Environmental Health Perspectives
Vol. 42, pp. 33-38, 1981

Extrapolation from Incomplete Data to Total or Lifetime Risks at Low Doses

by Marvin A. Schneiderman*

Both epidemiology and laboratory data can contribute to estimates of risks to humans of exposure to low doses of carcinogens. The sum of all these contributions does not permit us to make these estimates with certainty. In chronic disease epidemiology, in looking for possible excessive cancer risks, we sometimes fail to have an adequately long observation time or to observe a population sufficiently aged for cancers to appear in meaningful numbers. In studies of most human exposures, dose data are often lacking, beyond a vague "yes-no" or "lots, not much, hardly any." Thus, without a knowledge of what dose produced an observed result it becomes logically impossible to know what result some other (presumed) dose might yield.

Animal data show some promise of being useful in extrapolating to low doses in man. However, several problems exist: (a) man is not a tailless, two-legged mouse, or featherless chicken—that is, we do not know if man is more or less sensitive than the laboratory animal; (b) the mathematical model used for extrapolation leads to large differences in estimates of response; (c) man is genetically heterogeneous and is usually exposed to many more hazards than is the laboratory animal.

Thus, existing data, even from well-done studies, are inadequate if we want to make extrapolations in any detail or to apply to specific subgroups in the population. Any risk estimation we do may have to be stated in terms that point out the wide ranges of the estimates.

The theme of this paper derives from the lines in the spiritual that say, "Nobody knows all the troubles I've seen. Nobody knows but Jesus." I would claim this as my personal lament, except for the fact that I know that anyone who has attempted to conduct studies in the epidemiology of cancer shares it with me.

The issues as I see them are these: (1) What can we learn (or infer) about some humans from observing other humans living under different circumstances? (2) What can we learn (or infer) about humans, in general, from observing other mammals, or birds, or fish, or tissues, or cells, or subcellular particles?

The answer to both these questions is "a lot, but not nearly enough." The major problem that I see is an "incomplete data" problem and that comes, in large part, from our impatience. We want answers about lifetime effects from less than lifetime observations. We are not in the position of the materials testing engineers who can do accelerated life testing. We may speak of burning the candle at both ends, but we cannot speed up the life process for testing purposes in humans. If "senility" occurs in 20% of all persons 80 years old or over we do not think we will find 2% of 8-year-olds showing senility. Yet sometimes all our person-years of observation help us forget that biological processes ranging from pregnancy to cancer to senility have a minimum "maturation time." We know that nine women one month pregnant do not equal one woman pregnant nine months. Sometimes we forget that observing 400 people for one year when we are concerned with cancer may give us much less information than observing 20 persons for 20 years. Since I am concerned with cancer, I must deal with this maturation phenomenon. If the median age at diagnosis of bladder cancer is 68 years, I will be most unlikely to find anything of consequence in any current study about cancer in the "second generation" children now age 10-15 who are consumers of "diet" drinks, and whose mothers were also consumers of "diet" drinks. In fact we are scarcely likely to learn anything much about the mothers and cancer, considering when

*Clement Associates Inc., 1010 Wisconsin Avenue, N.W., Washington, D.C. 20007.

the large increases in the use of artificial sweeteners first started in this country.

We are concerned here today with dose-response phenomena, and I will try to deal largely with that. An obvious early question one asks about dose–response phenomenon in man is, "Is there a safe level?" This is another way of asking the question, "Is there a threshold for a carcinogen?" About two years ago, Charles Brown, Pierre Decouflé, and I assembled some evidence on this issue (*1*). Decouflé reviewed ten published occupational studies reported to have yielded "negative" results, i.e., to have demonstrated a threshold at the highest doses. He assembled some background data on these studies. These serve as examples of some of the problems epidemiologic studies can run into. Table 1 gives some characteristics of these studies.

The major deficiencies, or complaints of devidiencies, are obvious. Almost all the studies were short and thus did not allow for latent periods in cancer or were based on too young a population—a population unlikely to show any but very large effects—or there was dilution of the exposed group by a large number of (presumed) unexposed persons, or the authors really didn't know who was exposed or to how much. If these are common deficiencies in published studies asserting no effect, it is understandable that the Occupational Safety and Health Administration (*2*) should lay out rather rigid standards for acceptance of what they refer to as nonpositive epidemiologic studies: "The epidemiologic study involved at least 20 years' exposure of a group of subjects to the substance and at least 30 years' observation of the subjects after initial exposure;" "The group of exposed subjects was large enough for an increase in cancer incidence of 50% above that in unexposed controls to have been detected at any of the predicted sites."

Table 1. Characteristics of some "negative" occupational epidemiology (cancer) studies.

Characteristic	No. of studies
Nature of study	
Retrospective cohort	6
Forward cohort	0
Cross-sectional	3
Proportionate mortality	2
	11[a]
Nature of exposure	
Well defined (individual dose)	0
Some estimate (e.g. duration of employment) (general air levels)	4
Poorly defined:	
(a) yes-no only	1
(b) presumptive only	5
	10
Methodology problems	
Incomplete death reporting	1
Errors in cause-of-death	1
Young cohort	4
Short follow-up	3
Inappropriate group followed (dilution)	3
Small numbers	3
Identification problems	1
	16[b]

[a]One study in two categories.
[b]Several studies had more than one defect.

OSHA (*2*) supports these minimum 20-30 criteria by testimony from Richard Peto, Irving Selikoff, David Rall, Robert Hoover and John Berg, among others. The 50% criterion is further constrained in these remarks ". . . in theory, at least, an epidemiologic study must be sensitive enough to detect a 50% increase . . . at any site. . . . In practice, such a criterion could not be used, because no epidemiologic study of practical size could detect a 50% increase in a rare type of tumor. Thus OSHA is willing for practical reasons to consider studies which are sensitive enough to detect a 50% increase at all sites where the agent is judged likely to act in humans."

In at least one sense, this is a weak criterion. An excess risk of 50% is hardly trivial. However "to require greater sensitivity would place unreasonable demands. . . ." (*2*).

Dose-Response Phenomena

Decouflé's comments on the paucity or the inadequacy of dose data in putative, "negative" industrial epidemiology studies (*1*) apply almost as well to so-called positive studies. Rarely are there data on individual exposures. Rarely does one meet the 20-30 criteria. Yet, if positive results appear, clearly one need not wait for the 20-30 phenomenon. It is not that different levels of proof are demanded, but rather that the perceptions of "positive" and negative, must of necessity be different. "Positive" is taken to mean positive at any time, while "negative" is taken to mean negative at all times.

If we came to the conclusion that we have seen a positive result, i.e., that there is a relationship between some exposure and the appearance of disease, the immediate question that follows is "how much?" That is, we want the answer to the question "How many units of disease will be produced by adding X units of exposure?" Unfortunately, the question is almost unanswerable, because most of the time we do not know what a

"unit of exposure" is. Nonetheless, it is worth looking at some data that have been compiled relating exposures to consequences. I will mention the industrial data, mostly to point out problems, but these will be dealt with much more fully in the paper that follows by Professor Enterline. The major studies in which we have dose-response relationships in humans are the radiation studies from Hiroshima and Nagasaki (*3*) with which Dr. Land is most conversant, or among uranium miners (*4*, *5*), the smoking and lung cancer studies of Doll et al. (*6*), Hammond and Horn (*7*), and Dorn (*8*) and the studies of asbestos workers by Enterline (*9*) and McDonald (*10*). There are others, but these are enough to make the point at this time.

The radiation studies point up the problems most clearly, because we know most about radiation carcinogenesis and what we know is not enough. Having agreed that ionizing radiation is cancer-producing, we are immediately beset by problems in characterizing how carcinogenic. We recognize differences between high LET radiation and low LET radiation. We are uncertain about the effects of dose rate; we are even less certain about the joint effects of radiation and exposure to other identified carcinogenic activity. We have seen radiation carcinogenesis arising out of the one-time exposures (at high dose, to many people, at least) at Hiroshima and Nagasaki, and we have seen radiation carcinogenesis in chronically exposed radiologists. The dose-response relationships that many of us considered "conservative," the linear, no-threshold concepts are being challenged on both sides as both overstating and understating the risks.

What happens when we look into other physical carcinogens, and the (more likely) more complicated area of chemical carcinogens? The earlier work by Enterline and McDonald on cancer following asbestos exposures led them to consider seriously the possible existence of threshold and apparently did not lead them to any serious attempts at dose-response curve-fitting. I made some crude attempts (*11*) to bring their two sets of data together, and to try some curve-fitting (which I did not quantify). I saw less reason to believe there was a threshold when I looked at the data for lung cancer than when I looked at the data for the digestive system cancers (Figs. 1 and 2).

More recent (and more sophisticated) work by Enterline (*12*) and by Liddell in cooperation with McDonald (*13*) have led to curve fitting and dose-response curves. Both of these attempts have led to linear dose-response equations, Enterline's giving a response in respiratory cancer against cumulative dust exposure (mppcf-years) that led him to the equation (1):

$$\text{Predicted SMR} = 100.0 + 0.658 \text{ cumulative exposure}$$

and this remark, "[it is] remarkable that this empirical fit gives a y intercept of 100." He also notes small changes over two time periods (1941–69, compared to 1970–73), and concludes with this cautionary remark: "The numbers are much too small to speculate on the form this relationship takes [in recent times], however."

The Liddell data can also be fitted by a straight line (these data relate to Quebec asbestos miners). Liddell gives his results in terms of relative risk per mppcf-years. My fitting of the data yields the equation (2):

$$RR = 1 + 2.8 \times 10^{-3} \text{ mppcf-years}$$

which is substantially shallower than Enterline's "predicted curve"; i.e., it predicts lower risks at the low doses.

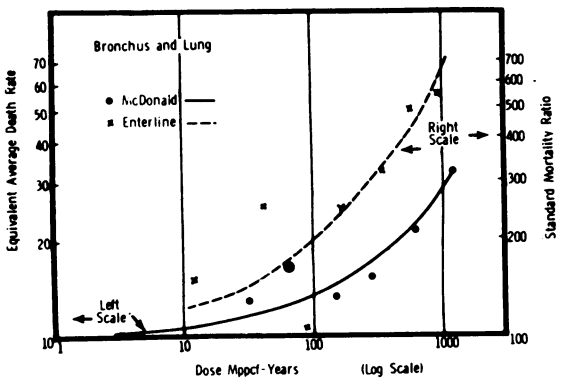

FIGURE 1. Cancer among workers exposed to asbestos: lung cancer.

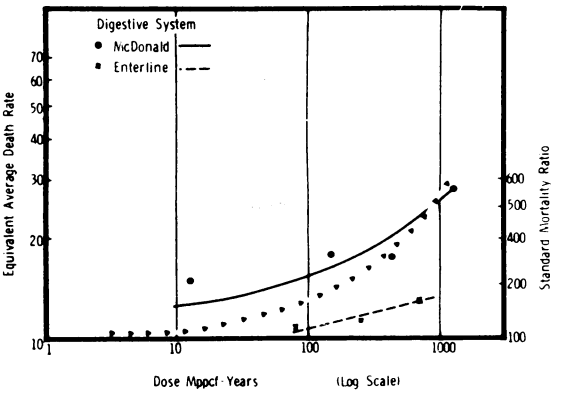

FIGURE 2. Cancer among workers exposed to asbestos: digestive tract cancers.

The different slopes raise very serious questions about what data can, or should, be used for prediction. Enterline points out, as do others, that the type of asbestos, and the conditions of work (whether exposure was intermittent or continuous) led to different SMR's for respiratory cancer. Cornfield (*14*, *15*) has shown that problems in the accurate measurement of doses and of response can lead to apparent linearity of what is truly a curvilinear response. Doll and Peto (*16*) re-examined the dose-response relationship in cigarette smoking and lung cancer (annual incidence rates of bronchiogenic cancer as related to dose and duration of exposures) and found that when various possible biases in reporting were removed, there was evidence for a distinct turning upward of the dose-response curve at the higher dose levels. This turning upward is consistent with a multistage hypothesis of cancer induction. At low doses, however, the linear term dominates.

The Doll-Peto re-examination of the smoking data points to an important area of needed interaction or interface. The mathematical models of dose-response must be consistent with the biological models of how the disease comes about. Thus, the Doll-Peto finding of (mild) positive curilinearity is consistent with the multistage model of cancer induction. This, of course, is only a first step in developing an appropriate prediction model since few, if any, of the models have yet built into them dose rates, duration of exposure, effects of cessation of exposure, age of the exposed persons, competing risks, etc.

One of the more recent attempts at extending the multistage model concept is a paper by Day and Brown (*17*) which considers the effects of stopping exposure. Their major finding is that the effect depends upon the stage in the multi-stage process which the carcinogen predominantly effects. They confirm the concern often expressed about the utility of human data as a means to prevention. They write ". . . by the time the human evidence that a hazard does exist becomes available, those already exposed may well have accumulated their fully effective dose."

The Day-Brown results have shown that the effect of exposures to early stage carcinogens is acquired in a relatively short time and then does not diminish rapidly, even after cessation of exposure. For example, for a first-stage exposure at age 40, five years of exposure would, by their calculations, produce 44% of the effect (excess cancer risk over background) that a "remaining life-time" of exposure would produce. On the other hand, similar five-year exposure to a penultimate stage carcinogen would give 11% of the excess cancer risk over background compared to a remaining life-time exposure.

These results have serious implications for extrapolation and for further experimentation; i.e., in attempts to identify where in the carcinogenic process a material fits and to identify materials other than complete carcinogens. I see the Brown-Day results as a possible unifying force in helping sort out initiators, promoters, pro-carcinogens, proto-carcinogens, facilitators, etc., etc. The Brown-Day results lend support to the suggestion that the control of late stage carcinogens can lead to relatively rapid decline in cancer rates. Continuing exposures to late stage carcinogens appears to be necessary until an irreversible cancer process is triggered. Ending this exposure should interrupt the process, so that, unless a preclinical cancer has already developed, cancer will then not develop. The experience with endometrial cancers and post-menopausal estrogens is consistent with this formulation of the cancer process. Day and Brown do not find it necessary to postulate the existence of thresholds for "promoters" to explain this behavior of late stage carcinogens.

Thus, we find such factors as type of radiation (high LET vs. low LET), type of asbestos (e.g., chrysotile, crocidolite, etc.), nature of exposure (intermittent vs. continuous), duration of exposure, effect of other exposures, effect of cessation of exposure all possibly modifying the dose-response curves and we have little way to take them into account. The human data rarely come in enough quantity and enough detail to allow us to do much more than describe a given situation, and that usually incompletely. When we want to take important details into account when forecasting what will happen, we rarely, if ever, have the data upon which to do it.

If our human data are so incomplete, we can have the hope that data from animal experimentation may fill the gap. In a sense, the mouse with the two to three year life-span may possibly supply the equivalent of accelerated life testing for man. David Rall recently wrote a succinct summary of the place of animal experiments in toxicity testing (*18*). The question he sees is this: "We are dealing here with large human populations with relatively small differences in exposure . . . . Can we use experiments with laboratory animals to project what is likely to happen in the human population?"

There is substantial dose-response data in animals so that if we were concerned with making estimates of the risks at low doses for some strains of mice or rats we might do moderately well. If we

want to go from mouse to man, we're on much shakier ground. Rall gives a table (Table 2) covering all the known carcinogens for which there are both human and animal dose-response data (from an NAS/NRC study on pest control practices). This table was constructed with doses compared on a milligram per kilogram basis. The comparisons were not as close when attempts were made to extrapolate on a surface-area basis, which, for certain other comparisons, seems to give much closer correspondence between human and animal responses.

The correct measure of dose in moving from one species to another is of some consequence. It has been looked at from time-to-time, but no simple solution seems to be at hand. As an example of the difficulties, Table 3, derived from the NRC/NAS review of the saccharin issue (*19*) is worth looking at. Two points thrust themselves at us from this table. First, the model makes an enormous difference. The range of estimates from model to model is about 5×10^6. Models will have to be justified on their biology, not just their goodness of fit. Second the dose-metameter is important, even within the same model. There is a 200-fold difference within the Mantel-Bryan scheme, for example, if one uses mg/kg/day vs. mg/kg/lifetime.

Table 2.

Material	Predicted human incidence based on most sensitive animal species relative to epidemiologic studies
Full data available	
Benzidine	Same
Chlornophthazine	Same
Cigarette smoke	Same
Aflatoxin B_1	10 × greater
Population still at risk	
DES	50 × greater
Vinyl chloride	500 × greater

Table 3. Estimated human risks from saccharin ingestion of 0.12 g/day (lifetime cases per million exposed).[a]

	Dose measure (rats to humans)		
Extrapolation method	Surface area	Mg/kg/ day	Mg/kg/ lifetime
One-hit (Hoel et al.)	1,200	210	5,200
Multi-hit (Food Safety Council)	0.001	0.001	0.001
Multistage (Brown et al.)	5	Not done	Not done
Probit (Mantel-Bryan)	450	21	4,200

[a]This table points out the two important problems in extrapolation: choice of the mathematical model, and choice of the dose metameter.

Thus, we find ourselves trying to estimate what will happen to people who are exposed to low doses of materials that we have previously shown to be toxic either to other humans, at much higher doses, or to animals. The issue is how we can use these high-dose data and/or these animal data to help us estimate what might happen at lower doses and under different circumstances. For this I see three principles emerging.

(1) Existing data even from well-done studies are inadequate if we want to make our extrapolations in any detail or to apply to specific subgroups in the population. Any risk estimation we do may have to be stated in terms that point out the wide ranges our estimates lead to.

(2) Experimental or epidemiologic studies are of themselves not sufficient. They need to be tied in with appropriate biological models of the diseases we are concerned with and these, in turn, must derive from other research.

(3) Principles of careful epidemiologic work and careful laboratory work will have to be carefully adhered to. The too small study, with the too short follow up, with the too little dosage information is too poor to use for risk estimation.

Public, political, and economic pressures will push us into making risk estimates. If we make it clear how poor these estimates can be perhaps we can push ourselves and our colleagues into doing better studies, and building biologically better models that might, sometime in the future, give us good enough risk estimates, so that political decisions can be made from them with some confidence. That day is not here yet.

REFERENCES

1. Schneiderman, M. A., Decouflé, P., and Brown, C. C. Thresholds for environmental cancer: biological and statistical considerations. Ann. N.Y. Acad. Sci. 329: 92-130 (1979).
2. Department of Labor, Occupational Safety and Health Administration. Identification, classification and regulation of potential occupational carcinogens. Federal Register 45(15): 5050-5060 (Jan. 22, 1980).
3. Land, C. E., and McGregor, D. H. Dose-response relationship in radiogenic breast cancer. J. Natl. Cancer Inst. 60: 728-729 (1978).
4. Jablon, S. Environmental factors in cancer induction: appraisal of epidemiologic evidence for leukemia, lymphoma, and radiation. In: Cancer Epidemiology, Environmental Factors, P. Bucalossi, Ed., American Elsevier, New York, 1975, pp. 239-243.
5. Lundin, F. E., Jr., Wagoner, J. K., and Archer, V. E. Radon Daughter Exposure and Respiratory Cancer: Quantitative and Temporal Aspects. National Institute for Occupational Safety and Health, National Institute of Environmental Health Sciences Joint Monograph 1, U.S. DHEW, Washington, D.C., 1971.
6. Doll, R., and Peto, R. Mortality in relation to smoking: 20

years' observations on male British doctors. Brit. Med. J. 2: 1525 (1976).
7. Hammond, E. C., and Horn, D. Tobacco and lung cancer. Proceedings of the Second National Cancer Conference, 1952, pp. 871-874.
8. Dorn, H. F. The increase in cancer of the lung. Ind. Med. Surg. 23(6): 253-257 (1954).
9. Enterline, P., Decouflé, P., and Henderson, V. Respiratory cancer in relation to occupational exposures among retired asbestos workers. Brit. J. Ind. Med. 30: 162-166 (1973).
10. McDonald, J. C. Cancer in chrysotile mines and mills. Paper presented at Conference on Biological Effects of Asbestos, Lyon, France, 1972.
11. Schneiderman, M. A. Digestive system cancer among persons subjected to occupational inhalation of asbestos particles: a literature review with emphasis on dose response. Environ. Health Perspect. 9: 307-311 (1974).
12. Enterline, P. Asbestos exposure: factors associated with excess cancer and respiratory disease mortality. Ann. N.Y. Acad. Sci. 330: 117-126 (1979).
13. Liddell, F. D. K., McDonald, J. C., and Thomas, D. C. Methods of cohort analysis: appraisal by application to asbestos mining. J. Roy. Statist. Soc. A140(4): 469-491 (1977).
14. Cornfield, J. Carcinogenic risk assessment. Science 198: 683 (1977).
15. Cornfield, J., Carlborg, F. W., and Van Ryzin, J. Setting tolerances on the basis of mathematical treatment of dose-response data extrapolated to low doses. Proceedings of the 1st International Congress on Toxicology, G. L. Plaa and W. A. M. Duncan, Eds., Academic Press, New York, 1978, pp. 143-163.
16. Doll, R., and Peto, R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and life-long non-smokers. J. Epidemiol. Commun. Health 32: 303-313 (1978).
17. Day, N. E., and Brown, C. C. Multistage models and the primary prevention of cancer. J. Natl. Cancer Inst. 64: 977-989 (1980).
18. Rall, D. P. Relevance of animal experiments to humans. Environ. Health Perspect. 32: 297-300 (1979).
19. NAS. Saccharin: Technical Assessment of Risks and Benefits. Report No. 1. Committee for a Study on Saccharin and Food Safety Policy. National Academy of Sciences, Washington, D.C., 1978.