

A Note on Chain-Binomial Models of Epidemic Spread: What Is Wrong with the Reed-Frost Formulation?

JOHN A. JACQUEZ

*Depts. of Biostatistics and Physiology, The University of Michigan, Ann Arbor, Michigan
48109*

Received 23 April 1987; revised 29 July 1987

ABSTRACT

Derivations of the Reed-Frost model are analyzed in terms of the assumptions about the probabilistic process used and in terms of internal consistency. Internally consistent derivations can be exhibited but require assumptions about the basic probabilistic process that are clearly unreasonable. A properly posed, more general model is derivable through the use of En'ko's model and the generating function for the number of contacts per person [6].

I. INTRODUCTION

During and subsequent to 1928, L. J. Reed and W. H. Frost developed some chain-binomial models of epidemic spread for use in teaching biostatistics and epidemiology at the Johns Hopkins School of Hygiene and Public Health. Though they did not publish their work, a number of others have reported on it [1-4]. Greenwood [5] presented a somewhat different chain-binomial model, and recently Dietz and Schenzle [6] have pointed out that a development of this type was published in the Russian literature in 1889 by En'ko [7]. The Reed-Frost model has been used widely [6] and has served as a basic model that stimulated the development of more detailed simulation studies [8].

The present work was initiated by my attempting a dimensional analysis on the Reed-Frost equation and was further stimulated by my reading the comments of Dietz and Schenzle [6] on the contribution of En'ko. It is concerned with the proper formulation of a chain-binomial model. In terms of reasonableness of the assumptions used in the model and its internal consistency, is the Reed-Frost model a good model?

II. CHAIN-BINOMIAL MODELS

For historical reasons I start with Soper's deterministic model; it played an important role in the development of epidemic theory and is related to the chain-binomial models. All of the models are for uniform mixing in a population. The chain binomial models as a group assume that the generations of infectives are separated by a significant latent period and time of infectiousness. Thus, they are applicable to diseases in which cases or groups of cases are separated in time well enough to allow identification of successive generations of infection.

1. *SOPER*

In 1929, Soper [9] proposed a finite difference model for measles in intervals of the incubation period. The population was treated as though it moved stepwise through generations of the incubation period; i.e., as though, after contact, the new infectives in $i + 1$ appeared all at once at the end of the i th period. The distinction between latent period and incubation period was not used; more accurately, the generation time is the latent period plus some period of infectiousness. If z_{i-1} is the number of cases and x_{i-1} is the number of susceptibles in $i - 1$, Soper writes Equation (1) for z_i :

$$\frac{z_i}{z_{i-1}} = \frac{x_{i-1}}{m} \quad (1)$$

The parameter m is the number of susceptibles required such that one case in $i - 1$ gives rise to one case in i ; it has the same dimensions as x_{i-1} . One could generalize this by letting m vary from period to period and writing the denominator on the right as m_{i-1} . Note that both z_i/z_{i-1} and x_{i-1}/m are dimensionless.

2. *REED-FROST*

The model generally used is that presented by Abbey [4] and Maia [3]. Let C_i be the number of infectives and S_i the number of susceptibles in the i th period for a population of size N , $C_0 + S_0 = N$. The infectives in the i th period are removed from the process in the $i + 1$ period, $C_{i+1} + S_{i+1} = S_i$.

The Reed-Frost model tries to take into account that multiple contacts between a susceptible and infectives can produce only one new case. The argument presented by Maia [3] and Abbey [4] runs as follows. Let p be the probability of an adequate contact between any two individuals; an adequate contact between a susceptible and an infective in i converts the susceptible into a case in $i + 1$. Then $q = 1 - p$ is the probability that two given individuals do not have adequate contact. The probability that a susceptible does not have adequate contact with any of the C_i cases in i is

then taken to be (2).

$$Q_{i+1} = q^{C_i} \tag{2}$$

That is the same as the probability of not developing the disease in $i + 1$. Therefore the probability that a susceptible has at least one adequate contact is shown in Equation (3); the expected number of cases in $i + 1$ is given by Equation (4).

$$P_{i+1} = 1 - q^{C_i} \tag{3}$$

$$E[C_{i+1}|C_i, S_i] = S_i(1 - q^{C_i}) \tag{4}$$

Equation (4) is the expected number of cases in picking a binomial variable from a population of S_i susceptibles, with probability P_{i+1} . The conditional probability of obtaining $C_{i+1} = x$ cases given C_i and S_i must then be equation (5).

$$P[C_{i+1} = x|C_i, S_i] = \binom{S_i}{x} P_{i+1}^x Q_{i+1}^{S_i-x} \tag{5}$$

The parameter P_i changes with i , thereby giving a chain-binomial process. Note that the probabilities Q_{i+1} and P_{i+1} and the expected value [see Equation (4)] are conditional on C_i and S_i .

I believe there is a problem with the way this is posed and that the difficulty pertains to the definition of p and the probabilistic process used in obtaining Equation (2). In Equations (2-4), q^{C_i} must be dimensionless or the equations are meaningless. Because probabilities are dimensionless, C_i in q^{C_i} must be dimensionless. Therefore C_i in q^{C_i} cannot be the number of infectives which has the dimensions of people or persons. Rather it must be a count of trials in a probabilistic process which happens to equal C_i . But why should the exponent equal C_i ?

Let us see if we can clarify the definition of p and the probabilistic process used in the development of the Reed-Frost chain binomial. First, how is p defined in the standard derivations [3, 4]? According to Abbey [4 p. 202], "If p is the probability of contact between any two specified individuals in the population in a given interval of time (the period of infectiousness), then $q = 1 - p$ is the probability of their not having contact. Contact, or adequate contact, as used by Reed and Frost, is contact such that, if it occurs between an infectious case and a susceptible, it will produce a new case." According to Maia [3 p. 168], "If K is the average number of adequate contacts per individual per time t , and the population size is N , then the probability of adequate contact between any two given individuals during time t will be $p = K/(N - 1) \dots$ " Maia's definition can be dismissed

immediately, for it presents problems with regard to dimensions. As it reads, K has dimensions defined by persons contacted per individual in t , and since $N - 1$ has dimensions of persons, $p = K/(N - 1)$ is not dimensionless. Furthermore K could be greater than $N - 1$ [10].

To return to Abbey's definition, if p were the probability of one adequate contact between two individuals, then q could not equal $1 - p$ because there might be nonzero probabilities of two or more contacts between the same two individuals. To make sense, Abbey must mean either that there can be only one contact between any two individuals or that p must be the probability of *one or more* contacts between two individuals. Let us use the latter, more general assumption. In fact, since we need only concern ourselves with contacts between susceptibles and infectives, p must be the probability of one or more contacts of one susceptible with one infective; then $q = 1 - p$ is the probability of no contact with that infective. This definition of p includes the probabilities of multiple contacts between the same two individuals. Hence, to obtain Equation (2), one must think of the number of possible contacts of a susceptible with different infectives as C_i independent Bernoulli trials. It follows that the probability of zero adequate contacts of one susceptible with the different infectives would be q^{C_i} . Viewed this way the derivation of the Reed-Frost model is clear and understandable. It separates the probability of multiple contacts between a susceptible and one infective from the probability of multiple contacts of a susceptible with different infectives. However, the step leading to Q_{i+1} is suspect. Is it really reasonable to model the possible contacts of a susceptible with the different infectives as C_i independent Bernoulli trials? I think not. Aside from the question of independence, it seems likely that the number of physically possible contacts of one susceptible is not necessarily equal to C_i ; for large C_i it is certainly less. On the other hand if C_i is small as in household epidemics, the number of physically possible contacts could be close to C_i and the Reed-Frost model might then be close though not necessarily exact.

K. Dietz has pointed out another way of deriving the Reed-Frost equation [10]. Assume that each of the C_i infectives makes one contact with another member of the population and that these contacts are made one by one, independently and with replacement. The basic process can be thought of as follows. Number the susceptibles $1, \dots, S_i$ and the infectives $1, \dots, C_i$. Take the C_i infectives in order; the probability that C_{ik} contacts a particular susceptible, S_{is} , is $[S_i/(N - 1)](1/S_i) = 1/(N - 1)$. Let p be the probability of transmission. Then the probability that S_{is} is not infected is $q = 1 - p/(N - 1)$. Because there are C_i such independent trials, the probability that S_{is} is not infected in C_i trials is q^{C_i} . Therefore the probability that S_{is} is infected in the C_i trials is $(1 - q^{C_i})$. Because this is true independently for each susceptible, the expected number of transmissions for S_i susceptibles

must be as given by Equation (6).

$$E[C_{i+1}|C_i, S_i] = S_i [1 - q^{C_i}] = S_i [1 - \{1 - p/(N - 1)\}^{C_i}] \quad (6)$$

The derivation is internally consistent and free of dimensional problems. However, the assumptions in the basic probabilistic process are unrealistic; it is just not believable that each infective makes exactly one contact in the population.

For historical completeness it is worth noting that Wilson and Burke [1] gave an account of Frost's 1928 Cutter lectures wherein they reported a derivation that differs from the standard one [3, 4], as follows. Wilson and Burke took the probability of a contact of an infective with a particular susceptible to be $p = 1/s_i$. This should be $1/(N - 1)$. Nevertheless, if there are k_i contacts between infectives and susceptibles, q^{k_i} is the chance that a susceptible escapes all contacts with infectives; hence the chance of at least one contact must be $1 - q^{k_i}$. The expected number contracting the disease in the next period would then be as given by Equation (7).

$$E[C_{i+1}|C_i, S_i] = S_i(1 - q^{k_i}) \quad (7)$$

Note that Equation (7) is similar to Equation (4) but that k_i , the number of contacts, is not equal to C_i . Properly k_i is a random variable. According to Wilson and Burke, Frost then assumed that the number of contacts is proportional to $C_i S_i$, i.e., $k_i = rC_i S_i$.

3. GREENWOOD

The model of Greenwood [5, 11] is similar to the Reed-Frost model. The chain binomial [Equation (8)] is given by equation (5), but with $P_{i+1} = P$ taken as a constant not dependent on C_i , which holds for $C_i > 0$.

$$P[C_{i+1} = x|C_i, S_i] = \binom{S}{x} P^x Q^{S_i - x} \quad (8)$$

The model of Greenwood is obtained if, instead of transfer by close contact, the transfer of infection occurs by contact of susceptibles with infectious material that is relatively widely spread, so that P_i is constant.

4. EN'KO

I follow the presentation of Dietz and Schenzle [6]. Given a contact of a susceptible in period i , the probability that it is with an infective is $p_i = C_i/(N - 1)$. The probability that the contact is not with an infective is $q_i = [1 - C_i/(N - 1)]$, and if an individual makes k contacts the probability that none are with an infective is $[1 - C_i/(N - 1)]^k$. The probability that at

least one of the contacts is with an infective must be $1 - [1 - C_i/(N-1)]^k$. Therefore the expected number of infectives in the $i+1$ interval is given by [9].

$$E[C_{i+1}|C_i, S_i] = S_i [1 - [1 - C_i/(N-1)]^k] = S_i [1 - q_i^k] \quad (9)$$

Equation (9) appears to be similar to Equation (4), but here q_i depends on i and the number of contacts k is constant. The basic probability p_i is more clearly defined in En'ko's model; note that En'ko's model does not distinguish between multiple contacts with the same individual and with different individuals. Actually the number of contacts is not constant and should be treated as a chance variable.

III. A RESOLUTION

Dietz and Schenzle [6] point the way to a more general solution by way of the generating function of the number of contacts an individual makes. In what follows I elaborate on Dietz and Schenzle's presentation.

Starting with En'ko's results for k contacts, let p_k be the probability an individual has k contacts. Note that p_k may be a function of N but not of C_i or S_i . Let $f(x)$ be the probability generating function of the number of contacts an individual makes:

$$f(x) = \sum_{k=0}^{\infty} p_k x^k \quad (10)$$

Recall that, given a contact, the probability that it is not with an infective is $1 - C_i/(N-1)$ for the i th period. Then $p_k [1 - C_i/(N-1)]^k$ is the probability of k contacts, none with infectives. Thus the probability of no effective contact with infectives, Q_i , must be as follows:

$$\begin{aligned} Q_i &= p_0 + p_1 [1 - C_i/(N-1)] + \cdots + p_k [1 - C_i/(N-1)]^k + \cdots \\ &= f(1 - C_i/(N-1)) \end{aligned} \quad (11)$$

It follows that,

$$E[C_{i+1}|C_i, S_i] = S_i [1 - f(1 - C_i/(N-1))] \quad (12)$$

Now one can set up the chain binomial and Q_i can be calculated at each stage if p_0, p_1, \dots are known. For cases in which only the first few p_i are expected to be nonzero, Q_i could be calculated easily. Dietz and Schenzle extended this approach to explicitly separate multiple contacts with the same person and those with different persons [10].

Soper's model [9] is a limit case of Equation (12). To verify this note that, as $C_i/(N-1)$ approaches zero, $f(1 - C_i/(N-1))$ goes to $f(1)$ and $f(1) = 1$. If it is given that $f(x)$ is analytic in x , expanding around $x = 1$ for small values of $C_i/(N-1)$ and retaining only first order terms yields

$$f(1 - C_i/(N-1)) = 1 - f'(1) C_i/(N-1) \quad (13)$$

Substituting this into Equation (12) gives Soper's model for $m = (N-1)/f'(1)$:

$$\frac{E[C_{i+1}]}{C_i} = \frac{S_i}{(N-1)/f'(1)} \quad (14)$$

Thus if $C_i/(N-1)$ is small, Equation (12) reduces to Soper's equation.

Now it is possible to catalog the forms of Equation (12) for different distributions of the number of contacts. Dietz and Schenzle [6] give the result for Poisson and geometric distributions.

1. POISSON

If the number of contacts has a Poisson distribution with mean number of contacts λ , the expected value for C_{i+1} is given by

$$E[C_{i+1}|C_i, S_i] = S_i [1 - \exp[-\lambda C_i/(N-1)]] \quad (15)$$

Here both $C_i/(N-1)$ and λ are dimensionless, for they are, respectively, the probability that a contact was with an infective and the mean number of persons contacted per individual (i.e., mean count in a Poisson process).

Dietz and Schenzle [6] relate this to the standard Reed-Frost model by pointing out that if one sets $q = \exp(-\lambda/(N-1))$ then Equation (15) becomes the Reed-Frost model with $p = 1 - \exp(-\lambda/(N-1))$. At first, one might rationalize that the standard Reed-Frost model is a special case of the general model with a Poisson distribution for the number of contacts and with $p = 1 - \exp(-\lambda/(N-1))$. However, this requires that both $\lambda/(N-1)$ and $C_i \lambda/(N-1)$ must be dimensionless. To interpret $\lambda/(N-1)$ as a dimensionless number means that it must be read as $\lambda(1/(N-1))$, in which $1/(N-1)$ is a probability of picking a particular person out of $(N-1)$. The term $C_i \lambda/(N-1)$ is dimensionless because $C_i/(N-1)$ is the probability of picking an infective from the population. But then converting the exponential term in Equation (15) to q^{C_i} requires that C_i suddenly change roles, from dimensioned number of infectives in Equations (12) and (15) to a dimensionless count of trials in a probabilistic process. As Dietz and Schenzle point out, this interpretation leads to the result, $\lambda = r(N-1)$, where r is dimensionless and the Reed-Frost probability $p = 1 - \exp(-r)$

becomes independent of population size, which does not seem to be reasonable.

Thus, while in a formal sense it is possible to derive the Reed-Frost equation by assuming that $q = \exp(-\lambda/(N-1))$, this assumption leads to dimensional inconsistencies.

2. GEOMETRIC DISTRIBUTION

If the number of contacts follows a geometric distribution with mean λ , the expected value for C_{i+1} is

$$E[C_{i+1}|C_i, S_i] = \frac{S_i \lambda C_i / (N-1)}{1 + \lambda C_i / (N-1)} \quad (16)$$

3. BINOMIAL DISTRIBUTION

Because both the Poisson and the geometric distributions are discrete distributions on the positive integers and zero, they allow for unbounded number of contacts. One could argue that the number of possible contacts in one period must be bounded, and therefore the admissible distributions should be restricted to discrete bounded distributions on zero and the first n positive integers, for some n . The following is an argument for using the binomial distribution.

Suppose we divide the infectious period into n equal subintervals, each small enough so that a person cannot have more than one contact in a subinterval. Let p be the probability of a contact in one subinterval and assume independent Bernoulli trials. Then the distribution of number of contacts is the binomial distribution for n trials, and the generating function for the number of contacts is given by

$$f(x) = \sum_{k=0}^n \binom{n}{k} p^k q^{(n-k)} x^k = (px + q)^n \quad (17)$$

Therefore

$$f(1 - C_i / (N-1)) = (1 - pC_i / (N-1))^n \quad (18)$$

and the expected value of C_{i+1} is then

$$E[C_{i+1}|C_i, S_i] = S_i [1 - (1 - pC_i / (N-1))^n] \quad (19)$$

IV. DISCUSSION

Schenzle [12] compared the performance of a number of chain-binomial models in fitting one set of data on household epidemics of the common cold

[13, 14]. The Reed-Frost model did better than the Greenwood model, and a number of more general models did slightly better than the Reed-Frost model. The best performance was from a Reed-Frost type of model in which q was assumed to decrease linearly for successive generations. The performance of some of the generating function models with this data set has not been published. However, comparisons of a number of theories on one data set are fraught with difficulties; a model could do particularly well on one set and poorly on others. A comparison using many or at least a number of data sets would be more instructive.

From the analysis presented here it is clear that the Reed-Frost model lacks the generality of the generating function models and that it makes some rather special assumptions that are unreasonable. As a result it is not found to be a generally sound model though it might provide an approximation for certain types of epidemics.

I thank Herb Hethcote and Ira Longini for their critical review of a rough draft of this paper. Dr. K. Dietz kindly provided a translation of En'ko's papers. Drs. K. Dietz, J. Hazelrig, H. Hethcote, J. Kemper, J. Koopman, I. Longini, R. May, L. Sattenspiel, and D. Schenzle commented on the second draft of this paper. I am particularly indebted to K. Dietz for his detailed critique of the penultimate version of the paper.

This work was supported in part by grant RR02176-01A1 from the NIH-DRR.

Note added in proof: LeLevre, C. and M.-P. Malice, *Math. Modeling* 7:785–92, 1986, have used the generating function approach to handle random numbers of contacts for an S-I-S model.

REFERENCES

- 1 E. B. Wilson and M. H. Burke, The epidemic curve, *Proc. Nat. Acad. Sci. (USA)* 28:361–367 (1942).
- 2 R. Serfling, Historical review of epidemic theory, *Human Biol.* 24:145–166 (1952).
- 3 J. de O. C. Maia, Some mathematical developments on the epidemic theory formulated by Reed and Frost, *Human Biol.* 24:167–200 (1952).
- 4 H. Abbey, An examination of the Reed-Frost theory of epidemics, *Human Biol.* 24:201–233 (1952).
- 5 M. Greenwood, On the statistical measure of infectiousness, *J. Hygiene* 31:336–351 (1931).
- 6 K. Dietz and D. Schenzle, Mathematical models for infectious disease statistics, in *A Celebration of Statistics: The ISI Centenary Volume* (A. C. Atkinson and S. E. Feinberg, Eds.), Springer, New York, 1985 pp. 167–204.
- 7 P. D. En'ko, The epidemic course of some infectious diseases, *Vrac* 10:1008–1010, 1039–1042, 1061–1063 (1889) (In Russian. Cited by Dietz and Schenzle [6]).

- 8 E. Ackerman, L. R. Elveback, and J. P. Fox, *Simulation of Infectious Disease Epidemics*, C. C. Thomas, Springfield, Illinois, 1984.
- 9 H. E. Soper, The interpretation of periodicity in disease prevalence, *J. Roy. Stat. Soc.* 92:34–73 (1929).
- 10 K. Dietz, Personal Communication.
- 11 N. T. J. Bailey, *The Mathematical Theory of Epidemics* (Ch. 6) pp. 75–108, Charles Griffin & Co., Ltd., London, 1957.
- 12 D. Schenzle, Problems in drawing epidemiological inferences by fitting epidemic chain models to lumped data, *Biometrics* 38:843–847 (1982).
- 13 F. W. S. Brimblecombe, R. Cruickshank, P. L. Masters, D. D. Reid, and G. T. Stewart, Family studies of respiratory infections, *Brit. Med. J.* 1:119–128 (1958).
- 14 H. A. Heasman and D. D. Reid, Theory and observation in family epidemics of the common cold, *Brit. J. Prev. Soc. Med.* 15:12–16 (1961).