Regulation of Carcinogens

Are Animal Tests a Sound Foundation?

Why does it matter if animal cancer studies are worthwhile or worthless or someplace in between? The answer to this question is that regulation of exposure to chemicals, including the intermittent exposure to trace elements to which the general public is subject, is largely based on interpretations of animal cancer bioassays. If these tests are reasonably accurate in predicting the probability, sites, and severity of human cancers, then regulation of chemicals suspected of causing cancer (carcinogens) is on firm ground. But if these animal cancer tests are weak or worse, so that one cannot reasonably predict human cancers from them, then regulation rests on quicksand.

Whether or not rodent tests predict human cancers, animal studies have many other important uses. Research into cancer mechanisms or problems of the immune system, for instance, may be furthered by introducing novel genes into small animals, such as transgenic mice, to discover better how life systems work (Hanahan 1989). There is no doubt that models based on research with animals have increased our understanding of metastasis, which is so important in the spread of cancer (Schirrmacher 1989). None of the many invaluable uses of animal cancer tests, however, tells us whether they can come close enough often enough to be a valid source of evidence in predicting human cancer.

The Right Questions

We would like to know how much damage to human populations is caused by different types and quantities of exposure to various substances. That

---

[0]The late Aaron Wildavsky was Class of 1940 Professor of Political Science and Public Policy at the University of California, Berkeley, and a Research Fellow for The Independent Institute.

The Independent Review, Vol.1, No.1, Spring 1996, ISSN 1086-1653, Copyright © 1996, pp. 29–54.

estimating requires answering subquestions about conditions. Routes of exposure may differ in that some come from breathing, others from eating, and still others, such as x-rays, through the skin, from natural sources and medical uses. Quantities may differ from a little to a lot to immense. Timing of exposure differs from all at once to over long periods of time. If a precise answer to the question of adverse effects cannot be given, we might be satisfied with knowing that there is a great deal of harm or moderate or very little or probably none (see Hattis and Kennedy 1986, 66).

How reliable are these tests? If the tests were repeated on the same species, would we get nearly the same results? If they were repeated on different animal species, would we come up with similar results? If chemicals are carcinogenic in several animal species, it is more likely that they are carcinogenic to mammals in general, including human beings, than if they cause cancer in a single species. It is also important to distinguish rates and sites of cancer by age, because cancer is largely, though not entirely, a disease of old age, and by sex, because men and women are affected differently.

Dioxin in large and continuous doses appears unfriendly to mammals, but it is the dose that matters. Regulatory agencies assume that chemicals carcinogenic at some dose in any animal are also carcinogenic to human beings. We want to find out if that assumption is true.

As precisely as possible we wish to answer Freedman and ZeiselÕs question, ÒAre chemicals that have been shown to be carcinogenic through experimental animals also carcinogenic to humans?Ó (Freedman and Zeisel 1988, 14). The reason for their inclusion of the modifier Òexperimental animalsÓ has to do with the particular conditions under which animals are tested. Therefore they also ask, ÒDo experimental animals (rodents, in particular) and humans have similar susceptibility to the carcinogenic effect of chemicals, or are rodents incomparably more susceptible than humans?Ó (14). The answer to the first question is: ÒsometimesÓ rodent cancers are cancers in humans too, but we do not know when. The same is true of one type of rodent to another. The answer to the second question is: Òyes, mostly, but not always.Ó If we know that a single LD50 dose for dioxin ranges from 2500 µg/kg in guinea pigs to 5000 in hamsters, a difference of 2500, does that give us confidence about rodent-to-human transfers?

Suppose we were to find that there is a 10 percent probability that a substance causing cancer in a mouse or rat at a given dose will do the same in a human being. From one point of view, 9 times out of 10 the extrapolation from mouse to man would be wrong. From another point of view, why take chances with human health if the probability of getting a cancer is that high?

Were we to find, however, that when we get the answers from the tests we would know within a factor of several hundred times to several hundred thousand times whether rodents predict to humans, that might not be a reasonable approximation. The question is not only whether we can get an

answer but what kind of answer we will get.

The Process of Animal Cancer Testing

Around 1915 or 1916 scientists learned that they could induce cancer in animals by treating them with certain chemicals. The methods of giving animals cancer vary greatly. Chemicals have been introduced into experimental animals by every orifice (orally, nasally, urethrally, vaginally, rectally), by various types of injections (intramuscular, intraperitoneal, intravenous, subcutaneous), by skin painting, by surgery, and by other methods (American Council on Science and Health 1984, 7).

Approximately 30 percent of the rodents get some form of cancer absent exposure to chemicals, though not all 30 percent die of it. This is one reason why a control group is essential. Because a chemicalÕs effects at high doses may not show up at low doses, it is necessary to further subdivide the animals into different dose groups. Given that sex plays an important role in cancer, a further subdivision is between male and female. Usually there are three dose groups (0, 0.5, or 0.1 the maximum tolerated dose [MTD], and the MTD) and two species. There are at least 12 groups of animals. By convention and by statistical necessity, there are usually 50 animals, most often rodents, in each group.

Though only a few facts about the process of animal cancer testing have been given, we are already in a position to understand three of its most basic aspects—its short time compared to human epidemiological studies, its high cost, and its essentially statistical character. A great advantage of rodent testing is that these animals live only about two years. Therefore one doesnÕt have to wait too long to get results. One can also test any chemical, including new chemicals, for which epidemiological evidence may not be available. But the task is not easy or cheap. It is costly to keep these animals under controlled conditions for up to two years. The painstaking work of examining animals for tumors requires pathologists. When each animal dies or, as the too-kind parlance states, is sacrificed, several pathologists must carefully examine about 40 sites within and around animal organs and tissues to search for tumors, some of which are so small they can be discerned only with high-powered microscopes.

That is why there is a team of pathologists who first work separately and then meet to resolve differences before their findings are accepted for further evaluation (Chu, Cueto, and Ward 1981). These pathologists consider whether the tumors or other abnormalities are actually induced by the chemical, an opinion based on what they know about the normal incidence of tumors and their experience. They ask themselves not only whether the incidence of tumors is higher but whether they are of a different size or shape or color or contain any other signs that might show them to be similar to or different from naturally occurring lesions (Chu, Cueto, and Ward

1981, 256–57).

In order to understand better whether the proper dose was administered, the animals have to be weighed to discern whether they have lost appropriate amounts of weight and examined to see whether the dose is either so large as to threaten their lives from causes other than cancer or so small as to make its effects unnoticeable.

Now we are in a position to understand why rodent cancer tests are so expensive. When one multiplies the time these tests take, roughly three years, by the cost of keeping 12 groups of animals in controlled conditions, then adds the costs of killing and dissecting them, preparing and examining 40 slides per animal, and reconciling differences, the substantial costs do not appear out of line (see the discussion in Rowan 1984). It is possible for a government regulator to conclude that the tests are inadequate or that the substance being tested is a carcinogen. But it is not possible under the rules to say that the substance is not, insofar as is known, a carcinogen; the closest government scientists are allowed to come is to say that Òthe compound has not been shown to be carcinogenicÓ (Chu, Cueto, and Ward 1981, 252–53). What, we may ask, is the meaning of classifying a substance as a suspected carcinogen? It is worth attending closely to Chu and his colleaguesÕ discussion:

> If malignant tumors or a combination of malignant and benign tumors are produced, then the compound is considered carcinogenic to the animals. If the significant result is only the production of benign tumors, then the compound may pose a potential health hazard and is termed a suspected carcinogen or a carcinogen, depending on the nature of the benign tumor. For example, 2,4-dinitrotoluene...was considered a suspected carcinogen since it induced only benign tumors (fibromas of the skin and subcutaneous tissue in male Fischer 344 rats and fibroadenomas of the mammary gland in females). Ideally, a distinction should be made between truly benign tumors, which never progress to malignancy, and tumors that are in a benign state according to histopathologic criteria at the time of diagnosis. Scientific judgments in this area are limited by inability to predict the biological behavior of a lesion on the basis of morphological criteria, but it appears that there are few, if any, truly benign tumors in rodents. (Chu, Cueto, and Ward 1981, 257-58)

The Òit appearsÓ in the last sentence above reflects a judgment that any tumor might turn bad. Is it in the interests of public safety to treat all tumors, however benign in appearance, as if they might turn malignant, because we do not know they wonÕt? Or is saying that they might turn malignant a way of prejudicing the outcome so that the chemicals will be found to induce cancers whether they do or donÕt?

In the Environmental Protection Agency (EPA) 1976 ÒInterim

Procedures and Guidelines for Health Risk Assessments of Suspected Carcinogens,Ó EPA Administrator Russell Train acknowledged that animal tests could not prove that a chemical would be carcinogenic in people, but that a substance would be considered a Òpresumptive cancer riskÓ if it Òcauses a statistically significant excess incidence of benign or malignant tumors in humans or animalsÓ (p. 21403). If benign is bad, what could be good?

Calculating Toxicity by the LD50 Test

In the field of pesticide regulation, lethality is calculated through the assignment of an LD50, the lethal dose for one-half of the test animals during the test period. The relevant number for aspirin would be 730 mg/kg, signifying that 50 percent of the test animals died when exposed to 730 mg of aspirin per kilogram of their body weight (Edwards n.d.). The larger the LD50, the more of a substance it takes to produce a toxic effect and the less harmful the chemical.

Among species most commonly used to carry out the LD50 test are fish, birds, rabbits, mice, and rats, although occasionally monkeys and dogs are used. Generally, about 60 animals of a particular species and a specific dosing method are used. The application is made by inserting a tube down the throat of the animal, by forcing injection of vapors, or by application to the skin (Paget 1970). The usual test lasts about two weeks, during which the animals either die or, at the end, are killed. The usual symptoms are bleeding from the mouth or eyes, convulsions, diarrhea, and what are exquisitely termed Òunusual vocalizations.Ó Rather than tolerate early death, according to the British Toxicological Society (1984), ÒThere is pressure on the toxicologists to allow the study to continue, even when the animals are in distress since their premature killing may alter the end-point of the study, and so possibly affect the classification of the material being tested.Ó Needless to say, animal rights advocates are not happy with this method.

Whether one believes that the LD50 test involves Òa ritual mass execution of animalsÓ (Rowan 1984, 207) or that Òthe main information they give is an indication of the size of dose required to commit suicideÓ (Baker 1969), or even whether most experts consider Òthe modern toxicological routine procedure a wasteful endeavor in which scientific inventiveness and common sense have been replaced by a thoughtless completion of standard protocolsÓ (Zbinden 1976), there is ample scientific doubt about the value of the LD50 test for the purpose of predicting effects on humans.[0] The basic difficulty is that enormous differences between different (even closely related) species are reported, ranging from 5 to 75 times, which renders

[0]. See Sharpe (1988) for a list of authorities with negative verdicts.

findings suspect.[0] If LD50 tests are useful in providing evidence to save human life from suffering, there would still have to be a debate over whether the animal suffering entailed can be justified. If, however, the observations are too unreliable to be useful, no such question arises.

Toxicity

The best explanation for laymen I have heard of the differences with which we have been concerned—between very large and very small exposures in different kinds of species—comes from reporter Richard Harris of National Public Radio (1992, 8–10), together with a number of cancer researchers and government officials. Their dialogue is instructive:

> PENELOPE FENNER CRISP (Environmental Protection Agency): WeÕre coming to discover that there are more differences between species than we had expected or, frankly, hoped that existed.
>
> HARRIS: It turns out that a great many chemicals that can cause cancer in one species donÕt seem to do anything at all in another species. HereÕs an analogy.
>
> [Excerpt from music from a CD]
>
> HARRIS: The difference between rodents and people can be as dramatic as the difference between this CD and an LP. You could drop this CD, get it dusty, even scratch it, you wouldnÕt necessarily hurt it.
>
> [Sound of record being scratched by a stereo needle and music played from an LP]
>
> HARRIS: But try the same thing with a record, and you can just hear the damage. To be sure, some things will damage either a CD or a record album—say a hot windowsill. Likewise, John Doull from the University of Kansas says some chemicals do cause cancer in all sorts of animals.... Now, nobodyÕs suggesting that these chemicals are harmless, but in some cases scientists believe that the standards may be vastly overstating the health risks. Again, this comes down to a necessary but flawed shortcut the EPA uses to size up a chemical. Scientists give a huge dose of chemicals to rats and then estimate the effects of that chemical at lower doses. By way of analogy, if you drop a bottle from 10 feet off the ground, itÕs pretty obvious whatÕs going to happen.

[0]. See the numerous examples in Rowan (1984).

[Sound of glass shattering]

HARRIS: This large drop is equivalent to a large dose of a chemical, and it can be deadly. But what if, instead of taking one bottle and dropping it from 10 feet, you take 10 bottles and drop them from one foot? ItÕs like giving many people a smaller dose of that toxic chemical. HereÕs what the EPA assumes will happen.

[Sound of several bottles hitting the ground and one of them shattering]

HARRIS: They figure one of the 10 bottles will break. The reasoning is that one-tenth the dose, or one-tenth the drop distance, will do one-tenth the damage. In reality, though, this is what happens.

[Sound of several bottles hitting the ground]

HARRIS: There is, in fact, a safe height you can drop a bottle from without breaking it, and John Doull from the University of Kansas says the same idea holds for toxic chemicals.

DOULL: It is the dose, not the compound, that determines its adverse effects....

HARRIS: So, recently, researchers like Swenberg have started to dig deeper and ask why some chemicals trigger cancer in some animals. ItÕs as though theyÕre trying to understand the difference between turntables and CD players. And Swenberg says one especially interesting example is unleaded gasoline. You may have seen the sticker at the pump warning that gasoline causes cancer in laboratory animals. Well, hereÕs the story with gasoline.

SWENBERG: It causes kidney cancer in male rats only, not in female rats and not in mice.

HARRIS: So whatÕs going on? Swenberg decided to find out by studying those animals, and he discovered that a chemical in gasoline binds to a naturally occurring protein thatÕs only found in the kidneys of male rats.

SWENBERG: And this results in a buildup of the protein and ultimately leads to the development of cancer. And since humans do not synthesize this protein, this is not likely to be a mechanism important to humans.

> HARRIS: Swenberg says dozens of other chemicals besides gasoline cause this specific kidney cancer in male rats, including copy machine toner, a bathroom deodorizer, and even a natural chemical called D-limonene.
>
> SWENBERG: It turns out that about two glasses of orange juice contains a carcinogenic amount of D-limonene for the male rat, but it has absolutely no effect on mice or on female rats, and IÕm sure it has no effect on humans.
>
> HARRIS: As a result of this research, the Environmental Protection Agency recently decided that if a chemical like gasoline only triggers this kind of kidney tumor in male rats and it doesnÕt do anything else bad, itÕs probably not going to cause cancer in people. So far there are just a handful of stories like this where scientists have actually figured out why a compound is causing tumors in certain animals. But there are a lot more studies in the works, including reassessments of dioxin, formaldehyde, and certain PCBs.

Knowledge of mechanisms yields far greater discriminatory power. With such knowledge scientists can determine whether there is a threshold below which there is no damage or whether harm occurs proportionate to the dose, however low that dose is. Without knowledge of the mechanism we cannot be sure of the dose-response relationship.

The War over the Dose-Response Threshold

There is disagreement over whether there is a dose-response threshold, such that below a certain level no harm occurs, or whether the damage is linear, such that harm from a chemical increases or decreases as a proportion of the dose. It is important, to start with, to ask why such an apparently technical matter has occasioned so much dispute. Because the field of toxicology is built on the principle that the poison is in the dose, the opposing linear (or proportional) principle—that there is no threshold dose below which damage cannot occur—is a challenge to toxicological science.

A common statement about dose-response levels is that no one really understands what happens when people are exposed to very low levels of chemicals (see, for instance, Marx 1990). There is no difficulty in finding substances, such as the heart medicine digitalis, that are helpful at low doses but can be fatal at large doses. But that does not answer the question of whether there are substances for which no threshold exists (see Smith and Sharp 1984, 3). Given that there is a considerable range of sensitivity among human beings, it can always be said that some hypersensitive people might be adversely affected. The traditional response has been to use a margin of safety to take care of the supersusceptible. Given also that chemicals may

interact with each other to create cancers that neither substance would alone, it cannot be said definitively that either is safe. By the same token, however, one chemical may render another harmless or less harmful (see Revision without Revolution 1984).

The regulatory response is that the dose-response relationship is linear. The rationale is that this provides a margin of safety for the public. The question is whether this assumption is true.

Going further into the Òfurious battle [that] rages around the 'threshold controversyÕÓ (Rowan 1984, 234–35), it will be instructive to read a semiofficial account by high-ranking Environmental Protection Agency officials published in a major journal, Risk Analysis. The models the EPA uses attempt to establish an upper bound, nearly the worst that could happen, on the basis of a no-threshold linear response. Anderson et al. (1983) are quite open in saying that Òrecognition that the lower bound may…be indistinguishable from zero stems from the uncertainties associated with mechanisms of carcinogenesis including the possibility of detoxification and repair mechanisms, metabolic pathways, and the role of the agent in the cancer processÓ (281). In short, for all the EPA knows, there may be no damage at low doses.

Furthermore, Ò[m]ost often there is no biological justification to support the choice of any one model to describe actual riskÓ (Anderson et al. 1983, 281). This task would be easy if there were data on actual environmental exposures in human beings, in which case an appropriate model could be fitted to the data. ÒIn the absence of such data a variety of models can be used to fit the data in the observed range, but these models differ sharply [in the danger estimates they produce] at low doseÓ (281). If the choice of model determines the results, because they Òdiffer sharply at low doses,Ó why bother with the experiment? Exactly. Employing the justification that nevertheless these models are the best available, Anderson et al. state that Ò[i]t should be clear from the preceding discussion that the linear non-threshold model has been used by the EPA to place plausible upper bounds on risk, not to establish actual riskÓ (281). This is a significant admission.

Use of the upper bound misleads people into thinking it is an actual estimate of hazard by an authoritative government agency when it is not. Use of Òworst caseÓ scenarios makes no sense, moreover, when the outcome may be zero and there is no biological sense in anticipating epidemic consequences.

Now we know that everything depends on which of the available statistical models are used and whether whichever one is chosen, in the absence of biological indications, tells us what we need to know. Does it?

The EPA claims that the linear, Òno dose-responseÓ model best fits knowledge about cancer causation. But its officials could not know this without knowledge of the mechanisms at work, in which case they would be able to choose a model they knew fit the causal relationship. At other times,

they acknowledge the real basis for their choice of model, the desire to choose the most conservative estimate so as to, as the saying goes, act on the side of safety. But are they so acting?

If it is true, as Anderson et al. (1983) say in their appendix, that Ò[t]here is no really solid scientific basis for any mathematical extrapolation model relating carcinogen exposure to cancer risks at the extremely low levels of concentration that must be dealt with in evaluating environmental hazardsÓ (289–90), then why make one? The answer must be that with going from rodents to people, most regulation of chemicals would lack a rationale that could be called scientific. No man-mouse extrapolation, no science; no regulation. The models that make this extrapolation plausible are the important thing. Extrapolations from animals dosed at very high levels to people exposed to far smaller levels make sense only in the context of the models of cancer causation into which they are meant to fit.

Multistage Models

Is this chemical hazardous? And to what degree at which dose and to whom? Interpretation of cancer causation depends on models, which we can think of as representations of theories, with numbers attached, that give meaning to data. Actually, there have to be two models in one: first, a model of the biology underlying cancer causation, and second, the statistical approximation of that model. Getting accurate results depends both on the predictive power of the biological model of cancer causation and on whether the statistical approximation captures the causal structure of the model. If the model does not well describe cancer causation in human beings and if on top of that the statistical approximation does not well describe the model, the errors in both models multiply to give unsatisfactory results. The task is a daunting one.

We can take the Armitage-Doll model as representative of those used by governmental agencies in regulation. It seeks to describe the relationship between exposures to chemicals and the incidence of cancer at various ages for men and women (see Moolgavkar 1986). The biological version portrays human cells as going through a number of stages that ultimately result in cancer. The hypothesis is that one or more cells receive an insult and then go through several changes that turn them into malignant cells, after which they proliferate. The times and the different stages are not specified. All these stages are probabilistic in that some cells under the same exposure will become cancerous and others will not. ÒPut another way, with multistage models,Ó Richard Peto (1977) tells us, Òwhen all the predisposing factors have been allowed for, luck has an essential role in determining who gets cancer and who does notÓ (1404). Thus the stages in the models are essentially probabilities, and the users do not know whether human cancer proceeds in those stages or according to those odds (Freedman and Navidi

1989, 72). In the field of economics, these would be called Markov chain models, which means essentially that every present stage depends on results of previous stages. The time spent in the various stages is assumed to be proportionate to the exposure of the affected individual.

The basic difficulty with multistage models, as the reader might imagine, is that there is little reason to believe they actually capture the biological process of cancer formation. At the same time, the statistical manipulations are very far from the causal requirements of the model, so that one has no idea what one has got when the result is cranked out (Moolgavkar 1986; Freedman and Navidi 1989). This brings us to the statistical interpretation of animal cancer studies, the most critical and least understood part of modeling cancer causation.

The existence of 12 different test groups shows that statistical inference is the essence of the matter: after all, conclusions are to be drawn by observing differences between the control group and other animals and between gender and dose groups. This is not something that is done by counting on the fingers of one hand. It requires methods based on statistical theory.

Biological Interpretation of Animal Cancer Tests

Fears et al. are wise in concluding that Ò[t]here is danger in relying solely on the finding of statistical significance without incorporating biological knowledge and corroborative evidence such as the presence of a dose-response relationship for experimentally consistent results in different species or sexesÓ (Fears, Tarone, and Chu 1977, p. 1941). But what if there is little or no biological knowledge?

In order to get accurate estimates of the probability that chemicals that cause cancer in animals also cause cancer in human beings, Salzburg (1983) recommends applying Òthe bioassay to a number of innocuous substances. There have to be some compounds that are not human carcinogens, or the whole exercise of looking for carcinogens makes no sense.Ó Yet after examining the literature, he finds that Òthis was never done for the [rodent] lifetime feeding studyÓ (65). His argument needs to be heard in full:

> Thus, it would appear that no attempt has ever been made to determine how well society can identify human carcinogens by feeding groups of 50 rats and mice, each, the suspect substance at maximum tolerated doses for their entire lives. Common scientific prudence would suggest that this assay be tried on a group of known human carcinogens and on a group of supposedly innocuous substances (such as sucrose or amino acids) before we either (1) believe that it provides some protection for society (sensitivity) or (2) believe it identifies mainly harmful substances (specificity). There is no substitute for such proper validation on any new bioassay. (Salzburg 1983, 63)

He believes Òthat we are confusing the effects of biological activity upon the old-age lesions of rodents with the thing we fear, cancerÓ (Salzburg 1983, 64, 65; see also Salzburg 1980 and Tomatis, Agthe, and Bartsch et al. 1978). There is also the claim that cancers found in rat autopsies are induced by test procedures that feed the rats at maximum tolerated dose.

Mitogenesis: Is Cancer Caused by the Test and Not by the Chemicals?

Proof that the tumors observed in animal cancer tests are due to the huge doses delivered at the MTD would be fatal to the no-threshold idea, for then the animal cancer tests themselves would be taking out only what they put in: cancer in, cancer out. Bruce Ames and Lois Gold, among others, claim that the chronic wounding induced by delivering heavy doses of a chemical promotes cancer by inducing cell division, a process called mitogenesis. As the animal is effectively wounded or poisoned, it grows replacement cells, a process known to increase chances of mutation and hence of cancer.

The theory was prompted by finds that while cancer is thought to be accompanied by mutation, alteration, or damage of DNA, a large proportion of chemicals that cause cancers in animal tests do not in fact damage genes in other tests. There are, as a paper by Ames and Gold is titled, ÒToo Many Rodent Carcinogens,Ó based on expectations flowing from knowledge of cancer and a belief there would be a great deal more cancer around if half the chemicals in the world caused this terrible sickness (Marsalis and Steinmetz 1990b, 10; Ames and Gold 1990a). Proliferation of cells with DNA damage is an important element in production of cancer in human beings. Cell proliferation is caused by chronic toxicity, by ionizing radiation, by chronic inflammation, and by hormones and viruses that cause infections that in turn lead to cells dying and hence to cell proliferation (Marsalis and Steinmetz 1990b, 10; see also Ames and Gold 1990b and 1991).

Support comes from Cohen and Ellwein (1990): ÒChemicals that induce cancer at high doses in animal bioassays,Ó they assert, Òoften fail to fit the traditional characterization of genetoxins. Many of these nongenetoxic compounds (such as sodium saccharin) have in common the property that they increase cell proliferation in the target organ.Ó They argue that Òthe increase in cell proliferation can account for the carcinogenicity of the nongenetoxic compoundsÓ (1007). Similarly, Daniel Krewski finds a Òfairly strongÓ correlation between carcinogenicity and toxicity, which one would expect to find when test animals are being wounded by being fed the maximum tolerated dose (Marx 1990, 744). If the mitogenesis theory is correct, then rodent tests run at mitogenic doses are invalid as predictors of human cancer from exposures below toxic levels.

Ames (1990) lays it right on the line: ÒWe think the current approach to cancer risk assessment is bankruptÓ (A38). A reply by Richard A.

Griesemer, who is head of the Division of Toxicological Research and Testing at the National Institute of Environmental Health, is: ÒThere are only two definitive ways to tell whether chemicals have potential to cause cancer. One is through epidemiological studies in humans....[T]he second way is to produce cancers in mammalsÓ (Ames 1990, A38). In other words, we are not to experiment on people, so we have to do so on animals. But do we have to experiment on animals if the results are meaningless?

The most sustained attack on the Ames-Gold thesis (that the animal bioassay, by feeding animals the MTD, is itself causing the appearance of excessive rates of cancer) is by Perera. Her first argument is that a variety of international agencies, including agencies in the United States, have Òadopted the general assumption of low-dose linearity for carcinogens— regardless of their presumed mechanism of action.Ó The rationale is that there must be something to it since so many agencies have gone in the same direction. Why? Scientists usually do not argue from authority when they process a theory they can validate with evidence.

The reasons, according to Perera, are a general lack of understanding of the mechanisms of cancer causation, especially those termed nongenetoxic; a lack of agreement on a safe threshold level below which exposure would not be harmful to a diverse population; and Òthe desirability of preventing cancer through the use of testing and model systems, obviating the reliance on epidemiological data in humans.Ó[0] The question is whether regulation should be undertaken as a replacement for the existing lack of knowledge. That is exactly what she advocates: ÒIn the meantime, EPA cannot ignore its responsibility to evaluate and control synthetic chemicals...since no one...has yet devised an acceptable alternativeÓ (Cogliano et al. 1991, 607). I disagree.

The argument that if we donÕt regulate weÕll count dead bodies is dead wrong. The predicted cancer rate at one in a million (even one in 100,000 or one in 10,000) is so low it will never be detected by epidemiology or any other method unless we know a lot more about the mechanism of cancer causation. ÒThe problem with...risk assessments...based an animal tests,Ó the Office of Technology AssessmentÕs Michael Gough tells students, Òis that their theories cannot be testedÓ (Gough, forthcoming).

False Positives or False Negatives

One way of looking at a test is to ask whether its error rate is low or whether there is a high proportion of false positives—those chemicals that do not cause cancer at the administered dose but are wrongly believed to do so (Fears, Tarone, and Chu 1977; Gart, Chu, and Tarone 1979). The opposite

[0]. Perera (1990, 1644). See Ames and Gold's response to Perera in the same issue (Ames and Gold 1990b).

error, an overabundance of false negatives—those chemicals that cause cancer at the administered dose and are wrongly believe to be benign—can also occur. Governmental agencies set their requirements so as to minimize the chances of false negatives.

Statistical phrases—such as Òlow-dose extrapolationÓ and Òlow-dose linearityÓ—determine regulation of chemicals in the United States. They are not sideshows; they are center ring. In 1985 an interagency committee working for the Executive Office of the President published 31 principles for conducting quantitative risk assessments (U.S. Office of Science and Technology Policy [OSTP] 1985). Appropriately, some of the principles emphasized the limitations of scientific knowledge about chemical carcinogenesis. For example, the committee acknowledged that existing knowledge was not up to determining whether chemicals that caused cancer at high doses have a carcinogenic effect at lower doses (OSTP 1985, 10376). The document accepted that Òno single mathematical procedure is recognized as the most appropriate for low-dose extrapolation in carcinogenesis,Ó but nevertheless endorsed linear extrapolation techniques: Òuncertainties are involved in the use of any of the commonly employed extrapolation models, [but] models…which incorporate low-dose linearity are preferred when compatible with the limited information [available]Ó (10378). Why preferred? No doubt because they are conservative, though no one can calculate how conservative they are.

To extrapolate from rodents to people, a number of basic assumptions must be made, of which three are of primary importance: (1) the biology of these mammals is sufficiently similar to justify the extrapolation; (2) there must be an adjustment for the huge size of people compared to rodents; and (3) the vast differences in doses given to animals, which we have seen is essential to make rodent tests feasible, must be taken into account. After all, perhaps the greatest controversy surrounding animal cancer tests is whether a chemical given to rodents in huge doses would actually produce cancer at much lower exposure.

Where might one find false positives? Site is important. Research reveals that false positives are more likely to occur at sites with a high rather than a low number of spontaneous tumors (Fears, Tarone, and Chu 1977, p. 1941). It is also known that rare tumors are less likely to be false positives than are common ones. Thus knowledge of the spontaneous tumor rate is essential, especially if it is above 5 percent, because it then becomes difficult to tell the natural from the chemically caused tumor (Chu, Cueto, and Ward 1981, 259–62). The type of chemical also interacts with the type of animal; for instance, some rat organs, when exposed to chemicals, are pretty good predictors of tumors in mice, but mouse liver is a very poor predictor of tumors in rats. There is also a striking difference between chlorinated and nonchlorinated chemicals in regard to the sensitivity of mice versus other animals (Gold et al. 1989, 218; see also Marsalis and Steinmetz 1990a, 156).

The most important defect of animal cancer studies, as Freedman and Zeisel demonstrate, is that the choice of statistical models overdetermines the results: in speculation regarding the effects of the low doses to which human beings are subject in nonoccupational exposures, the choices of statistical models produce outcomes that vary by hundreds, thousands, ten of thousands, and occasionally millions of times. Yet without knowledge of the biological mechanisms of cancer causation, there is no way of choosing among these models.

In the study of the grain fumigant EDB, for instance, the probability that an individual would get cancer from eating food in which tiny amounts of EDB were present varies over a million times, depending on the model (Hattis and Kennedy 1986, 65). Using the same animal data but different statistical models in regard to saccharin, to take a well-known but extreme instance, led to differences of some five million times (National Academy of Sciences 1978). As Table 1 shows, statistical models that depend on different assumptions predict harms that vary by a factor of 200 for DDT, 800 for dioxin, and 40,000 for aflatoxin, a naturally occurring toxic substance found in various foods, especially peanut butter.

With results this far apart using different models, and without a plausible biological reason for preferring one model over another, the results are no better than guesswork. No, they are probably worse; I doubt that educated guesses would produce results so far apart. Would the reader accept the result in anything that mattered if that result was anywhere between 100 and 40,000 times off and the reader could not know which? Even if there were no alternative sources of information, we would do better ignoring such empty data.

Echoing lines from literature, my students often ask whether animal cancer tests could not be considered second best and therefore better than nothing. I ask them whether it would be second best if they want to go from Berkeley to Oakland (they are contiguous) or Brooklyn to Queens or Minneapolis to St. Paul and are advised to travel via Beijing. Some alternatives are so bad they should not be dignified by the designation Ònth best,Ó but rather should be recognized as too bad to be used.

Table 1: The Impact of the Model on Low-Dose Risk Estimates

Substance	One-hit	Multistage	Weibull	Multihit
Aflatoxin	1	30	1,000	40,000
Dioxin	1	400	400	800
DMNA	1	700	700	2,700
Dieldrin	1	3	200	1,000

DDT	1	2	70	200

Notes: From Food Safety Council (1980, Table 4). The virtually safe dose is estimated (from each of the four models) as that dose giving a risk of one in a million. The column for the multistage model shows the ratio of the estimated from one-hit model, for each of the five substances. Likewise for the Weibull and the multihit (Freedman and Zeisel 1988, 11).

In retrospect, to be sure, we can see that heading due west from Spain wasnÕt such a great route for Columbus to follow if he wanted to get to India. But with the more limited knowledge of ColumbusÕs day, he was justified in sailing. Had he continued, more to the point, he would ultimately have gotten there. This much cannot be said of animal cancer tests. They can be performed forever without improving our ability to predict cancer in human beings. Had Columbus been even 200 times off, he would have been on another planet.

Alternatives

What are we to do to protect people against cancer? Eat fruits and vegetables? Stop smoking? These behavioral changes help a lot. But our question concerns the effects of chemicals. Epidemiology would be best in that it is the most accurate in predicting rates of cancers from different levels of exposure to various chemicals. It is also morally appropriate in that human beings are tested to protect human beings. But there are shortcomings. We could collectively decide to accept the lesser sensitivity of epidemiological studies to gain their accuracy and reliability, and at the same time seek the mechanisms of causation that alone hold promise of effective intervention. But that would leave many people worried.

Currently, the alternative to epidemiology is animal cancer tests at the MTD with interpretation based on multistage models dependent on administering the MTD. The advantages are better control over conditions in the laboratory and the ability to get short-term results. But the defects are insuperable, for without knowledge of the biological mechanisms through which cancer is caused, there is no good way to interpret results. When all we know is that the potential link between exposure to a given chemical by rodents is dozens, hundreds, thousands, or tens of thousands of times away from human exposure, we know nothing of value. Animal cancer tests for the purpose of predicting the effects of small, intermittent, nonoccupational human exposures are not better than choosing at random. What, then, are we to do about the continuing stream of chemicals being used and concocted?

The demand either to use animal cancer studies or to find a substitute depends on the belief that very small amounts of chemicals cause significant amounts of damage to human beings and other living creatures. If large doses are required to cause large effects, epidemiology would do. The belief in the power of small doses is reinforced by the related view that there is no level below which damage does not occur.

If no level is low enough to be safe, then we would expect, with the introduction of so many new chemicals from the 1950s onward, that deleterious effects would show up in health statistics. But they do not. Neither a general cancer epidemic nor the once-feared asbestos epidemic has materialized. Once we control for cigarette smoking and hence lung cancer, for age (because cancer is largely a disease of old age), and for AIDS-related cancers, overall cancer rates either are falling or have leveled off. Life expectancy has gone up decade by decade from the turn of the century.

This should not be surprising. The chemicals regulated are so small in amount and so far from people that they could hardly do much damage unless, through unknown mechanisms, very small exposures are doing significant harm. In a seminal study, Michael Gough (1990) showed that if everything the EPA claimed for its regulation proved out, the most that could happen would amount to a 1 percent or smaller reduction in cancer rates. The EPA criterion for regulation limits exposure to some 374,000 times less than a dose shown to cause harm in experimental animals (Gaylor 1989). The health benefits from limiting exposure to such tiny amounts are likely to vary from insignificant to nonexistent.

In an article titled ÒInformation Value of the Rodent Bioassay,Ó Lester Lave and his co-authors (1988) conclude that

> [f]or almost all of the chemicals tested to date, rodent bioassays have not been cost-effective. They give limited and uncertain information on carcinogenicity, generally give no indication of mechanism of action, and require years to complete. Instead, some of these resources should be devoted to improving the sensitivity, specificity and cost-effectiveness of alternatives to the long-term bioassay, such as in vitro [literally, Òin glass,Ó i.e., in the tube] and in vivo short-term test schemes.[0]

I agree. I also agree with Salzburg that

> [p]resently the lifetime feeding study pre-empts the field. As long as it is considered to be useful in detecting human carcinogens this very expensive and time-consuming procedure will continue to drain the toxicological resources of society. This report questions its usefulness and suggests that it is time to seriously consider the

[0]. Lave et al. (1988, 633). As the NMAS (National Medical Advisory Service) asserts: ÒAfter more than 15 years of utilizing the B6C3F1 mouse as a mainstay animal on which to perform cancer risk assessment studies, many in the scientific community are calling for a methodology review. At issue is whether or not using this particular test mouse results in inaccurate conclusions. It was specifically bred to be sensitive to cancer causing agents, and it has a high rate of spontaneous tumors (25 to 30 percent). The theory behind the creation of this type of mouse was that if the tests were being performed on a very sensitive animal, the data produced would be conservative, therefore setting very cautious levels of exposure. However, there is a groundswell of opinion today which recognizes that this test mouse has produced results that are overly cautious, and perhaps an inaccurate base upon which misleading risk assessments are being conductedÓ (NMAS 1992, 1).

alternatives. (Salzburg 1983, 66)

There are alternatives. Some in vitro tests seek to detect damage to DNA within the cell nucleus. The Ames bacterial test detects mutagens. None of these in vitro tests predict human cancers (Goldberg and Frazier 1989).

Another of many possibilities is pharmacokinetics, in which quantitative mathematical modeling is used to estimate such things as how much of a given chemical gets to a particular kind of tissue and the absorption, metabolism, and distribution of the chemical in the human body (Wilson 1991). This is a theory-building exercise. How it can be related to theory testing is not yet decided. The conclusion has to be that while alternatives are promising, they have yet to fulfill their promise.

If epidemiology is too insensitive and animal cancer tests are invalid, the questions remain: What should be done to reduce cancer rates? How should the multitudes of chemicals be treated until we possess the knowledge to eliminate or restrict those that cause human cancers at low doses? Should our collective decision be to cut through the complexity by severe regulation? Would such a policy actually improve our health? What type of strategy is suited for our current ignorance? What do we actually know about the sources of carcinogenic chemicals to which human beings are exposed? The two categories of interest are synthetic carcinogens produced by industry and natural carcinogens produced by plants to ward off their predators.

It is no small matter to read a report by Washington Post science writer Malcolm Gladwell beginning with the headline ÒNew Panel Questions Traditional Carcinogen Testing: Cancer Experts Respond to Growing Doubts about Massive-Dose Experiments in AnimalsÓ (8 September 1990). The article starts out by saying: ÒThe nationÕs top experts on assessing whether chemicals cause cancer say that traditional methods are sometimes misleading and that improvements or entirely new methods should be developedÓ (A5). At a meeting lasting three days, a National Academy of Sciences panel on risk assessment stated, in GladwellÕs words: ÒThe use of rats ands mice to test potential carcinogens—a practice that has formed the basis for regulating chemicals in the United States for more than twenty-five years—should be brought under sharp scrutiny. Many scientists said the studies are too unreliable and too inaccurate to form the basis for evaluating risks to humansÓ (A5). The panel was particularly critical of feeding animals at the maximum tolerated dose, because it was so much greater in proportion to weight than that to which human beings were subject.

Reform versus Revolution in Risk Assessment

There we have it: rodent studies are speedier but too inaccurate, while mechanistic studies are exceedingly accurate but very slow. Epidemiology lies in the middle on both counts: it is far more accurate than animal tests

(being tried on humans at normal doses) but not accurate enough to detect effects at low doses, especially with smaller populations. My associate, Leo Levenson, recommends a return to the traditional method of controlling the consequences of toxic substances without making special cases out of those that might conceivably cause cancer. For this purpose he would accept the results of rodent cancer tests but would then apply a 100–fold reduction from the level that caused cancer as determined by rodent bioassays. If there were knowledge or experience that led to greater concern, he would then multiply by 10 to reach a level 1,000 times below the animal test level. Were there less reason for concern a level only 10 times under might be applied.

The virtue Levenson sees in this traditional approach is, first and foremost, that it has worked well in the past. It also has other advantages: it is relatively speedy, so it can be applied to new chemicals, and it is relatively straightforward. This traditional Òsafety factorsÓ approach would also end fascination with and stultification by vanishing small levels of chemicals.

Arguments Against the Use of Safety Factors for Potential Carcinogens[0]

The 1985 federal interagency committee that published principles for conducting quantitative risk assessment gave four reasons for continuing to treat animal carcinogens differently from other chemicals and rejecting a Òsafety factorsÓ approach:

1. The Baseline for Applying Safety Factors (NOAEL or LOAEL) is too Sensitive to the Particular Experimental Design.
2. Safety Factors Fail to Use All the Information from Dose-Response Curves.
3. Safety Factors Imply Absolute Safety.
4. Safety Factors Imply Thresholds.

The funny thing about these arguments is that they all apply equally well to the quantitative cancer risk extrapolation methods that the interagency committee endorsed. LetÕs look at how the interagency report presented each of the arguments in turn.

1. Sensitivity to Experimental Design: The interagency report argued that Ò[i]n spite of its common use, there are a number of potential problems associated with the safety factor approach. The observation of no treatment-related effects at a given dose may depend, at least in part, on the number of animals exposed at the particular levelÓ (OSTP 1985, 10439).

[0]. This section was written by Leo Levenson.

Thus the fewer animals you use, the harder it will be to see a statistically significant increase in tumors or other effects attributable to the chemical exposures. However, experimental design makes a great difference with the linear extrapolations as well. Regulatory agencies must establish protocols for the minimum number of animals and experimental conditions that can constitute an acceptable study no matter what technique is finally used to characterize health risks posed by the chemicals.

2. Safety Factors Do Not Make Use of Information from Dose-Response Curves: According to the interagency report, Ò[t]he determination of a NOAEL ignores the shape of the dose-response curve, even though it would seem that a curve that has a shallow slope in the experimental NOAEL region potentially represents a greater toxicological hazard than one that rises steeply in this regionÓ (OSTP 1985, 10439). In other words, when you apply a safety factor to a dose that appears harmless, you are failing to make use of information from the relationship between higher doses and increased cancer rates. True. There is also no way to know whether a linear risk extrapolation technique can make good use of the high-dose relationships either. Since there is no Òadequate biological rationaleÓ for any extrapolation technique, it is hard to see how quantitative acrobatics that incorporate extra-high-dose information are likely to improve low-dose risk estimates.

3. Safety Factors Are Arbitrary: The interagency report complains that Òthere is no biological justification for the general use of any specific safety factorÓ (OSTP 1985, 10439). Safety factors are always arbitrary—but at least they are transparently arbitrary, and there can be an informed debate about what safety factor to choose without anyone maintaining the illusion that there is one ÒcorrectÓ margin of safety. By contrast, the linear extrapolation methods sanctioned by the committee contain a large number of equally arbitrary assumptions about how to use tests involving high chemical doses in animals to predict risks to people at much lower chemical doses, and the assumptions are effectively concealed from lay people. When a model result is announced stating, ÒSuch and such a dose allows a maximum risk of 1–1 million,Ó most people have no idea what they are hearing.

4. Safety Factors Imply Thresholds: ÒAnother important consideration that would argue against the use of a safety factor approach in cancer risk assessment is the fact that this approach assumes the existence of a true population threshold below which no adverse effects can occur. Even if the concept of the individual thresholds could be supported, the well-recognized genetic variability in the human population would effectively prevent the estimation of a general population threshold valueÓ (OSTP 1985, 10439). In other words, the use of a safety factor would give people a false sense of security about potential residual cancer risks for sensitive individuals. By implication, quantitative risk estimation procedures are more

honest in admitting that there could always be a potential risk, no matter how low the exposure. Is this so?

We agree that the government should never promise, and should never accept the responsibility, to eliminate all risk. The EPA and FDA can respond to the unanswerable questions ÒIs this standard absolutely safe?Ó or ÒIs there still a risk of cancer from drinking this water?Ó with the honest reply, ÒWe think that to the best existing knowledge the chemicals in the water protect human health more than would their diminution or removal.Ó But there is no reason the regulators need to promise that the use of a safety factor implies absolute safety. In contrast, the use of seemingly precise quantitative risk estimates gives the illusion that regulators know more than they really do about cancer risks, if any, from low-dose exposures.

Conclusions

Here are LevensonÕs and my conclusions regarding human safety standards for chemicals:[0]

- The less of a potentially toxic chemical people are exposed to, the less likely they will get sick from the chemical. This includes chemicals that may cause cancer. It would be better if we talked about carcinogenic or toxic ÒdosesÓ of chemicals, rather than calling the chemicals themselves ÒcarcinogenicÓ or Òtoxic.Ó

- If the only evidence about toxic effects in a chemical is from high doses, there is no good reason to apply the effects at lower doses to people or other animals. Numerical extrapolations will all be statistical games; they cannot provide insight into real risks. These ideas hold true for all types of chemicals—both carcinogenic and noncarcinogenic. There is no guarantee that any chemical dose will be Òabsolutely safe.Ó But we can make good guesses that a particular dose will be small compared to other potential disease factors.

- CongressÕs attempts to ban chemical carcinogens in the nationÕs food supply appear to stem from beliefs that chemicals could be easily divided into those that Òcause cancerÓ and those that do not, and that the public health benefits of eliminating Òcancer-causingÓ chemicals entirely had to be greater than the expense. In fact, the categories of ÒcarcinogenÓ and Ònot carcinogenÓ are fundamentally flawed—many chemicals may help

[0]. All of these principles relate to standards designed to protect human health. Additional factors must be considered if concerns have been raised regarding effects of chemicals on animals, plants, or ecosystems.

> cause cancer at very high doses but would not cause cancer at lower doses. The cost of reducing chemical residues all the way down to nondetectable amounts turns out to be increasingly high as scientific technology for detecting tinier amounts of chemicals improves. And the public health benefits of eliminating tiny amounts of synthetic chemical residues have been called into question given the defense systems that people have, which work to prevent damage from low levels of natural chemical carcinogens to which ordinary people are exposed in their diets. Cost, however, is not the strongest argument against the criteria used to regulate chemical exposures in the United States today; the strongest argument is that there are no health benefits.

The result of adopting safety-factor criteria would be a reasonable level of protection in the light of existing knowledge, while greatly facilitating regulation. Our expectation (Levenson is a former EPA project director) would be far fewer Superfund sites, greater willingness to clean up the remaining sites, and greater capacity to clean up at the stipulated levels. LawyersÕ fees and waiting lists would decline precipitously, and accomplishments would rise.

Cut the Gordian Knot: Reject Regulation Based on Weak Causes and Weaker Effects

As theories and evidence against the validity and predictive value of animal tests accumulate, as I am confident they will, a return to the traditional safety-level rule-of-thumb approach will become much more politically feasible. Nevertheless, given a choice, I would recommend a rejection of the existing risk-assessment approach. While acknowledging the importance of political feasibility, I would rather advocate the approach I believe best enhances human (and, for that matter, animal) health and safety.

My objection to a return to the time-tested safety-factor method is that it would be based, as Levenson stipulates, on rodent cancer tests, which I believe to be worthless in predicting cancer in human beings (or, indeed, other species). There is something wrong with recommending an invalid method.

Instead, I propose that the government of the United States use its resources and those in the private sector it regulates to enhance two approaches that show promise of developing a knowledge-based policy of cancer control—epidemiology and discovery of actual cancer-causing mechanisms in humans and other species.

LetÕs take another look at the weaknesses of epidemiology. Studies of human populations do not reveal possible harm from small doses of chemicals even if they exist. This weakness could be diminished by putting resources into doing larger studies and developing better statistics, but as

long as regulatory action is conducted at terribly small risk levels, such as one in a million, the weakness would remain. Over time, mechanistic understanding is the only way of distinguishing those small insults that are harmful from those that the human body successfully defends against. What should be done in the meantime?

I propose that small harms from small causes be ignored until we learn how to identify them more reliably so that the harm done by generating so many false positives is exceeded by the health gains from discovering true positives. We do not give up much by ignoring small harms for two reasons: one is that they are small (or epidemiology would pick them up); the other is that there is no good reason, given the invalidity of animal studies, to believe that the models actually pick up real dangers except by accident. In fact, we are suffering those harms now, if they exist, because we do not know their causes and, hence, are unable to take effective preventive measures. A desire to prevent cancers—even more, a desire to show the public that their government is trying to protect its citizens—is not the same as actually providing protection. The pretense of protection, however, is expensive not only in its loss of money but in its loss of the very health and safety it is supposed to defend. I will not be a party to a method of risk assessment and regulation that makes people sicker in the name of keeping them healthier, and that is exactly what is happening.

The other shortcoming of epidemiology is that it takes a long time, because the latency periods of cancers may be decades long. True, but not, I think, conclusive. A long time is a long time; if a disease takes so long, it cannot be striking people down at early ages, because the evidence would already have shown up. We could argue over whether preventing deaths when people are already quite old should be part of governmental policy. There is no need to do that, however, because short-term harms perpetrated by preventive measures are palpable, while long-term gains, in the absence of knowledge about cancer causation, are dubious.

While the long-latency-period argument is cogent for occupational exposures, it has much less force for the general citizenry exposed only to intermittent small doses. For the population at large, moreover, shorter-term evidence from epidemiological studies has value in that if workers exposed to comparatively large doses, even for only a few years or a decade, show no ill results, the probability that small and sporadic exposures would be harmful is low. In the same way, when we learn that symptoms decline or disappear when doses are reduced, this dose-response relationship gives us a pretty good idea that we can identify the cause.

Perfection is not for this world. People in industrial democracies, I am arguing, should accept the modest imperfections they know of while striving to do better, rather than the imperfections they can hardly imagine from utterly invalid animal tests.

My advice is to cut the Gordian knot of chemical regulation by requiring a standard parallel to that set for medical drugs; where evidence of effi-

cacy in promoting health is properly demanded of medical drugs, so should evidence of harm to health be demanded for regulation of chemicals. Unless and until the existing reversal of causality (negative evidence that a chemical does no harm) is replaced by the straight standard—evidence of harm— regulation will continue to harm health in the name of protecting it.

It is not this or that lack of scientific knowledge that lies at the heart of our difficulties in protecting public health against harmful doses of chemicals. Nor is it the inevitable disagreements that are characteristic of science until the rare occasions when strong knowledge exists and is widely accepted. Rather, it is the demand of negative proof in regard to weak causes of minuscule effects that expands disagreements. If proof positive were required, and de minimus dangers ignored, existing knowledge would be far more adequate to the task.

References

American Council on Science and Health. 1984. Pamphlet, Of Mice and Men: The Benefits and Limitations of Animal Cancer Tests.

Ames, Bruce N. 1990. More on Carcinogen Test Processes. Los Angeles Times (31 August): A38.

Ames, Bruce N., and Lois Swirsky Gold. 1990a. Too Many Rodent Carcinogens: Mitogenesis Increases Mutagenesis. Science 249 (31 August): 970–71.

———. 1990b. Response to Perera. Letter to the Editor. Science 250 (21 December): 1645–46.

———. 1991. Endogenous Mutagens and the Causes of Aging and Cancer. Mutation Research 250 (1/2): 3–16.

Anderson, Elizabeth A., and the Carcinogenic Assessment Group of the United States Environmental Protection Agency. 1983. Quantitative Approaches in Use to Assess Cancer Risk. Risk Analysis 3 (4): 277–94.

Baker. 1969. Quoted in Rowan.

British Toxicological Society. 1984. Human Toxicology 3:85–92. Quoted in Robert Sharpe, The Cruel Deception. London: Thorson's Publishing, 1988.,94–95.

Chu, Kenneth C., Cipriano Cueto, Jr., and Jerrold M. Ward. 1981. Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays. Journal of Toxicology and Environmental Health 8 (1–2): 251–80.

Cogliano, Vincent James, et al. 1991. Carcinogens and Human Health: Part III. Letter to the Editor. Science 251 (8 February): 606–7.

Cohen, Samuel M., and Leon B. Ellwein. 1990. Cell Proliferation in Carcinogenesis. Science 249 (31 August): 1007–11.

Edwards, Gordon J. N.d. Worried about Pesticides in Food and Water? Here Are the Facts. Pamphlet distributed by National Council for Environmental Balance, Inc.

Fears, Thomas R., Robert E. Tarone, and Kenneth C. Chu. 1977. False-Positive and False-Negative Rates for Carcinogenicity Screens. Cancer Research 37 (July): 1941-45.

Freedman, D. A., and William C. Navidi. 1989. Multistage Models for Carcinogenesis. Environmental Health Perspectives 81:169–88.

Freedman, David A., and H. Zeisel. 1988. From Mouse to Man: The Quantitative Assessment of Cancer Risks. Statistical Science 3 (11): 3–56.

Gart, J. J., K. C. Chu, and R. E. Tarone. 1979. Statistical Issues in the Interpretation of Chronic Bioassay Tests for Carcinogenicity. Journal of the National Cancer Institute 62: 957–74.

Gaylor, D. W. 1989. Preliminary Estimates of the Virtually Safe Dose of Tumors Obtained from the Maximum Tolerated Dose. Regulatory Toxicology and Pharmacology 9: 101-8.

Gold, Lois Swirsky, Leslie Berstein, Renae Magaw, and Thomas H. Sloane. 1989. Interspecies Extrapolation in Carcinogenesis: Prediction between Rats and Mice. Environmental Health Perspectives 81: 211–19.

Goldberg, Alan M., and John M. Frazier. 1989. Alternatives to Animals in Toxicity Testing. Scientific American 261 (2): 24–30.

Gough, Michael. 1990. How Much Cancer Can EPA Regulate Away? Risk Analysis 10 (1): 1–6.

———. Forthcoming. Chemical Risk Assessments Is Not Science. To appear in Chemistry, an American Chemical Society publication for undergraduates.

Hanahan, Douglas. 1989. Transgenic Mice as Probes into Complex Systems. Science 246 (December): 1265 ff.

Hattis, Dale, and David Kennedy. 1986. Assessing Risks from Health Hazards: An Imperfect Science. Technology Review 89 (4): 60–71.

Lave, Lester B., Fanny K. Ennever, Herbert S. Rosenkranz, and Gilbert S. Omen. 1988. Information Value of the Rodent Bioassay. Nature 336 (15 December): 631–33.

Marsalis, Jon C., and Karen L. Steinmetz. 1990a. The Role of Hyperplasia in Liver Carcinogenesis. In Mouse Liver Carcinogenesis: Mechanisms and Species Comparisons, edited by D. S. Stevenson, J. A. Popp, J. M. Ward, R. M. McClain, T. J. Slaga, and H. C. Pitot. New York: Wiley-Liss.

———. 1990b. Scientists Question Use of Rodent Tests in Risk Assessment. Chemecology (October): 10.

Marx, Jean. 1990. Animal Carcinogen Testing Challenged. Science 250 (9 November): 743–45.

Moolgavkar, Suresh H. 1986. Carcinogenesis Modeling: From Molecular Biology to Epidemiology. Annual Review of Public Health 7:151–69.

National Academy of Sciences. 1978. Saccharin: Technical Assessment of Risks and Benefits. National Research Council, Committee for Study on Saccharin and Food Safety Policy, Washington, D.C. Chap. 3., pp. 72 and 61 ff.

National Medical Advisory Service (NMAS). 1992. NMAS Advisor IV (4).

National Public Radio. Morning Edition. Transcript. Washington, D.C. Thursday, 12 March 1992, 10:00 A.M. EDT.

Paget, G. E. ed. 1970. Methods in Toxicology. Boston: Blackwell Scientific Publications.

Perera, Frederica P. 1990. Carcinogens and Human Health: Part 1. Letter to the Editor. Science 250 (21 December): 1644–45.

Peto, Richard. 1977. Epidemiology, Multistage Models, and Short-Term Mutagenicity Tests. In Human Risk Assessment. Book C of Origins of Human Cancer, edited by H. H. Hiatt, J. D. Watson, J. A. Winsten. Cold Spring Harbor (N.Y.) Laboratory.

Revision without Revolution in Carcinogen Policy. 1984. Regulation 8 (4): 5–7.

Rowan, Andrew H. 1984. Of Mice, Models, and Men: A Critical Evaluation of Animal Research. Albany: State University of New York Press.

Salzburg, David. 1980. The Effects of Lifetime Feeding Studies on Patterns of Senile Lesions in Mice And Rats. Drug & Chemical Toxicology 3:1–33.

Salzburg, David. 1983. The Lifetime Feeding Study in Mice and Rats—An Examination of Its Validity as a Bioassay for Human Carcinogens. Fundamental and Applied Toxicology 3:63–67.

Schirrmacher, Volker. 1989. Immunobiology and Immunotherapy of Cancer Metastasis: Ten-Year Studies in an Animal Model Resulting in the Design of an Immunotherapy Procedure Now under Clinical Testing. Interdisciplinary Science Reviews 14 (3): 291–303.

Sharpe, Robert. 1988. The Cruel Deception. London: ThorsonÕs Publishing.

Smith, Allen H., and Dan S. Sharp. 1984. A Standardized Benchmark Approach to the Use of Cancer Epidemiology Data for Risk Assessment. Typescript Presented at the Environmental Protection Agency Symposium on Advances in Health Risk Assessment for Systemic Toxicants and Chemical Mixtures, Cincinnati, Ohio, 23–25 October.

Tomatis, L., C. Agthe, and H. Bartsch et al. 1978. Evaluation of the Carcinogenicity of Chemicals: A Review of the Monograph Program of the International Agency for Research on Cancer (1971 to 1977). Cancer Research 38:877–85.

U.S. Environmental Protection Agency (EPA). 1976. Interim Procedures and Guidelines for Health Risk Assessments of Suspected Carcinogens. Federal Register 41:21402–21405.

U.S. Office of Science and Technology Policy (OSTP). 1985. Chemical Carcinogens: A Review of the Science and its Associated Principles, February 1985. Federal Register 50:10372–10442.

Wilson, Alan E. G. 1991. Encouraging the Development and Application of Pharmacokinetic Modeling in Risk Assessment. American Industrial Health Council Quarterly (Summer): 6–9.

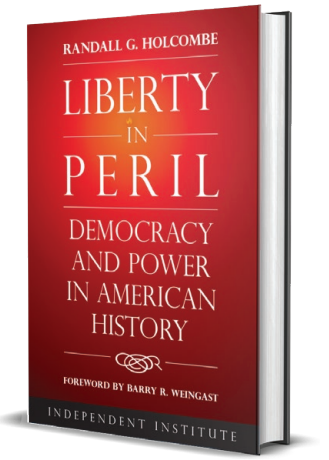
Zbinden. 1976. Quoted in Rowan.

SUBSCRIBE NOW AND RECEIVE A FREE BOOK!

"***The Independent Review*** does not accept pronouncements of government officials nor the conventional wisdom at face value."
—**JOHN R. MACARTHUR**, Publisher, *Harper's*

"***The Independent Review*** is excellent."
—**GARY BECKER**, Nobel Laureate in Economic Sciences

Subscribe to ***The Independent Review*** and receive a free book of your choice such as ***Liberty in Peril: Democracy and Power in American History***, by Randall G. Holcombe.



Thought-provoking and educational, ***The Independent Review*** is blazing the way toward informed debate. This quarterly journal offers leading-edge insights on today's most critical issues in economics, healthcare, education, the environment, energy, defense, law, history, political science, philosophy, and sociology.

Student? Educator? Journalist? Business or civic leader? Engaged citizen? This journal is for YOU!

Order today for more *FREE* book options

SUBSCRIBE

The Independent Review is now available digitally on mobile devices and tablets via the Apple/Android App Stores and Magzter. Subscriptions and single issues start at $2.99. **Learn More.**

Download on the App Store | GET IT ON Google Play | Available on MAGZTER

INDEPENDENT INSTITUTE, 100 SWAN WAY, OAKLAND, CA 94621 • 1 (800) 927-8733 • ORDERS@INDEPENDENT.ORG