epi-
demic within such a center and throughout a
community is much greater, all other things
being equal. However, most day care centers
in Albuquerque do not have a very localized
distribution of children, and data from 1979
seem to indicate that other factors are more
important.

One factor may be social links between cen-
ters. One family in Albuquerque, New Mex-
ico, owns and runs five different centers in
the city. There is evidence that some children
attend two or more of these centers on a
regular basis. It is quite likely that there is
also exchange of staff among the centers.
These social links provide a means for trans-
mission of a disease among the centers. Dur-
ing the hepatitis A epidemic of 1979, all five
centers had at least one case of hepatitis A
among the students, staff, or household con-
tacts, and there was at least one documented
case where the disease spread from one of
these centers to another because of an in-
fected child who attended both centers

(Bernalillo County Health Department, un-
published). In addition, four of the six largest
outbreaks of hepatitis A (> 10 clinical cases)
during the city-wide epidemic occurred in
these centers.

The simulations described above have
shown that patterns of disease spread which
are reasonable approximations of actual pat-
terns can result. Simulation has the added
advantage that incorporation of greater com-
plexity into a model is relatively easy. One
of the most critical areas for future study is
the effects of population structure on the dy-
namics of disease spread. All human popula-
tions are structured in many different ways,
and this structuring determines the interac-
tions of different individuals and therefore
has a critical influence on the relative risks
of different individuals. Models such as these
hepatitis A models can be used to study a
variety of diseases in human populations. For
example, models could be developed to study
the spread of AIDS within the United States,
by looking at the patterns of interaction of
individuals within and between well-defined
groups, such as homosexual males, intrave-
nous drug users, or heterosexuals with no
known risk factors.

The simulations presented in this paper
are applied to the spread of a particular dis-
ease in a particular population. However, the
incorporation of population structure into
models of disease spread is a very general
problem. In addition to this study, there have
been several models developed for other spe-
cific diseases, as well as a few more general
models (for an exhaustive review see Acker-
man et al., 1984). Much more work needs to
be done with such models so that a better
understanding of the factors that influence
the risk of different individuals within a het-
erogeneous population is possible.

LITERATURE CITED

Ackerman, E, Elveback, LR, and Fox, JP (1984) Simula-
tion of Infectious Disease Epidemics. Springfield, IL:

Charles C. Thomas.

Bailey, NTJ (1967) The simulation of stochastic epidemics in two dimensions. Proc. 5th Berk. Symp. Math. Stat. Prob. 4:237–257.

Bailey, NTJ (1975) The Mathematical Theory of Infectious Diseases. New York: Hafner Press.

Baroyan, OV, Genchikov, LA, Rvachev, LA, and Schaschkov, VA (1969) An attempt at large-scale influenza modelling by means of a computer. Bull. Int. Epidemiol. Assoc. 18:22–31.

Baroyan, OV, Rvachev, LA, Basilevsky, UV, Ermakov, VV, Frank, KD, Rvachev, MA, and Shaskov, VA (1971) Computer modelling of influenza epidemics for the whole country (USRR). Adv. Appl. Prob. 3:224–228.

Bartlett, MS (1961) Monte Carlo studies in ecology and epidemiology. Proc. 4th Berk. Symp. Math. Stat. Prob. 4:39–55.

Benenson, AS (1981) Control of Communicable Diseases in Man, 13th edition. Washington, DC: The American Public Health Association.

Black, FL (1966) Measles endemicity in insular populations: Critical community size and its evolutionary implication. J. Theor. Biol. 11:207–211.

Cliff, AD, Haggett, P, Ord, JK, and Versey, JR (1981) Spatial Diffusion. Cambridge: Cambridge University Press.

Elveback, LR, Ackerman, E, Gatewood, L, and Fox, JP (1971) Stochastic two-agent epidemic simulation models for a community of families. Am. J. Epidemiol. 93:267–280.

Ewy, W, Ackerman, E, Gatewood, L, Elveback, L, and Fox, JP (1972) A generalized stochastic model for simulation of epidemics in a heterogeneous population (Model VI). Comput. Biol. Med. 2:45–58.

Forsythe, GE, Malcolm, MA, and Moler, CB (1977) Computer Methods for Mathematical Computations. Englewood Cliffs, NJ: Prentice Hall.

Hadler, SC, Erben, JJ, Francis, DP, Webster, HM, and Maynard, JE (1982) Risk factors for hepatitis A in day-care centers. J. Infect. Dis. 145:255–261.

Hadler, SC, Webster, HM, Erben, JJ, Swanson, JE, and Maynard, JE (1980) Hepatitis A in day-care centers. N. Engl. J. Med. 302:1222–1227.

Kermack, WO, and McKendrick, AG (1927) Contributions to the mathematical theory of epidemics, part I. Proc. R. Soc. Lond., A. 115:700–721.

Longini, IM (1986) The generalized discrete-time epidemic model with immunity: A synthesis. Math. Biosci. 82:19–41.

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

Lueck, M, Orr, AC, and O'Connell, M (1982) Trends in Child Care Arrangements of Working Mothers. United States Bureau of the Census, Current Population Reports, Series P-23, No. 117. Washington, DC: U.S. Government Printing Office.

Sattenspiel, L (1984) The Spread of Disease in Subdivided Populations. Ph.D. dissertation, University of New Mexico.

Sattenspiel, L (1987) Population structure and the spread of disease. Hum. Biol. 59:411–438.