Annals N.Y. Acad. Sci. 246, 237-48 (Jan 31, 1975)

FROM MOUSE TO MAN—OR HOW TO GET FROM THE LABORATORY TO PARK AVENUE AND 59TH STREET*

Marvin A. Schneiderman

Associate Director for Field Studies and Statistics, DCCP
National Cancer Institute
Bethesda, Maryland 20014

Nathan Mantel

Biostatistics Center
7979 Old Georgetown Road
Bethesda, Maryland 20014

Charles C. Brown

Biometry Branch, FS&S
National Cancer Institute
Bethesda, Maryland 20014

INTRODUCTION

If the title of this paper sounds like a guide to a road map, that is what it is intended to be, with information about detours, chuckholes, swamps, quagmires, and dead ends. In fact, our description of the problems may lead some people to conclude that you really cannot get there from here. We think you can, however —perhaps not *exactly* there, but certainly into the neighborhood.

When we extrapolate moderate or high dose animal toxicity data to presumed effects in man, we must do it like the animal trainer entering the lion's cage—very cautiously. The first caution in this careful operation is to extrapolate from mouse to mouse from the higher dose levels where it is possible to see some results with small or moderate-sized experiments to the very low doses, where even very large experiments are not likely to show much. We really do want to extrapolate and we do not want to do these very large experiments because they are so full of problems that they probably cannot be done well. We will discuss this later. Weinberg, in his paper on Trans-Science,[1] refers to this sort of difficulty that the mega-experiment poses as his first kind of trans-science problem.

MODELS FOR IN-MOUSE EXTRAPOLATION

The usual way to extrapolate to responses at low doses is to assert the existence of some mathematical model of dose-response and then compute the response at the desired dose or compute the dose for the desired response. Weinberg suggests this as a possible way out of the mega dilemma.

Presently two classes of mathematical models are advocated for in-mouse extrapolation. First, there are the models that deal with yes-no data, the dichotomous data models. These are concerned with whether or not an effect occurred.

* The Working Group met at the Delmonico Hotel, Park Avenue and 59th Street in New York City.

The data presented by Professor Maltoni in his experiment BT-1 after 52 weeks of exposure and 130 weeks of observation are an example of the yes-no type of data:

Dose (ppm)	*Angiosarcomas (positive animals/total animals)*
0	0/68
50	0/64
250	4/67
500	7/67
2,500	13/74
6,000	14/72
10,000	7/69

The most common models used to fit such data are:

1. *The Probit Model*—which assumes a normal (Gaussian) distribution of sensitivities in the exposed population to the material tested, usually against the logarithm of the dose.
2. *The Logit Model*—which assumes a response mechanism similar to first-order chemical kinetics with a damping effect at upper response levels as "receptor sites" (places on, or in, the cell where the reaction takes place), etc., are used up or become occupied.
3. *The "Hitness" Models*—including the one-hit and multi-hit models assume that an effect occurs when a vulnerable portion of the cell receives n hits, n being some integer, from 1. Some multi-hit models have been used to describe the rapid increase with age of cancer incidence in man.[2] The assumption is made that n assaults on the cell are necessary to begin the irreversible changes that later show themselves as clinical cancer. The one-hit model is often used in describing the dose-response for radiation effects.
4. *The Extreme-value Model*—attempts to describe rare events such as the number of floods that will exceed a given height during some long time interval. In a limiting form, the extreme-value model becomes the one-hit model and is, in fact, a generalized "hitness" model without the constraint that the n be an integer.

The major difference among these models lies in the estimated effects at the very low doses in which we are interested. The doses necessary to achieve low response levels (i.e., 10^{-6}) are substantially different depending on the model used.

The different extrapolations that would arise from the Maltoni data (including some later data) assuming several of these models are shown in TABLE 1. For the logit model a slope of 2.303 (in logarithms to the base 10) was chosen because this corresponds to the one-hit model. A slope 50 percent larger (3.454) was also used to show the large effect of change in slope. FIGURE 1 shows these comparisons. TABLE 2 shows estimated response rates for the Maltoni experiment at several lower doses.

The other class of models is concerned with time to response. The use of these models to estimate safe doses is based on the obvious fact that everyone must die of something at some time and a new cause of death is important only if it will occur early enough in the lifespan to shorten life. A letter in *Science*[3] summed up the arguments this way: ". . . the only relevant parameter in such discussions is how the average lifespan of a person within a given population may be affected by . . . exposure. How can . . . 'extra deaths' per year . . . be translated into reduced lifespan?"

TABLE 1
EXTRAPOLATED "SAFE" DOSE LEVEL (PPB) USING MALTONI'S LIVER ANGIOSARCOMA DATA (99% ASSURANCE LEVEL)*

Extrapolation Model	Risk Level				
	10^{-6}	10^{-5}	10^{-4}	10^{-3}	None specified
Probit (slope = 1, Mantel-Bryan)	251 (225)	81.6 (73.0)	29.2 (26.1)	11.3 (10.1)	
Logit (slope = 3.454)	574 (553)	124 (119)	26.7 (25.7)	5.74 (5.53)	
Logit (slope = 2.303) and "one-hit"	23 (21.4)	2.3 (2.14)	0.23 (0.214)	0.023 (0.0214)	
"Classical"					500
Weil					(not possible)
					125-50
					(25-10)

* The data as of 130 weeks of observation up through dose level 500 ppm, but not higher, were used for extrapolations since the response curve flattened out at higher levels. The numbers in parentheses show "safe" levels computed from data developed after 135 weeks of observation, which included 1 liver angiosarcoma at a dose of 50 ppm.

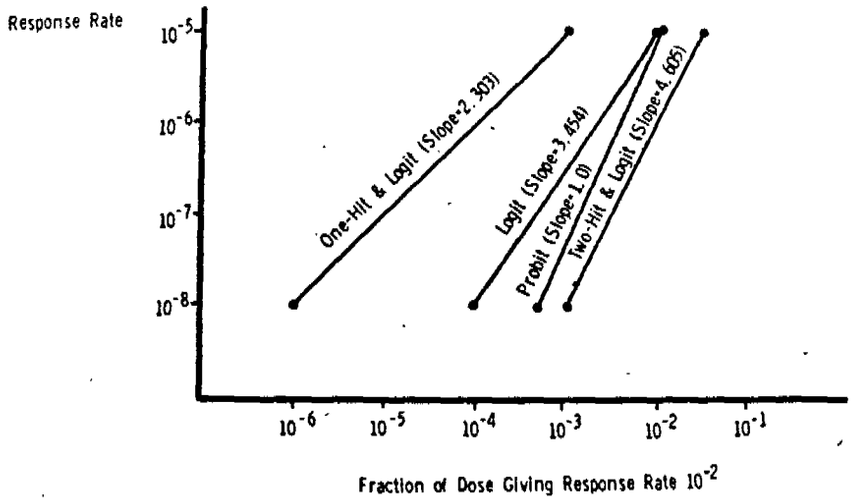

FIGURE 1. Fraction of experimental dose giving response rate 10^{-2} using several extrapolations.

This problem has been considered by several authors. In cancer research Druckrey[4] first looked at the latent periods for the induction of cancer in experimental animals. Albert and Altschuler[5] extended the Druckrey model working with the average time to death (or incidence) from cancer, if there were no other causes of death. Hoel and Walburg[6] attempted to lay out the appropriate statistical analysis of survival experiments. Gail[7] has constructed a time-to-occurrence model in which he has three "measures of merit": reduction in life-span at birth, reduction in life-span of persons affected, and reduction in life-span up to some given age (e.g., working years of life lost or life lost before age 65).

There are difficulties with both classes of models for low dose effects. In the midrange of doses where the experiments are usually done, the data are usually adequately fit by any one of the yes-no models. An example constructed by Cornfield[8] shows how similar the three major models are over a 256-fold range of doses. If given perfect data, we cannot tell these models apart in the 2–98 per-

TABLE 2
ESTIMATED RESPONSE RATES: LIVER ANGIOSARCOMA
(99% UPPER CONFIDENCE LIMIT IN PARENTHESES)*

Model	Dose (ppm)			
	50	25	5	1
Probit (slope = 1)	0.012 (0.028)	0.0055 (0.013)	0.00059 (0.0018)	0.000041 (0.00015)
Logit (slope = 3.454)	0.0041 (0.0085)	0.0015 (0.0030)	0.00013 (0.00027)	0.000012 (0.000024)
Logit (slope = 2.303) and "one-hit"	0.012 (0.023)	0.0061 (0.012)	0.0012 (0.0023)	0.00025 (0.00047)

* Based on data from Maltoni, experiment BT-1, 52 weeks of exposure and 135 weeks of observation.

cent range of response; this says that the data you can get often tell you too little about the data you cannot get, i.e., you still will not know what model to use.

Mega-Mouse Manipulations

Under these circumstances it would seem reasonable to attempt to conduct experiments in the dose range in which we are interested and where the usual models give widely different answers. This has led to suggestions for the mega-mouse experiment. The experiment will have to be mega-mouse to be capable of detecting low levels of damage (say of the order of 1 in 10^6) and/or small reduction in survival, etc. To demonstrate, at a single dose level, that the response rate is less than 1 in a million, with say 95 percent confidence and with no spontaneous tumors in untreated animals, would require 3 million animals each in both a treatment and a control group. For a carcinogenesis experiment these animals would have to be maintained on the appropriate regimens for a lifetime (about 18 months to 2 years for mice). Obviously millions of animals cannot be started on experiments all at once, so the experiments would have to be conducted sequentially with great care that animals are appropriately randomized and identified at each stage of the separate subexperiments and that appropriate controls are set up at each subexperiment, etc. To make certain that one had sufficiently sensitive animals it probably would be necessary to set up (relatively small) parallel positive controls—animals treated with known carcinogens at suitable doses, to show that the experiment was biologically capable of detecting materials at least as carcinogenic as known carcinogens. The numbers added would not be great, but added precautions have to be taken when known carcinogens are being handled. If we take into account the fact that the untreated animals may develop spontaneous tumors, then to detect an increase of 1 in 10^6 over the spontaneous incidence could require a study several thousand times larger than one in which there were no spontaneous tumors.

Even with unlimited funds to conduct these experiments, the probability that they could be conducted with no mistakes in handling, feeding, loss of animals, or any other of the myriad of common laboratory errors that ruin the best planned experiments of mice by men is very low. Purely logistical problems might guarantee failure. Such laboratory errors, considered as noise, could tend to overwhelm any signal present in the data, making it still more difficult to differentiate between treated and control animals. The statistical remedy is to enlarge the experiment by yet another large factor increasing the risk of further blunders that could wipe out the gains of the increased experiment size. In this kind of experiment if one wanted to show "safety" of a material, quite possibly unconscious pressures would exist to do a poor job. The more errors made, the less likely it is to show a difference between treated and controls, between signal and noise. Finally, at low dose levels results may be ambiguous, with confusion possible between spontaneous and induced tumors. Thus, we think the mega-mouse experiment is liable to end in a ditch. It does not have as much to commend it on second thought as it looked to have at first.

If Not Mega-Mouse, Then What?

In view of these problems, what are the prospects that the experiments suggested earlier in this meeting by the MCA will provide useful information? We have computed, for Professor Maltoni's BT-1 data, that a dose as low as 1 ppm is

almost certain to have a risk of less than 1 in 10,000. What could an experiment at a dose of 1 ppm using several hundred animals show? Hardly anything. We would urge the MCA people to review their design plans. If they cannot design a potentially more productive practical experiment perhaps they should abandon the whole thing. Maybe they cannot get there from here.

Difficulties arise in applying any mathematical model at very low doses. Does the slope of the response curve in the mid-dose range hold for very low doses? Is the slope likely to become steeper or shallower at very low doses? If steeper, then any dose suggested as "safe" will be even safer than we think. On the other hand, if the slope is shallower, then any "safe" dose will be less safe than we think. Armitage[9] showed that at high doses the curves flattened out—that there seemed to be highly resistant animals. Is there symmetry in the dose-response relationship? If there is then we should look for flattening at the low end of the response curve, too. Mantel *et al.*[10] seem to have showed this.

Mantel-Bryan Extrapolation

One of us (N.M.) has suggested that in view of these many problems in extrapolation it might be useful, in attempting to compute a safe dose, to fix upon one of the models for which there seems to be adequate experience and a good biological basis and to fix upon some suitably shallow slope, and with appropriate 95 or 99 percent upper limit of response calculations, extrapolate to a level which can be agreed upon as a "virtually safe" level. This approach is embodied in the Mantel-Bryan procedure[11] and in recent modifications.[12]

Mantel and Bryan suggest using the log probit model, with the slope of 1 probit per 10-fold dilution. Both have been criticized. With respect to the model Peto *et al.*,[13] among others, have suggested that it has little biological basis and that in the field in which most work has been done at low dose levels—radiation carcinogenesis—a one-hit model is appropriately conservative.[14] Some of Peto's own work recently would imply that a two-hit model would fit the data better. Such a model is less conservative than the probit with slope = 1. The criticism of the Mantel-Bryan choice of slope runs in both directions. There are workers who think it is too shallow. A recent review of about 180 papers in the carcinogenesis dose-response literature by the Franklin Institute[15] has turned up many examples of shallower slopes implying that a slope of 1 is too steep.

Latent Period Models

The models involving time to response need more development. Most of them are concerned with mean or median times to appearance of cancers while clearly what we need is the time for 1 or 0.1 percent (or some smaller number perhaps) to develop cancer. If this event did not occur until a late age then the risk, however looked at, would obviously be small. Gail[7] has pointed out that with one particular model of the distribution of time to appearance, a cancer with a median age of appearance (if there were no other causes of death) of 300 years (e.g., lung cancer) would nonetheless reduce life-span for those developing it by about 8 years.

There is also a psychological problem in working with these models that perhaps someday will be overcome. It is the problem of dealing with an unfamiliar currency. Is something which costs 72 rials expensive or cheap? If you are an Iranian you will have little trouble working this out, but if you're an American tourist you will have to do some translating. How important is a cause of death

which shortens average life-span at birth by 20 min, or 2 months, or 2 years? How do you translate that into your own experience? As an aid, or as the first line in the tourist's phrase book, please notice that all cancer, the second leading cause of death in the United States, reduces life-span at birth, by less than 2 years.

The time-to-appearance models are not free of model assumptions. The Albert-Altschuler model assumes a log-normal distribution of occurrence times. Gehan[16] has shown that this model implies unusual biological behavior with a nonmonotone hazard rate. The model that Peto and Pike and their colleages prefer is the so-called Weibul model which is much better behaved and does conform better to what is currently known about the biology.

In addition the models developed so far are what the statistician calls "fitting" or "graduation" models. They fit or graduate the observed data, but it is not clear whether they can be extrapolated safely far outside the observable range; they do not incorporate a "guarantee" that the response will be below a certain level. The Mantel-Bryan procedure does attempt to produce such a "guarantee."

The latent period models have not only inherent experimental (observational) problems but also some unsolved mathematical problems. It is not hard to observe superficial tumors in experimental animals, but what is the time-to-appearance of an internal tumor that kills the animal? That does not kill the animal? Do we have to develop experiments with serial sacrifice designed into them? Again Peto[17] has some suggestions on how to handle some of these data.

Gail points out one further difficulty that the time-to-appearance models have brought more vividly to our attention. These are the "competing risk" problems. Suppose a material can lead to death from any one of a group of causes, not just cancer. If this material shortens life, so that all the exposed individuals (animals or humans) die early before cancer is likely to occur it may even look like a cancer preventive. This is in the class of the cartoonist, Webster's, advice on how to keep from growing old—walk in front of a bus. More statistical sophistication is needed to handle the problem of competing risks. With yes-no models survival data as well as the simple "yes-no" response must be considered. Did the animals live long enough to develop cancer? Or, was the material so effective in eliminating all other causes of death that all animals died of cancer at a very late age?

When first looked at it appears that there are large differences in concept between yes-no models and the time-to-appearance models. However, Hoel and Chand[18] have shown how they are related. The Albert-Altschuler models, using a log-normal time to appearance of tumors, correspond to the Mantel-Bryan choice of the probit model where observations would be made at one instant in time for experimental animals exposed to different doses. Similarly the Weibul-based models, favored by Peto and Pike correspond to the "hitness" models, when animals are all sacrificed at one time and observed for presence or absence of tumors.

More Troubles

No matter the model used, there still remain a substantial group of unsolved problems. One is the problem of interactions of materials. Beginning with the early initiator-promoter studies of Berenblum and Shubik,[19] research workers have been aware that materials affect each others' effects and often do not operate in an "independence vacuum." Doll[20] has shown that in two important environ-

mental exposures, asbestos and radon (in the exposed uranium miners), the interaction with cigarette smoking was multiplicative rather than additive. This is a frightening prospect which implies that the problems that follow the addition of new materials to man's environment may be much more serious than just adding another drop to the pool. Some work has been done on allowing for the "natural" incidence of tumors in control groups but there is not agreement on the appropriate procedure to use—simple subtraction, the Abbott's correction, or a slightly more complicated "functional" method proposed by Albert and Altschuler.[5] From the experimentalist's point of view the most unfortunate effect of natural tumor incidence is that it substantially increases the experiment sizes needed to show differences between the treated and control groups.

Then there are questions of dose. Should an animal experiment be a lifetime feeding experiment? A series of pulsed-doses, with zero exposure in between? An exposure for a short time, followed by no further exposure? Exposure *in utero?* Exposure only relatively late in life? Each of these is meaningful and could have a counterpart in human exposure. How one extrapolates from one to the other is not known.

Establishing Acceptable Risk Levels

Applying the extrapolations from almost any model will entail problems and sometimes hardship. When people talk about levels they "can live with," they must be reminded that these could be levels that other people might die from. We therefore recommended that the appropriate "safe" level (i.e., a risk of 1 in a billion, or 1 in a million, etc.) be chosen after considering the following:

1. Is the material already in the environment? If so, how extensively is it used? What are its economic uses? Its health uses (i.e., is it a medicine for use in a very serious condition? a mild condition? a self-limiting condition?)?
2. Is it a new material? What will it be used for? What are its economic uses? Its health uses?

For example, with materials now in the environment and of substantial economic importance, limiting risk of 1×10^{-6}, reducing in subsequent years to 1×10^{-7} and then to 1×10^{-8} may be appropriate. For new materials not yet in the environment, the limiting risk could be set at 1×10^{-6}, unless there are overwhelming health or economic reasons to indicate higher risk levels. The setting of a limiting risk level should be an "open market" operation.

In the early application of the arithmetic of the models, consideration can be given to how many persons are or will be exposed to the material. If there are limited numbers of persons exposed to the material, as an initial "safe" level, possibly a risk level as high as 1×10^{-5} may be permitted. This level will then have to be progressively reduced to a country-wide or worldwide universal lower risk. This means a first risk of 1×10^{-5} for those "exposed," and perhaps 1×10^{-10} for those apparently "unexposed" (but who may even be exposed to a "rub-off" level much as the nonsmoking person may be exposed to the smoke from someone else's cigarette). This proposal involves a double standard and raises serious ethical issues as to whether some portions of the population should be assigned higher risk—suffer the losses—so that other lower risk portions of the population might gather the gains, another "open-market" question.

Finally, ways are needed to include nonexperimental evidence in setting up safe doses. For example, if man has been exposed to a material for a long time (over 30 years is a long time) and no untoward irreversible effects are seen, a

way is needed to incorporate these data into the arithmetic. No objective procedures now exist for doing this. Perhaps, as an interim measure until objective procedures are developed, experts could be called upon to modify the "safe" dose. This might be done by a change in the allowable risk level, i.e., a change from 10^{-8} to 10^{-6}. Perhaps the experts should be restricted in making changes in the risk level to no more than 100-fold, i.e., from 10^{-8} to 10^{-6}.

Out-Mouse Extrapolation

After the in-mouse extrapolation problem comes extrapolation from mouse to man—the out-mouse problem. David Rall,[21] in a recent symposium on Statistics and the Environment, at the National Academy of Sciences, dealt with some of the issues very well. We will not attempt here to do more than summarize parts of his discussion, in places using Rall's own language. He points out the difficulties in extrapolating "from an inbred mouse strain to a genetically heterogenous population, such as man." It appears reasonable (to us) that an inbred strain is likely to have a steeper dose-response curve than would a more heterogeneous population. Rall asks, "If a heterogeneous population is a collection of inbred strains, each with its own dose-response curve, what then would the dose-response curve be for this heterogeneous population?"

Rall also speaks of problems of relative size of the experimental animals and man which lead him to recommend that doses are more likely transferrable on the basis of relative surface area (or two-thirds power of weight) than on a milligram per kilogram basis. He notes the differences in metabolic rates, and in the ratio of blood volume to circulation time which would lead to materials being retained by man relatively longer than they are retained by the smaller experimental animals. If this is true, then equivalent doses of carcinogens could be more carcinogenic for man than for the mouse. Rall also raised questions of plasma protein and tissue binding with the remark that small mammals tend to bind compounds less extensively than man. Because hepatic metabolism is often of major consequence Rall notes that "in general, small mammals metabolize compounds more rapidly than large ones; in general herbivores metabolize compounds more rapidly than carnivores. . . . The entire area of metabolic disposition of foreign compounds is of major importance."

The generation time of cells, their rates of repair processes, etc., obviously will affect malignant or mutagenic change. "In the mouse . . . the generation time of the rapidly proliferating cell population is about half of that in man, so that a cancer in a mouse is likely to appear earlier in a mouse's life span than in a man's." This observation has considerable importance with respect to the time-to-appearance models.

We must recognize that experimental animals usually live in a controlled environment, with relatively uniform diets, temperature, humidity, exposure to other agents, etc. Man, as a free li\ ıg animal, has a much more diversified experience—some of which will make some of him more sensitive, more susceptible, and more likely to succumb to environmental assaults. Some experimenters have argued that the appropriate experimental animal would be a non-inbred animal. Using non-inbred animals reduces the reproducibility of experiments, requiring larger experiments to give answers with assurance. Another suggestion is to conduct experiments in several species using inbred animals of each species and using results in the most sensitive of the species to extrapolate to man.

In spite of all these difficulties, it still seems possible to extrapolate from

mouse to man—at least from the median mouse to some hypothetical median man. To allow for responses of the nonmedian man, for the more sensitive individuals of the species, will require introducing additional factors. It is less than reasonable to take a no-effect level in a relatively small animal experiment (for the kinds of effects we are considering here, 100 animals at a dose is a small experiment) and translate it to a threshold limit value or some similar "safe" standard.

Cost-Benefit, Or Whose Ox Is Gored?

The remaining issues in this paper are ones in which the statisticians have no special competence. On the other hand there is nothing in their training to make them more than normally incompetent, so we think we can talk about these issues as well (or as badly) as other scientists. These issues are the ones usually lumped together under the heading of "cost-benefit," or, in less elegant language "whose ox is gored?" There were some direct references to this earlier in this meeting when Mr. Mazzocchi asserted that issues of safety would have to be resolved by adversary proceedings. This disturbed some people who then proposed the alternative to adversary activity, cooperation between science, industry, and the labor unions.

It seems to us that we really have no alternative to the adversary-type activities when different people have different interests. The scientist who believes that he can resolve all through the use of his science alone impresses us as being out of touch. Does this mean that no cooperation is in order? Of course not. As statisticians, we could not work without the data on who was exposed, how much he was exposed to, and for how long he was exposed. These data must come from industry. To follow up exposed people we need data from industry and union and government sources. Evaluating animal results requires cooperation with our laboratory science colleagues. But cooperation does not mean suspension of critical faculties or suspension of scientific disbelief. If our laboratory colleague has designed a nonproductive experiment, we must tell him and try to help him design one that will yield something. If we have misinterpreted or misused data, somebody must tell us. (Usually somebody does.) And the telling need not be too gentlemanly, either. We would urge saying a lot of these things out in the open where other people can see and hear, too, because these "other people" have a stake in all this—their lives, possibly. We do not find it too dreadful that two letters discussing an industrial exposure to a carcinogen in a recent *New England Journal of Medicine*[22] read like exchanges of "you're an expletive deleted." Some good might come of the exchange.

As for "whose ox is gored," we have already referred to the ethical issue of choosing some people to be at risk so that some others (or "Society") might benefit. We believe a good shibboleth in cost-benefit discussions might consist of the Cornfield questions: "Cost to whom? Benefit to whom?" The decisions on the balancing must be made with affected parties included no matter how hard it is to do. It is enough to sit in on a hearing on the possible banning of a pesticide and to see the Department of Agriculture on one side and the Department of Health, Education, and Welfare on the other to know that "The Government" is not a monolith and cannot make the decision, either.

Summary

Where does this all leave us?

It leaves us able to develop rather good animal data at dose levels that do not

really interest us. That is a first-class highway that takes us where we do not want to go. It leaves us unlikely to be able to develop good data at "realistic" doses. To extrapolate animal results to man exposed at these "realistic" doses today requires assuming a mathematical model of dose-response in the animal and conservative use of this model. Then we have to jump from one species to another in ignorance of the terrain of the landing site, i.e., the many species differences. However, we will have knowledge of some important species similarities and that makes the jump a lot less dangerous. With respect to costs and benefits we are just beginning to understand some of the implications of the arithmetic. We have begun to see that there are few, if any, good ways of totaling the costs or computing the benefits.[25] Cost-benefit may be another blind alley.

Tomorrow and the next day we must do the appropriate research on species differences in metabolism and in the mathematics of the modeling and extrapolation—as a minimum. The socially related issues, such as what is an acceptable risk, what are the costs, what are the benefits, must be discussed in the open, freely. This implies recognizing that someone's costs may be someone else's benefits. (Our medical costs are our physician's source of living.) The inputs to the cost-benefit algebra are not well worked out. Our ways of working must include the adversary approach as well as the pleasanter way of cooperation. And today, we must get to precautionary decisions for man's safety and health based on the road maps from animal data—inadequate as they are. We have gotten to the neighborhood of Park Avenue and 59th Street and we can probably one day get to a lot of other places.

References

1. Weinberg, A. 1972. Science and trans-science. Minerva **10:** 209–222.
2. Armitage, P. & R. Doll. 1954. The age distribution of cancer and a multi-stage theory of carcinogenesis. Brit. J. Cancer **8:** 1–12.
3. Trefall, H. 1973. Extra deaths. Science **182:** 4114.
4. Druckrey, H. 1967. Quantitative aspects of chemical carcinogenesis. *In* UICC Monograph Series, Potential Carcinogenic Hazards from Drugs (Evaluation of Risks), Vol. 7: 60–78. Springer-Verlag, New York, N.Y.
5. Albert, R. E. & B. Altshuler. 1973. Considerations relating to the formulation of limits for unavoidable population exposures to environmental carcinogens. *In* Radionuclide Carcinogenesis. J. E. Ballou, *et al.*, Eds. pp. 233–253, AEC Symposium Series, CONF-72050, NTIS, Springfield, Va.
6. Hoel, D. G. & H. E. Walburg. 1972. Statistical analysis of survival experiments. J. Natl. Cancer Inst. **49:** 361–372.
7. Gail, M. 1974. Measuring the benefit of reduced exposure to environmental carcinogens. J. Chronic Diseases. In press.
8. FDA Advisory Committee on Protocols for Safety Evaluation. 1971. Panel on Carcinogenesis Report on Cancer Testing in the Safety Evaluation of Food Additives and Pesticides. Toxicol. Appl. Pharmacol. **20:** 419.
9. Armitage, P. 1959. An examination of some experimental cancer data in the light of the one-hit theory of infectivity titrations. J. Natl. Cancer Inst. **23:** 1313–1330.
10. Mantel, N., W. E. Heston & J. M. Gurian. 1961. Thresholds in linear dose-response models for carcinogenesis. J. Natl. Cancer Inst. **27:** 203–215.
11. Mantel, N. & W. R. Bryan. 1961. "Safety" testing of carcinogenic agents. J. Natl. Cancer Inst. **27:** 455–470.
12. Mantel, N., N. R. Bohidar, C. C. Brown, J. L. Ciminera & J. W. Tukey. 1974. An improved "Mantel-Bryan" procedure for "safety" testing of carcinogens covering the case of heterogeneous data. Submitted for publication.
13. Peto, R., P. N. Lee & W. S. Paige. 1972. Statistical analysis of the bioassay of continuous carcinogens. Brit. J. Cancer **26:** 258–261.
14. Personal correspondence.
15. Craig, P. Carcinogen Dose-Response Relationship. Progress Report to the Na-

tional Cancer Institute by the Franklin Institute Research Laboratories, Philadelphia, May 7, 1974, Contract No. 43268 (Unpublished).
16. GEHAN, E. A. 1969. Estimating survival function from the life table. J. Chronic Diseases **21:** 629–644.
17. PETO, R. 1974. Guidelines on the analysis of tumour rates and death rates in experimental animals. Brit. J. Cancer **29:** 101.
18. HOEL, D. & N. CHAND. 1974. A comparison of models for determining safe levels of environmental agents. Submitted for publication.
19. BERENBLUM, I. & P. SHUBIK. 1947. A new, quantitative, approach to the study of the stages of chemical carcinogenesis in the mouse's skin. Brit. J. Cancer **1:** 383–391.
20. DOLL, R. 1971. The age distribution of cancer: implications for models of carcinogenesis. J. Royal Stat. Soc. A **134:** 133–166.
21. RALL, D. P. 1974. Problems of Low-Doses of Carcinogens. Presented at the NAS Symposium on Statistics & the Environment. Publication in Proceedings of the Symposium. In Press.
22. BEAVERS, E. M. 1974. Lung cancer in chloromethyl methyl ether workers (Letter to Editor); W. G. FIGUEROA & W. WEISS. 1974. Reply to above letter. New Engl. J. Med. **290:** 971–972.
23. SCIENCE AND TECHNOLOGY POLICY OFFICE, NATIONAL SCIENCE FOUNDATION. 1973. Chemicals and Health. U.S. Government Printing Office, Washington, D.C. No. 3800-00159.