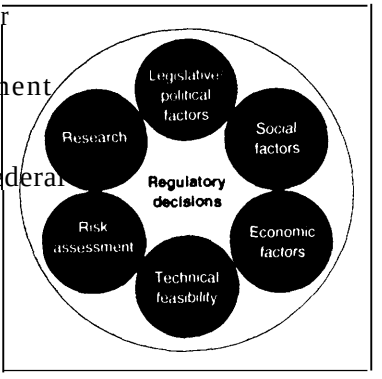
Research and Decisionmaking 5

Science and policymaking are uneasy partners. In an address at the press conference to release the 1983 National Research Council report on risk assessment in the Federal Government, the former administrator of the U.S. Environmental Protection Agency (EPA), William D. Ruckelshaus, said "The main reason for the uneasiness lies, I think, in the conflict between the way science really works and the public's thirst for certitude that is written into EPA's laws. Science, as you all know, thrives on uncertainty" (Ruckelshaus, 1983).

Yet despite that uneasy relationship, the primary criterion for health risk assessment research is that it be useful for decisionmaking. With that observation in mind, the Office of Technology Assessment (OTA) examines three interrelated questions in this chapter:

1. How has research influenced Federal risk assessment guidelines and risk assessment practices?
2. What impact has research had on decisionmaking?
3. How can research be designed to make risk assessment more useful in decisionmaking?



To answer those questions, we review the evolution of Federal risk assessment guidelines and risk assessment practices and the comments and criticisms made regarding them. The analysis focuses on Federal activities in this area in part because the record of Federal regulatory decisionmaking is more accessible than the record of decisionmaking in the private sector.

Research findings from epidemiology and toxicology provide the primary database for health risk assessment. But those data

are seldom extensive enough for answering many of the questions that arise in regulatory decision-making. Weinberg (1972) characterized such issues as "transience questions"--questions that "can be asked of science and yet . . . cannot be answered by science. " Agencies frequently confront them, and because science cannot answer them, agencies adopt so-called science policy assumptions in order to make decisions. The assumptions can be divided into two general types: those that are used to bridge gaps in scientific knowledge and those that compensate for a lack of agent-specific data (NRC, 1983).

IMPACT OF RESEARCH ON RISK ASSESSMENT GUIDELINES AND DECISIONMAKING

EPA is the main player in developing and revising risk assessment guidelines. Although the Consumer Product Safety Commission (CPSC), the Food and Drug Administration (FDA), and the Occupational Safety and Health Administration have also published health risk assessment guidelines in the ***Federal Register, only*** EPA has completed scientific reviews of some of its guidelines and formally modified them in response to new scientific information. This chapter considers three of EPA's guidelines. The agency first adopted guidelines for assessing the risks of ***carcinogens*** in **1976;** it formally modified those guidelines in 1986 and is now revising them further. It adopted its first guidelines for ***developmental toxicants*** and for ***estimating exposures*** in **1986,** modifying the developmental toxicants guidelines in 1991 and the exposure guidelines a year later. Also discussed in this chapter is the International Agency for Research on Cancer's most recent revisions of its procedures for evaluating the risks to humans posed by carcinogens.

Reviewing the EPA risk assessment guidelines and their revisions makes it clear that the agency has changed relatively few of its science policy assumptions in response to new scientific information. The impact of research on the guidelines is more evident in EPA's increased attention to identifying all of the relevant scientific questions that the guidelines should address. Often, new questions reveal new uncertainties that have to be bridged with assumptions.

Guidelines for Risk Assessments of Carcinogens

Reflecting society's concern about cancers and the Federal Government's regulatory focus on them, extensive scientific research has been and is being conducted to identify the causes of cancers and the mechanisms of carcinogenesis. To date, that research has had only a modest effect on efforts to revise the EPA's carcinogen risk assessment guidelines. It has, however, had a substantial impact on chemical-specific risk assessments and consequently on regulatory actions. In addition, it is currently generating considerable debate as EPA considers new revisions to its 1986 cancer policy (U.S. EPA, 1988b, 1992a). In general, research has had greater impact in displacing assumptions that EPA adopted to bridge inadequacies in the data than in changing assumptions to compensate for theoretical uncertainties.

EPA's 1976 interim guidelines for carcinogen risk assessment (U.S. EPA, 1976) discussed the assumptions underlying the agency's regulatory approach, but they provided no explicit list of agency science policy positions. In contrast, EPA's 1986 guidelines on carcinogen assessment (U.S. EPA, 1986a) detailed several major science policy positions that guide the agency's interpretation of incomplete or uncertain data (see box 5-A).

Among the most controversial agency policy positions is the use of the results of animal tests to predict human effects. The 1981 OTA report ***Assessment of Technologies for Determining Cancer Risks from the Environment*** (pp. 124-

127) discussed a number of objections to the use of animal bioassay data:[1]

1. The doses given to test animals are too high, and the results do not predict carcinogenic effects in humans.
2. Routes of exposure in test animals are not the same as routes of exposure in humans.
3. Some life processes of test animals (e.g., physiology and metabolism) are so different from those of humans that the test results may not be relevant for predicting human cancer risks.
4. Some test animals and animal organs are so sensitive to certain chemicals which induce tumors that the test results do not predict cancer risk in humans.
5. Benign tumors in test animals do not predict human cancer risk.

Twelve years later, such objections are still being raised—most often, of course, when tests indicate that a commercially important chemical causes cancer in animals and the chemical's manufacturer, distributors, and users face regulation. As long as the bioassay remains the basic source of information for evaluating carcinogenic risks to humans, those objections will be raised.

The sensitivity of test animals and test organs is an issue because certain chemicals cause tumors only in the liver of male B6C3F1 mice (and not in other mouse organs), female mice, or in rats. EPA's 1986 guidelines describe the problems posed by agents that cause cancer only in certain organs in a single species. In discussing the evaluation of animal test results for assessing human risk, EPA has stated that it accords more weight to conclusions based on results showing that a chemical causes cancer in more than one species. The agency classifies the evidence for carcinogenicity in a descending scale that runs from "sufficient," through "limited," "inadequate," and ''no data, ' to ''no evidence. ' EPA defines "sufficient" as follows:

> Sufficient evidence of carcinogenicity, which indicates that there is an increased incidence of malignant tumors or combined malignant and benign tumors [footnote]: (a) in multiple species or strains. . . (U.S. EPA, 1987b, pp. 1-11).

But EPA's statement does not mean that it always accords less weight to results that are obtained in only one species. In fact, the two footnotes in the definition of sufficient' dismiss two arguments that scientists have not been able to resolve. The first footnote states:

> An increased incidence of neoplasms that occur with high spontaneous background incidence (e.g., mouse liver tumors and rat pituitary tumors in certain strains) generally constitutes ''sufficient' evidence of carcinogenicity, but may be changed to "limited" when warranted by the specific information available on the agent (U.S. EPA, 1987b, pp. 1-11).

The footnote not only gives the agency the flexibility to judge sufficiency of evidence in the absence of positive results from two species, but it also allows agency staff to ignore the great uncertainty that many scientists attach to any conclusion based on a chemical's causing only mouse liver tumors. As EPA's guidelines note:

> For a number of reasons, there are widely diverging scientific views about the validity of mouse liver tumors as an indication of potential carcinogenicity in humans when such tumors occur in strains with high spontaneous background incidence and when they constitute the only tumor response to that agent. These Guidelines take the position that when the other conditions for a classification of "sufficient' evidence in animal studies are met . . . the data should be considered as 'sufficient' evidence of

[1] Bioassay is a term used for long-term (e.g., 2 years for rodents) experimental studies for cancer induction. Rodent bioassays generally employ both sexes of rats (Fischer 344/N) and mice (B6C3F1 hybrid), using two or three exposure levels plus untreated controls in groups of 50 animals for 2 years.

Box 5-A–Major Science Policy Positions for EPA Cancer Guidelines

- Benign tumors should generally be combined vvith malignant tumorsfor risk estimates unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin.
- Agents that are positive in long-term animal experiments and also show evidence of promoting cocarcinogenic activity in specialized tests should be considered as complete carcinogens unless there is evidence to the contrary because it is, at present, difficult to determine whether an agent is only a promoting or cocarcinogenic agent.
- These guidelines take the position that when the only tumor response is in the mouse liver and when other conditions for a classification of "sufficient" evidence in animal studies are met (e.g., replicate studies, malignancy; see section IV), the data should be considered as "sufficient" evidence of carcinogenicity.
- Because it is possible that human sensitivity is as high as the most sensitive responding animal species, **in the** absence of evidence to the contrary, the biologically acceptable data set from long-term animal studies showing the greatest sensitivity should generally be given the greatest emphasis in estimating human carcinogenic risk
- Where two or more significantly elevated tumor sites or types are observed in the same study, extrapolation may be conducted on selected sites or types. To obtain a total estimate of carcinogenic risk animals with one or more tumor sites or types showing significantly elevated tumor incidence should be pooled and used for extrapolation. The pooled estimates will generally be usedin preference to risk estimates based on single sites or types.
- In the absence of adequate information to the contrary, the linearized multistage procedure will be employed for estimating human carcinogenic risks.
- In the absence of comparative toxicological, physiological, metabolic, and pharmacokinetic data for a given suspect carcinogen, the Agency takes the position that the extrapolation on the basis of surface area is considered to be appropriate because certain pharmacological effects commonly scale according to surface area.
- Unless there is evidence to the contrary in a particular case, the cumulative dose received over a lifetime, expressed as average daily exposure prorated over a lifetime, is commended as an appropriate measure of exposure to a carcinogen. That is, the assumption is made that a high dose of a carcinogen received over a short period of time is equivalent to a corresponding low dose spread over a lifetime.
- In characterizing the risk due to concurrent exposure to several carcinogens, the risks are combined on the basis of additivity unless there is specific information to the contrary.

The basic source of information for most risk assessments is the carcinogenesis bioassay, which is **a test** of the suspect carcinogen's capacity to cause tumors in laboratory animals, generally rats and mice. Because of reliance on the bioassay, much of the cancer risk assessment guidelines are devoted to the execution **and** interpretation of the bioassay.

SOURCE: US. Environmental Protection Agency, "Guidelines for Carcinogen Risk Assessment," *Federal Register* 51(1986):33992-34003.

carcinogenicity. It is understood that this classification could be changed on a case-by-case basis to ''limited," if warranted, when factors such as the following are observed: an increased incidence of tumors only in the highest dose group and/or only at the end of the study (U.S. EPA, 1987b, pp. 1-5).

EPA's science policy thus considers results that show that a chemical causes tumors only in mouse livers as sufficient evidence of carcinogenicity. The agency's discussion of this policy admits that not all scientists agree with that decision, but it does not address the arguments of those who disagree. Instead, the guidelines discuss the replication of test results and consistent

test design. But no matter how many times tests are replicated or how well they are done, those that generate false signals can do no more than that.

Many observers believe that EPA persists in treating mouse liver tumors as sufficient evidence because the agency has regulated several chemicals on the basis of such findings. For example, the agency banned the organochlorine pesticide DDT ostensibly on the basis of potential liver carcinogenicity (Dunlap, 1988). To back away from the sufficient classification might open up some of EPA's past actions to renewed scrutiny and criticism.

The second footnote in the paragraph about sufficient evidence deals with benign tumors:

> Benign and malignant tumors will be combined (to arrive at a count of total turners) unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin (U.S. EPA, 1987b, pp. 1-11).

The language of the footnote indicates that EPA will consider other evidence in deciding how to count benign tumors. But it stops short of saying what kind of evidence would be considered, much less what kind would be considered convincing. As a practical matter, it is probably impossible to demonstrate that a "potential to progress to the associated malignancies" does not exist.

The guidelines about sensitive organs and sensitive test animals retain the features that caused objections more than a decade ago. Those features, like most parts of the guidelines, are designed to protect health, especially in circumstances of few or no data. That position is not likely to change. Nevertheless, EPA has altered some of its science policy assumptions to accommodate new information.

Extrapolating from the results of tests in animals to predictions of risk to humans is at the heart of most risk assessments, and recent years have seen three examples of altered approaches to it. The first is a change from a general ''default' extrapolation method to a particular ''default' factor derived from specific data and applicable to all chemicals. The second example describes the process used to develop specific data for a chemical to replace a default approach. The last example notes a few incidence in which specific information about a chemical altered risk assessment decisions,

SCALING FACTOR FOR CROSS-SPECIES EXTRAPOLATIONS

In extrapolating test results, allowances must be made for the differences in size, shape, life-span, physiology, and biochemistry between test animals and humans. When toxicologists and risk assessors lack specific information for making those adjustments, they have traditionally used one or the other of two standardized ''scaling factors. ' One scaling factor is body weight; thus, a dose of 1 milligram per kilogram body weight in a rat is treated as equivalent to 1 milligram per kilogram body weight in a human, The other factor is body surface area. Surface area is difficult to measure, but it can be approximated by raising body weight to the 2/3 power. Although that approximation is used, it has been challenged as likely to be in error (Slone, 1993).

Regardless of the certainty of either factor being appropriate, FDA uses the body weight scale, and EPA uses the surface area scale, with the result that EPA predicts higher risks than FDA. Given the same data about toxicity, EPA would predict a risk 14 times higher than the risk FDA would predict when the tests are done in mice and a risk 6 times higher when the tests are done in rats (U.S. Congress, OTA, 1981).

Recently, EPA proposed, for itself and on behalf of FDA and CPSC, changing its choice of scaling factor. The proposal was made in response to three events: an analysis of all available interspecies scaling data (Travis and White 1988), a reassessment of the rationale

underlying the use of the surface area factor (Travis et al., 1990), and a political effort to harmonize the approaches of different Federal regulatory agencies. The new scaling factor that EPA has proposed lies between the body weight and surface area scaling factors. It will result in risk estimates about midway between the estimates that would be produced by the two older methods.

USING METABOLIC AND PHARMACOKINETIC DATA FOR CROSS-SPECIES EXTRAPOLATION

Researchers have long believed that studies in pharmacokinetics and metabolism provide important information for understanding the mechanisms by which agents evoke toxicity (see Slone, 1993). Pharmacokinetic studies examine the rates of absorption, distribution, metabolism, and excretion of a compound. They also examine the time-dependent features of those processes, as they link to the compound's toxicological effects. Metabolic studies examine the coordinated reactions and pathways that transform the compound into reactive or inactive intermediates that can be toxic entities. Information from pharmacokinetic and metabolic studies provides a basis for determining the internal dose and the validity of various extrapolations from the level of exposure to the expected response.

The case of methylene chloride is an example of the use of pharmacokinetic and metabolic data, by government and industry scientists to improve animal-to-human extrapolation. Three factors seem to have influenced the collaboration. First, influential members of the scientific community were interested in using pharmacokinetic information to obtain better estimates of the doses of methylene chloride necessary to produce turners in test animals and doses that may affect humans. Second, staff from various agencies wanted to work with industry and each other to evaluate metabolic data and determine how such information could be used in risk assessment. Third, the EPA Science Advisory Board encouraged a thorough review of the data obtained using pharmacokinetic methods (Preuss, 1992).

The story begins in the early 1980s, when the U.S. Air Force supported a research project at the Wright-Patterson Air Force Base Inhalation Toxicology Laboratory to develop methods to better assess the risks associated with human exposure to organic solvents. The Air Force-supported scientists first developed a generic, physiologically based pharmacokinetic (PBPK) model to assess exposure of humans to organic solvents through inhalation. Following a 1986 review of EPA's risk assessment of methylene chloride, members of the advisory board concluded that the PBPK model was a valid alternative to other approaches for estimating the dose of methylene chloride that can cause human toxicity based on animal study results. As a result of that encouragement, scientists from EPA, CPSC, FDA, Dow chemicals, ICI (a U.K. fro), and the European Council of chemical Manufacturers Federation began to hold periodic meetings to discuss their research needs and share information. The meetings identified gaps in the database for methylene chloride, produced a commitment to initiate further studies, and provided a forum to share and discuss protocols and experimental results.

The effort to understand the mechanisms by which methylene chloride induces cancer stimulated research in PBPK modeling and enthusiasm for improved interspecies extrapolations. For methylene chloride specifically, both EPA (1987c) and the California Department of Health and Human Services (1988) have concluded that additional data are necessary to clarify metabolic differences between rodents and humans. The agencies expect some of these data to come from ongoing Navy-supported research at Wright-Patterson (Gearhart, 1992). Additional information will be generated by a research program at the National Institute of Environmental Health Sciences to better characterize the tumorigenic response in mice exposed to methylene chloride.

The ongoing research on methylene chloride is an example of collaborative research that has led

to additional studies. To date, these investigations have not resolved to EPA's satisfaction the question of how potent a carcinogen methylene chloride is in humans. They have, however, pointed to additional projects that may clarify questions of human risk. More generally, the methylene chloride experience underlines the difficulties of animal-to-human extrapolation and should caution risk assessors against too-ready an acceptance of generalized approaches (e.g., scaling factors) in those extrapolations.

RESEARCH AND CHANGES OF PRESUMPTIONS

The fundamental premise of toxicology is that animals and humans respond similarly to chemicals. Some see accruing evidence for that premise (Huff, 1993), For others, substantial problems are surfacing in extrapolating results from animal models to human populations (Ames and Gold, 1990a,b; Cohen and Ellwein, 1991, 1992). In carcinogenic risk assessments, regulators and scientists have established criteria for evaluating the effects of chemicals in animals to help in deciding whether the chemicals present carcinogenic hazards to humans (IARC, 1992a; U.S. OSTP, 1985). Once a substance has been classified as a potential human carcinogen, U.S. regulatory agencies consider it appropriate to use linear, no-threshold extrapolation models to estimate the magnitude of human risk (see ch. 2). Arguing against these general procedures is the increasing recognition that biological mechanisms, physiology, and biochemistry or differences in routes of exposure may affect the agent's interactions with the target tissue. Those effects could alter the toxicological consequences in humans compared with those observed in animals.

Demonstrating the validity of generalized extrapolation procedures requires information about mechanisms of toxicity in sufficient detail such that no important gaps remain. That's seldom the case. Instead, limited mechanistic data become the focus of controversies about interpretation. One such debate arises in assessing the risk of cancer in humans from exposure to agents that are not DNA-reactive but that test positive for carcinogenicity in animals. The mechanisms by which the diverse array of nonmutagenic, carcinogenic chemicals induce cancer are little understood. As a result, new generic approaches to interpret those tests and project their results to human risk estimates are not likely to be available soon.[2] Nevertheless, specific studies of some chemicals have led to deviations from the general presumption of risks to humans and the application of the linear, no-threshold model.

Some substances induce cancer through indirect mechanisms that operate at high doses in test animals but are unlikely to operate at lower doses. Or they may operate through mechanisms that do not exist in humans. An example of the former mechanism, which renders linear, no-threshold extrapolation inappropriate, is ethylene thiourea. That chemical disrupts hormone levels in the rodent thyroid gland only at high doses (U.S. EPA, 1988c). An example of a situation in which positive findings in animals are not applicable to humans is the induction of kidney tumors in male rats through the interaction of d-limonene and a protein present in male rats but not in humans (U.S. EPA, 1991a). Some chemicals that cause bladder cancers in test animals (e.g., melamine, aliette, and saccharin) are considered unlikely or impossible human carcinogens because exposures in humans are unlikely to reach the extremely high levels required to cause urinary calculi, which are the proximal cause of those cancers induced at very high doses (see Huff, 1992, for a review). These examples demonstrate the importance of chemical-specific mechanistic data and the growing importance of understanding mechanisms of toxicity for more realistic extrapolations from animals to humans.

2 In contrast t. U.S. regulatory agencies, which use no-threshold models for all carcinogens, regulatory agencies in some other countries assume a safe level of exposure exists in estimating risks from nongenotoxic carcinogens (see appendix A).

Guidelines for Developmental Toxicity

EPA first formulated science policy assumptions for developmental toxicity in 1986 (U.S. EPA, 1986c) and made few changes when the guidelines were revised in 1991. Yet in response to criticism from its Science Advisory Board, EPA did modify one assumption. The agency agreed to use results from a species most relevant to humans for estimating potentially toxic effects on human development whenever data were available, rather than automatically selecting results from the most sensitive responding species.

The most significant change, however, was the agency's decision to merge the hazard identification and dose-response phases of the risk assessment process to "reflect hazard within the context of dose, route, duration and timing of exposure' (U.S. EPA, 1991 b). The 1991 guidelines also abandoned a proposed weight-of-the-evidence scheme that would have classified substances as having "definitive," "adequate," or "inadequate' evidence of toxicity during human development (U.S. EPA, 1989b). The approach that EPA adopted takes into account extensive scientific research that suggests that manifestations of developmental toxicity depend strongly on species and exposure. EPA's new approach thus addresses scientific concerns about dependence of developmental toxicity on the contexts of test animal and duration and timing of exposures. But it also clearly expands the extent of scientific analysis required by the agency even to identify a substance as a potential developmental toxicant.

There are both scientific and policy arguments for replacing EPA's current approach to noncancer risk assessment (the no-observed-adverse-effects level divided by uncertainty factors; see ch. 2) (Pease et al., 1991). Most of the alternatives, however, demand significantly more data and analysis. Looking beyond the relatively simple modeling of the benchmark-dose approach, EPA is supporting research to develop ''models that are more biologically based to provide more accurate estimates of low-dose risk to humans" (U.S. EPA, 1991 b).

Guidelines for Exposure Assessment

The changes in EPA's guidelines for assessing exposures to toxic chemicals exemplify expanding scientific examination in the risk assessment process. EPA's initial 1986 guidelines (U.S. EPA, 1986b) were relatively brief (8 pages), "laying out a set of questions to be considered in carrying out an exposure assessment in order to help avoid inadvertent mistakes of omission. ' In contrast, the 1992 revisions present a detailed discussion (45 pages) of the scientific foundation of exposure assessment and state that extensive data acquisition and analysis are necessary to produce good assessments (U.S. EPA, 1992b).

Probably the most important change in agency guidelines has been the development of methods for displacing ''worst-case' assumptions about exposure with more reasonable estimates of "high-end" exposures. "The concept of high-end exposure is fundamentally different from terms such as worst case, in that the estimate is by definition intended to fall on the actual exposure distribution" (U.S. EPA 1992b). This change has the advantage of basing risk estimates and potential regulation not on the maximum possible exposure but on the exposures that are likely to be occurring to some members of the actual population. However, this approach requires considerably more data and analysis to characterize those exposures.

In contrast to other guidelines, the exposure assessment guidelines require relatively few assumptions. Both EPA's 1986 and 1992 exposure assessment guidelines contain the same fundamental science policy assumption: in the absence of measurement data, exposure assessment may be based on mathematical models.

The Limited Impact of Research

Given the length and breadth of EPA's guidelines, the agency has made few changes in the

assumptions and procedures included in those directives. Three interacting factors account for the limited impact of new scientific research on EPA's science policy assumptions.

First, the nature of the assumption is a factor. An assumption that bridges a specific, well-defined information gap that can be resolved experimentally is more likely to be displaced when the needed information is generated. In contrast, an assumption bridging broad areas of scientific uncertainty, especially gaps in scientific knowledge, and understanding is less likely to be replaced.

Second, the relative importance of the assumption to the paradigm underlying the predominant regulatory approach is a major factor in whether new research changes science policy. The no-threshold assumption for carcinogens (no amount of exposure, however small, is not without an effect) is central to EPA's cancer risk assessment guidelines (U.S. EPA, 1986a). In comparison, the assumption of surface area as the appropriate default scaling factor among species is more peripheral. Displacing the no-threshold assumption for carcinogens has proved difficult, not only because important aspects of the question cannot be resolved experimentally but also because displacement requires developing an alternative model of carcinogenesis that can command a scientific consensus. Because it lacked such a consensus, EPA failed in the early 1980s to modify the agency's cancer policy by separating carcinogens into two classes based on their mechanism of action, which would have allowed threshold-based risk assessment for nonmutagenic carcinogens (Rushefsky, 1986).

The third and final factor influencing the agency's response to new scientific research involves the policy reverberations associated with changing specific default assumptions. The policy consequences of changing some agency default positions are slight: shifting interspecies scaling factors reduces EPA's risk estimates slightly and raises FDA's slightly, but it also encourages interagency consistency and does not conflict with public expectations. In the area of extrapolating from high to low doses, however, existing agency guidelines have been significantly influenced by a policy commitment to err on the side of public safety in order to fulfill statutory mandates. Changing the default assumption of low-dose linearity can dramatically reduce estimates of risk, as well as conflict with public beliefs that no carcinogenic exposure is safe. To a large extent, the predominant role policy considerations played in their adoption explains the resistance to changing EPA's default assumptions for establishing the relevance of animal cancer to humans and estimating low-dose risks.

New scientific findings that are promoted because they result in less "conservative" estimates of risk are likely to be strongly contested because they may be perceived as undermining the government's or a particular agency's commitment to protect public health in the face of uncertainty. Rushefsky (1986) referred to these inferential choices as being either ''risk averse' or "risk tolerant, ' depending on the choice of assumption made (see table 5-l). In general, public health agencies are risk averse and more protective of public health to compensate for the inherent uncertainties in the process. Whenever changing a traditional assumption causes substantial policy reverberations, the change confronts very high hurdles indeed.

Which interest group introduces new scientific information into the regulatory arena and how that information is used in the political process may affect the fate of those findings in the risk assessment process. Often, research sponsored by industry is perceived as less-than-objective science and as a self-interested effort to undercut regulation. Federal agencies are not immune to such suspicions. When the Office of Management and Budget selectively used scientific information to attack regulatory risk assessment during the Reagan and Bush Administrations, it may have stigmatized pharmacokinetic and mechanism-based modeling as procedures that weaken regu-

Table 5-l—Patterns of Inferential Choices in Developing Cancer Policy

	Position	
Controversy	Risk-averse	Risk-tolerant
Evidence	Bioassays sufficient	Epidemiology only
Bioassays	Indicate human carcinogenicity	May not be accurate
Positive/negative	Positive more important	Negative may indicate species sensitivity y
Conflicting studies	Positive more important	All evidence should be weighed
Benign/malignant	Benign sufficient	Only malignant significant
Dose levels	High dose provides qualitative evidence	At least three levels should be used; iow doses show reversibility; high doses may overwhelm defense mechanisms
Mathematical models	Linear	No-observed-effects level for epigenetic carcinogens
Thresholds	Insuffident evidence	May exist for promoters and epigenetic carcinogens
initiators/promoters	Cannot demonstrate distinction	Distinction important
Genotoxic/epigenetic	Cannot demonstrate distinction	Distinction important

SOURCE: M.E.Rushefsky, *Making Cancer Policy* (Albany, NY: State University of New York Press 1986), p. 41.

lation rather than improve risk assessment (Pease, 1992).

IARC Procedures for Classifying Carcinogens

The International Agency for Research on Cancer (IARC), a World Health Organization, publishes monographs on the evaluation of carcinogenic risks to humans. The monographs are used widely as source material for assessing risks from exposure to carcinogens. Currently, IARC classifies chemical agents into one of four groups (see box 5-B).

IARC convenes working groups from time to time to evaluate the need for procedural changes in its methods for evaluating carcinogenic risks. (See appendix A for a description of IARC and its policies.) In 1983, a working group concluded that classifying carcinogens according to their mechanism of action could be neither exhaustive nor definitive. Nine years later, in 1991, another working group noted that mechanistic data have always been used in determiningg human carcinogenic risks. It then considered whether the procedures that were currently in use might be revised (IARC, 1992b).

The group concluded that it was impossible to formulate definitive guidelines for all of the possible situations in which mechanistic data may influence the evaluation of carcinogens. Nevertheless, it identified two circumstances in which alternative criteria could be considered for deciding to which category a chemical belongs.

Mechanistic data, according to the working group, can be considered in deciding whether an agent belongs in group 1 or group 3. For group 1, the category ‘‘may be extended to include agents, mixtures, or exposure circumstances for which evidence of carcinogenicity in humans is less than sufficient but there is sufficient evidence of carcinogenicity in experimental animals and strong evidence in exposed humans that the agent acts on a relevant mechanism of carcinogenesis." The working group agreed that group 3 “maybe extended to include agents for which there is sufficient evidence of carcinogenicity in animals and strong evidence that the mechanism of carcinogenicity in animals does not operate in humans.

Box 5-B—IARC's System of Classifying Carcinogenic Risks to Humans

The International Agency for Research on Cancer (IARC) classifies carcinogenic risks into four groups:

Group l—The agent (mixture) is carcinogenic to humans. The exposure circumstance entails exposures that are carcinogenic to humans.

This category is used when there is sufficient evidence of carcinogenicity in humans. Exceptionally, an agent (mixture) may be placed in this category when evidence in humans is less than sufficient but there is sufficient evidence of carcinogenicity in experimental animals and strong evidence in exposed humans that the agent (mixture) acts through a relevant mechanism of carcinogenicity.

Group 2—This category includes agents, mixtures, and exposure circumstances for which, at one extreme, the degree of evidence of carcinogenicity in humans is almost sufficient, as wel as those for which, at the other extreme, there are no human data but there is evidence of carcinogenicity in experimental animals. Agents, mixtures, and exposure circumstances are assigned to either group 2A (probably caranogenic to humans) or group 2B (possibly carcinogenic to humans) on the basis of epidemiologic and experimental evidence of carcinogenicity and other relevant data.

Group 2A-The agent (mixture) is probably carcinogenic to humans. The exposure circumstance entails exposures that are probably carcinogenic to humans.

This category is used when there is limited evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in experimental animals. in some cases, an agent (mixture) may be classified in this category when there is inadequate evidence of carcinogenicity in humans and sufficient evidence of carcinogenicity in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent, mixture, or exposure circumstance maybe classified in this category solely on the basis of limited evidence of carcinogenicity in humans.

Group 2B-The agent (mixture) is possibly carcinogenic to humans. The exposure circumstance entails exposures that are possibly carcinogenic to humans.

This category is used for agents, mixtures, and exposure circumstances for which there is limited evidence of carcinogenicity in humans and less than sufficient evidence of carcinogenicity in experimental animals. it may also be used when there is inadequate evidence of carcinogenicity in humans but there is sufficient evidence of carcinogeniaty in experimental animals. in some instances, an agent, mixture, or exposure circumstance for which there is inadequate evidence of carcinogenicity in humans but limited evidence in experimental animals together with supporting evidence from other relevant data may be placed in this group.

Group 3--The agent (mixture or exposure circumstance) is not classifiable as to its carcinogenicity to humans.

This category is used most commonly for agents, mixtures, and exposure **circumstances** for which the evidence for carcinogenicity is inadequate in humans and inadequate or limited in experimental animals.

Exceptionally, agents (mixtures) for which the evidence of carcinogenicity is inadequate in humans but sufficient in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans.

Agents, mixtures, and exposure circumstances that do not fall into any other group are also placed in this category.

Group 4-The agent (mixture) is probably not carcinogenic to humans.

This category is used for agents or mixtures for which there is evidence suggesting lack of carcinogenicity in humans and in experimental animals. in some instances, agents or mixtures for which there is inadequate evidence of carcinogenicity in humans but evidence suggesting lack of carcinogenicity in experimental animals, consistently and strongly supported by a broad range of other relevant data may be classified in this group.

SOURCE: International Agency for Research on Gamer, *Monograph*, vol. 55 (Lyon, France, 1992), pp. 34-35.

THE INTERDEPENDENCY OF RESEARCH AND DECISIONMAKING

As is evident from the above discussion of the effects of research on EPA's guidelines for risk assessments and assessment practices and the procedures used by IARC, the speed with which science influences assessment procedures and science policy assumptions is very slow indeed. Besides the factors noted above, other barriers exist to incorporating the results of research into agency actions.

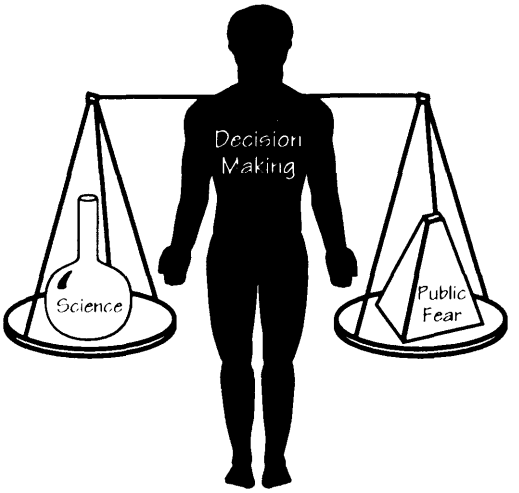
In its 1989 report on improving risk communication, the National Research Council included a guidance paper by B. Fischoff in which he described the interrelationship of science and policy:

> Science is a product of society; as such, it reflects the values of its creators. Conversely, society is partly a product of science. And understanding these inter-dependencies is essential to, on the one hand, discerning the objective content versus inherently subjective science and, on the other hand, directing science to serve socially desired ends. An understanding of these relationships is also necessary to appropriately interpret the conflicts between lay and expert opinions that constitute the visible core of many risk controversies (NRC, 1989),

OTA has chosen two examples of environmental decisionmaking that illuminate this interdependency. Each shows that research is driven by public concern about risk and incomplete risk data for decisionmaking. The dioxin case also demonstrates that researchers and analysts can produce decisionmaking tools in the absence of desired information.

Power-line Electromagnetic Fields and Cancer

Recent public apprehension about the health risks of exposure to power-line electromagnetic fields (EMFs) has been driven by widespread dissemination of the outcomes of epidemiologic studies, even though the evidence was, and still is, considered inconclusive by many scientists. In response to the situation, EPA prepared a draft report in 1990 evaluating the potential carcinogenicity of EMF exposures. But EPA's Science Advisory Board and the White House Committee on Interagency Radiation Research and Policy Coordination were critical of parts of the report (an outcome that may have resulted in part from inadequate analysis and imprecision in the writing of that report). The White House Committee requested a review of the literature on EMF and cancer by the Oak Ridge Associated Universities, which assembled an expert panel.



The panel reported that the epidemiologic findings about EMF and cancer were inconclusive, inconsistent, and without a plausible mechanism (ORAU, 1992; Young, 1993). It also concluded that, given the ever-decreasing resources for basic health and science research, further research investigation of this topic should not receive high priority. Subsequent to the Oak Ridge report, new epidemiologic data appeared (Feychting and Ahlbom, 1992; Floderus et al., 1992), which the interested parties interpreted to support their original positions. The net effects were polarization of the affected groups and heightened public concern.

Earlier, public concern about EMFs had spurred political, legal, and market reactions. In 1989, a background paper prepared for an OTA study proposed prudent avoidance as a policy option if it could be achieved without significant cost or inconvenience (U.S. Congress, OTA, 1989). But such a scenario was and still is unlikely, given that prudent avoidance often involves reconfiguring, rerouting, or burying transmission lines.

The conflict between the public's apprehension about the potential risk posed by EMFs and society's need for reliable, inexpensive electricity elicited political action. After a number of congressional hearings and legislative deliberations, the 102d Congress passed legislation that provided funds for research and dissemination of information to the public (P.L. 102-486, H.R. 776, Oct. 24, 1992). The Department of Energy (DOE) was designated as the lead agency to coordinate the Federal research effort. About $60 million was authorized over 5 years, including $5 million for information dissemination (CRS, 1993).

Interestingly, in the 1960s and 1970s, the U.S. Navy was a major source of finding for studies of the biological effects of extra-low-frequency (76 Hz) because of EMFs associated with submarine communication systems. The research found no significant biological impacts, and the available data were judged inconclusive and controversial (NRC, 1977). Since the 1970s, DOE has been the primary source of support for research on the health effects of exposure to power-line EMFs. The Electric Power Research Institute, an industry-supported, private nonprofit organization, has also been active in the field since the mid- 1970s. Studies of the health effects of EMFs are continuing, but many observers expect that legislative initiatives will be considered in the near future to regulate or otherwise limit exposures. Court cases in which nearby residents are claiming damages to their health and losses in property values are another force driving decisions about power lines,

Dioxins

The polychlorinated dibenzo-p-dioxins ("di-Oxins," or PCDDs) and polychlorinated dibenzofurans ("furans," or PCDFs) make up a family of 210 structurally related chemical compounds. Many researchers believe that these substances that often coexist as contaminants in various materials produce similar effects on health. The chemical most often called dioxin, 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), is the most thoroughly studied and most toxic of this group of chemicals. Formerly, it was inadvertently produced during the manufacture of 2,4,5-trichlorophenol, which was a precursor in the production of some important disinfectants and herbicides, including one of the herbicides in Agent Orange, It is also formed as a byproduct of combustion and chlorine bleaching of paper and pulp.

Concerns about dioxins and furans are rooted in toxicity studies that found TCDD the most potent rodent carcinogen ever studied. Nevertheless, despite more than a decade of epidemiologic studies, there is still no convincing proof that dioxin causes cancer in humans (Bailar, 1991; Gough, 1992/1993). Yet carcinogenicity is not the only concern: TCDD causes adverse effects to every organ system in at least one test animal species (Gough, 1986). Is humankind more like the most sensitive or the least sensitive species tested-or is it somewhere in between? The dioxin problem has prompted intense research *to* support decisionmaking about TCDD as well as about dioxins and furans in general. That research has enhanced scientists' ability to detect and measure dioxins and furans in environmental and biological materials and has expanded the body of knowledge about dioxin's effects and its mechanism of action. Nevertheless, an understanding of the sensitivity of humans to dioxin remains elusive.

The size of the dioxin research effort can be measured by the number of research papers it has produced. The frost dioxin symposium, held in

Rome in 1980, saw 50 papers presented. The 12th symposium, held in Research Triangle Park in 1992, included 10 times that number. Both government and industry have contributed to the more than 20,000 dioxin-related papers and presentations in circulation, and that number continues to increase linearly (Gallo, 1993).

The expansion of dioxin research is due in part to a belief by both scientists and decisionmakers that intense research efforts could resolve the problems that dioxin raises. Because dioxin is a contaminant and has no commercial value, no one has a vested interest in keeping it in the market, and much attention is being given to ways to control its release and to mitigate TCDD already in the environment. But arguments continue to arise about how much such mitigation is worth and who should pay for it.

In 1985, EPA prepared a health assessment document on PCDDs (U.S. EPA, 1985). Because IARC (1987) and EPA (1985) considered the evidence for carcinogenicity in humans inadequate, EPA based its risk estimate on the Dow Chemical Company's study of dioxin and cancer in rats (Kociba et al., 1978). In 1988, pathologists reevaluated the Dow pathology data using revised criteria from the National Toxicology Program for classifying liver tumors. In response to the reevaluation, EPA prepared a draft report that proposed revising the "potency" estimate for TCDD (U.S. EPA, 1988a). The report also proposed methods to derive average risk estimates with models using different mechanistic underpinnings. EPA's Science Advisory Board rejected the methodology, however, and the document remains in draft form while research continues.

Prompted by publication of a new epidemiologic study of cancer mortality in workers exposed to TCDD (Fingerhut et al., 1991) and the conclusions about dioxin toxicity mechanisms from the 1990 Banbury Conference (Banbury, 1991), EPA announced that in 1991 it was reassessing dioxins. The agency held workshops in September of 1992 to review its draft reports and continues to work on the report, with release expected soon. To involve the public and invite its participation, EPA has held public meetings as part of its assessment process.

Studies of TCDD have produced almost all of the current knowledge of dioxins and furans, but most human exposures are to mixtures of dioxins and furans, about which very little is known. The need to address the risks posed by other dioxins and furans has stimulated development of an interim procedure, the toxicity equivalency factor (TEF) procedure. This interim method is being used to estimate risks from exposure to mixtures of PCDDs and PCDFs in the absence of specific information about their specific toxicities. The TEF method is a science-based response to a regulatory need—to estimate the toxicity of dozens of chemicals that have not been tested. In many respects, it represents the response of scientists to the demands of regulators.

The TEF approach is a numerical procedure based on scientific data and scientific judgment. It was first considered for use in the late 1970s and early 1980s when data began to reveal the relative toxicity of some of the different dioxins and furans. In 1986, EPA's Risk Assessment Forum requested that the Science Advisory Board convene a panel of experts to review the TEF methodology and its scientific support. The board's panel of scientists accepted the procedure with some reservations. First, the panel emphasized that the procedure should be considered an interim method. Second, it noted that the procedure lacked scientific validity and therefore needed validation. Third, it accepted the report with the understanding that EPA would fulfill its commitment to periodically review and update the procedure. The frost report on dioxin and furan toxicity equivalency factors was published by the Risk Assessment Forum in 1987 (U.S. EPA, 1987a); it was updated by the forum and republished in 1989 (U.S. EPA, 1989a).

Because of the universality of the dioxin problem, the North Atlantic Treaty Organization (NATO) adopted a TEF procedure for use in

Europe. The NATO committee that adopted a TEF scheme expressed some of the same reservations that the EPA Science Advisory Board noted. It said that the procedure should be considered an interim one and recommended that a ''vigorous program of research be conducted to address areas of uncertainty, to test, refine, or replace the TEFs" (U.S. EPA, 1989a).

Research Into Feedback From Decisionmakers to Researchers

At its most basic level, the relationship between research and decisionmaking can be seen as a feedback loop: one-half of the loop is the impact of research on decisionmaking, and the other half is the impact of decisionmaking on the research that needs to be done. Taken together, the relationship provides a panoply of options, not only for possible decisions but for research priorities as well.

The first half of the loop, which has already been discussed, allows the results of risk assessment (and by extension, risk assessment research) to provide the range of options to be considered in decisions about how to manage particular risks. The past decade witnessed the increasing use of risk assessment in decisionmaking, whether for standard-setting or as a tool for screening and priority-setting (Rosenthal et al., 1992). As described earlier in this chapter, the impact of research on risk assessment may be felt slowly and can be difficult to measure. Nevertheless, changes in agency guidelines for risk assessment and in case-specific regulatory decisions have been observed.

Few inquiries have been made about the other half of the feedback loop, in which policy decisions influence priorities for research. Based on meager evidence, some analysts argue that the poor record of applying analytical thinking to developing research priorities will change as research resources dwindle. In that case, studies will be planned, ranked, and ultimately funded in the light of the decisions on which they will have an impact (Finkel, 1993). Agencies, in turn, will give high priority to the research that is most likely to reduce compliance costs, minimize controversies, and reduce the health toll that hazardous agents may pose. In such circumstances, the value of information to the decisionmaker, not just the increase of knowledge, should influence what research will be given high priority.

The few researchers who are examining feedback from decisionmakers to researchers expect to obtain some insight into this portion of the feedback loop. The notion that information has a measurable value allows decisionmakers to use a quantitative process to evaluate both the need for additional information and the nature of that information.

An additional line of research is the exploration of the relative value and costs of various types of studies for decisionmaking. Lave and Omenn (1986), for example, developed a framework for decision analysis that examines the cost-effectiveness of short-term tests as predictors of carcinogenicity. Their framework estimates both the direct cost of testing and the total social costs of correctly classifying true-positive and true-negative carcinogenic chemicals; it also calculates the costs of misclassifying chemicals. Similar work has been done on the value of animal bioassays and the information they provide (Lave et al., 1988; Taylor et al., 1993). Based on this modeling, decisionmakers can decide on the relative worth of different testing schemes.

THE LIMITS OF SCIENCE IN SOCIAL DECISIONS

Whatever is expected of risk assessment in any given set of circumstances, it is only one of the elements in the formulating regulatory actions. Legislative mandates, social values, technical feasibility, economic factors, and the achievements or shortcomings of the research that feeds into risk assessment may assume a more prominent role than expert projections of risk (figure

5-1). The case study about regulating radon in drinking water in the next chapter illustrates some of the interplay between science and decisionmaking. Scientific research can provide a more solid foundation for the decisionmaker in choosing among alternatives to manage risk, but by itself it will not necessarily influence decisions so as to control the most significant risks. Moreover, the capacity of science to inform decisions even on many technical risk-related issues is limited.

The limits of science manifest themselves at a variety of levels. Uncertainty in measurements and observations constrains science at the most fundamental level, and the scientific underpinnings of risk assessment are more subject to that limitation than are experimental sciences. At a higher level of complexity, the interpretation of data and observations to predict outcomes introduces additional uncertainties. And risk management actions can themselves produce uncertainty.

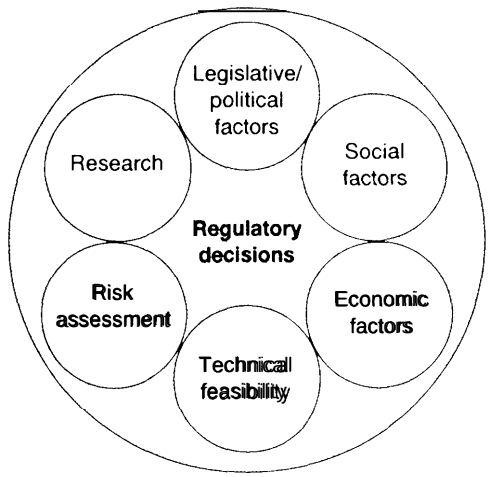
Measurements and Observations

Information for assessing risks to human health comes from epidemiologic observations, animal toxicity testing, various laboratory studies, and measures and estimates of human exposures. As detailed in this report, guidelines are available for the use of this information, but all measurements and estimates are subject to the limits of the methods used and uncertainties. There are technical bounds to the experimental methodologies and equipment as well as limitations that prevail in interpreting and analyzing data.

Epidemiologic Data for Assessing Risks

The availability of epidemiologic data eliminates the problem of extrapolating from animal data to humans, but epidemiologic data can suffer from a number of substantial limitations. Some are methodological in nature and may be overcome with new techniques. For instance, many investigators attempt to couple epidemiologic studies with cellular-molecular techniques so that the results will provide information about the mechanism of carcinogenesis.

Figure 5-l-Research as an Element in Regulatory Decisionmaking



SOURCE: U.S. Congress, Office of Technology Assessment, 1993.

The lack of exposure measurements can limit the usefulness of results from epidemiologic investigations because it forces researchers to rely on estimates and can lead to errors. For example, scientists who are studying health effects in the "Ranch Hands," the men who sprayed dioxin-contaminated Agent Orange in Vietnam, used records of job assignments and recall by the men to estimate exposure to the chemical. When techniques became available for measuring exposure directly, researchers found flaws in the classification of exposures on the basis of job category, a standard practice in many epidemiologic studies (Air Force Health Study, 1991a). In a related study, Needham (1991) reported few, if any, correlations between activities around dioxin-contaminated soils and measured exposure levels.

More generally, "confounders" complicate the design and interpretation of epidemiologic studies. As an example, Air Force scientists reported that diabetes is more common in Ranch Hand veterans with higher dioxin levels (Air Force Health Study, 1991 b). But a connection between that disease and dioxin exposures was

confounded by the finding that more obese veterans, who are more likely to have diabetes, also tend to have higher dioxin levels. The connection between obesity and dioxin levels is largely explained by differing rates of dioxin elimination from Ranch Hand veterans: dioxin persists longer in more obese men (Wolfe et al., 1993). Therefore, although the details remain to be sorted out, the connection between obesity and dioxin metabolism confounds the interpretation that can be put on the observation that diabetes is more common in men with higher levels of dioxin.

More common confounders are exposure to multiple agents in the environment, the effects of different lifestyles-including eating, drinking, and smoking habits—and genetic differences. All of these factors complicate extrapolating from study data to estimations of risk to the general population.

Results From Animal Toxicity Testing

Most often, information about the toxicity of substances comes from tests in animals, and all tests are compromises. To compensate for the small number of animals (usually rodents) that can be tested, rodents are exposed to higher concentrations of the agent than the levels that humans are expected to experience. There maybe differences in response to high and low doses or to particular routes and patterns of exposure, and between species.

Some of these limitations can be overcome by appropriate pharmacokinetic studies and analysis, which relates exposure to time-dependent distribution of the chemical in the body, and by pharmacodynamic analysis, which examines internal doses and effects at the organ, tissue, cellular, and molecular levels and relates them to the development of toxic effects. Almost always, however, comparative metabolic and pharmacokinetic data are incomplete; frequently, they are simply unavailable. Even when data are available, sites of toxic action can vary among species, further complicating interpretation and prediction. Unless the target organ is known, physiologically based pharmacokinetic modeling is of limited value.

A further complication is that a chemical frequently causes several effects in test animals. In evaluating animal studies, toxicologists must decide whether the response caused by the agent is well within the range of normal physiological adjustments (homeostatic response) or an abnormality that constitutes a toxic response. In many instances, homeostatic and toxic responses represent different parts of the same continuum. Put more simply, the question is, when does a response represent an adverse effect on health?

In carcinogenesis, a process that involves many etiologic factors, more than one mechanism may be operative in each of the steps or stages, which are often called initiation, promotion, and progression. Sometimes, information on the mechanism of carcinogenic action of a chemical or product are developed years after it is found to be carcinogenic in rodent studies. Sorting through the possible mechanisms can involve a broad range of issues and fields-for example, direct interactions with DNA, disturbance of hormonal balances, changes in cell organelles, organ-specific cytotoxicity, immunomodulation, perturbation of DNA methylation, peroxisome proliferation, and inhibition of intercellular communication. In some (perhaps many or most) cases, the explanation may be found in combinations of those mechanisms. When the information about a substance is incomplete, interpretations that draw connections between animal data and estimates of human cancer can produce great disagreement among scientists. As far as scientific understanding of the mechanisms of carcinogenicity has come, it remains far short of certainty.

Information on Exposure

After a chemical is released into the environment, it maybe transported or transformed, it may persist, enter, and be concentrated in the food

chain, or it may be degraded or deposited where humans cannot come in contact with it. It can reach humans through the air they breathe, the food they eat, or the water they drink, or by contact with their skin. There are many ways to predict, **estimate,** and measure the exposures of humans to environmental agents, but it is a complex, uncertain undertaking.

Typically, researchers measure levels of pollutants at the sources of their discharge into the environment and then use models to predict the concentrations reaching humans. Personal monitoring devices produce more realistic measurements of human exposure. Even so, age, physical activity, nutritional conditions, and other factors related to lifestyle can affect the body's uptake of the pollutant, leading to uncertainty about the dose that any one individual receives.

The most direct measure of human exposure is through biological monitoring of body fluids or tissues. But these techniques are expensive and not without risk if they require biological samples. Generally, **estimates** of exposure are generated by reconstruction of behavior and surveys of recall as in early parts of the Ranch Hand study mentioned above.

Social and Political Factors in Decisionmaking

As research identifies the potential adverse health effects of toxicants to which humans may be exposed, the public conveys its concerns to Congress, and Congress considers and often passes laws to address those concerns. This reactive mode may limit the capacity of agencies, such as EPA, to structure long-term solutions **that are** both efficient and effective. In 1991, then EPA Administrator William K. Reilly stated:

> For 20 years we have established goals on a pollutant-by-pollutant and medium-by-medium basis without adequately considering broader environmental quality objectives. We have seldom if ever been directed by law to seek out the best opportunities **to** reduce environmental risks, in toto, or to employ the most efficient, cost-effective procedures.

Regulatory decisions are often made with inadequate data and in response to statutory mandates. This limits the capacity of science to support regulatory decisions. Furthermore, scientific input is but one element in the formulation of regulatory decisions. As in all kinds of human activity, change is difficult, and various factors—risk perception, economic impact, social values, lack of **trust** between the public and industry, and less than complete confidence in government—play a role in decisionmaking. In such **a** context, new facts from science may have little impact.

Some analysts and scientists (e.g., Abelson, 1993; Gori, 1992; Moolenaar, 1992) maintain that more scientific information is needed to support environmental rulemaking by the agencies. Openly critical, they contend that advances in the biological and biomedical sciences make their way too slowly into regulatory decisionmaking and that those decisions remain mired in the science of the past two decades. Jasanoff (1990) characterizes the contention that better decisionmaking will result from more and better scientific information as the ''technocratic viewpoint. '

Not everyone shares that view. First, it is difficult to prove that better (or "more," as some detractors say) science has improved decisionmaking. Since the risks that most regulations address are below the limits of detection by epidemiology, it is impossible to know if one approach or the other produced better results in protecting health. Understandably, few examples of improved decisionmaking exist, and the social and political implications of decisionmaking may mask any effect science has on the process (Jasanoff, 1990).

A more basic point is that risk assessment is contentious because scientific data are seldom definitive and consensus on some issues appears unlikely. A recent National Research Council report on risk assessment included rare majority and minority recommendations; the issue with no

agreement was whether toxic effects observed at the maximum tolerated dose are predictive of human risk. This topic has been debated for decades (NRC, 1993).

Moreover, research findings can complicate risk assessment (Huff, 1993). Research takes time, and more research can be used to serve the political objective of delaying regulatory action (O'Brien, 1993; Olson, 1984; Silbergeld, 1993). Finally, risk assessment may be the last point at which science is considered because of the power of policy mandates that place more weight on the side of safety (Graham, 1991).

Tradeoffs, Teamwork, Trust, and Leadership

An optimistic view is that the field of health risk assessment is still young. With the advances being made in the biological and biomedical sciences, the field of toxicology will evolve to provide better data; combinations of epidemiologic and laboratory-based investigations will produce more revealing information; and measurements of exposure to environmental chemicals will sharpen risk assessments.

Nevertheless, it is unrealistic to expect research to resolve all uncertainty and eliminate all differences in interpreting data. Solving the problems in environmental risk assessment goes beyond more and better science: it requires building trust among government, industry, and citizens. It also requires leadership in setting realistic goals and arranging collaborations of researchers from various disciplines and sectors of society.

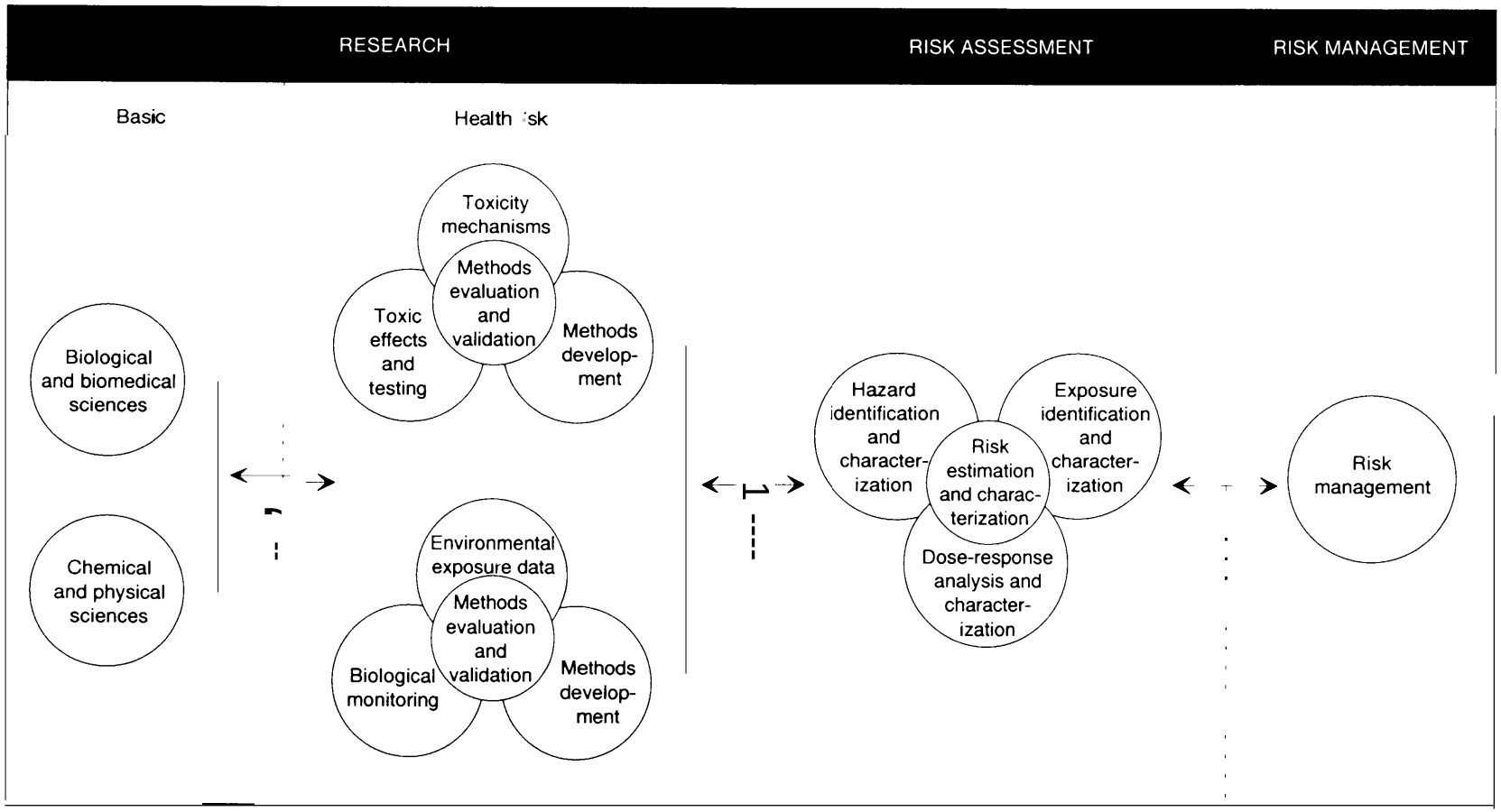
In 1983, the National Research Council Committee on the Institutional Means for Assessment of Risks to Public Health called for separating risk assessment and risk management (NRC, 1983). This recommendation was taken up by the agencies, which separated the functions of scientists and decisionmakers to prevent 'subtle value judgments" from influencing empirical analyses before they reached the regulators. Now, a decade later, some scientists argue that opportunities are being lost in an inefficient system whereby those making decisions are unaware of the process generating the information on which their decisions are based. One of the principal deficiencies of that system, according to Finkel (1990), is that decisionmakers remain insulated from the inherent uncertainty in the process.

Still, the division between risk management and risk assessment has never been complete. All kinds of policy judgments reach back into the risk assessment process. For example, the decision to accord greater weight to public health than to industrial output greatly influences the default position that estimates cancer risk using a no-threshold model. However rigid or flexible the boundary between risk assessment and management, it is clear that the relationships between the two are under discussion and perhaps in flux. Moreover, as Congress expresses interest in risk assessment, its discussions and mandates will influence both the process itself and risk assessment research.

LINKING HEALTH RISK RESEARCH TO DECISIONMAKING

The relationship between research and decisionmaking is complex. OTA developed figure 5-2 to describe the relationships among the various research activities in health risk assessment and decisionmaking. Previously, the National Research Council (NRC, 1983) depicted a unidirectional flow of information from research to decisionmaking, which emphasized the compartmentalization of the process for providing public transparency (see figure 2-1). Yet information sharing throughout the process is important to increase the efficiency of research for decisionmaking. Thus, figure 5-2 highlights the bidirectional flow of information as well as the integration and synthesis of information from the various disciplines and types of research. The evaluation and validation of methods can serve as the focal point for integrating all the areas of health risk

Figure 5-2—Linking Research on Health Risks to Decisionmaking



SOURCE: U.S. Congress Office of Technology Assessment, 1993.

research, given that a new model or method must be examined and compared with methods of known and established veracity. Figure 5-2 also indicates OTA's stress on the interdependency of research activities, the risk assessment process, and policymaking.

The link between health effects research and the basic biological, chemical, and physical sciences has often been neglected in discussions about health risk assessment research. Now, however, bridges are being constructed between basic and health risk research in response to calls from Congress (Brown, 1993; U.S. Congress, House Committee, 1992) and the private sector (Carnegie Commission, 1992) for linking science to social needs. Although some basic scientists may respond grudgingly at first, later they may actually find it rewarding to modify and redirect their research to serve health risk assessment. Much as research on AIDS or cancer links social needs and unexplored avenues of research, improving risk assessment can similarly endow toxicological research with an objective that transcends the purely scientific.

Health risk assessments, to be valid, require the participation of scientists from many disciplines. Those from the toxicological and biomedical sciences are best qualified to critique the validity of the scientific underpinnings of assessments. But when data are lacking, assumptions and policy positions with embedded value judgments are used, arguably, as tools to complete the assessment, their selection may benefit from involving practitioners of disciplines other than the biological, chemical, or physical sciences. Jasanoff (1993) argues for "bridging the two cultures of risk analysis' '—the ''hard, ' or quantitative, sciences and the ''soft, ' or nonquantitative, disciplines, such as the behavioral and political sciences.

Although OTA did not address whether risk assessment itself should be formally considered a scientific endeavor, making risk assessment an active field of research may well be what is needed to facilitate the application to health risk assessments of the new biological understanding of diseases and toxicological mechanisms. From that perspective, OTA sees risk assessment as involving the analysis and synthesis of all that is known about the risk at hand, such as a specific chemical or class of chemicals. For example, risk assessments use findings from epidemiologic studies or results from animal toxicity tests to generate hypotheses about risks to human health. A substantial amount of reasoning and judgment is required in determining whether the composite data on toxic effects, exposure, and dose-response characteristics as a whole make the risk hypothesis tenable.

In contrast to the approach described here, risk assessments are all too frequently performed by merely stacking up the positive findings that imply that risk exists, without careful attention to conflicting results and alternative interpretations. A more iterative process of questioning can reveal the strengths and weaknesses of the case for the existence of risk and identify the need for further research at each step. In that way, gaps in the data can be recognized and research conducted in response. Not every alternative interpretation need be considered or presented every time. Risk characterization and communication can serve as a focus for the iterative process of risk assessment and abridge between risk assessors and decisionmakers.

Risk characterization summarizes and interprets the information available about a given risk for risk managers and the public. There is general agreement that methods of risk characterization are poorly developed and that efforts to improve them have been neglected (Gray, 1993). Research to improve risk characterization is directed toward developing methods to describe more completely the uncertainties and assumptions and to express the full range to plausible estimates of risk.

For example, one of the methods frequently proposed for distinguishing among alternative estimates of risk is the distributional approach, which is an outgrowth of uncertainty analysis in

the decision sciences (Morgan and Henrion, 1990). It uses explicit expert judgment to analyze and quantify the plausibility of alternative interpretations of available data (Otway and von Winterfeldt, 1992). This method can capture the range of scientific evidence and opinion on key biological uncertainties. Proponents of describing ranges of risk believe that it avoids the focus on a single numerical estimate of risk and requires risk managers to confront qualitative uncertainties, such as the likelihood that a compound is or is not a carcinogenic hazard to humans (Gray, 1993).

The purpose of risk characterization is to help risk managers and others understand the results of complex risk assessments. For a risk manager, good risk characterization will aid decisionmaking. But for that to occur, a risk manager must understand the basis for risk estimates including the scientific, analytical, and policy choices that underlie the assessment. Improving risk characterization will also help legislators, journalists, and the public understand the nature and magnitude of the day-to-day risks citizens in this country face.

Two Federal programs support research on risk communication and decisionmaking. EPA's Office of Policy, Planning, and Evaluation conducts and supports research in risk communication. In addition, the agency holds workshops and offers training in risk communication and decisionmaking. The National Science Foundation funds research on decisionmaking through the Decision, Risk, and Management Science (DRMS) program in the Division of Social and Economic Sciences. For the past decade, DRMS has operated a competitive grants program that supports research to develop new methods in the field of decision theory and methods to optimize the technical handling of risk probability. The program funds research on social factors that influences risk assessment, and it seeks to distinguish between technical and social definitions of risk (Cantor, 1993).

SUMMARY

Research has had only a modest effect on efforts to revise the science policy assumptions adopted in EPA's risk assessment guidelines. It has, however, had a substantial impact on chemical-specific risk assessment and consequently on regulatory actions, and it is currently generating considerable debate as EPA considers revisions to its 1986 cancer risk assessment guidelines.

Three interacting factors account for the limited impact of new scientific research on the science policy assumption adopted in EPA's risk assessment guidelines. The nature of the assumptions, the importance of the assumption to the paradigm underlying the regulatory approach, and the policy reverberations associated with changing specific default positions all jointly limit any expedient change of the agency's science policy assumptions based on new knowledge.

Health risk research and decisionrnaking are interdependent. As research identifies potential adverse effects on health, the public conveys its concern to Congress, and Congress considers and passes laws to address those concerns. This reactive mode limits the capacity of agencies to structure long-term solutions including appropriate research.

Although the sciences can provide solid foundations for choices about reducing health risks, their contribution is limited because measurements and interpretations of data are inherently uncertain. In addition, science is only one of the elements in regulatory decisions. Legislative mandates, social values, technical feasibility, and economic factors may assume more prominent roles, depending on the specific issue. Solving the problems in health risk assessment goes beyond more and better science, it requires building trust among government, industry, and citizens. It also requires leadership in setting realistic goals and encouraging collaboration in research.

CHAPTER 5 REFERENCES

Abelson, P.H. 1993. Editorial: regulatory costs. Sci*ence* 259:159.

Air Force Health Study. 1991a. *Serum dioxin: analysis of 1987 examination results,* vol. 1. Epidemiology research division, Brooks Air Force Base.

Air Force Health Study. 1991b. *Serum dioxin analysis of 1987 examination results,* vol. 4. Epidemiology research division, Brooks Air Force Base.

Ames, B. N., and Gold, L.S. 1990a. Chemical carcinogenesis: too *many* rodent carcinogens. *Proceedings of the National Academy of Sciences USA* 87:7772-7776.

Ames, B. N., and Gold, L.S. 1990b. Too many rodent carcinogens: mitogenesis increases mutagenesis. *Science* 248:970-971.

Bailer, J. 1991. How dangerous is dioxin? New *England Journal of Medicine* 324:260-262.

Banbury. 1991. *Biological basis for risk assessment of dioxins and related compounds,* M.A. Gallo, R.J. Scheuplein, and K.A. Van Der Heijden (eds.). Banbury Report 35 (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Brown, G.E. 1993. Science's real role in policymaking. *Chemical and Engineering News,* May31, pp. 9-110

California Department of Health and Human Services. 1988. Health effects of methylene chloride, N. Gavitz (cd.) Sacramento, December.

Cantor, R. 1993. Director, Decision, Risk, and Management Science Program, National Science Foundation. Personal communication.

Carnegie Commission. 1992. *Environmental Research and Development: Strengthening the Federal Infrastructure.* A report of the Carnegie Commission on Science, Technology, and Government.

Cohen, S. M., and Ellwein, L. 1991. Letters: carcinogenesis mechanisms—the debate continues. *Science* 252:902-903, May 17.

Cohen, S. M., and Ellwein, L.B. 1992. Risk assessment based on high-dose animal exposure experiments, *Chemical Research in Toxicology* 5(6):742-748.

CRS (Congressional Research Service). 1993. *Health Effects of Power-Line Electromagnetic Fields (EMFs),* C.S. Redhead and C.H. Dodge. CRS Issue Brief IB9211 8. Science Policy and Research Division. Feb. 8.

Dunlap, T. 1988. *DDT: Scientists, Citizens, and Public Policy* (Princeton, NJ: Princeton University Press).

Feychting, M., and Ahlbom, A. 1992. *Magnetic Fields and Cancer in People Residing Near Swedish High Voltage Power Lines* (IMM-rapport, June 1992, Stockholm: Karolinska Institute).

Fingerhut, M. A., Halperin, W, E., Marlow, D. A., et al. 1991. Cancer mortality in workers exposed to 2,3,7,8-tetrachorodibenzo-p-dioxin. New *England Journal of Medicine* 324:212-218.

Finkel, A.M. 1990, *Confronting Uncertainty in Risk Management: A Guide for Decisionmakers* (Washington, DC: Resources for the Future),

Finkel, A. 1993. Fellow, Resources for the Future. Personal communication,

Floderus, B. et al. 1992. *Occupational Exposure to Electromagnetic Fields in Relation to Leukemia and Brain Tumors: A Case-Control Study* (Solna, Sweden: National Institute of Occupational Health) (P. M.: ed.).

Gallo, M. 1993. Professor. Department of Environmental and Community Medicine, Robert Wood Johnson Medical School. Personal communication.

Gearhart, R. 1992. California Department of Health and Human Services. Personal communication.

Gori, B.G. 1992. Whither risk assessment? An overall perspective. Presented at the annual meeting of the American Industrial Health Council, Washington, DC, Dec. 1.

Gough, M. 1986. *Dioxin, Agent Orange (New* York and London: Plenum Press).

Gough, M. 1992/1993. chemical risk assessment is not science. *Chemistry,* December/January :23-26.

Graham, J.D. 1991. The failure of agency-forcing: the regulation of airborne carcinogens under Section 112 of the Clean Air Act. *Duke Law Journal 1985:100-150.*

Gray, G. 1993. The challenge of risk characterization. National Technical Information Service. PB93-218857.

Huff, J.E. 1992. Chemical toxicity and chemical carcinogenesis: is there causal connection? A comparative morphological evaluation of 1,500 experiments. In: *Mechanisms of Carcinogenesis in Risk Identifica-*

tion, H. Vainio (cd.), IARC Scientific Publications 116 (Lyon, France: International Agency for Research on Cancer), pp. 437-475.

Huff, J. 1993. Issues and controversies surrounding qualitative strategies for identifying and forecasting cancer-causing agents in 'the human environment. *Pharmacology and Toxicology* 72(Suppl.1):12-27.

IARC (International Agency for Research on Cancer). 1987. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Overall Evaluations of Carcinogenicity. An updating of IARC monographs, vol. 1 to 42, supplement 7, pp. 350-354.

IARC. 1992a. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Solar and Ultraviolet Radiation.* IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (Lyon, France), Feb. 11-18.

IARC. 1992b. *Mechanisms of Carcinogenesis in Risk Identification,* H. Vaino, P.N. Magee, D.B. McGregor et al. (eds.). Scientific Publication 116 (Lyon, France).

Jasanoff, S. 1990. *The Fifth Branch: Science Advisors as Policymakers* (Cambridge, MA: Harvard University Press).

Jasanoff, S. 1993. Bridging the Two Cultures of Risk Analysis. *Risk Analysis* 13:123-129.

Kociba, R., Keyes, D., Beyer, J. et al. 1978. Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8 -tetrachlorodibenzo-*p*-dioxin (TCDD) in rats. *Toxicology and Applied Pharmacology* 46:279.

Lave, L. B., and Omenn, G.S. 1986. Cost-effectiveness of short-term tests for carcinogenicity. Nature 324(6092):29-34, Nov. 6.

Lave, L. B., Ennever, F. K., Rosenkranz, H.S. et al. 1988. Information value of the rodent bioassay, *Nature* 336:631-633.

Moolenaar, R.J. 1992. Overhauling carcinogen classification. *Issues in Science and Technology* 8(4):70-75.

Morgan, M. G., and Henrion, M. 1990. *Uncertainty: A Guide to Dealing with uncertainty in Quantitative Risk and Policy Analysis* (Cambridge: Cambridge University Press).

Needham, L.L. 1991. Remarks reported in: Human exposures from dioxin in soil, M. Gough. *Journal of Toxicology and Environmental Health* 32:205-245.

NRC (National Research Council). 1977. *Biologic Effects of Electric and Magnetic Fields Associated with Proposed Project Seafarer.* Committee on biosphere effects of extremely-low-frequency radiation (Washington, DC: National Academy Press).

NRC. 1983. *Risk Assessment in the Federal Government: Managing the Process* (Washington, DC: National Academy Press).

NRC. 1989. *Improving Risk Communication* (Washington, DC: National Academy Press).

NRC. 1993, Scientific debate continues over cancer risk assessment methods. Press release, Jan. 28.

O'Brien, M.H. 1993. Alternatives to risk assessment. New *Solutions* Winter:39-42.

Olson, E.D. 1984. The quiet shift of power: Office of Management and Budget supervision of Environmental Protection Agency rulemaking under Executive Order 12,291. *Virginia Journal of Natural Resources Law* 4(1):3-80.

ORAU (Oak Ridge Associated Universities). 1992. *Health Effects of Low-Frequency Electric and Magnetic Fields,* ORAU 92/F8 (Oak Ridge, TN), June.

Otway, H., and von Winterfeldt, D. 1992. Expert judgment in risk analysis and management: process, context, and pitfalls. *Risk Analysis* 12:83-93.

Pease, W. 1992. Environmental Health Policy Program, School of Public Health, University of California at Berkeley. Personal communication.

Pease, W., Vandenberg, J., and Hooper, K. 1991. Comparing alternative approaches to establishing regulatory levels for reproductive toxicants: DBCP as a case study. *Environmental Health Perspectives* 91:141-155.

Preuss, P. 1992. Office of Science, Planning, and Regulatory Support, U.S. Environmental Protection Agency. Personal communication.

Reilly, W.K. 1991. Why I propose a national debate on risk. *EPA Journal* March/April:2-5.

Rosenthal, A., Gray, G. M., and Graham, J.D. 1992. Legislating acceptable cancer risk from exposure to toxic chemicals. *Ecology Law Quarterly 19(2):269-362,*

Ruckelshaus, W.D. 1983. Science, risk, and public policy. Presentation to the National Academy of Sciences, Washington, DC, June 22.

Rushefsky, M.E. 1986. *Making Cancer Policy (Albany*: State Universities of New York Press), pp. 107-150.

Silbergeld, E. 1993, Guest editorial: the risks of risk assessment. New *Solutions,* winter:43-44.

Slone, T.H. 1993. Body Surface Area Misconceptions. *Risk Analysis* Vol. 13, No. 4.

Taylor, A., Evans, J., and McKone, T. 1993. The value of animal testing information in environmental control decisions. *Risk Analysis* 13:403-412.

Travis, C. C., and White, R. 1988. Interspecies scaling of toxicity data. *Risk Analysis 8:119-125.*

Travis, C.C. et al. 1990. Interspecies extrapolation of pharmacokinetics. *Journal of Theoretical Biology* 142:285-304.

U.S. Congress, House Committee on Science, Space, and Technology. 1992. *Report of the Task Force on the Health of Research* (Washington, DC: Government Printing Office).

U.S. Congress, OTA (Office of Technology Assessment). 1981. *Assessment of Technologies for Determining Cancer Risks From the Environment,* OTA-H-138.

U.S. Congress, OTA. 1989. Biological effects of power frequency electric and magnetic fields. BP-E-53.

U.S. EPA (Environmental Protection Agency), 1976. Interim procedures and guidelines for health risk and economic impact assessment of suspected carcinogens. *Federal Register* 41:21402-21407.

U.S. EPA. 1985. *Health Assessment Document for Polychlorinated Dibenzo-p-dioxins.* Final Report. 01-IQ EPA/600/8 -84/014F. September.

U.S. EPA. 1986a. Guidelines for carcinogen risk assessment. *Federal Register* 51:33992-34003.

U.S. EPA, 1986b. Guidelines for exposure assessment. *Federal Register 51*:34042-34154.

U.S. EPA. 1986c. Guidelines for the health assessment of suspect developmental toxicants. *Federal Register* 51(185):34028-34040.

U.S. EPA. 1987a. *Interim Procedures for Estimating Risks Associated with Exposures to Mixtures of Chlorinated Dibenzo-p-dioxins and -Dibenzofurans (CDDs and CDFs).* Risk Assessment Forum, EPA/625/3-87/012.

U.S. EPA. 1987b. *The Risk Assessment Guidelines of 1986,* EPA/600/8-87/045.

U.S. EPA, 1987c. Technical analysis of new methodologies and data regarding dichloromethane hazard assessment. External review draft.

U.S. EPA. 1988a. Cancer risk: specific dose estimate for 2,3,7,8-TCDD. External review draft (EPA 600/6/88/007 Aa).

U.S. EPA. 1988b. Intent to review guidelines for carcinogenic risk assessment. *Federal Register 53:32656-32658.*

U.S. EPA. 1988c. *Thyroid follicular cell carcinogenesis: mechanistic and science policy considerations.* EPA 625/3-88/014A. Risk Assessment Forum.

U.S. EPA. 1989a. *Interim Procedures for Estimating Risks Associated with Exposures to Mixtures of Chlorinated Dibenzo-p-Dioxins and -Dibenzofurans (CDDs and CDFs) and the 1989 Update.* Risk Assessment Forum. EPA 625/3-89/016.

U.S. EPA. 1989b. Proposed amendments to the guidelines for the health assessment of suspect developmental toxicants. *Federal Register* 54:9399.

U.S. EPA. 1991a. *Alpha -2u-globulin: association with chemically induced rend toxicity and neoplasia in the nude rat.* Risk Assessment Forum. EPA 625/3-91/019F.

U.S. EPA. 1991b. Guidelines for developmental toxicity risk assessment. *Federal Register* 56:63798-63826.

U.S. EPA. 1992a. Working paper for considering draft revisions to the U.S. EPA guidelines for cancer risk assessment (Washington, DC: EPA 600/AP-92-O03).

U.S. ERA. 1992b. Guidelines for exposure assessment. *Federal Register* 57:22888-22932.

U.S. OSTP (Office of Science and Technology Policy). 1985. Chemical carcinogenesis: a review of the science and its associated principles. *Federal Register* 50: *10372-*1040O.

Weinberg, A. 1972. Science and trans-science. *Minerva* 10:209-222.

Wolfe, W.H., Michalek, J.E., Miner, J.C. et al. 1993. Determinants of TCDD half-life in veterans of operation ranch hand. Submitted for publication.

Young, F, 1993. Deputy Assistant Secretary for Health/Science and Environment, U.S. Department of Health and Human Services. Personal communication.