A stochastic distributed-delay model of disease processes in dynamic populations

H. Scott Hurd\textsuperscript{a,b}, John B. Kaneene\textsuperscript{a} and James W. Lloyd\textsuperscript{a}

\textsuperscript{a}Population Medicine Center and Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, MI 48824, USA
\textsuperscript{b}National Animal Health Monitoring System, USDA/APHIS/VS, Fort Collins, CO, 80521, USA

ABSTRACT


A simulation model that is applicable to infectious and non-infectious disease is proposed. This paper describes a model for simulation of infectious and non-infectious disease processes in dynamic populations, and compares its behavior to a stochastic version of the Reed–Frost model for a hypothetical infectious disease. A distributed-delay model is applied.

Monte-Carlo simulations of both modeling approaches produced epidemics of randomly determined sizes. Although both models demonstrated the characteristic bimodal distributions of total number of cases per epidemic, the shape of the distributions was slightly different. Separation between the two peaks was not as great with the distributed-delay model as with the Reed–Frost model. The tail was slightly more extended than the Reed–Frost, and there were more epidemics in the 50–100 case range. Both models produced similar average attack rates.

INTRODUCTION

Simulation modeling has been used in epidemiology since the nineteenth century (Susser, 1985). It is useful for generating hypotheses regarding disease processes, for quantitatively describing the behavior of a complex system, and for determining sensitive areas in the system for management input. Virtually all of the work in epidemiologic simulation modeling has taken place in the area of infectious diseases (Bailey, 1975; King and Soskolne, 1988).

Infectious-disease models have been built on the premise that disease can be viewed as the movement from the state of susceptibility, to latency, to incubation, to infection, to recovery, death, or immunity (Nokes and Anderson, 1988). Using this perspective it was straightforward to the development...
of various state-transition models that would simulate the movement of individuals from one state to the next. The driving force for the critical transition from susceptible to infected was the interaction between current infectious and susceptible individuals (Abbey, 1952; Ackerman et al., 1984).

Non-infectious disease processes move through a similar series of states. For example, cancer development can be seen in terms of the onset of early cellular changes followed by undiagnosed disease which progresses to clinical manifestations, followed by death or recovery (Morrison, 1979). However, very few models of non-infectious disease have been developed.

The two general modeling techniques used for infection-disease modeling (mass action and chain binomial) have certain limitations. Some of the limiting assumptions of mass-action models are that they assume random and homogeneous mixing of susceptible individuals (Fine, 1982), and that there is a linear relationship between the incidence rate and the number of cases. They can be deterministic or stochastic (Bartlett, 1953), and they can be discrete time (Soper, 1929) or continuous time (Bailey, 1955). However, mass-action models are always continuous entity. Also, the exit rate from these first-order differential equations has an exponential distribution over time.

In order to overcome these limiting assumptions, the chain-binomial models were developed (Greenwood, 1946). In these models, new cases of disease occur in a series of stages. The number of cases at any stage will have a binomial distribution depending on the number of infectious and susceptible individuals at the previous stage (Bailey, 1975). These models are fully stochastic, discrete time and discrete entity. These models assume that the period of infectiousness is relatively short and of constant duration, and that there is a constant probability of infection in each serial interval (Fine, 1982). There are at least four types of chain-binomial models; the Greenwood type (Greenwood, 1946), Reed–Frost, the Elveback type, and Markov models.

The objectives of this paper are to describe a distributed-delay model (DDEL) for the simulation of infectious and non-infectious disease. The behavior of the model will be compared with that of a Reed–Frost model. The output distributions and average attack rates for a hypothetical infectious model will be evaluated.

MATERIALS AND METHODS

The distributed delay used in this paper is a Euler numerical integration (Hamming, 1962) of the kth order differential equation shown in eqn. (1). (The Quick Basic (Microsoft, 1988) subroutine for simulating this equation is available from the senior author.)

\[ a_k \frac{d^k y(t)}{dt^k} + a_{k-1} \frac{d^{k-1} y(t)}{dt^{k-1}} + \ldots + ay(t) = x(t) \]  

(1)
where \( x(t) \) is the input at time \( t \), i.e. new susceptible individuals, \( y(t) \) is the output at time \( t \), i.e. new cases, \( k \) is the order of the defining differential equation, \( a_k \) is the \( k \)th specific parameter defining the response of \( y(t) \) to \( x(t) \), e.g. \( 1/k \) of the delay \( E(\tau) \).

**Model parameters**

To implement the distributed-delay model, the two parameters, \( E(\tau) \) and \( k \), must be defined. When applied to disease modeling, \( E(\tau) \) can be defined as the average waiting time spent in a given state. For the susceptible state, \( E(\tau) \) is derived as the inverse of the incidence rate (Morrison, 1979; Rothman, 1986). For the clinical state, \( E(\tau) \) is the average duration of infection or illness. In eqn. (1), \( a_k \) is defined as \( 1/k \) of \( E(\tau) \).

For the susceptible state in an infectious-disease model, \( E(\tau) \) is a function of the number of individuals in both the infected and susceptible states, and their contact rate (Riley et al., 1978; Fine, 1982). For this infectious-disease model, \( E(\tau) \) was defined as in eqn. (2). In eqn. (2) the waiting time varies as the model progresses and the number of infected individuals changes over time. For a non-infectious model, the waiting time can be set as the inverse of the expected incidence rate for that disease. It can also be set to vary over time as might be required for seasonal differences in disease rates.

\[
E(\tau) = \frac{1}{1 - (1 - \beta)I(t)} \tag{2}
\]

where \( I(t) \) is number of infectious individuals at time \( t \), \( \beta \) is effective contact rate.

The value used for \( k \) is a function shape of the output distribution that is desired. The setting of \( k \) determines which Erlang family distributions will be simulated (Manetsch, 1966). For example, if a group of individuals was simultaneously added to a DDEL, Fig. 1 shows how the output of the DDEL will vary with different settings of \( k \). The value of \( k \) can only be determined by assumption or by observation of population events. For example, if a susceptible population was exposed to a point-source exposure, the distribution of cases over time would approximate the appropriate \( k \) for simulating this disease.

In the proposed distributed-delay model, stochasticity is implemented by allowing \( \beta \) to vary randomly. This method represents the uncertainty of the \( \beta \) value as well as the random error for any individual in a population. The values for \( \beta \) were varied along a triangular distribution. (The triangular distribution serves as an approximation for the normal, where a minimum, maximum and mode can be identified.)
Model comparisons

The Reed–Frost computer model outlined in Ackerman et al. (1984, p. 34) was compared with the proposed distributed-delay model. Monte-Carlo runs of the models were implemented for 100 epidemics. Different values for β were used. The goal was to determine if the DDEL could produce the characteristic bimodal distributions of total number of cases per epidemic observed with the Reed–Frost model. Also, the average attack rate (total cases per epidemic per total population at risk) for 100 epidemics was compared between models.

Settings of $k$ from 1 to 20 were evaluated. Values of $k=1$ did not display as pronounced a bimodal distribution, and values of $k=20$ produced computational problems (a result of the lack of conservation of flow). The model performed best when a setting of $k=6$ was used. The output distributions approximated those expected in an early outbreak of an infectious disease. Model runs were carried out on a microcomputer with Quick Basic (Microsoft, 1988) as the programming language for the distributed-delay model and Pascal for the stochastic Reed–Frost model (Foster, 1984). All populations started with 1000 susceptible individuals and two infectives.
RESULTS AND DISCUSSION

A comparison of average attack rates and number of epidemics with total cases less than 50 for different settings of $\beta$ are shown in Table 1. The average attack rates for the Reed–Frost and the distributed-delay model are similar (Table 1). A comparison of the two model outputs shows that the number of epidemics with total cases less than 50 cases tended to be slightly higher with the DDEL model than with the Reed–Frost. Total cases less than 50 is a means of measuring the bimodal nature of the distributions. This parameter demonstrates the threshold effect discussed by McKendrick (1926). According to his theorem, there is a threshold number of susceptibles required in a population. If the number of susceptibles is below this threshold, then the epidemic cannot be sustained.

The observation that the average attack rates are similar between the two models demonstrates that the DDEL is modeling the overall disease frequency similar to the Reed–Frost model. Although the number of epidemics with less than 50 cases is slightly higher with the DDEL model, it is still simulating the threshold behavior comparable with the same frequency as the Reed–Frost. The reason for the higher number of epidemics with more than 50 cases is that the DDEL model retains individuals in the non-clinical (susceptible) state longer than the Reed–Frost does. The DDEL model allows for some delay, after exposure, before the individual exits to the infected state. The Reed–Frost model, however, is based on the assumption that an individual is infected in the subsequent time period after exposure. The number of epidemics with cases less than 50 may have been more similar if the model had been set with $k=1$ for the susceptible state.

Histograms comparing the frequency distributions for the Reed–Frost and distributed-delay models of $\beta=0.016$ and $\beta=0.012$ are shown in Figs. 2 and

<table>
<thead>
<tr>
<th>Mean $\beta$</th>
<th>Reed–Frost model</th>
<th>Distributed-delay model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average attack</td>
<td>No. of epidemics</td>
</tr>
<tr>
<td></td>
<td>rate $^1$</td>
<td>(out of 100) with</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$&lt; 50$ cases</td>
</tr>
<tr>
<td>0.0016</td>
<td>0.58</td>
<td>10</td>
</tr>
<tr>
<td>0.0012</td>
<td>0.15</td>
<td>50</td>
</tr>
<tr>
<td>0.001</td>
<td>0.03</td>
<td>78</td>
</tr>
<tr>
<td>0.0008</td>
<td>0.013</td>
<td>95</td>
</tr>
</tbody>
</table>

$^1$Average attack rate $= \frac{\text{no. of cases/population size}}{\text{no. of runs}}$. 
3, respectively. Comparison of the histograms from the two models shows similar bimodal distributions. However, the detailed DDEL plots do differ from the Reed-Frost model. Specifically, with $\beta=0.016$, the separation between the two peaks is not as great as with the Reed-Frost (Fig. 2). This

Fig. 2. Frequency distributions (100 epidemics) of distributed-delay (DDEL) (a) and Reed-Frost (b) stochastic disease models for $\beta=0.0016$, initialized with 1000 susceptible and two infective individuals.
observation is not surprising because the distributed-delay model is a continuous-entity model, i.e. it uses real numbers instead of integers. The use of real numbers leads to the unrealistic representation of individuals as partial entities, resulting in a smoother curve. For $\beta=0.012$ (Fig. 3), the tail is slightly
more extended than the Reed–Frost, and there are more epidemics in the 50–100 case range.

The bimodal distribution of epidemics is predicted mathematically. The Reed–Frost model has been used to produce predicted bimodal distributions. However, there is no empirical evidence to suggest that the exact shapes of the Reed–Frost distributions are closer to reality than those produced by the DDEL. Therefore, it is not possible to say which model produces more realistic results.

The exact biological meaning of ‘k’ is not clear. Statistically, it is a parameter of an Erlang distribution as shown in Fig. 1. The higher the value of k, the more Gaussian the resulting distribution. Such distributions allow for more realistic models than most mass-action models that assume an exponential distribution of infection duration, i.e. $k=1$ (Bailey, 1975). For example, if the mean infection duration is 7 days ($E[\tau]=7$) and $k=6$, some individuals will start to exit the state in only 2 days and others will stay longer. This is similar to the real world where infection duration varies for different individuals. In the Reed–Frost model, however, all individuals exit the infected state at a given number of days.

Some disadvantages of the DDEL compared to the Reed–Frost model are that they (1) may be difficult to stabilize computationally, (2) do not allow for tracking of individuals with their unique characteristics, (3) involve some aggregation error.

Advantages of the DDEL model are that it (1) is useful for both infectious and non-infectious diseases, (2) represents the stochastic nature of waiting times until disease occurrence, (3) allows flexibility in defining the distribution of the outputs by altering the ‘k’ value, (4) can realistically accommodate other vital dynamics, such as birth or migrations into the populations, (5) allows the computational speed of the model to be unaffected by the number of individuals in the population.

A continuous-time model can be developed that produces the bimodal distribution previously found only in discrete-time models such as the Reed–Frost and its progeny. This distributed-delay model will provide a generic model to simulate infectious and non-infectious disease processes in dynamic populations.

REFERENCES


Foster, D., 1984. Program ReedFrost. University of Michigan, School of Public Health, Ann Arbor, MI.


