PREDICTIONS OF THE EFFECTS OF ENERGY PRODUCTION ON HUMAN HEALTH

R. T. Lundy and D. Grahn, Argonne National Laboratory

In recent years, it has been discovered that there are certain risks to human health associated with various facets of an affluent industrialized society. As a a consequence regulations have been promulgated with the intent of protecting our health, and numerous studies have been done to determine whether additional protection is needed. As many of the effluents from energy systems are among those identified as hazardous, those who must plan and analyze various energy options must perforce take into account the health effects anticipated in any situation being considered.

Some effluents are now regulated; others may be in the future. Pressures are occasionally brought to bear from industry to relax or eliminate regulations once instituted. The person trying to foresee the effects of a given policy needs to have some way to guess the likely course of future regulations, which are among the major economic and engineering constraints that must be considered. Future regulations may in part be projected on the basis of health effects. Also, there are economic tradeoffs to be made, whatever the constraints, and these, too, require a realistic estimate of the health consequences. All of this analysis requires an appropriate quantitative model.

For purposes of analysis, it is desirable to be able to project what will happen; how much of it will happen; when it will happen; and to whom it will happen. This information can be expressed from several perspectives. The most important of these are the "personal" and the "real population" perspectives. The "personal" perspective expresses risks as seen by an individual -- "what will happen to my personal chances of survival?" It is from this point view that insurance premiums are (ideally) calculated. This is in many cases a useful point of departure, but it carries with it some important assumptions which are not always apparent: For example, it assumes that , if your expected days in the hospital are raised by 20%, that the hospital facilities will be available for your use and the doctors will be there to treat you. Such an assumption may be reasonable if we are talking about a relatively small occupational group within the larger society, in which case the situation would fall within the normal variation in the usage of the facilities available to society as a whole. The same assumption may not be reasonable when the group at risk is essentially the whole society. To deal effectively with that case, we must look from the perspective of the "real population" to determine from society's point of view what the potential demand for health-related facilities might be. This determination is made by integrating the "personal" risks over the distribution of persons at various types of risk. At this level, estimates are often made by defining a single does-response coefficient and applying it to an estimate of the total population at risk. This, however, requires that the distribution of persons within the population at risk remain constant. This is not a safe assumption. Also, the use of an independently derived population estimate can lead to the theoretical death of the same person more than once in the course of a projection.

Frequently a dose-response coefficient derived from one population is applied to another population whose composition and characteristics are so different that the results become unreliable. For example, such an error would involve projecting health effects in a general population by using dose-response coefficients from a study of asthmatics. Although such a blatant error has never to our knowledge been made, a more subtle form of this error occurs whenever dose-response coefficients derived from one population are applied to a population whose age profile differs significantly from the one from which the coefficient was derived. This error can occur even if the population appears at first glance to be the same. For example, a study examining hypothetical health effects expected in the population of the U. S. in 1970 would make such an age distribution error if the dose-response coefficients used had been generated from the U. S. population in 1960, since the age distribution shifted markedly in that decade as a result of changing fertility levels over the previous 40 years. Consequently, a model must carry out two functions:

1. It must project the response to an exposure as a function of level and duration of exposure, and of the age, sex, and any other predisposing factor associated with a definable class of person.

2. It must project the distribution of such people during the period of time to be covered by the analysis.

Projecting Distributions of Persons at Various Levels of Risk: The Demographic Module.

Most major risk factors are associated

with age and sex. The susceptibility of most people to the ill effects of exposure to a toxic material tends to increase exponentially with age though there is an additional peak in susceptibility in the first year of life. Thus, one can go a long way towards projecting the risk level distribution simply by projecting the age and sex distribution.

The problem of projecting the future population has concerned demographers for over a century. A number of procedures, some of them quite sophisticated, have been devised to deal with it. The most appropriate procedure for any particular population will depend on its particular characteristics. However, for illustrative purposes, the component projection model developed by Whelpton, generalized by Leslie, and described by Keyfitz (1), will be presented here. The procedure by which the model is extended to project deaths as well as living population may be applied to any projection scheme.

Let

x = exact age index.

n = the length of an age interval or projection interval.

i = the age group definition index.

$= \text{trunc}\left(\frac{x}{n}\right) + 1$

${}_s k_i^t$ = population in age group i and sex group s.

${}_s S_i$ = probability of survival from age group i to age group i+1 during an n-year interval.

${}_s F_i$ = expected number of children that will be born to a woman starting in age group i during n years.

We can then assemble the ks into a column vector of population K and the F and S terms into a square matrix M such that the population vector at time t+n is related to the population vector at time t by

$$\underset{\sim}{K}^{t+n} = \underset{\sim}{L} \times \underset{\sim}{K}^{t} \qquad (1)$$

The cells in the projection matrix are customarily estimated by assuming that the age distribution within each age group is similar to that in a stationary population in which case the subdiagonal survival terms are given by

$$S_i = \frac{{}_n L_{x+n}}{{}_n L_x}$$

where ${}_n L_x$ = numbers in life table age age distribution aged x to x+n.

The reproductive terms are given by

$$F_i = 2.5\ ({}_n b_x + S_i\ {}_n b_{x+n}) \qquad (2)$$

where ${}_n b_x$ is the yearly birth rate for women age x to x+n.

By interdicting the computations at the appropriate point, it is possible to estimate the distribution of deaths by age group as well. The total deaths in a cohort starting in age group i and surviving to age group i+1 is given by

$$d^*_i = k_i - k_{i+1} \qquad (3)$$

We can define a factor Z_i such that

$Z_i \cdot d^*_i$ people die in age group i

$(1.0 - Z_i) \cdot d^*_i$ people die in age group i+1

If we assume that the age distribution of the deaths as well as the population within each 5-year age group during the passage between one age group and the next is the same as in the life table, then

$$Z_i = \frac{{}_n d_x}{({}_n d_x + {}_n d_{x+n})} \qquad (4)$$

where ${}_n d_x$ is the number of life table deaths in age group x to x+n.

The number of deaths in age group i during the projection interval is then

$$D_i = (1 - Z_{i-1})\ d^*_{i-1} + Z_i\ d^*_i. \qquad (5)$$

Projecting Changes in Health: The Dose/Response Module.

It is in the area of dose/response relationships that most other modeling efforts are concentrated. A dose/response function is a relationship between the degree of exposure to a toxic substance and the degree of excess risk that can be observed as a consequence of that exposure. In its simplest form, the function states that:

$$du = Bdp,$$

where

u = risk of death

p = exposure index

B = proportionality factor

It is apparent, however, that the change in the risk of death as a consequence of any given change in

exposure will not be the same for all persons exposed. Also, the consequences of the pollutants with which we will be dealing tend to show a prolonged latent period before the full effects can be seen. Consequently, the function should be disaggregated to whatever degree is necessary to assure reasonable homogeneity within groups, and it should be made duration-specific as well. At the current stage of development , this disaggregation is limited, as is the demographic module, to age (in 5-year groups) and sex.

Let us now focus our attention on the response function, and the effluents to which it refers.

In the area of energy production and public health, one class of effluents is of particular importance: airborne combustion products and the by-products which they give rise to in the course of their travels through the atmosphere. There is a bewildering array of them, and almost all can be found in any given sample of polluted air. For purposes of analysis, however, most investigators have chosen to index air pollution levels on one or two of the more prominent, easily measured, or otherwise interesting components. The most commonly used of these are total suspended particulates (TSP), sulfur dioxide (SO_2), or suspended sulfates (SO_4). The next thing traditionally done in such studies is to focus attention on a carefully selected subgroup of the population, usually chosen on the basis of a compromise between high a priori susceptibility and large numbers.

Most existing models had their origins in studies in which the major interest was in the derivation of qualitative estimates of relationships (e.g., is SO_2 bad or isn't it?), or in estimating in retrospect what the cumulative quantitative effects had been. Epidemiological models especially tend not to consider explicitly that the composition of the population being studied can (and usually will) change markedly with time. They generally refer to populations defined so broadly that their internal structure can change drastically with respect to many factors often confounded with pollution-related health effects (i.e., age, socioeconomic status, % suffering from morbid conditions, etc.), while still remaining within the original definition of the study population (e.g., "total," "whites 35 years of age and over," "employees hired in 1950-55," etc.)

There is, however, one major source of air pollution associated with combustion products that has been studied very throughly indeed: the cigarette. It is not, of course, usually considered in the context of fossil energy sources, although the number of BTU's of cigarettes burned each hour in the United States is the approximate equivalent of 12-15 tons of coal, an amount great enough to operate a 26 MWe power plant.

Unlike the epidemiological studies of air pollution, in which neither duration nor magnitude of exposure are easily measured, the investigators of smokers have been able to do reasonably well-controlled prospective studies in which age at onset, degree of exposure, and outcome are all defined with reasonable accuracy. Assuming cigarette smoke, then, to be just another air pollutant, let us look at the function relating increments in age-specific death rates to exposure measured in number of cigarettes smoked per day, as shown in Table 1. The same data are graphed against age in figure 1.

We note that above age 50, the semilog plot of the response curve constitutes for both sexes (males particularly) a reasonably straight line indicating a constant exponential increase in damage, while below that age the curve drops away from this line and presumably would, if extended properly, hit 0 at around the mean age at which each sex begins to smoke. This is close to 15 for males and 20 for females. Why should this be so?

If we assume any of several models indicating that the ability of mammalian organisms to withstand the ravages of their environment declines in inverse proportion to their age, the familiar Gompertz law of exponentially increased risk would be expected

$$g(x) = a\, e^{bx} \qquad (6)$$

On the other hand, it can be shown that whenever some increment of damage occurs to an organism, various repair mechanisms are brought into play. In a situation of constant exposure to a toxic agent, the amount of repair taking place tends to rise in direct proportion to the damage accrued. Under this assumption, one would expect the damage function to rise asymtotically to some constant value as the incremental damage and repair effects reached equilibrium over a period of time. Under this assumption, furthermore, the change in the damage function would follow the logistic function

$$v(x) = \frac{1}{1 + ce^{-d(x-x_o)}} \qquad (7)$$

where x_o is the age at onset of exposure.

Both effects would appear to be operating simultaneously in the present case. The constant exposure to the toxic agent would be initiating a process

whereby the damage function would attempt to rise over a period of several years to an equilibrium value; at the same time, however, this equilibrium value would be changing with the advancing age of the exposed organisms according to the Gompertz law. Hence, the damage function, which describes the data of Table 1, ought to have the form

$$B(x,x_o) = \frac{a\, e^{bx}}{1 + c\, e^{-d(x-x_o)}} \qquad (8)$$

This function has been fitted to the cigarette data, as shown in Table 2, and is shown superimposed on the data points in figure 1.

Applicability of the Cigarette Model to Other Forms of Air Pollution

Cigarette smoke and coal smoke differ markedly in some respects. In particular, carbon monoxide is found in far higher concentrations in cigarette smoke than in coal smoke. (Its presence indicates inefficient combustion, anathema to engineers.) The cigarette model can be justified, however, on two grounds: First, the kind of damage done by air pollutants is not specific by causative agent; sulfuric acid droplets, fly ash particles, NO_2, O_3, and SO_2 all cause the same kinds of damage to the lung in appropriate concentrations, as indeed do most of the aldehydes, ketones, and other noxious organics likely to be encountered under similar circumstances. Second, the response curves for smoking seem to fit those for air pollution data reasonably well.

It will be noted in the fit of the cigarette data given in Table 2 that the age at onset of exposure is around age 15 for males and 20 for females. When one is dealing with other airborne pollutants, however, it is immediately clear that no decision on the part of the person involved, other than migration, will prevent exposure from commencing at birth.

During the time when the data for the most recent studies of air pollution and health were collected, it is probably fair to assume that the then current level had prevailed long enough for the latency effects to have worked themselves out some time previously. Therefore, a response function was calculated from the cigarette model of equation 13 assuming constant exposure to the effluent of interest from birth. For comparative purposes, the response function for the case where SO_2 and TSP are incremented equally for whites on the basis of the regressions derived by Lave and Seskin and presented in Finch and Morris (4) were plotted against the mean of the age groups considered in their analysis as derived from the 1960 U. S. population. Figure 2 compares the cigarette function thus adapted with a plot of the Lave-Seskin points.

It can be seen that while the response pattern seems to fit very well for males, the fit for females is not as close. Three possible explanations suggest themselves : First, the exposure data used in the Lave analysis may not be as well matched to the female population in his sample as it is to the male population. There is reason to believe that the males in the SMSAs (Standard Metropolitan Statistical Areas) treated were more likely than the females to spend a significant part of their day in areas close to the locations of the sampling stations from which the the exposure data were derived. These stations were for the most part in the central urban areas, and males are more likely to commute into these areas for work than females, whose lower labor force participation rates and and differing array of employment opportunities would tend to keep them out of the relatively more polluted areas. Second, it might well be the case that there is a strong interaction between the effects of air pollution and smoking history, in which case the later onset of smoking among females might contribute significantly towards the pattern seen here. Third, one of the points in the cigarette data, specifically the one for females in the 70-79 age group may be a spurious point. If this point is eliminated from the calculations, the resulting function fits the air pollution data far more closely, although it seems not to fit the cigarette data quite as well at the lower ages. The curve derived when this point is eliminated is shown as the dashed line in figures 1 and 2.

Fitting the Cigarette-Derived Model to Air Pollution Data

A number of studies have been done that give response coefficients for various population subgroups exposed to various pollutants. We would now like to fit these data into the framework of our model. Fitting is most conveniently done by simulating the particular study and determining the cigarette-equivalent dose needed in the current model to reproduce the effect of a given pollutant dose. For example, Finch & Morris (4) have determined that the response function implied in Winkelstein's study of air pollution in Buffalo, NY, as indexed by Total Suspended Particulated (TSP) for white males 50-69 years of age is about 14 deaths per 10^5 population per ug/m^3/day incremental long-term exposure. Starting with the life table and age distribution

of U.S. white males in 1960, one finds that an assumed increment of 0.35 cigarettes per day will have the same effect in that age group. Consequently, to convert the cigarette model into a TSP model one needs only multiply the a coefficient in Table 2 by .35. Similar fits for data from other studies are given in table 3.

This fitting procedure yields a further dividend in that it gives us a method for projecting the results of some studies beyond their original age boundaries.

Merging the two components

The complete model, then, operates as follows: The exposure level of the index pollutant and the initial population are both defined. Then the population is projected forward in time, with the projection matrix being modified at each cycle according to the dose-response function.

Example

The question might be raised, what advantage do we derive from using such an elaborate model? How will its results differ from those obtained with one of the simpler methods, e.g., OBERS or other projections of the total population size, and using a simple response coefficient? How important are these distributional factors? Table 4 compares the results obtained with a single-coefficient procedure and with the one proposed here. Estimates were calculated on the assumption that fertility levels in the 30-state region would be the same as the 1971 level through the year 2020, and that the 8.95 $\mu g/m^3$ increment in suspended SO_4 was instituted in 1970. The pattern of deviations between the two systems is striking. The simple model grossly overestimates the number of excess deaths in 1985 due to the latency factor. In 2000, the latency effect has passed, and, by coincidence, the age distribution estimated for that year leads to a reasonably close concordance between the results of the two models. By 2020, however, the simple coefficient suggest 23% fewer excess deaths than does the model because the population at that time will have a decidedly older average age profile than in either 2000 or in 1960, the time at which the current age distribution was used to fit the simple coefficient to the response function.

Discussion

We have defined here a model system for projecting the excess mortality that might be observed in a population exposed to an increment of environmental insult. It avoids many of the pitfalls found in most current approaches.

The system here presented is not meant to be the last word on the subject. Among the features not considered here, but which deserve attention, are:

- The effects of constantly changing exposure levels on the response function.
- The effects of migration into and out of a polluted area.
- The effects of reductions, as opposed to increases, in exposure levels. Cigarette data would suggest that for some phenomena, particularly cardiovascular disease, the recovery rate once exposure has ceased is far faster than would be anticipated on the basis of the current equations. This phenomena could have a strong impact when investigating the policy implications of tightening air quality standards.

References

1. Keyfitz, Nathan, Introduction to the Mathematics of Population, Addison-Wesley, New York, 1968.
2. Lave, Lester, and E. P. Seskin, table cited in Finch & Morris below.
3. Carnow, B. W., and P. Meier, "Air Pollution and Pulmonary Cancer," Archives of Environmental Health, vol. 27, Sept. 1973, pp. 207-218.
4. Finch, S. J., and S. C. Morris, Consistency of Reported Health Effects of Air Pollution, Brookhaven National Laboratory, BNL-21808.
5. Hammond, E. C. "Smoking in Relation to the Death Rates of One Million Men and Women," in W. Haenzel, ed., Epidemiological Study of Cancer and Other Chronic Diseases, NCI Monograph 19, Washington, D.C., 1966.

Table 1
Effects of smoking on death rates for both sexes by age. Data derived from Hammond(5) especially appendix tables 2 and 3.

Age Group	Mean Cigs/day All Smokers Female	Male	Increase in Death Rate/Cig. Female	Male
35-39	20.6	28.5	0.34	3.9
40-44	20.3	28.8	1.3	7.1
45-49	20.0	28.9	3.1	14.6
50-54	19.5	28.6	6.0	19.2
55-59	18.7	27.3	9.5	33.3
60-64	17.6	25.4	12.2	46.1
65-69	16.4	23.4	28.8	72.3
70-74	14.9	21.0	51.1	94.6
75-79	14.2	18.0	19.8	139.2
80-84	12.0	17.4	172.4	188.3

Table 2
Fitted coefficients of equation using data of table 1.

Coef.	Females	Males
a	6.24×10^{-7}	9.14×10^{-6}
b	8.84×10^{-2}	6.44×10^{-2}
c	100.0	100.0
d	0.2	0.2
x_o	20	15

Table 3
Coefficients to convert $\mu g/m^3$ pollution exposures into cigarette/day equivalents.

Study	Index Pollutant	Crude Response	Conv. Coef.
Winkelstein	TSP	.00014	.35
Morris & Novak	SO_4	.000033	.21
Lave & Seskin	TSP	.835	.09
	SO_2	.715	

Table 4

Projected premature deaths in a hypothetical population with a resemblance to that of the North Central and Northeastern regions in 1985, 2000, and 2020 assuming that the mean suspended sulfate exposure rises to 8.95 $\mu g/m^3$ of air starting in 1970.

Year	Est. Pop. $\times 10^6$	Simple Estimate (Morris & Novak)	Model Estimate
1985	201	48,000	12,000
2000	225	66,400	62,500
2020	256	75,600	92,800

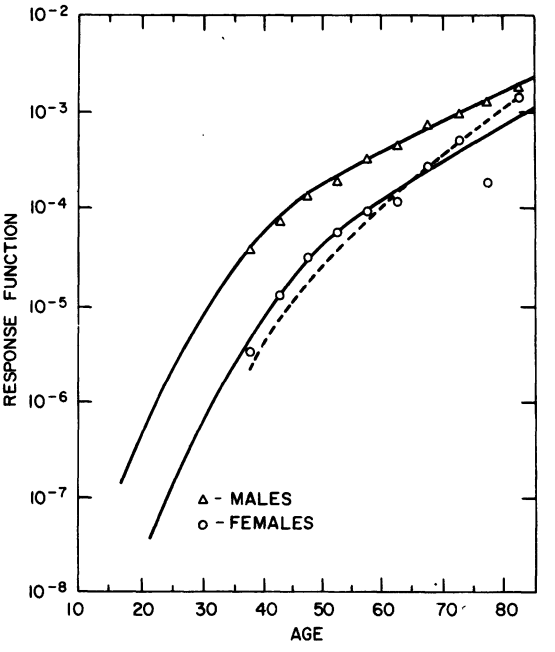

Figure 1

Increment in death rates per cigarette plotted with fitted response function.

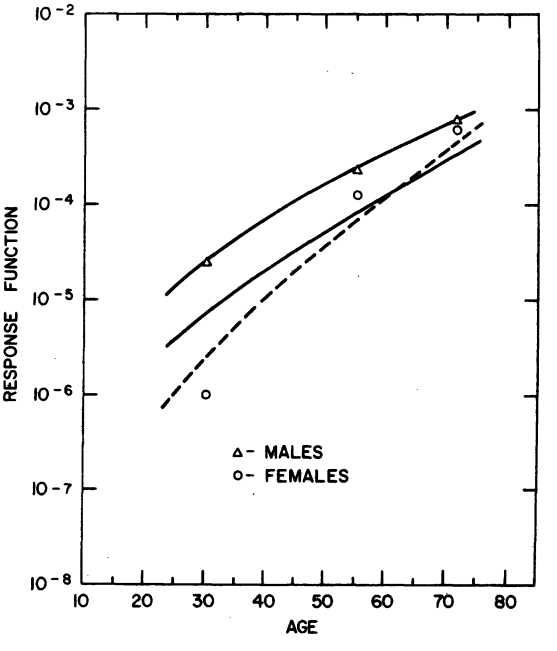

Figure 2

Increment in death rates per unit of polluted air plotted with adapted cigarette response function.