Policy Study No. 205
May 1996

KEY ISSUES IN ENVIRONMENTAL RISK COMPARISONS
REMOVING DISTORTIONS AND INSURING FAIRNESS

by
George M. Gray, Ph.D.

Table of Contents

I. OVERVIEW

Risk assessment has developed from an arcane practice of interest to a few regulators, academics and specialists in industry to a major factor in evaluating sources and sizes of risks to health. Comparative risk, substitution analysis, risk-based priority setting, and benefit/cost analysis are all parts of this new risk management strategy. What is risk assessment? How has it been used in the past? What are the new uses of risk assessment? What is the promise and what are the pitfalls of using risk assessment to protect human health and the environment? This study will explore these questions and, in the end, sound a warning.

Current methods of risk assessment were developed for the purpose of setting standards—safe levels of exposure to hazardous materials. Central to these methods was the concept of conservatism: deliberately inflating estimates of risk in order to avoid setting levels that might not be safe. The new uses of risk assessment, which involve comparison of risks from different chemicals, different types of risk and, often, risks to different people, are seriously compromised by the use of conservative risk assessment methods. It is critical for risk comparison that no particular risk has a thumb on the scale in its favor.

Risk assessment is a method of using scientific information to make informed decisions. Risk assessment for carcinogens is the most prominent in environmental regulation, and will be the focus of this article, but risk assessment for non-cancer health effects and ecological hazards is gaining in importance. Risk assessment plays a key role in such environmental decisions as setting drinking water standards, setting cleanup goals at Superfund sites, and setting standards for pesticide residues on food. The most influential agency in terms of the practice of risk assessment is the Environmental Protection Agency (EPA). Risk assessment touches on virtually all of the two percent of Gross National Product that the United States spends annually on environmental protection.

The growing acceptance of the principles of risk assessment has expanded its use beyond simple regulatory-standard setting. The new uses of risk assessment involve making *comparisons* of health risks. Comparative risk projects have been undertaken to help states, localities, and even regulatory agencies rank large sources of risk to citizens. Risk-based priority setting uses the results of risk-ranking efforts to allocate resources for protecting health, attacking the worst problems first. Risk communication benefits from comparison between different types of risks to life so the public can understand the relative size of the risks of accidents, diseases, and environmental threats. Industry is beginning to compare different pollution-prevention alternatives to determine which generate the greatest reduction in risk. Finally, there is a growing emphasis, both nationally and at the state level, on comparisons of the benefits and the costs (in terms of risk reduction) of environmental as well as health and safety regulations.

The use of carcinogen risk assessment in standard setting led to the use of conservative assumptions in estimating risk. Whenever there was scientific uncertainty these conservative assumptions meant that risk assessors assumed the worst, deliberately inflating risk estimates in a quest for safety.

The perception that conservative risk assessment is skewing regulatory priorities, misleading the public about the relative size of different sources of risk to their health, and leading to large expenditures to generate very small reductions in risk has brought risk assessment onto the radar screen of the U.S. Congress. The past two congressional sessions have seen the introduction of a number of bills which would force regulatory agencies, primarily EPA, to develop best estimates of risk to accompany the "plausible upper bound" on risk calculated with conservative risk assessment assumptions.

This study will argue that the changing uses of risk assessment require that conservatism be removed from risk assessment procedures. Comparing risks means that methods which deliberately inflate the risks posed by certain environmental chemicals cannot last. In order to give regulators a clear ranking of health risks, to help citizens understand sources of risks to their health, for industry to make real risk-reducing pollution-prevention decisions, and to ensure the costs and benefits of regulations are fairly evaluated, carcinogen risk assessment must abandon

methods designed to deliberately overestimate risk. This will require legislative efforts and increased scrutiny by the journalists, legislators, regulators, and citizens who use risk assessments. Fair comparisons among risks requires the best science and the best methods of risk assessment to be used in efforts to protect public health.

II. A BRIEF HISTORY

Carcinogen risk assessment began as a tool for establishing the safety of food additives.[1] Scientists at the Food and Drug Administration (FDA) attempted to use information from animal studies to find levels of exposure so that no harm would be expected when people consumed preservatives, veterinary medications, or other additives in their food. The EPA was the first agency to adopt guidelines for cancer risk assessment in 1976.[2]

Then, as now, risk assessment relied on incomplete information. Toxicological and epidemiological data were often scarce. The relationship between the response of animals to high doses of a substance and humans to much lower

[1] A more detailed discussion of the history of risk assessment can be found in *An Historical Perspective on Risk Assessment in the Federal Government*, Harvard Center for Risk Analysis, Harvard School of Public Health, 1984. See also J.V. Rodericks, "Origins of Risk Assessment in Food Safety Decision Making," *Journal of the American College of Toxicology* (1988), pp. 7:539–542.

[2] U.S. EPA, "Interim Procedures and Guidelines for Health Risks and Economic Impact Assessments of Suspected Carcinogens," *Federal Register* (May 25, 1976), 41:21402.

doses was uncertain. The amount of additives that humans consumed was only roughly known. In the face of this uncertainty, the decision was made to be conservative.[3] In this case, conservative means that efforts were made to avoid any chance of underestimating the risk. Exposure estimates were deliberately inflated. It was assumed that chemicals which induced tumors in rodents would also pose a danger to humans at levels of exposure thousands of times lower. In estimating human response based on animal experiments, the relationship that made the risk look highest was chosen.[4]

These conservative methods of human health risk assessment now form the basis of EPA efforts to estimate potential harm from environmental threats. Risk assessment forms the underpinning for virtually all environmental regulation aimed at protecting human health.[5] It is also used to identify and clean up Superfund sites. Risk assessment is used in setting drinking water standards, deciding if and how a pesticide may be used, setting clean air and water standards and to determine whether the manufacturing, processing, and distribution of a chemical pose an unreasonable risk.

EPA eventually codified conservative methods of risk assessment in its 1986 *Guidelines for Carcinogen Risk Assessment.*[6] For the many situations in which a risk assessor faced a choice—which animal bioassay to use, how much exposure to assume, the form of the dose-response function—EPA mandated choosing the option which made the risk look highest. Key factors include the decisions to:

[3] For an inside perspective on the development of cancer risk assessment guidelines at EPA during the 1970s see R.E. Albert, "Carcinogen Risk Assessment in the U.S. Environmental Protection Agency," *Critical Reviews in Toxicology* (1994), pp. 24:75–85.

[4] See E.L. Anderson and the Carcinogen Assessment Group of the U.S. Environmental Protection Agency, "Quantitative Approaches in Use to Assess Cancer Risk," *Risk Analysis* (1983), pp. 3:277–295, for a discussion of the history and rationale of early risk assessment methods.

[5] For a discussion of the use of risk assessment by EPA see A. Rosenthal, G.M. Gray, and J.D. Graham, "Legislating Acceptable Cancer Risk from Exposure to Toxic Chemicals," *Ecology Law Quarterly* (1992), pp. 19:269-362.

[6] "Guidelines for Carcinogen Risk Assessment," U.S. Environmental Protection Agency, 51 FR33994-33997 (1986).

- assume all animal carcinogens are human carcinogens;

- use conservative estimates for calculating exposure; and

- to use a linear dose response model patterned on the behavior of powerful carcinogens like ionizing radiation.

In the *Guidelines,* EPA states that this prescribed method of risk assessment generates "plausible upper bounds on risk consistent with some proposed mechanisms of carcinogenesis" and warns that "the true value of the risk is unknown and may be as low as zero."

Conservative risk assessment has been defended by EPA and others on several grounds.[7] For example, given the scientific uncertainty in risk assessment, it is better to assume the worst rather than potentially expose people to a significant risk. In addition, some are concerned that although conservative assumptions are made, variability in response among humans or exposure from many sources of pollution are not always explicitly accounted for, producing risk estimates that may not be very conservative at all. In general, it can be said that the use of conservative methods of risk assessment has been justified on the grounds of "better safe than sorry."

[7] For example A.M. Finkel, "Is Risk Assessment Really Too Conservative?: Revising the Revisionists," *Columbia Journal of Environmental Law* (1989), pp. 427:427–467.

The early uses of risk assessment have one thing in common: they involve setting standards. The standard may be the amount of a pesticide that is allowed to be on a crop, the maximum contaminant level goal (MCLG) for drinking water, a cleanup level for soil at a Superfund site, or other numerical targets. Although not without controversy, the use of conservative estimates of risk for standard setting has generally been supported by the idea that even if we are exaggerating the risks, we can at least be sure that standards protect health.

Many forces are combining to drive evolution in the use of risk assessment. All involve actually making broader use of risk-based thinking. For example, the 103rd and 104th congressional sessions have seen the introduction of bills to require the use of benefit/cost analysis in environmental regulation, to require comparative risk assessment as a risk communication tool, and to require risk-based priority setting by regulatory agencies. Benefit/Cost legislation, as proposed, would require reporting and justification of the costs of a regulation in terms of its health and environmental benefits. Concern over public misperception of health risks generated proposed requirements that every environmental regulation be accompanied by a list of risks of similar size addressed by the agency and risks commonly encountered by citizens. In an effort to ensure that the worst problems are addressed first, legislation requiring regulatory agencies to rank the problems under their jurisdiction from highest to lowest risk has been proposed. A more strict proposal is to allocate resources to agencies, and programs within agencies, based on the size of the risks they address.

The role of risk assessment is expanding even without legislation. Many states and localities are conducting their own comparative risk projects, trying to rank different sources of health risk to residents. Companies are turning to risk assessment to guide pollution prevention efforts, making sure that chemical substitutions or process changes really reduce risk.

But, as more and more is asked of risk assessment, the practice must evolve. Risk assessors can no longer ignore basic scientific research about different mechanisms of cancer formation for different chemicals. Risk assessors can no longer be satisfied with worst-case estimates of risk. Comparing risks will require best estimates of risk using the best science. To do otherwise is to invite misleading and incorrect comparisons. When these comparisons provide the basis for efforts to protect human health, we cannot afford to deliberately delude ourselves with overly conservative estimates of chemical risks.

III. SOME SOURCES OF CONSERVATISM IN RISK ESTIMATES

There are many sources of conservatism in current carcinogen risk assessment. The following is a discussion of just a few of the most important sources in the three risk assessment steps: hazard identification, dose-response evaluation, and exposure assessment.

- Hazard identification is the process of determining whether an "agent" (defined loosely as an industrial chemical, a natural product in the environment, or a particular lifestyle) increases a person's risk of developing cancer.

- A dose-response relationship reveals how the likelihood of cancer changes with the amount of exposure. A risk assessor might estimate, for example, how the probability of lung cancer changes with the number of cigarettes smoked.

- The process of exposure assessment quantifies the amount (or dose) of the carcinogen to which people may be exposed. This may be the number of cigarettes a person smokes, the amount of a chemical in the air near a factory, the concentration of radon in the basement of a home, or the amounts of pesticide residues on various foods consumed per day by an individual.

After the quantitative aspects of a risk assessment have been determined, the numbers are combined to yield an overall estimate of risk in a risk characterization, the final step of risk assessment. This is usually a numerical characterization of the incremental lifetime risk of cancer due to a particular agent at a specific exposure level (*i.e.*, incremental risk). Current risk characterization suffers from many weaknesses but most important is the way numerical estimates of risk have been characterized. Procedures designed to develop upper bounds on risk are routinely treated as generating best estimates, and rarely are key assumptions and uncertainties in risk assessment fully acknowledged. The important role of choice of data and extrapolation model, for example, is rarely made clear. Thus, users of risk assessments cannot know the scientific plausibility of risk predictions.

A. Conservatism in Hazard Identification

The primary conservatism in hazard identification is the assumption that all chemicals found to increase tumors at some site in rodents have the potential to cause tumors in humans. In addition, chemical-induced decreases in tumor rates are ignored, animal tests which find no evidence of carcinogenicity are ignored if a positive result is available, and the hazard identification process hides important scientific information about the strength and consistency of tumor responses in animals.

The most direct way to determine if a compound can cause human cancer is through the science of epidemiology. Cancer epidemiology attempts to establish associations between human exposure to a suspected cancer-causing agent and the frequency of cancer in the human population (usually workers). The major drawback of relying on epidemiology to assess cancer risks is that we want to identify human carcinogens now, before they can be discovered by epidemiology.

From a technical perspective, cancer epidemiology is fraught with interpretive difficulty. Cancer is a disease with a long latency period that arises from many causes, only some of which are known. Human exposures to potential carcinogens are often complex, uncertain, and poorly documented. Moreover, epidemiological studies are often plagued by confounding factors (*e.g.,* smoking), a lack of suitable control groups, and alternative interpretations of data. Due to practical limitations on the size of studies, epidemiologists cannot usually detect modest cancer risks that might still be of social concern. Roughly 25 specific chemicals, including asbestos, benzene, and several chemotherapeutic drugs, have been identified as human carcinogens by epidemiology.[8]

In light of the limits of epidemiology and the need to identify hazards before they become problems, experiments are conducted with animals to identify agents that are potential human carcinogens. The true workhorse test of hazard identification is the long-term rodent bioassay, which is conducted on the assumption that a rodent carcinogen may also be a human carcinogen. In addition, short-term laboratory tests of the biological properties of chemicals provide information which may help inform judgments concerning potential for human carcinogenicity.

When making a judgment about whether a particular agent is likely to be a human carcinogen, EPA states that all available data concerning the potential carcinogenicity of the compound should be reviewed and considered in a weight-of-the-evidence approach. Critics of EPA argue that, in fact, agency scientists give undue emphasis to positive results from long-term bioassays and do not incorporate mechanistic data and negative test results, and hence EPA's method is not truly based on the weight of the evidence. In addition, EPA scientists focus on anatomical sites in which chemical treatment causes tumor rates to increase, ignoring the common phenomenon of *decreases* in tumor rates in some tissues. For example, dioxin (2,3,7,8-tetrachlorodibenzodioxin) caused a clear increase in liver and lung tumors in female rats. However, rates of mammary cancer and uterine cancer were lower in rats treated with dioxin. Overall, animals at the highest dose tested did not have higher tumor rates than untreated control animals, and those at lower doses actually had fewer tumors than controls. EPA's hazard identification,

[8] L. Tomatis, A. Aitio, J. Wilbourn, and L. Shuker, "Human Carcinogens So far Identified," *Japanese Journal of Cancer Research* (1989), pp. 80:795-807.

however, focused only on the sites where tumors increased.

Other tests, on single-cell organisms, isolated tissues, cells, or cellular components are also considered important in judging the carcinogenic hazard of a chemical. These are often called short-term tests. Most of these tests try to find out if the chemical can interact with and alter DNA: that is, is it mutagenic? It is widely believed by cancer researchers that mutagenic chemicals are a greater carcinogenic hazard than nonmutagens and that there may be no threshold—no absolutely safe level of exposure—to a mutagenic chemical. Nonmutagens, on the other hand, are generally believed to act through mechanisms which have strongly nonlinear dose response relationships or even have a threshold for their carcinogenic effect. Examples of tests use to identify mutagens include the Ames Test and the mouse micronucleus test, among a host of others.

According to EPA guidelines, in carcinogen classification (the main function of hazard identification) the evidence for carcinogenicity in humans and animals, both positive and negative, is considered. The judgments about these two types of evidence are then combined, and the agent is given a preliminary assignment to an EPA carcinogen classification. The preliminary assignment can then be changed, to a higher or lower likelihood of human carcinogenicity, on the basis of the results of short-term tests, metabolic, and toxicokinetic studies, although such changes rarely occur. EPA scientists then place the compound in one of five categories, listed below, accompanied by official verbal descriptions used in communicating carcinogenic hazard to the public.

- Group A - Carcinogenic to humans

- Group B - Probably carcinogenic to humans

- Group C - Possibly carcinogenic to humans

- Group D - Not classifiable as to human carcinogenicity

- Group E - Evidence of non-carcinogenicity for humans

Group B is further subdivided into categories B1 and B2, which differ by the availability of positive epidemiological data. B1 chemicals have limited human evidence supporting, although not establishing, a finding of carcinogenic hazard to humans while B2 (as well as C and D compounds) are classified solely on the basis of animal data due to inadequate human evidence.

This method fails to make adequate distinctions between chemicals. There is no place in this method for studies showing no effect and, in theory, one positive experiment has more weight than any number of negative experiments of equal quality. It has been argued that, for this reason, EPA actually uses a strength-of-the evidence, not weight-of-the-evidence approach.[9]

For the majority of compounds which have no epidemiological data, a finding of carcinogenicity in an animal test results in placement in category B (actually B2) or C. If an agent has a positive response, this is taken as indisputable evidence of human risk. Yet there are several known animal carcinogens which likely act via species-specific mechanisms and would not pose a cancer threat to humans.[10] There are also animal carcinogens thought to act by dose-specific mechanisms, causing tumors at the very high doses employed in animal tests but not at the doses encountered by humans.[11] The rodents in carcinogenicity tests are allowed to eat all they want and are usually quite obese compared to wild rodents. It is well known that obese animals have more tumors than their leaner counterparts, even in the absence of chemical treatment.[12] Overfed animals also may respond differently to chemical treatments.[13] The use of very high doses in rodent bioassays also casts doubt on the interpretation of these studies in the minds of many toxicologists. All of these factors together make EPA's assumption that any rodent carcinogen should be treated as a human carcinogen suspect. In fact, a recent study asked a random sample of members of the Society of Toxicology to agree or disagree with the statement, "If a scientific study produces evidence that a chemical causes cancer in animals, then we can be reasonably sure that the chemical will cause cancer in humans." Almost 60 percent of the responding toxicologists disagreed with this statement.[14]

In addition, EPA's system hides some important information about the response of test animals to chemicals. Some chemicals have a consistent response across tested sexes and species, but most do not. An analysis of the results of a number of rodent bioassays conducted on male and female mice and rats found that of all positive studies 28 percent are positive in all four sex/species (male and female mice, male and female rats), 16 percent are positive in three out of four, 35 percent are positive in two out of four (usually the same species) and 21 percent are only positive in one out of four. In general, EPA hazard identification does not distinguish between chemicals which give a consistent response across all tested animals and those which do not. Some have caused tumor increases in many anatomical sites in many different tests. Others cause tumors in only a single site. One example, vinylidene chloride, has been tested 18 times, in tests of varying quality, and only found positive once (and only in mice: it was negative in rats in the same experiment), yet it is considered a possible human carcinogen, and risk estimates

---

[9] J. Ashby, et al., "A Scheme for Classifying Carcinogens," *Regulatory Toxicology and Pharmacology* (1990), pp. 12:270–295.

[10] W.G. Flamm, and L. Lehman-McKeeman, "The Human Relevance of the Renal Tumor-Inducing Potential of d-limonene in Male Rats: Implications for Risk Assessment," *Regulatory Toxicology and Pharmacology* (1991), pp. 13:70–86.

[11] S.M. Cohen, and L.B. Ellwein, "Cell Proliferation in Carcinogenesis," *Science* (1990), pp. 249:1007–1011.

[12] J.D Thurman, et al., "Survival, Body Weight, and Spontaneous Neoplasms in Ad Libitum-Fed and Food-Restricted Fisher-344 Rats," *Toxicologic Pathology* (1994), pp. 22:1–9.

[13] M.W. Chou, et al., "Effect of Caloric Restriction on the Metabolic Activation of Xenobiotics," *Mutation Research* (1993), pp. 295:223–235.

[14] N. Kraus, T. Malmfors, and P. Slovic, "Intuitive Toxicology: Expert and Lay Judgments of Chemical Risks," *Risk Analysis* (1992), pp. 12:215–232.

for it are conducted in exactly the same way as those for a known human carcinogen like radon.

B. Conservatism in Dose-Response

Dose-response is the key source of conservatism in cancer risk estimates for many chemicals. Huge exaggeration of risk potential occurs when a dose-response model designed for certain types of chemicals is inappropriately applied to all chemicals.

Dose-response evaluation for carcinogens is different than traditional toxicology because it is typically assumed that there is no threshold dose for a carcinogenic response. In addition, it is assumed that the response at high doses (like in the rodent bioassay) is roughly proportional to the response at the much lower environmental levels to which people are exposed. This is perhaps the most important source of conservatism, at least for some chemicals.

To illustrate this assumption imagine a hypothetical experiment. First, we will give 10,000 tablets of a chemical to rats every day for their entire life time. Now assume that this treatment causes 50 out of 500 (10 percent) exposed animals to develop tumors when untreated animals had none. Roughly, the proportional dose-response assumption would generate the responses predicted in Table 1. In general, if our human exposure to the test chemical were like most environmental chemicals, we would be attempting to predict what would happen to people exposed to less than one tablet per day based on an experiment in which rats received 10,000 tablets per day. In any other type of toxicology, no attempt would be made to link very high dose responses to those at lower doses. We would never try to predict what would happen to a person taking one aspirin tablet based on the response of rats to 10,000 aspirin.

Table 1: Proportional Dose-Response in Laboratory Rats

Exposure (# tablets)	# of Rats	# Predicted w/Tumors
10,000	**500**	**50**
1000	500	5
100	500	0.5
10	500	0.05
1	500	0.005

The focus on linearity and the lack of a threshold in cancer dose-response arises from early research on radiation-induced cancer. Various types of ionizing radiation are known, from studies of atom bomb survivors and others, to increase tumor rates in people. Like animal studies, however, most data are from very high exposures to radiation, there are little or no data for typical levels of radiation exposure. However, the data that are available show little evidence of a response threshold.[15] In addition, ionizing radiation is generally considered to act by directly inducing mutations in DNA in a random fashion. More exposure is believed to increase the likelihood of an interaction with DNA, lower exposure to decrease the likelihood.

These two observations about radiation-induced cancer greatly influenced thinking about chemically induced tumors. When assessors began to look for ways to estimate cancer risk from chemicals, they seized upon the linear, no-threshold models from radiation. The implication—that there is no safe level of exposure to a carcinogenic agent because there is no threshold—is not widely accepted in the scientific community. In the study referred to above, the Society of Toxicology members were also asked to agree or disagree with the statement, "There is no safe level of exposure to a cancer-causing agent." This time more than 75 percent of the responding toxicologists *disagreed*.[16]

The linear, no-threshold approach may make sense for some chemical carcinogens. Some strongly mutagenic chemicals act very much like ionizing radiation in the body. For this group of chemicals these radiation-based models may be quite appropriate. There may be little conservatism due to dose-response modeling in risk estimates for these chemicals. However, for many other chemical carcinogens, especially nonmutagens, assuming radiation-like behavior makes no sense. These chemicals cause tumors not by interacting directly with DNA but by disrupting hormonal balances, causing cytotoxicity and compensatory cell proliferation, or interfering with cell-to-cell communication and control.[17] There is no reason to believe that these chemicals will exhibit low-dose linearity proportional to the responses at high doses. Many may even have thresholds. Risk estimates for these chemicals based on low-dose linear models may overstate risks by a factor of 100, 1000, or even more. In some cases, the overestimate is infinite because the best risk estimate is zero.

---

[15] Low-dose linearity is also often justified on the grounds that any effect which adds to a process already occurring in the body must exhibit low dose linearity. Since there are obviously cancer causing processes occurring in the absence of chemical exposure, carcinogens must add to this background. This justification for low dose linearity, however, does not justify proportionality with high-dose responses.

[16] Kraus, Malmfors, and Slovic, "Intuitive Toxicology," pp. 12:215–232.

[17] For a discussion of the different modes of action of nonmutagenic chemicals and one suggestion for their risk assessment see B.E. Butterworth, R.B. Conolly, and K.T. Morgan, "A Strategy for Establishing Mode of Action of Chemical Carcinogens as a Guide for Approaches to Risk Assessment," *Cancer Letters* (1995), pp. 93:129–146.

C. Conservatism in Exposure Assessment

Many standard risk assessments deliberately overstate people's exposure to environmental contaminants. This source of conservatism is rarely apparent when the resulting risk assessment is used. Conservatism in exposure assessment also applies to non-cancer risk assessment.

Exposure assessment is the phase of a risk assessment that determines just how much of a carcinogen people are actually exposed to. Exposure can occur through different routes, including inhalation, dermal absorption, and ingestion in food or water. In some situations, a pollution source will cause human exposure to occur through more than one exposure route (pathway), although multiple pathways are not always considered in EPA risk assessments. More recently, risk assessments have begun to account for as many sources and routes of exposure as possible.

In many risk assessments, exposure is not estimated for the average person but instead for the Maximally Exposed Individual (MEI). The MEI is the person, usually hypothetical, who is predicted to receive the greatest lifetime exposure from a particular source. For example, the MEI may be the resident living closest to a factory which emits the suspected carcinogen, or, in the case of groundwater contamination, the resident whose well for drinking water is next to a Superfund site which may be leaking one or more suspected carcinogens. Use of the hypothetical MEI to design standards is an extremely contentious issue. Critics of MEI-based standards argue that it is unsound to regulate, often at very great cost, on the basis of an inflated exposure scenario that never occurs.[18]

---

18 B.D. Goldstein, "The Maximally Exposed Individual: An Inappropriate Basis for Public Health Decision Making," *The*

It is generally recognized that the MEI calculation is conservative, giving an upper bound on the true lifetime exposure. The exposure of the MEI is usually calculated on the basis of predictive models rather than direct measurements. In the case of a resident at a factory fenceline, a mathematical dispersion model might estimate the air concentration of the carcinogen 200 meters from the source, considered to be the fenceline of the factory and the residence of the MEI. In addition, it is often assumed that the MEI is outdoors breathing air at this predicted concentration 24 hours a day for 70 years. These assumptions are used, though no one spends their entire life outdoors at the fenceline of the factory, and few factories produce the same products, or even exist, for 70 years. Although none of these assumptions is outrageous when considered alone, the combination of these factors leads to conservative exposure estimates.

It has been estimated that standard MEI calculations overestimate the exposure to a true maximally exposed individual by a factor of 10 to 100.[19] These estimates of exposure may be thousands of times larger than those faced by the average exposed individual.

Another example comes from pesticide risk assessment. When estimating exposure to pesticides for the general public EPA would like to know, in effect, the dose of pesticides "on the dinner plate." However, this type of data is rarely available, so exposure must be estimated.

There are three primary ways that EPA can estimate the public's exposure to pesticides. In order of increasing accuracy they are: 1) theoretical maximum residue concentrations (TMRC); 2) farm gate data; and 3) residue monitoring. Theoretical maximum residues are usually used by the agency in the absence of other data. This method assumes that every acre of a particular crop has the highest possible allowed level—the tolerance level—of the pesticide on it, and this level does not decrease with time, storage or cooking. This method gives an upper bound on exposure to the pesticide. This is the process used in most pesticide risk assessments.[20] Farm gate, or field data, use the results of experiments submitted by the pesticide manufacturer at the time of registration. These field experiments measure the levels of pesticide on a crop after it has been treated at the maximum allowable rate, and had the minimum required preharvest time interval. These levels, often referred to as farm gate levels because they are the levels that could be expected to be on the produce when it leaves the farm, may be adjusted with experimentally determined processing, washing, or cooking factors to give a more realistic estimate of consumer exposure. The final type of exposure estimate, residue monitoring, is based on measurements of pesticide residues for raw and processed produce at the grocery store, where the consumer comes into contact with it. Residue monitoring data reflect actual agricultural practices, such as different preharvest intervals, the effects of time and storage, and different pesticide application rates. Few farmers use pesticides at the highest allowable level. Although residue monitoring is expensive and sometimes difficult, it provides the most accurate estimate of consumer exposure to pesticide residues.

Environmental Forum (November/December 1989).

[19] N.C. Hawkins, "Conservatism in Maximally Exposed Individual (MEI) Predictive Exposure Assessments: A First-Cut Analysis," *Regulatory Toxicology and Pharmacology* (1991), pp. 14:107–117.

[20] National Research Council, *Regulating Pesticides in Food: The Delaney Paradox* (National Academy Press: Washington, D.C.).

Table 2: Variation in 3 Methods of Estimating Pesticide Residues on Celery		
	Chlorothalonil (ppm)	% of tolerance
TMRC	15.0	100.0
Field Data	4.07	27.1
Residue Monitoring	0.12	0.8

Source: Eilrich, "Tracking the Fate."

The degree of conservatism due to the TMRC assumption varies from crop to crop but, in general, leads to overestimates of exposure by a factor of 10 to 100. Table 2 shows the levels of the pesticide Chlorothalonil (Bravo®) found on celery.[21] Clearly, conservatism in exposure assessment can lead to risk estimates more than 100 times higher than risks based on more accurate exposure assessment.

IV. ACTUARIAL VERSUS PROBABILISTIC ESTIMATES OF RISK

When making risk comparisons, an important distinction must be kept in mind, that between *probabilistic* and actuarial risks. Most risk predictions that people see, outside of the environmental arena, are actuarial risks. They are based on predictions of the future behavior of a phenomenon about which much is known. For example, in making estimates of the number of Americans who will be killed in car accidents next year a researcher would look at data showing

[21] Gary L. Eilrich, "Tracking the Fate of Residues from the Farm Gate to the Table," *Pesticides and Food Safety*, ed. B.G. Tweedy, et al., American Chemical Society (Washington, D.C., 1991).

how many people had died in car accidents in past years. Then, making adjustments for trends (for example, auto death rates per vehicle mile traveled are declining) a prediction could be made for next year. Now, no one expects that the number will be exactly correct, the true toll could be higher or, as we hope, lower. In any case, this prediction can be made with a great deal of precision and confidence. Of course, this is an estimate of risk for an average person. It will be higher for those who drive more and lower for those who drive less. Nonetheless, it is virtually impossible to imagine that this risk estimate could be wrong by more than 50 percent, and it more likely to be within 10 to 20 percent. Many public health risk estimates are based on actuarial data. These include accident risks, injury risks, and disease rates.

Cancer risk assessment, on the other hand, estimates future danger from a combination of data, theory, and models. This means that cancer risk predictions are based on extrapolated probabilities, not on past frequencies. There are several reasons for this. For instance, there are very few chemicals known to cause cancer in people, even at very high doses, so tests on animals form the basis for estimating risk. We are often concerned with exposures far below those at which a cancer hazard was identified. Levels of cancer risk from environmental pollution are, even in the worst situations, so small that they cannot be detected by epidemiology. Therefore, models are used to estimate risk from available information. Because the causes of cancer are complex and often unknown, there is considerable uncertainty in risk models. As a consequence, predictions of cancer risk cannot be made with a high degree of precision. This means that there is a great deal of uncertainty in estimates of risk at Superfund sites, for example. Any prediction of the number of people who may be harmed will be quite imprecise. Because of assumptions made in the risk assessment, uncertainty about the true relationship between exposure and dose, and other factors, cancer risk estimates for an average individual, or estimates of population risk, may be wrong by factors of 100, 1000, or more. Of course, because of conservative procedures these errors are on the side of making risks look bigger.

V. COMPARING RISKS: THE EVOLUTION OF RISK MANAGEMENT

The use of conservative methods for estimating risk has been contentious from the start.[22] From the scientific perspective, the seeming inflexibility of the risk assessment process—refusing to acknowledge the difference between carcinogenic chemicals—has given the process a very bad reputation in the scientific community. At the same time, the uses of risk assessment are attracting more and more attention. The use of conservative risk estimates to set Superfund cleanup levels or water standards has given rise to charges that risk assessment is forcing the expenditure of enormous amounts of money for very little benefit. There is also concern that inflating estimates of environmental harm is misleading the public about the sources and sizes of health risks. Since public opinion has been found, by EPA itself, to play a large role in setting regulatory priorities, public health officials are concerned that too much attention is being placed on small risks to health while larger risks are ignored.

One solution to conservative risk assessment procedures is to require changes in the way risks are calculated and reported. During the 1990s, several attempts have surfaced to use legislation to reform the way in which risk assessments are conducted. Various bills would have required EPA to use "the best available science" in risk assessment and to provide best estimates of risk to accompany upper bounds. By early 1996, this legislation had gone nowhere. A primary reason is that almost any attempt to use the best science, or provide best estimates of risk, will result in the risk looking smaller than was thought with conservative methods of risk assessment—often significantly smaller. Rather than being perceived as common sense reform, the natural result of having used worst case risk assessment methods in the past, this has been portrayed as rolling back environmental protection, and risk-assessment reform has stalled. Legislation to limit risks is still necessary, although new comparative risk management strategies championed by Congress, EPA, the states and private industry may help drive improvements in risk assessment and spell the death of conservatism in the future. The question is whether the cost

[22] J. Cornfeld, "Carcinogenic Risk Assessment," *Science* (1977), pp. 198:693-699; R.A. Squire, "Ranking Animal Carcinogens: A Proposed Regulatory Approach," *Science* (1981), pp. 214:877-880.

of misleading comparisons and misplaced priorities due to conservative risk assessment is too high to wait for such gradual changes.

As we have seen, the old uses of risk assessment involved setting standards for food additives, pesticides, soil cleanup levels at Superfund sites, air pollutants, water pollutants, and drinking water contaminants. Notions of safety meant conservatism simply made standards even safer. The new risk management roles for risk assessment all involve comparisons, and many are closely related. For example, major efforts are underway in many cities, states, tribal areas and regions to compare risks to health from many different sources (often called comparative risk).[23] This information can then be used for risk-based priority setting and in risk comparisons designed to help the public understand the size of different risks to health. Comparison of benefits (in risk reduction) with costs is being widely advocated as a measure to ensure that health-protection resources are wisely allocated. Finally, substitution analysis is emerging as a concern as risk assessors realize that any risk-reducing effort, whether a regulation or a chemical substitution for pollution prevention, must be closely examined to ensure overall risk reduction. These new risk management techniques are entirely appropriate, and risk assessment is the right tool for the job. In each case—comparative risk, benefit/cost analysis, or studying substitution risks—the perils of conservative risk assessment loom large. We will see that only when science-based best estimates of risk are used can we be confident in making comparisons.

23 R.A. Minard, Jr., "Comparative Risk and the States," *Resources* (Winter 1996), pp. 6–10.

VI. HOW CONSERVATIVE RISK ASSESSMENT DISTORTS RISK COMPARISONS

A. Comparative Risk

What are the biggest risks to public health? Major efforts have been undertaken across the United States, by regulatory agencies, states, and cities to rank health risks. In every case the goal is a ranking of the risks posed by threats to human health and the environment. In the early days of the EPA there was generally agreement on the direction of efforts to protect human health and the environment. For example, it was clear that many waterways needed cleaning, and air pollution was a problem in many cities. As the environmental problems that could be identified with the eyes or the nose were fixed, and as EPA's mandate grew to include abandoned hazardous waste sites, industrial chemical releases and other issues, questions began to arise about the relative effort given to different environmental protection programs.

Under the leadership of Administrator Lee M. Thomas, EPA undertook one of the first comparative risk projects in the mid-1980s. Agency scientists and outside experts ranked sources of environmental and health risk and compared problems to the amount of agency attention they received. In a candid assessment, the project's report "*Unfinished Business: A Comparative Assessment of Environmental Problems*," stated that, "Overall, EPA's priorities appear more closely aligned with public opinion than with our estimated risks."[24]

In the last few years at least 20 states and numerous cities and localities have undertaken comparative risk projects looking at environmental risks. The EPA provides financial and technical support for many of these efforts. Most projects use both technical experts and members of the public in ranking risks. These projects usually result in several lists of risks, usually for both human health and ecological risks, divided into high, medium and low concern. In most cases, similar to *Unfinished Business*, rankings run counter to common wisdom.

Importantly, these rankings are based not only on the technically derived risk numbers but also on expert and citizen expressions of the value that is attached to a risk. For example, use of a chemical in the workplace and its release into the environment may pose similar risks to exposed people, but expert and citizen values may rank the environmental release as a greater risk since people will be exposed without their knowledge. It is important to stress, however, that any ranking must be informed by technical analysis and that ranking may be of interest to a decision maker as well. It has been suggested that broader risk comparisons, for example across all federal agencies responsible for protecting health, safety and the environment, would be even more useful than those which focus only on the environment.[25]

Many advocate the use of these risk rankings to set priorities for public and regulatory attention. The results of the *Unfinished Business* study provoked a great deal of concern in the public health community: if the nation was focusing attention and resources on smaller problems while bigger ones were ignored then we, as a nation, were not doing all we could to save lives. In fact, one could demonstrate that this was akin to "statistical murder" in the words of John D. Graham of the Harvard Center for Risk Analysis. If we instead attacked the biggest problems first, critics of the status quo suggested, we could do the most good with available resources. Various legislative proposals would have required EPA, for example, to rank the risks that it addresses and to tell Congress how it was going to allocate resources to address those risks. Others have gone even further, suggesting that resource allocation by Congress should be directly tied to agency's risk rankings.

The other important notion that came from *Unfinished Business* was that public misperceptions about sources of

[24] U.S. EPA, *Unfinished Business: A Comparative Assessment of Environmental Problems* (Office of Policy Analysis and Office of Policy, Planning and Evaluation:1987), p. XV.

[25] Harvard Group on Risk Management Reform. "Reform of Risk Regulation: Achieving More Protection at Less Cost," from the Harvard Center for Risk Analysis (1995).

health risk were driving America's environmental program. As studies were conducted it was found, as in Figure 1, that the public believed that "things in the environment" played a major role in damaging health.[26] Yet scientists were finding that environmental contaminants play a very minor role in human diseases such as cancer (Table 3) and that factors such as a poor diet and lack of exercise were much more serious threats. To address this risk communication problem it has been suggested, and legislation has in fact proposed, that environmental regulations be accompanied by comparison of the size of the risk being addressed to other risks with which the public is familiar. These might be risks addressed by the same agency or more general types of risks. For example, a drinking water standard to protect the public from cancer risk from a chemical might be accompanied by a comparison of the risk of that chemical compared to the cancer risk of a pesticide residue on food or the mortality risk from particulate air pollution. More broadly, the risk might be compared to other risks of death from natural disasters, accidents, a poor diet, or diseases. The goal of these suggestions is to communicate to the public the relative size of different risks to their health.

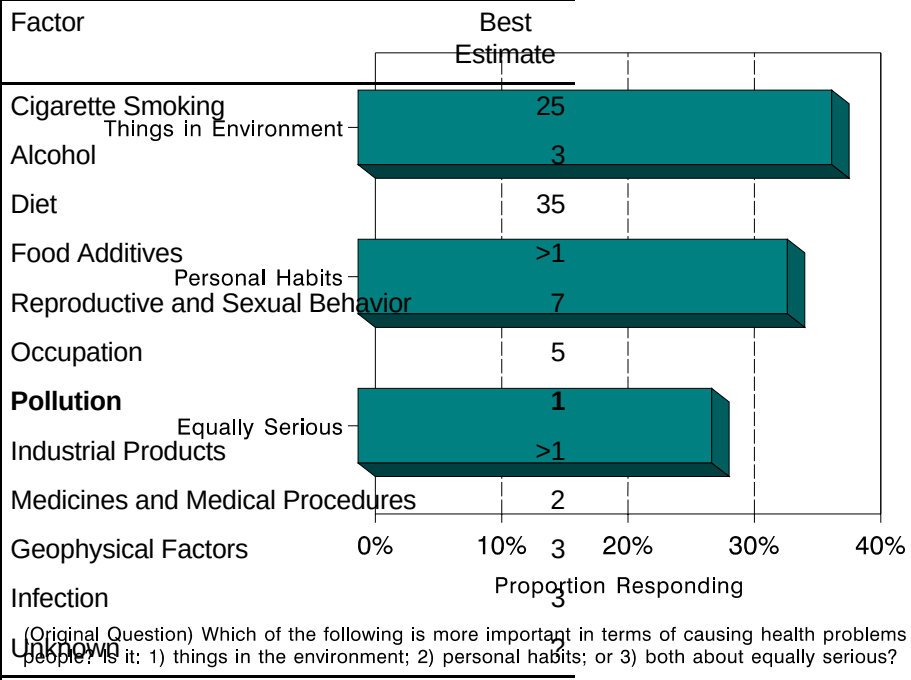
26 Roper Organization, *Roper Reports*, 90-3 (New York: The Roper Organization, 1990).

Table 3: Proportion of Cancer Deaths Attributable to Different Factors

Factor	Best Estimate
Cigarette Smoking	25
Alcohol	3
Diet	35
Food Additives	>1
Reproductive and Sexual Behavior	7
Occupation	5
Pollution	**1**
Industrial Products	>1
Medicines and Medical Procedures	2
Geophysical Factors	3
Infection	3
Unknown	9

R. Doll, and R. Peto, *The Causes of Cancer* (Oxford University Press: Oxford, 1981).

Figure 1: Which Causes More Health Problems in People?



(Original Question) Which of the following is more important in terms of causing health problems in people? Is it: 1) things in the environment; 2) personal habits; or 3) both about equally serious?

Risk-based priority setting is simply an extension of comparative risk. Given limited resources, which types of health or environmental threats deserve the most attention? Can our resources be spent in other ways that might do more for public health without costing more? Recent efforts to stimulate systematic evaluation of public health risks have identified many situations in which tiny, often hypothetical, risks receive a great deal of attention and bigger risks are ignored.[27] For example, the hypothetical, and even if true, tiny, risks of pesticide residues on food garner a much larger share of regulatory resources and effort than the much larger risks to farmers and farm workers using agricultural chemicals. The goal of risk-based priority setting is to use the valuable information gained from comparative risk studies to focus attention and resources where they can do the most to protect human health and the environment.

Legislation has been proposed at both the state and national levels to require regulatory agencies to rank the risks that they address. The goal is to focus attention on the largest opportunities for reducing risks and to make these actions a priority. Going further, some proposals would require agencies to actually make resource allocation decisions to address the worst problems first.

The perils of conservatism for comparative risk are quite obvious. If estimates of risk for substances suspected of causing cancer are systematically inflated then other risks will suffer in comparison. Risks of accidents or diseases are best estimates based on actuarial risks. Is it fair to use worst-case estimates of environmental harm when comparing and ranking these risks? Risk-based priority setting, even within an agency like EPA, suffers from the problem of differential conservatism in risk estimates for different substances. The prospect of different programs within an agency like EPA striving to justify their efforts in relation to other programs illustrates why risk ranking will drive conservatism from risk assessment. Imagine the people responsible for the indoor radon program attacking the rankings of the pesticide program because the pesticide risk assessments are based on inflated exposure estimates and inappropriate dose-response models. Institutional competition could, in effect, drive risk assessments toward best estimates of risk.

B. Substitution Risks

Lynn Goldman, Assistant Administrator for Prevention, Pesticides, and Toxic Substances at EPA, has been quoted saying, “Before I make a decision [to phase out a pesticide], I’d like to know what are the substitutes...and are those

[27] Much of this discussion is based upon the report of the Harvard Group on Risk Management Reform called *Reform of Risk Regulation: Achieving More Protection at Less Cost* from the Harvard Center for Risk Analysis (1995).

substitutes going to be more or less risky than what I'm phasing out?"[28] This is an example of substitution analysis, a growing use of risk assessment. Risk managers are coming to understand that actions to reduce one risk may cause another type of risk to appear, may shift risk from one population to another, or may change the form of the risk.[29] Concern that efforts to reduce risk may be hampered by unacknowledged substitute risks has prompted federal agencies, private industry, and even Congress to look closely at this problem. Several recent regulatory reform bills in Congress even contained language requiring substitution analysis for major regulations.

As an example of substitution analysis, let's consider the pesticide regulation problem faced by Goldman. An EPA decision to ban a pesticide could give rise to risks from the substitute pesticide (banning the pesticide doesn't make the pest go away) including transfer of risks from consumers to farm workers or a change from chronic cancer risk to the risk of neurotoxicity, from the pest itself, and even from increased prices for food.[30] Clearly, risk assessment can play a key role in evaluating the health and environmental outcome of a regulatory decision.

---

28 *Toxic Materials News* (December 6, 1993), pp. 411–413.

29 For a discussion of "risk tradeoffs" see *Risk vs. Risk: Tradeoffs in Protecting Health and the Environment*, ed. J.D. Graham, and J.B. Weiner, (Harvard University Press: Cambridge, MA, 1995).

30 For a more detailed discussion of pesticide risk tradeoffs see G.M. Gray, and J.D. Graham, "Regulating Pesticides," *Risk vs. Risk: Tradeoffs in Protecting Health and the Environment*, ed. J.D. Graham, and J.B. Weiner, (Harvard University Press: Cambridge, MA, 1995).

For private industry, a greater emphasis is being placed on pollution prevention, rather than end-of-pipe controls to reduce human and environmental health risks. A key tool for pollution prevention is the substitution of one chemical used in a process for another. Currently, many companies focus on simply reducing their use of chemicals on regulatory lists like the Toxics Release Inventory. Some are becoming concerned that this list-based decision-making process may be doing little to reduce risk.[31] There are many cases in which firms blindly made substitutions only to later find that the replacement chemical was more dangerous. For example, even though a replacement chemical was less toxic much more might be needed to achieve production requirements. Or when the new chemical was thoroughly tested, it too was might be considered a hazard. The tools of risk assessment are crucial to ensuring that these chemical substitution decisions truly reduce risk.

The perils of conservatism for substitution arise from both hazard identification and dose-response evaluation. Let's examine two examples from pesticide regulation.

A number of important regulatory implications result from an agricultural chemical being deemed a suspect carcinogen by the EPA. It may become subject to the Delaney Clause of the Food, Drug, and Cosmetic Act which bans the use of any carcinogenic pesticide under certain conditions. In general, the chemical may have a more difficult time being registered, as standard conservative risk assessment methods may lead to unacceptable risks at the lowest levels of use at which it is still effective. It certainly will also attract more attention from environmental and anti-pesticide groups. Remember, according to EPA, a finding of increased tumor rates in any test makes a chemical a suspect carcinogen. Some pesticides are clearly carcinogenic to animals, increasing tumor rates in many different tests. Others are not so clearly carcinogenic. For example, permethrin is an insecticide that is considered by EPA to be a possible human carcinogen. It has been the subject of six long-term studies and has proven positive in only one, increasing the rates of lung tumors in female mice. It shows no evidence of mutagenic potential and so cannot directly alter DNA. Finally, it is chemically related to the pyrethroid insecticides, isolated from chrysanthemums, that are used on organic farms. There is a big difference in the evidence for carcinogenicity of permethrin and other clearly carcinogenic agents, but conservative methods mean this information can play little role in hazard identification.

Substitution analysis for pesticides requires relevant science to be used, not ignored. Yet all too often standard conservative methods of risk assessment are applied to all chemicals whether appropriate or not. Consider captan and maneb, two fungicides currently used on apples. According to EPA analyses, both pose about the same lifetime cancer risk from residues left on apples, about two in a million increased risk. If one were to be banned, it is likely that the other would take its place. Should we care which one increases in use? Captan is a mutagenic compound that has increased the rates of rare intestinal tumors in several rodent studies. In addition, there are concerns about potential teratogenicity (birth defect) risks with captan. Maneb is a member of a family of fungicides indicted for carcinogenicity based on a breakdown product, ethylene thiourea (ETU). There is no evidence for mutagenicity of maneb. In addition, many scientists believe that the tumors caused by ETU are due to disruption of the thyroid hormone system that occur at very high doses. At lower doses there is no evidence of disruption, and tumors would not be expected. This makes maneb a very poor candidate for the standard conservative no-threshold, linear-at-low-doses dose-response model used by EPA. It is not mutagenic or a direct acting carcinogen. Captan may well be. This means that the risk estimates for captan have some scientific plausibility while those for maneb have little. The true risk of maneb residues on apples may well be zero. This is a case where uniformly applying a conservative risk assessment approach would not buy us safety, because we would be misled about the relative risk of two pesticides. Ignoring the science could lead us to make a bad decision, banning maneb and substituting captan, that might increase real risk while getting rid of a substance that posed no risk at all.

These questions of hazard identification and dose-response evaluation plague every attempt to do substitution

[31] See G.M. Gray, and J.K. Hartwell, "The Chemical Substitution Tree: Using Risk Analysis to Understand the Relative Risks Posed by Alternative Chemicals," *Pollution Prevention Review* (Spring 1995), pp. 7–17.

analysis based only on conservative risk assessment methods. We need science based best estimates of risk, or substitution analysis may be rendered useless, or even harmful, by misleading comparisons.

C. Benefit/Cost Analysis

Environmental regulations are designed to reduce risk: cancer risk to people, risk of other diseases, risks to wildlife habitat or to ecosystems. Achieving these benefits costs money. The tools of benefit/cost analysis were developed to help examine the size of benefits in comparison to costs. Concern that environmental regulations, often based on conservative risk assessments, have cost a great deal with relatively small benefits has led many to call for closer examination of benefits and costs. As an example, Table 4 lists the average cost per life year saved for different federal agencies charged with reducing risks. Saving lives is not the only goal of regulation at EPA, but it is not the only goal of the other agencies either. The question is, are the corollary benefits achieved by EPA big enough to explain the huge disparity in resource allocation?

Table 4: Median Cost per Life Year Saved

Federal Aviation Administration	$23,000
Consumer Products Safety Commission	$68,000
National Highway Transportation Safety Administration	$78,000
Occupational Safety and Health Administration	$88,000
Environmental Protection Agency	$7,600,000

Source: T.O. Tengs, *et al.*, "Five Hundred Life-Saving Interventions and Their Cost-Effectiveness," *Risk Analysis,* Vol. 15, pp. 369–390, 1995.

The use of benefit/cost analysis in environmental regulation has been advocated for many years. There has been an executive order requiring benefit/cost analysis of all regulation from every president since Jimmy Carter. Many believe that too often the provisions of these executive orders are ignored. Legislation requiring benefit/cost analysis for environmental regulations has been proposed by the 103rd and 104th Congresses. In spite of the bipartisan history of presidential support, congressional efforts have stalled. Few doubt, however, that the philosophy of benefit/cost analysis will not come to play a larger and larger role in regulation in the future.

Conservative risk assessment distorts benefit/cost analysis. Deliberately inflated estimates of risk make the benefits of any action look bigger. Remember, most of the lifesaving interventions by EPA in Table 4 were based on conservative risk estimates. Science-based best estimates of risk would make the cost per life year saved even higher. As benefit/cost analysis is applied to more regulations and programs across the government, all programs must use science-based best estimates to predict benefits.

VII. CONCLUSION

Risk comparisons have great promise for improving the health of Americans. Ranking of risks and risk-based priority setting can ensure that public health resources are used to protect the greatest number of people. Risk comparisons can help create an informed citizenry, helping people identify the actions they can take to safeguard their health. Substitution analysis will make sure that risk-reducing efforts are carefully planned and account for potential creation of risks, transformation of risks, or transfer of risks. Benefit/cost analysis can make sure that we are getting the best results for health protection investments. Resources that are squandered cannot be used for other valuable health-protective efforts.

Risk comparisons are distorted by conservative risk assessment methods. Important scientific information is ignored. Consequent differences between chemicals are papered over with a one-size-fits-all-approach. Key assumptions and choices differ between risk assessments, yet their influence is not clear to anyone comparing risks. The solution to these problems is a science-based procedure that acknowledges the uncertainties present in risk assessment. The results of this procedure must be well characterized so important information is communicated to users of the assessment. At the same time, those responsible for risk comparisons must become better consumers of risk information, asking hard questions about the reliability and precision of estimates.

A. Improve Risk Characterization

The key to making better use of risk assessment for risk comparisons is abandoning conservative procedures and making better use of science in estimating risk. Risk assessment is a valuable tool; but one that is subject to significant scientific uncertainty. Consumers of risk assessments, especially those comparing risks, must have knowledge of the scientific plausibility of different estimates of risk. We must improve risk characterization.[32]

Improved risk characterization means presenting risk estimates characterized by alternative assumptions and

[32] G.M. Gray, *Complete Risk Characterization*, Harvard Center for Risk Analysis Risk in Perspective, 2(4), (November 1994).

methods. But all estimates are not equal. We must make use of scientists and the range of expertise and data they possess in assessing risks. An example of a risk assessment that relies on science, rather than conservative assumptions, has recently been published.[33] The result is not a single estimate of risk, as conservative methods encourage, but a range of risk estimates based on different data and assumptions but weighted by plausibility judged by scientists. This reflects the uncertainty inherent in any attempt to estimate cancer risk from environmental exposures.

Better risk characterization is difficult but it will have several benefits. It should lead to a better appreciation of the strengths and limitations of the risk assessment process for informing risk comparisons. It will contribute to the scientific credibility of the risk assessment process as scientists see more of their data used in risk estimates. Finally, it will improve our ability to compare risks by taking the invisible (but heavy!) thumb off the scale of environmental carcinogens when risks are ranked.

B. Be Critical Consumers of Risk Information

Anyone comparing risks, as part of a state comparative risk project, a company examining substitute chemicals, or an analyst doing a benefit/cost study, should ask a series of questions about risk estimates:

[33] J.S. Evans, et al., "Use of Probabilistic Expert Judgment in Uncertainty Analysis of Carcinogenic Potency," *Regulatory*, Toxicology and Pharmacology (1994), pp. 20:15–36.

- Are they based on actuarial data or probabilistic modeling?
- How certain are we of the effect?
- How much uncertainty is there in the predicted size of the risk?
- Is this a best estimate of risk or a worst-case upper bound estimate?

Acknowledging uncertainty in risk estimates does not make them useless. In fact, it increases our confidence in risk comparisons because important factors and choices are not hidden. Admitting uncertainty is not the same as admitting ignorance. The risk assessor we should be most wary of is the one who claims to predict risk with certainty.

C. Making it Happen

Wider use of risk assessment and risk ranking will help defeat conservatism. As more people are using risk assessment and competing interests are closely monitoring ranking exercises it will become evident that conservatism distorts the process. Better understanding of risk assessment by the public, legislators, and journalists will lead to hard questions and better assessments. We cannot, however, count on this process alone. The cost, in terms of wasted resources and misplaced priorities, is just too high.

Comparing risks means that methods which deliberately inflate, to different degrees, the risks posed by environmental chemicals must not endure. In order to give regulators a clear ranking of health risks, to help citizens understand sources of risks to their health, for industry to make real risk-reducing pollution prevention decisions, and to ensure that the costs and benefits of regulations are fairly evaluated, carcinogen risk assessment must abandon methods designed to deliberately overestimate risk. In order to ensure fair comparisons strong action is needed.

- Legislation to require the development of best estimates of risk using the best available science is necessary. Risk assessors must be required to report a central estimate of risk, perhaps accompanied by upper and lower bounds.

- The opportunities and incentives for regulatory agencies to "game" risk ranking and comparison exercises means that a method of oversight is necessary. An excellent model was recommended by the Harvard Group on Risk Management Reform which suggested the establishment of an oversight function in the Office of Science and Technology Policy, under the president's science advisor.[34] This group would have the scientific knowledge to evaluate risk assessments from across the government and could present a forum for challenging the scientific bases of a risk assessment.

Risk assessment is no longer simply an exercise of interest to technocrats and regulated companies. Increasing use of risk assessment for setting priorities, communicating with the public, reducing pollution and evaluating regulatory benefits and costs will bring needed scrutiny. Concerns about the conduct of risk assessment have already led to several legislative proposals to improve the analyses. Risk assessors must be required to develop best estimates of risk using the best available science and fully acknowledging the uncertainty in the analysis. Third party oversight will ensure that no one is making selective use of data or applying inappropriate methods in evaluating risks. Risk assessment is a valuable tool that can help make the most of our resources to protect public health through risk comparisons, but fair comparisons require that conservatism in risk assessment be eliminated.

ABOUT THE AUTHOR

George M. Gray is Deputy Director and Regulatory Toxicologist of the Harvard Center for Risk Analysis. He holds a B.S. in Biology from the University of Michigan, and M.S. and Ph.D. degrees in Toxicology from the University of Rochester. Dr. Gray's research focuses on improving the scientific basis of risk assessment and the use of risk analysis to evaluate pollution prevention options. His policy work advocates complete risk characterization to improve decision making. Dr. Gray has worked with many companies and trade groups as well as federal and local government groups to increase understanding of the size and sources of health and environmental risks.

[34] Harvard Group on Risk Management Reform, *Reform of Risk Regulation: Achieving More Protection at Less Cost* from the Harvard Center for Risk Analysis, (1995).