REDUCING MORTALITY WHEN DISEASES ARE INTERDEPENDENT**

by

H. N. Barnum*

78

CENTER FOR RESEARCH ON ECONOMIC DEVELOPMENT
The University of Michigan
Ann Arbor, Michigan 48109

Discussion Paper 78

August 1978

REDUCING MORTALITY WHEN DISEASES ARE INTERDEPENDENT**

H.N. Barnum*

R. Barlow*

August 1978

*Center for Research on Economic Development, University of Michigan

**This research was sponsored under a contract with the United States Agency for International Development.

ABSTRACT: REDUCING MORTALITY WHEN DISEASES ARE INTERDEPENDENT

H.N. Barnum and R. Barlow

Operational research studies in the area of public health often involve an objective function which predicts morbidity or mortality reduction to be achieved by proposed health programs. These studies often ignore important interrelationships among diseases which could have an impact on the evaluation of the cost effectiveness of alternative health activities.

This paper presents a simple approach to health modelling which allows for interaction among diseases. Also, to allow a clear distinction between the impact of preventative activities affecting morbidity and curative activities affecting case fatality rates, the model separates mortality into morbidity and case fatality components.

REDUCING MORTALITY WHEN DISEASES ARE INTERDEPENDENT

I. INTRODUCTION

Many operational research models in the health area are concerned with allocating resources among alternative health programs so as to minimize morbidity or mortality. In these applications the focus is on a single disease[1] or at a level of aggregation which ignores relationships among diseases. But these relationships often affect the outcome of specific health activities, and hence their cost-effectiveness.

Comparatively little work has been done on the modeling of interdependent diseases. A recent study by Correa is typical. In developing a series of models for use in health planning, he explicitly assumes that no interdependence exists on the grounds that "fortunately, when the plan refers to short time periods, this interdependence is not likely to be important."[2] But it would not be difficult to argue that interdependence might well be quite important, even in the short run. A few examples of the phenomenon can be given from the area of childhood diseases:

- gastro-enteritis leads to malnutrition, and vice versa, both conditions being major causes of child mortality in low-income countries.[3]
- measles often causes respiratory problems.[4]

[1] For instance, Feldstein, M., Piot, M. A. and Sundaresan, T. K., Resource Allocation Model for Health Planning: A Case Study of Tuberculosis Control, WHO, Geneva, 1973.

[2] Correa, H., Population, Health, Nutrition and Development, (1970). D.C. Heath, Lexington, Mass. p. 134.

[3] Heller, P. S. and Drake, W. D. Malnutrition, child mortality, and the family decision process, (1976). Center for Research on Economic Development, University of Michigan, Discussion Paper No. 58.

[4] Rashmi, A., Guha, D.K. and Khanduja, P.C. Postmeasles pulmonary complications in children, (1971). Indian Pediatrics 8, pp. 834-38.

- malaria produces a general debilitation, and increases the death rate from a variety of respiratory and digestive diseases.[1]

The purpose of this paper is to describe a simple approach to health modelling which allows for this kind of interdependence. First, some features of existing mortality risk models (which do not allow for interdependence) are reviewed. Next, a general approach to incorporating interdependence is outlined. A specific model based on two interdependent diseases is then developed, with some numerical illustrations. Finally, we discuss how the approach might be used in the area of public health planning.

II. MODEL OF COMPETING RISKS

Suppose that in a population of a given age the general death rate (from all causes) is 20 percent. The death rate from a certain specific disease is 5 percent. That disease is then eliminated. What happens to the general disease rate?

One approach to this question says that the answer will definitely exceed 15 percent. This is the theory of competing risks which points out that some of the 5 percent who would have died from the eliminated disease will now die from some other cause (during the period in question). Several models of competing risks have been developed. For example, refer to the models of Chiang and Kimball.[2]

---

[1] Newman, P., Malaria Eradication and Population Growth, (1965). Bureau of Public Health Economics, Ann Arbor, Michigan.

[2] Chiang, C.L., "On the probability of death from specific causes in the presence of competing risks," (1966). Fourth Berkeley Symp. IV, 162-80, and "Competing risks and conditional probabilities," (1970). Biometrics 26, 767-76; and, Kimball, A.W., "Models for the estimation of competing risks from grouped data," (1969). Biometrics 25, 329-37. For a summary of other models see Gail, M., "A review and critique of some models used in competing risk analysis," (1975). Biometrics 31, 209-222.

Interdependence has not typically been allowed for in models of competing risks. When one disease is eliminated, it is essentially assumed that the probability of catching any other disease is unaffected. With this assumption Chiang (1961) derives the following formulation for answering our opening question about death rates:

$$p_2' = 1 - (1 - p_1 - p_2)\ p_2 \ / \ (p_1 + p_2)$$

where p_1 = probability of death from Disease 1 (later eliminated)

p_2 = probability of death from all other causes, before elimination of Disease 1.

p_2' = probability of death from all other causes, after elimination of Disease 1.

(The notation follows that used in the version of Chiang's model presented by Kimball, 1969). Using our numbers, $p_1 = 0.05$ and $p_2 = 0.15$. Chiang's model therefore predicts that the general death rate after the elimination of Disease 1 (i.e., p_2') will be 0.154.

Kimball offers a constrasting model based on what he terms a "multinomial approach". Still assuming independence of risks, he derives the following formulation for the general death rate after the elimination of Disease 1:

$$p_2' = p_2\ (1 - p_1)$$

With our numbers for p_1 and p_2, p_2' equals 0.158, slightly different from Chiang's prediction.

In a comment on Kimball's contribution, Chiang (1970) recognizes the importance of interdependence between causes of death, and suggests that a proper method of allowing for interdependence is to specify

probabilities of morbidity as well as of mortality. Models of competing risks have been deficient in not specifying morbidity rates:

> " We must recognize the fact that the death of an individual is usually preceded by an illness (condition, disorder). It is not realistic to speak of a person's chance of dying from tuberculosis when he is not even affected with the disease. Also competition of risks of death depends on the health condition of an individual: a person affected with a disease (say, cardiovascular-renal (CVR) diseases) probably has a probability of dying of a second disease different from a person who is not affected with CVR. Therefore, a mortality study is incomplete unless illness is taken into consideration".

Following Chiang's suggestion, we develop in the next section an alternative approach to an analysis of mortality which explicitly incorporates morbidity. This second approach, which is based on the notion that many diseases are interdependent, raises the possibility that the answer to the question posed at the beginning of this section might well be less than 15 percent. Because of interdependence there may be a tendency for the death rate from the noneradicated diseases to decline, and this tendency may offset the opposite effect of competing risks on the general death rate.

III. A GENERAL APPROACH TO THE MODELLING OF MORTALITY RISKS WITH DISEASE INTERDEPENDENCE

The objective is to derive a procedure for the specification of a model which will predict the mortality effects of specific health improvements while allowing for interdependence between diseases. This can be done in four logical stages as follows:

a) Specify a morbidity function for each disease (or group of diseases), showing for the population in question the determinants of the morbidity rate:

$$p_i = p_i\ (A_i,\ p_j) \qquad i, j = 1,2,\ldots, n; \quad i \neq j \qquad (1)$$

where p_i is the fraction of the population (defined by age, sex, or other characteristics) who have disease i at some time during a defined time interval. There are n diseases. Each morbidity rate p_i is determined by a set of socioeconomic, environmental, and policy variables represented by the vector A_i and also by the prevalence of each other disease p_j. Degrees of interdependence are measured by the partial derivatives $\partial p_i/\partial p_j$. The impact of policy variables is measured by the partial derivatives $\partial p_i/\partial A_i$.

b) Define all possible combinations of diseases, and derive from the morbidity functions the fraction of the population having each combination:

$$c_k = c_k\ (p_i) \qquad k = \bar{1}23\ldots n,\ 1\bar{2}3\ldots n,\ 12\bar{3}\ldots n,$$

for a total of $(2^n - 1)$ such terms; $(\bar{1}23\ldots n)$ is read as "having all diseases except Disease 1" (2)

where c_k is the fraction with disease-combination k. If there are n diseases, there will be $(2^n - 1)$ desease-combinations. For example, in a two-disease system a person may have Disease 1 but not Disease 2,

may have Disease 2 but not Disease 1, or may have both diseases, for a total of three possible disease combinations. For an individual in the given population c_k can be interpreted as the joint probability of having disease combination k. Clearly the c_k are mutually exclusive (unlike the p_i), and the following relationship also holds true:

$$\sum_k c_k + c_w = 1 \qquad (3)$$

where c_w is the fraction of the population who remain well (or only have diseases excluded from the analysis) throughout the entire interval. In addition the generated rates must be consistent with the elementary postulates of probability theory.[1]

c) For each disease-combination, specify a fatality function, showing what determines the fatality rate among those with a given disease-combination:

$$f_k = f_k (B_k) \qquad (4)$$

where f_k is the fraction of those with disease-combination k who die during the specified time interval, and B_k is a vector of socioeconomic, environmental, and policy variables. Specifying fatality functions in this way

---

[1] See Dalkey, N. C., "An Elementary Cross Impact Model," Technological Forecasting and Social Change 3, American Elsevier, for a summary of the postulates which must be satisfied by a cosistent system of probabilities.

allows for the sort of phenomenon noted by Chiang in the quotation cited above: that a person who has cardiovascular-renal disease in combination with some other disease may have a probability of dying different from that of a person who has only that other disease.

d) Multiplying a fatality rate ($\underline{f}_k$) by the corresponding disease-combination rate ($\underline{c}_k$) then gives the fraction of the entire population who die from that particular disease-combination during the interval in question ($\underline{d}_k$):

$$d_k = f_k \cdot c_k \qquad (5)$$

And because the $\underline{c}_k$'s are mutually exclusive, the sum of the d_k's equals the general death rate D:

$$D = \sum_k d_k \,[1] \qquad (6)$$

Using this approach, we note that an intervention designed to reduce mortality, like mass vaccination against measles, will cause changes in the vectors $\underline{A}_i$ or $\underline{B}_i$. A simulation exercise with the model could therefore show the effects of the intervention on the general death rate D, with full allowance being made for both interdependence and competing risks. A heuristic advantage of the approach, it might be noted, is that public health interventions involving "prevention" affect $\underline{A}_1$, while those involving "treatment" affect $\underline{B}_i$.

---

[1] This derivation of D can be compared with that appearing in a series of health planning models developed by Correa (1975, pp. 117-69), whose expression for D, using our notation, is as follows:

$$D = \sum_{i=1}^{n} (f_i \cdot p_i)$$

But since the morbidity rates $\underline{p}_i$ are not mutually exclusive, this formulation obviously overstates $\underline{D}$.

IV. PROBLEMS IN SPECIFYING A MORBIDITY-MORTALITY MODEL

Several steps must be taken in order to use the approach outlined above for the specification of a morbidity-mortality model in a given application. First, a form must be chosen for the functions generating the marginal morbidity rates (equations 1). Second, assumptions must be made and a procedure developed which will allow the transition between the marginal disease rates to the mutually exclusive joint disease rates (equations 2). Third, a form must be chosen for the fatality rate functions (equations 4). The second of these steps presents some difficult conceptual problems.

a. Marginal Morbidity Rates

To derive a functional form for the marginal morbidity rates we commence with the general form (equation 1) which we assume to be continuous and differentiable and note that changes in the rate of disease i are related to changes in other disease rates and to changes in policy variables by the differential equation:

$$dp_i = \sum_{j \neq i} \frac{\partial p_i}{\partial p_j} \, dp_j + \frac{\partial p_i}{\partial A_i} \, dA_i \qquad (7)$$

Since the morbidity rates, p_i, are interpreted as probabilites, equations (1) and (7) must satisfy the following boundary conditions: if $p_i = 1$ or 0 then dp_i must equal zero for changes in p_j and A_i. It can be demonstrated[1] that an

---

[1] See Turoff, M., "An Alternative to Cross Impact Analysis," Technological Forecasting and Social Change, No. 3, American Elsevier, for a demonstration and the development of a cross impact analysis based on a logit specification of behavioral probability functions.

algebraicly simple specification which can satisfy these conditions is the logit,

$$p_i = \frac{1}{1 + \exp(-\gamma_i - \alpha_i A_i - \sum_{j \neq i} \beta_{ij} p_j)} \qquad (8)$$

where $\frac{\partial p_i}{\partial p_j}$ is a function of β_{ij}, and the policy impact derivatives, $\frac{\partial p_i}{\partial A_1}$, are functions of the elements of the vector α_i.

b. The Transition from Marginal to Joint Disease Rates

i) Consistency

In general it is not possible to go from a marginal distribution to a joint distribution without assumptions regarding the kind and degree of relationships among the events whose distributions are being considered. For example, the simplest assumption one might make is that the events are independent; it then follows that the joint rates can simply be obtained as the product of the marginal rates. However, if there are interdependencies among events, as there are with many disease events, the generation of joint relationships quickly becomes more complicated. One of the problems is that the probabilities generated must be consistent with the elementary postulates of the calculus

of probability.[1] Given a particular choice of functional form, the consistency requirement imposes relationships on the parameters of the marginal probability functions. In concept, consistency can be achieved either 1) by solving analytically for the required relationships between parameters or 2) by specifying functional forms which are consistent for all possible parameter values or 3) by the use of parameter values which are locally consistent in the context of a particular problem. When the model has more than two or three diseases, the analytical solution for consistent parameters or the derivation of consistent functional forms appears to be intractable. For our example we use the third alternative, that is parameter values and a functional form are chosen which are at least locally consistent.

ii) Calculation of Joint Disease Rates

The generation of the joint disease rates is accomplished through their decomposition into components distinguished by the presence or lack of a causal influence between diseases. The problem has been simplified by assuming that there is no mutual causation, i.e., the probability that i leads to j and j to i in a given individual over a given time period is zero. It is also assumed that joint causation is negligible, i.e. the probability that i and j together cause k is zero.

---

[1] The problem is analogous to one which occurs in the estimation of demand systems in economics. There the estimated system should be consistent with the elementary theorems of utility theory. In practice one of two procedures can be used, either 1) a system of equations yielding consistent solutions is used for estimation or 2) a less restrictive system of equations is used and the estimated parameters are tested for consistency.

Finally the probability of more than three diseases is set to zero under the assumption that the loss in accuracy is negligible. An exposition of the derivation of the joint disease rates for a four disease model is given in the appendix. The derivation for a two disease model is given in the next section.

V. AN EXAMPLE: A TWO-DISEASE MODEL OF INTERDEPENDENCE

We now specify an interdependence model with a view to providing a numerical illustration, choosing the simple case of only two diseases or disease-groups. The numbering of the equations in this specific model follows that used in the general model above. It will be noted that the most complex part of the specific model lies in the derivation of the disease-combination rates $\underline{c}_k$.

The morbidity rate for Disease 1 ($\underline{p}_1$) is determined by the rate for Disease 2 ($\underline{p}_2$) and by an index of other factors ($\underline{a}_{10}$):

$$p_1 = \frac{1}{1 + e^{-(a_{10} + a_{12}\, p_2)}} \tag{1a}$$

And similarly for Disease 2:

$$p_2 = \frac{1}{1 + e^{-(a_{20} + a_{21}\, p_1)}} \tag{1b}$$

$\underline{a}_{12}$ and $\underline{a}_{21}$ being indexes of interdependence. As we noted above, the particular functional form chosen for these equations (the logit form) has the property of containing the morbidity rates to lie between zero and one.

The disease-combination rates $\underline{c}_{\bar{1}2}$ and $\underline{c}_{1\bar{2}}$ are readily derived as follows:

$$\underline{c}_{\bar{1}2} = p_2 - c_{12} \tag{2a}$$

$$\underline{c}_{1\bar{2}} = p_1 - c_{12} \tag{2b}$$

where c_{12} is the fraction of the population having both diseases. If the two diseases were independent, it would be reasonable to assume that this fraction was equal to the product of the two morbidity rates:

$$c_{12} = p_1 \cdot p_2 \tag{2c}$$

But with interdependence, equation (2c) is not valid. With interdependence, $\underline{c}_{12}$ must be measured by adding together (i) cases where Disease 1 causes Disease 2 ($\underline{c}_{1 \to 2}$), (ii) cases where Disease 2 causes Disease 1 ($\underline{c}_{2 \to 1}$), and (iii) cases where both diseases occur but without any causal link between them ($\underline{c}_{(12)}$). The point is clarified in Table 1. Each disease may be caused by the other, not be caused by the other, or not be present at all. Considering both diseases, there are then nine categories for the population to fall into. Three of these are eliminated on logical grounds.[1] The remaining six are mutually exclusive.

---

[1] First, a person who "does not have Disease 1" cannot have during the same time interval "Disease 2 caused by Disease 1." (Lagged effects, as when Disease 1 occurring in one period causes Disease 2 to occur in the next, can be readily incorporated into a dynamic version of the model.) Second, a person who "does not have Disease 2" cannot have "Disease 1 caused by Disease 2." Third, cases of mutual causation are ruled out, where a person has "Disease 1 caused by Disease 2" as well as "Disease 2 caused by Disease 1." It is assumed that within the given time interval, one of the diseases will occur first, and cannot therefore be caused by the other.

By examination of Table 1 we therefore conclude that:

$$c_{12} = c_{1 \to 2} + c_{2 \to 1} + c_{(12)} \tag{2d}$$

Separate expressions must now be obtained for the three components of $\underline{c}_{12}$. The fraction of the population where Disease 1 causes Disease 2 can be measured as the <u>difference</u> between (i) the fraction actually having Disease 2 ($\underline{p}_2$) and (ii) the fraction who would have Disease 2 if Disease 1 did not exist (i.e. if $\underline{p}_1 = 0$). Symbolically:

$$c_{1 \to 2} = p_2 - \frac{1}{1 + e^{- a_{20}}} \tag{2e}$$

Similarly:

$$c_{2 \to 1} = p_1 - \frac{1}{1 + e^{- a_{10}}} \tag{2f}$$

The fraction of the population where both diseases exist without causal connection ($\underline{c}_{(12)}$) can be reasonably estimated as the product of two independent frequencies: (i) the frequency of having Disease 1 not caused by Disease 2, and (ii) the frequency of having Disease 2 not caused by Disease 1. That is:

$$c_{(12)} = (p_1 - c_{2 \to 1}) \cdot (p_2 - c_{1 \to 2}) \tag{2g}$$

The three disease-combination rates $\underline{c}_{\overline{12}}$, $\underline{c}_{1\overline{2}}$, and $\underline{c}_{12}$ having been determined, the next step is to specify the corresponding fatality rates $\underline{f}_{\overline{12}}$, $\underline{f}_{1\overline{2}}$, and $\underline{f}_{12}$. To simplify this version of the interdependence model, the fatality rates are assumed to be fixed. We then have equations for the disease-combination-specific death rates $\underline{d}_{\overline{12}}$, $\underline{d}_{1\overline{2}}$, and $\underline{d}_{12}$:

Table 1.

DISEASE - COMBINATIONS

WITH TWO INTERDEPENDENT DISEASES

Entries in the matrix represent the fraction of a population with the characteristics indicated

	Has Disease 1 caused by Disease 2	Has Disease 1 not caused by Disease 2	Does not have Disease 1	Row Sum
Has Disease 2 caused by Disease 1	0	$c_{1 \to 2}$	0	$c_{1 \to 2}$
Has Disease 2 not caused by Disease 1	$c_{2 \to 1}$	$c_{(12)}$	$c_{\bar{2}1}$	$p_2 - c_{1 \to 2}$
Does not have Disease 2	0	$c_{1\bar{2}}$	c_w	$1 - p_2$
Column Sum	$c_{2 \to 1}$	$p_1 - c_{2 \to 1}$	$1 - p_1$	1

$$d_{\bar{1}2} = f_{\bar{1}2} \cdot c_{\bar{1}2} \tag{5a}$$

$$d_{1\bar{2}} = f_{1\bar{2}} \cdot c_{1\bar{2}} \tag{5b}$$

$$d_{12} = f_{12} \cdot c_{12} \tag{5c}$$

and for the general death rate D:

$$D = d_{\bar{1}2} + d_{1\bar{2}} + d_{12} \tag{6a}$$

The operation of the model is illustrated with a numerical example in Table 2. It is imagined in Column (B) that information is available on the morbidity rates p_1 and p_2, for example from a morbidity survey of households. Information also exists on the partial derivatives $\partial p_1/\partial p_2$ and $\partial p_2/\partial p_1$, which can be interpreted as the likelihood that one disease will lead to the other (within the given time interval). This information may be obtainable from analysis of individual medical histories. Given p_1, p_2, and the two derivatives, the a terms in the morbidity functions can be derived, as well as the disease-combination rates (c_k).[1] With fatality rates (f_k) also supplied, the general death rate D can then be calculated.

In Column (c), the effects of eliminating Disease 1 are shown. To eliminate the disease, the value of a_{10} is changed to $-\infty$.

---

[1]The calculation of the a terms involves using the following for the two derivatives:

$$\partial p_1/\partial p_2 = a_{12} e^{-(a_{10} + a_{12} p_2)} \left[1 + e^{-(a_{10} + a_{12} p_2)}\right]^{-2} \tag{7a}$$

$$\partial p_2/\partial p_k = a_{21} e^{-(a_{20} + a_{21} p_1)} \left[1 + e^{-(a_{20} + a_{21} p_1)}\right]^{-2} \tag{7b}$$

Table 2.

HYPOTHETICAL EFFECTS OF DISEASE ELIMINATION ON MORTALITY IN A TWO-DISEASE SYSTEM WITH AND WITHOUT INTERDEPENDENCE

Values marked (*) are assumed; remaining values are derived from the equation(s) indicated

		With interdependence		Without interdependence	
	Equation(s) from which derived	Both diseases present	Disease 1 eliminated	Both diseases present	Disease 1 eliminated
	(A)	(B)	(C)	(D)	(E)
1. p_1	1a	.1500*	0	.1500*	0
2. p_2	1b	.4000*	.3488	.4000*	.4000
3. $\partial p_1/\partial p_2$	7a	.1000*	0	0	0
4. $\partial p_2/\partial p_1$	7b	.3500*	-	0	0
5. a_{10}	1a, 7a	-2.0483	$-\infty$*	-1.7346	$-\infty$*
6. a_{12}	1a, 7a	.7843	.7843*	0*	0*
7. a_{20}	1b, 7b	-.6242	-.6242*	-.4055	-.4055*
8. a_{21}	1b, 7b	1.4583	1.4583*	0*	0*
9. $c_{1 \rightarrow 2}$	2e	.0512	0	0	0
10. $c_{2 \rightarrow 1}$	2f	.0358	0	0	0
11. $c_{(12)}$	2g	.0398	0	.0600	0
12. $c_{\overline{12}}$	2a	.2667	.3488	.3400	.4000
13. $c_{1\overline{2}}$	2b	.0167	0	.0900	0
14. c_{12}	2d	.1333	0	.0600	0
15. $f_{\overline{12}}$		.2500*	.2500*	.2500*	.2500*
16. $f_{1\overline{2}}$		.3000*	.3000*	.3000*	.3000*
17. f_{12}		.4000*	.4000*	.4000*	.4000*
18. $d_{\overline{12}}$	5a	.0667	.0872	.0850	.1000
19. $d_{1\overline{2}}$	5b	.0050	0	.0270	0
20. d_{12}	5c	.0533	0	.0240	0
21. D*	6a	.1250	.0872	.1360	.1000

The other $\underline{a}$ terms remain the same as in Column (B), and the morbidity rate for the remaining disease is recalculated. The disease-combination rates are also affected, and the general death rate falls by 30 percent.

Columns (D) and (E) show what happens if the two diseases are independent. The same model is used (except that zero values are assumed for the interdependence terms $\underline{a}_{12}$ and $\underline{a}_{21}$), and the same initial values are assumed for the morbidity and fatality rates. The elimination of Disease 1 reduces the general death rate by 26 percent. As would be expected, this reduction is smaller than the result in an interdependent system.

Table 2 also suggests how to answer the question about death rates which was posed at the beginning of this paper. It is evident that the answer depends on how cases with both diseases are handled in the statistics on causes of death. If deaths of persons with both Disease 1 and Disease 2 are officially attributed to Disease 1, then the elimination of Disease 1 will cause the general death rate to fall by an amount <u>less than</u> the official death rate for Disease 1, regardless of any interdependence. With this kind of record-keeping, the death rate from Disease 1 is $(\underline{d}_{1\bar{2}} + \underline{d}_{12})$. The general death rate is always $(\underline{d}_{\bar{1}2} + \underline{d}_{1\bar{2}} + \underline{d}_{12})$. Eliminating Disease 1 reduces $d_{1\bar{2}}$ and d_{12} to zero but necessarily raises $\underline{d}_{\bar{1}2}$ (because of competing risks). Hence the fall in the general death rate will be less than $(\underline{d}_{1\bar{2}} + \underline{d}_{12})$, or the official death rate from Disease 1.

If, however, deaths of persons with both diseases are officially attributed to Disease 2, the elimination of Disease 1 could lower the general death rate by <u>more</u> than the official death rate for Disease 1.

Table 2 shows this happening both with and without interdependence. When the death rate for Disease 1 is defined as $\underline{d}_{1\bar{2}}$, the elimination of that disease will lower the general death rate ($\underline{d}_{\bar{1}2} + \underline{d}_{1\bar{2}} + \underline{d}_{12}$) by more than the death rate for Disease 1 if the rise in $\underline{d}_{\bar{1}2}$ (due to competing risks) is less than $\underline{d}_{12}$.

VI. APPLICATION OF THE INTERDEPENDENCE MODEL

The approach devised here is sufficiently general to be applied in a variety of contexts. Obvious applications exist in the area of public health planning. As an example consider a model of child mortality consisting of three diseases - malnutrition, diarrhea and measles. Suppose that a public health agency with given resources has the objective of minimizing infant mortality and that the agency knows how the $\underline{A}$ and $\underline{B}$ vectors in the morbidity and fatality functions will be affected by interventions (such as vaccination, nutritional programs, sanitation or clinical treatment) with given resource requirements. Programming techniques can then be applied to the model to determine what particular set of interventions would minimize the general death rate $\underline{D}$.

An advantage of the interdependence model is that it would give a more realistic evaluation of interventions affecting diseases which have low fatality rates (such as malnutrition) but which may be important contributors to higher morbidity rates for other diseases (such as diarrhea) with notably high fatality rates. The model would also differentiate between joint disease states with high fatality rates (such as malnutrition combined with measles) and other states (such as malnutrition or measles alone) with low fatality rates. Thus it might

be found that nutritional programs would lower the mortality from diarrhea and measles. And similarly, because of the simultaneous relationship which may exist between diarrhea and malnutrition it might be found that sanitation interventions would lead to lowered malnutrition and lowered mortality from measles.

A difficulty in the application of this type of model is the general lack of morbidity and mortality data for estimating the simultaneous relationships among diseases. Although the health literature contains considerable analysis of intervention impacts, it is only recently that potential simultaneous relationships between diseases have begun to be measured. The major obstacle to further analysis is that diagnostic records are needed which are more complete than those normally found. The morbidity data must specify all of the diseases present in the individual surveyed, and mortality data must include all of the diseases present at the time of death. The conventional practice of recording only a single cause of death is not helpful for the analysis of interdependence between diseases.

APPENDIX

Algebraic Summary of the Morbidity - Mortality Model

This appendix presents another version of the general morbidity-mortality model discussed above. This version is developed for four diseases, and leaves the morbidity and fatality functions in general form. The objective in the design of the model is to allow the calculation of a set of mortality rates for mutually exclusive morbidity categories which include all possible combinations of diseases. There are four distinct segments to the system. The first involves the specification of a set of marginal probabilities as functions of interventions as well as other variables. This segment can be referred to as that of the structural (or causal) morbidity equations. The second segment is a set of algebraic relationships that allows the calculation of the probabilities of being in the mutually exclusive joint disease states. The third segment involves the specification of fatality probabilities for each of the mutually exclusive categories. This segment can be called the structural fatality equations. Finally, the fourth segment uses the fatality rates and joint morbidity rates to calculate the mortality rate.

I. Structural Morbidity Equations

The model is developed for four diseases. The extension to n diseases is straightforward although tedious. We specify the following marginal morbidity functions for diseases 1, 2, 3 and 4.

$$p_1 = p_1 (A_1, p_2, p_3, p_4) \qquad (1a)$$

$$\vdots \qquad \vdots$$

$$p_4 = p_4 (A_4, p_1, p_2, p_3) \qquad (1d)$$

II. Calculation of Mutually Exclusive Joint Disease Rates

a) Preliminary defnitions:

i. p_{ij} is the probability of having $\underline{i}$ and $\underline{j}$ jointly.

ii. p_{ijk} is the probability of having $\underline{i}$, $\underline{j}$ and $\underline{k}$ jointly.

iii. $\tilde{p}_{ij}$ is the probability that i occurs caused by j, that is,

$$\tilde{p}_{ij} = p_i\ (A_i, p_j \cdot \cdot \cdot p_\ell) - p_i(A_i, p_k \cdot \cdot \cdot p_\ell)$$

iv. $\tilde{p}_{ijk}$ is the probability that i occurs caused by either j or k, that is,

$$\tilde{p}_{ijk} = p_i\ (A_i, p_j \cdots p_\ell) - p_i\ (A_i, p_\ell)$$

v. p^*_{ij} is the probability that i occurs and is not caused by j, that is,

$$p^*_{ij} = p_i - \tilde{p}_{ij}$$

vi. p^*_{ijk} is the probability that i occurs and is caused by neither j nor k, that is,

$$p^*_{ijk} = p_i - \tilde{p}_{ijk}$$

b) Preliminary assumption:

It is assumed that intersection probabilities of order higher than three are zero, i.e. $p_{1234} = 0$.

c) Preliminary calculation: two-disease intersections --

There are two kinds of cases where two diseases can occur simultaneously:

i) no causal links between diseases

ii) the two diseases are linked causally (either i causes j, or j causes i, or both).

Thus the two-disease intersection for i and j can be written using the notation introduced above as

$$p_{ij} = \tilde{p}_{ij} + \tilde{p}_{ji} - \tilde{p}_{ij} \cdot \tilde{p}_{ji} + p^*_{ij}\, p^*_{ji}$$

d) Preliminary calculation: three-disease intersections -- there are eight kinds of cases where the three diseases can occur simultaneously:

		Diagrammatically
i.	no causal links between the 3 diseases	
ii.	i and j are causally linked (either i causes j, or j causes i, or both), but k is not causally linked with either i or j	
iii.	i and k linked, j independent	
iv.	j and k linked, i independent	
v.	i and j linked, also i and k, but not j and k	
vi.	i and j linked, also j and k but not i and k	
vii.	i and k linked, also i and k, but not i and j	
viii.	i and j linked, also i and k, also j and k	

The eight groups being mutually exclusive, p_{ijk} is the sum of all eight. Keeping the eight groups in order, and defining

$$x_{ij} = \left[\tilde{p}_{ij} + \tilde{p}_{ji} - \tilde{p}_{ij} \cdot \tilde{p}_{ji}\right]$$

the three-disease intersection can be written as the sum of eight terms.

$$p_{ijk} = p^*_{ijk} \cdot p^*_{jik} \cdot p^*_{kij}$$

$$+ x_{ij}\, p^*_{kij}$$

$$+ x_{ik}\, p^*_{jik}$$

$$+ x_{jk}\, p^*_{ijk}$$

$$+ x_{ij}\, x_{ik}\, (1 - x_{jk})$$

$$+ x_{ij}\, x_{jk}\, (1 - x_{ik})$$

$$+ x_{ik}\, x_{jk}\, (1 - x_{ij})$$

$$+ x_{ij}\, x_{ik}\, x_{jk}$$

e) Given the preliminary calculations above, the rates for the mutually exclusive disease states can be derived as follows:

i. $c_{1\overline{234}} = p_1 - p_{12} - p_{13} - p_{14} + p_{123} + p_{124} + p_{134} - p_{1234}$

$c_{2\overline{134}} = p_2 - p_{12} - p_{23} - p_{24} + p_{123} + p_{124} + p_{234} - p_{1234}$

.
.
.

and so on for $c_{3\overline{124}}$ and $c_{4\overline{123}}$.

ii. $c_{123\bar{4}} = p_{12} - p_{123} - p_{124} + p_{1234}$

.
.
.

and so on for $c_{13\bar{2}4}$, $c_{14\bar{2}3}$, $c_{23\bar{1}4}$, $c_{24\bar{1}3}$, and $c_{34\bar{1}\bar{2}}$.

iii. $c_{12\bar{3}\bar{4}} = p_{123} - p_{1234}$

.
.
.

and so on for $c_{12\bar{4}\bar{3}}$, $c_{13\bar{4}\bar{2}}$ and $c_{23\bar{4}\bar{1}}$.

iv. the states defined above + p_{1234} + p_{well} will sum to one. (We have assumed that p_{1234} is negligible.)

III. Structural Fatality Functions

Fatality functions are specified for each of the joint disease states.

$$d_1 = d_1(B_1)$$

.
.
.

$$d_H = d_H(B_H)$$

where H is the total number of disease states ($H = 2^n - 1 = 15$).

IV. The Mortality Rate

The mortality rate can be derived as the sum of the mortality rates for the individual joint disease states.

$$D = \sum_{h=1}^{H} d_h \, c_h$$

REFERENCES

Chiang, C. L. [1961]. On the probability of death from specific causes in the presence of competing risks. Fourth Berkeley Symp. IV, 162-80.

Chiang, C. L. [1970]. Competing risks and conditional probabilities. Biometrics 26, 767-76.

Dalkey, N. C. An elementary cross impact model, Technological Forecasting and Social Change 3, American Elsevier.

Feldstein, M., Piot, M.A. and Sundaresan, T.K. [1973]. Resource Allocation Model for Health Planning: A Case Study of Tuberculosis Control, WHO, Geneva.

Gail, M. [1975]. A review and critique of some models used in competing risk analysis, Biometrics 31, 209-222.

Heller, P. S. and Drake, W. D. [1976]. Malnutrition, child morbidity, and the family decision process. Center for Research on Economic Development, University of Michigan, Discussion Paper No. 58.

Kimball, A. W. [1969]. Models for the estimation of competing risks from grouped data. Biometrics 25, 329-37.

Newman, P. [1965]. Malaria Eradication and Population Growth. Bureau of Public Health Economics, Ann Arbor, Michigan.

Rashmi, A., Guha, D. K. and Khanduja, P. C. [1971]. Postmeasles pulmonary complications in children. Indian Pediatrics 8, 834-38.

Turoff, M. An alternative to cross impact analysis, Technological Forecasting and Social Change, No. 3, American Elsevier.