

When Risk Assessment Came to Washington: A Look Back

Joseph V. Rodricks¹

Dose-Response:
An International Journal
January-March 2019:1-15
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1559325818824934
journals.sagepub.com/home/dos



Abstract

Federal regulatory agencies had, by the 1970s, been charged with enforcing a host of new laws requiring that they establish controls on human exposures to chemicals necessary to protect health. The agencies relied upon a methodology introduced in the 1950s to identify safe levels of exposure to chemicals known to display toxicity. During the 2 decades prior to the 1970s, federal authorities had come to treat carcinogens as distinct from other toxic agents, and to regard them as unsafe at any level of exposure, and no systematic methods had been developed to deal with the rapidly increasing numbers of carcinogens. Beginning in the mid-1970s, some scientists and policy makers in regulatory agencies, including the present author, began to propose adopting emerging quantitative methods to evaluate the risks of carcinogens and introduced new notions of safety based on explicit consideration of risk. Quantitative risk assessment rose to prominence in the decade reviewed in this article (1974-1984) and began to replace the unsystematic approaches that provided no view of how well health would be protected under various regulatory controls. This article offers the author's recollections of that important decade.

Keywords

carcinogens, regulation, introduction of risk assessment, early controversies, red book

I spent a good part of the period from 1974 to 1984 attempting to understand chemical carcinogenicity, the quantification of carcinogenic risk and its scientific underpinnings, and the value of risk assessment for public health and regulatory decision-making. Although scientific and policy debates pertaining to these matters continue today, they were of a somewhat different kind during the decade in which risk assessment, as it is currently defined, was beginning to achieve prominence. The debates that were central to the emergence of risk assessment during that critical period, and the ways in which some of them were resolved can, I believe, reveal much about how and why risk assessment came to be the force it is today and about why certain difficulties pertaining to its conduct and uses persist. The editors of this journal have invited me to offer my recollections and perceptions of that somewhat combative but nevertheless highly productive period, and I hope what I present here will contribute to an understanding of how the present world of risk assessment and risk-based decision-making came to be.

At the beginning of the decade about which I write, quantitative risk assessment had no role in regulation or other public health efforts. When the decade ended, it had become central to many programs at the federal and state levels. The decade also ended with the publication of the famous National Academies

report: "Risk Assessment in the Federal Government: Managing the Process."¹ The Red Book, as it came to be called, offered a clear path forward for risk-based decision-making, the importance of which is still underappreciated today.

A Fungal Metabolite and An Endocrine Disruptor

For the first 6 years of the decade I shall discuss, I was employed as a scientist at the US Food and Drug Administration (FDA). I had entered the agency in 1965 as a lab scientist, assigned to study the important fungal metabolite and food contaminant known as aflatoxin (actually a group of closely related compounds). Gerald Wogan, at the Massachusetts Institute of Technology (MIT), and other scientists in England had demonstrated the animal carcinogenicity of aflatoxin and had

¹ Ramboll, Arlington, VA, USA

Received 29 August 2018; accepted 30 August 2018

Corresponding Author:

Joseph V. Rodricks, Ramboll, 4350 North Fairfax Drive, Suite 300, Arlington, VA 22203, USA.

Email: jrodricks@ramboll.com



shown that one of these substances, known as B1, was capable of producing malignancies at doses lower than those at which any other known carcinogen displayed similar activity. (Even today, aflatoxin is surpassed in this respect only by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. It is interesting to compare and contrast these 2 very potent animal carcinogens, which exhibit substantially different kinetic and dynamic behaviors.)

By 1970, aflatoxins were found to be common contaminants of certain human foods, their carcinogenicity had been reproduced in several animal species, and suggestive evidence of carcinogenicity emerged from epidemiology studies in certain populations experiencing relatively large exposure through foods. The FDA had placed limits on the allowable levels of aflatoxins in foods, based on what were considered to be the limits of analytical detection—initially 30 µg/kg and reduced to 20 µg/kg in the early 1970s as analytical methods improved. Although I was involved in implementing the regulatory scheme, it struck me as an odd approach. I understood that the FDA was required to ensure the safety of food, but it was obvious that analytical detection limits were not a measure of safety. Moreover, detection limits continued to decline and reached 2 µg/kg by 1972. Strict application of this approach to regulation would result in loss of very large fractions of some affected crops, most especially peanuts.²

As I began to become more involved in issues related to the uses of science in regulation, I was asked to assist with another compound, one used in food production since about 1950, that came under intense scrutiny in the early 1970s. The compound was diethylstilbestrol (DES), a synthetic compound with estrogenic properties approved for use in women who were unable to maintain pregnancy because of natural estrogen deficiency. In 1970, DES was identified as the highly probable cause of vaginal adenocarcinomas in young women whose mothers had used DES during pregnancy.³ The Food and Drug Administration immediately acted to prohibit the use of DES in human medicine. As the evidence for DES-induced carcinogenicity emerged, it ignited new concerns regarding the widespread use of DES as a growth-promoting agent in animals used for human food. This use, approved in 1949, was known to result in low levels of DES residues in meat; some evidence of the carcinogenicity of DES in animals had been developed in the 1960s, but only after the human findings emerged in the early 1970s did the FDA act against the animal drug uses. The Food and Drug Administration's actions in this matter were hampered by the fact that the applicable food law actually permitted the use of drugs such as DES if that use could be shown to result in "no residue" of the drug in human food. Here, as in the case of aflatoxin, we encounter the issue of analytical detection limits as the determinant of the amount of carcinogen permitted to be present in food. Although the law dictates that there should be "no residue," the only way to determine compliance with such a standard is to sample food and apply an analytical method to determine whether a residue of the carcinogen can be detected. Every analytical method has, of course, a detection limit, so that, even if "no residue" is found, it can be concluded only that there is no residue greater than the limit of detection (LOD) of

whatever method of analyses is used. The LOD was a de facto safety standard.⁴

In the case of drugs used in food-producing animals, FDA is required to approve the analytical method to be used and, therefore, must specify the LOD to be achieved. As I had observed in the case of aflatoxin, safety was in the hands of analytical chemists. Food residues of a highly potent carcinogen might be permitted at higher levels than those of a far less-potent carcinogen, simply because the LOD for the highly potent compound turned out to be at a higher concentration than the LOD for the less potent one. This made no sense and the LOD approach to regulation could not be claimed to satisfy any criteria for safety.

In the late summer of 1974, I had occasion to meet with Leo Friedman, director of the FDA's toxicology division. I raised with Friedman my concerns about reliance upon analytical detection limit as the basis for regulations. This approach made some sense for aflatoxins because they were food contaminants and could not be readily controlled in the way an intentionally introduced substance such as DES could be controlled. But in neither case did possible risks to human health seem to enter the FDA approach to regulation. I had laid out this issue in a 3-page memo, which I delivered to Friedman.

Although Friedman was a scientist, he started our conversation by reminding me that the FDA's activities were governed by the Federal Food, Drug, and Cosmetic Act and that certain amendments introduced in 1958 that pertained to food additives contained the requirement that no substance could be legally introduced into food unless it was shown to be safe under its conditions of use. The language of the law then went on⁵:

Provided, that no additive shall be deemed to be safe if it is found to induce cancer when ingested by man or animal . . . ,

This is named the Delaney Clause, introduced by Representative James Delaney of New York, chairman beginning in 1950 of the *House Select Committee to Investigate the Use of Chemicals in Food Products*. Although the Committee heard from many experts before adopting the new food additives amendment in 1958, perhaps the most significant testimony came from Arthur Fleming, then Secretary of the Department of Health, Education, and Welfare (now Health and Human Services). Fleming offered the following statement from a National Cancer Institute report:

No one at this time can tell how much or how little of a carcinogen would be required to produce cancer in any human being, or how long it would take for cancer to develop.

This "no safe level" concept, although directed at the time at substances intentionally added to food, had much broader influence in the world of chemical regulation.

I nevertheless persisted in my argument that reliance upon analytical detection limits for regulation, without considering risks to human health, seemed an inadequate approach to

decision-making, especially in light of the powerful dictate of the Delaney Clause. The latter may not be legally applicable to a food contaminant such as aflatoxin, but it seemed clear that it should be applicable to an intentionally introduced substance such as DES.

Leo Friedman told me that my 3-page memo echoed some ideas he had been thinking about for several years. He told me that the method for establishing safe levels for most chemicals, devised by his predecessor at the FDA, Arnold Lehman, and another FDA toxicologist, O. Garth Fitzhugh, was based on a widely accepted view, held by most toxicologists, that the toxic properties of most chemicals expressed themselves only after a threshold dose was exceeded and that concept provided a scientific basis for establishing safe levels for humans. The Lehman–Fitzhugh approach, published in the 1949 to 1955 period, relied upon the application of what were then called “safety factors” to data obtained from experimental toxicology studies and, in some cases, epidemiology studies.⁶ But Friedman was aware that there was a community of experts working in the areas of chemical and radiation carcinogenesis who had developed quite different views of the biological actions of agents having carcinogenic properties and that their notions of thresholds, reversibility, and dose–response were radically different from those of the traditional toxicologists.⁷

Leo Friedman pulled from his file a thin folder containing half a dozen publications he asked me to study. He felt it was time for the FDA to find a more scientifically satisfactory way to deal with carcinogens. I promised to return after I had gone through those few papers.

“Virtually Safe Doses”

I studied the papers Leo Friedman had given me but never discussed them with him. Leo died of heart failure not long after our meeting, at age 52. He was one of the most thoughtful and creative people I have known in the field of toxicology, and his early death was a significant loss to our community.

The single paper in the Friedman collection that I found most valuable was one published in 1961 by the prominent National Cancer Institute (NCI) biostatistician, Nathan Mantel, and an associate of his, William Bryan. I arranged through one of my office directors for Mantel to present his work at a seminar at our FDA offices, and he agreed to do that. Mantel demonstrated how carcinogenicity dose–response data on a series of polycyclic aromatic hydrocarbons could be described with a simple probit model, and how the tail of the model could be extrapolated downward from the low end of the observed dose–risk relationship (which typically describes lifetime risks of tumor development no less than about 1 in 10), to estimate doses corresponding to some extremely low and completely unmeasurable risks. Mantel’s risk target was 1 in 100 million, and the calculated doses corresponding to that excess lifetime risk he labeled “virtually safe.” I should add that Mantel, in extending the tail of the probit model, imposed an artificial slope on it that he thought would place an “upper bound” on the low-dose risk, so that risk would not be underestimated; it

was intended to impose a “conservative” element into the procedure for extrapolation into the unknown.⁸

It seemed to me that the approach described by Mantel provided a way to approximate the magnitude of the health risks carcinogens might pose at low (human) doses, and to set standards based on the notion that once risks reached some very low levels, we could declare that exposures in these ranges and below were not a health threat. Decision-making would not be deterministic (“safe/not safe”) but rather probabilistic (ie, risk-based).

It would also become possible to gauge the magnitude of risk reductions achieved as regulatory standards were tightened, so that policy makers could examine the important question of whether the public health benefits achieved by the imposition of various control technologies were worth the economic and social costs of achieving them. It appeared that a systematic means for dealing with the increasing numbers of carcinogens that could be found in the environment was available and that it should be developed for practical application. Decision-making would be linked to the risk characteristics specific to individual carcinogens, together with other factors that dictate the practical limits of risk reduction technologies.

In late 1974, I spent several months scrutinizing many animal studies then available on the carcinogenicity of aflatoxins and exploring, together with 2 FDA biostatisticians, the implications of applying the Mantel–Bryan procedure for low-dose extrapolation; we applied other statistical procedures as well. We estimated how human intakes of aflatoxin would decline if regulatory tolerance levels were to be made more restrictive, and we then estimated corresponding risk reductions. I worked with FDA policy officials to begin to craft a risk-based tolerance level for aflatoxins in peanut products.⁹

Because of my work with aflatoxins, I was asked to join another agency effort to move toward risk-based decisions. This effort concerned not food contaminants such as aflatoxins, but rather the class of intentionally added substances represented by DES. The uses of these veterinary drugs would result in their presence as “residues” in meat, milk, or eggs. Because they were intentionally added substances, the original form of the Delaney Amendment applied to any such drug that was found to be carcinogenic. During the 1960s, however, our Congress modified the law to allow the use in food production of animal drugs that were carcinogenic. That modification permitted such use if, as I noted earlier, “no residue” of the drug could be detected in human food. This modification of the law came to be known as the “DES Proviso.”¹⁰

It turns out that at this time, the FDA was blessed with a very astute general counsel who had great foresight, Peter Barton Hutt. Hutt had come to lead the effort to put risk assessment into the regulatory equation for this class of added food ingredients. After extensive discussions with me and other scientists, Hutt proposed that “safe doses” for carcinogens such as DES could be defined as those associated with lifetime risk levels of less than 1 in 1 million, when these risks were estimated using a linear, no-threshold model (several publications had demonstrated that the Mantel–Bryan approach could not be counted

on to place an upper bound on risk at low doses but that a linear, no threshold model could—see later). Carcinogenic animal drugs would be acceptable only if it could be shown that their uses led to food residues at levels no greater than the level that corresponded to the “safe dose,” as defined in the foregoing. Those seeking drug approval would be required to develop analytical methods capable of reliably detecting the safe residue level and demonstrating that “no residue” could be found in food, under the intended conditions of drug use, when that analytical method was applied. The “Sensitivity of the Method” regulation became the first formally to adopt a risk-based approach for carcinogens.¹¹

The FDA, after a protracted administrative hearing, acted to extinguish the veterinary uses of DES in 1979, based in part on the fact that levels of residues detected in food did not meet these new risk-based safety criteria.⁴

Although it was estimated in a somewhat different way than that proposed by Mantel and Bryan, their “virtually safe dose” became Peter Hutt’s “safe dose” (FDA, he used to say, did not permit doses for added carcinogens that were only “virtually” safe). The selection of a lifetime risk level considered sufficiently low to ensure safety was not a scientific, but rather a policy choice. Thus, emerged what later came to be labeled a risk management decision, distinguishable from risk assessment.

The EPA

I do not recall how I came to be contacted by the EPA, but I believe publication of the “Sensitivity of the Method” regulation in 1977 attracted the attention of that agency. The EPA, responding to the large number of new laws and regulatory requirements that emerged in the 1970s, had to deal with many carcinogens, some, such as pesticides, intentionally introduced, but most occurring as widespread contaminants of air, water, and soils. The EPA and its precursor agencies had used the traditional methods of toxicology, based on the threshold concept, to establish health-based standards but had no methodology to deal with carcinogens. The agency had, for pesticide residues in food, to enforce the Delaney Clause but was faced with a “no residue” requirement similar to the one I have described for veterinary drug residues. The EPA was relying upon analytical chemistry criteria and technological achievability for contaminants, in much the same way FDA had for aflatoxins and other carcinogenic contaminants of food (PCBs became during the late 1970s another outstanding example of this problem). I learned that some EPA scientists well understood that the absence of any systematic way to deal with the health risks posed by carcinogens was a serious impediment to decisions having the primary purpose of protecting human health. Earlier agency efforts proposing complete bans on exposures to carcinogens had incurred serious criticisms from many quarters, and, in the end, were rejected as impractical. Achieving zero risk on a wide scale was not an available option.¹

I learned from discussions with Elizabeth Anderson, the EPA scientist having the responsibility for establishing risk-based approaches in the agency, and the agency’s superb consultant on this matter, Roy Albert of New York University, that EPA was moving quickly on this topic. The agency had published written guidelines on the conduct of carcinogenic risk assessment¹² and had major efforts underway to implement these guidelines. EPA was, I thought, moving in the right direction, with what seemed to me much greater internal support than I and my fellow risk assessment advocates were receiving within the FDA. Indeed, many toxicologists within FDA were uncomfortable with this new approach to safety assessment.

Elizabeth and Roy wanted to speak with me because they were concerned about the possibility that, as different federal agencies moved to adopt quantitative risk assessment approaches for carcinogens, inconsistencies might arise in the methods proposed for use. Such inconsistencies might weaken the credibility of federal efforts to incorporate risk assessment into decision-making.

One of the reasons for this concern was the proliferation, during the mid-to-late 1970s, of statistical models proposed for low-dose extrapolation. The various models discussed in the literature during this time equally well described the observed dose-risk relationships, but predicted large differences in low-dose risks for the same carcinogens, and the existence of these differences could be (and were) used to cast doubt on the validity of the risk assessment methodology. I shall deal with this important dilemma later but for now simply note that the EPA was relying upon Kenny Crump’s important 1976 publication, illustrating the use of the “linearized” multistage model for low-dose extrapolation.¹³ I assured Elizabeth and Roy that the FDA, after reviewing the Mantel-Bryan approach and several others, had elected to adopt the linear model.

Many other questions arose in connection with carcinogenic risk assessment, but this one issue of low-dose extrapolation was, at this time, the most controversial and the most likely to threaten this new and badly needed approach to dealing with carcinogens.

Interagency Collaboration 1977 to 1980

The 5 federal regulatory agencies with responsibilities for chemical regulation all had new leadership under the Carter administration: Douglas Costle (EPA), Eula Bingham (Occupational Safety and Health Administration [OSHA]), Susan King (CPSC), Carol Tucker Foreman (United States Department of Agriculture [USDA])—Food Safety and Inspection Service, and Donald Kennedy (FDA). I moved into a science advisor position under Kennedy in early 1977. These new agency heads got together and formed what was called the Interagency Regulatory Liaison Group (IRLG), having the purpose of ensuring consistent approaches to various common scientific and policy issues. The new regulatory leadership was in part reacting to increasing public attacks on perceived regulatory failures. In early 1977, for example, the FDA proposed to ban the hugely popular noncaloric sweetener, saccharin, on

the basis of bladder cancer findings in rats administered extremely high doses (percent levels in the diet) and the requirements of the Delaney Clause. Attacks on this proposed action played out in all major media, with much ridicule and scientific criticism of animal studies. The agencies knew they could not succeed in their responsibilities without reliance upon animal toxicity studies, and assembled an interagency committee under the IRLG, involving experts from the regulatory agencies and from federal public health agencies, to develop and bolster scientific support for the use of animal data in decisions. Several other committees were put to work on various topics, including the Work Group on Risk Assessment.⁷

Eula Bingham, head of the OSHA, was the IRLG member chosen to oversee this group, and I was asked to chair the group. All the agencies contributed members, including Elizabeth Anderson of the EPA, and David Gaylor, a statistician from the FDA's National Center for Toxicological Research, someone I found to be an excellent guide to my own thinking. Our work received superb assistance from Roy Albert, David Hoel (then at NIEHS), and Umberto Saffiotti and Marvin Schneiderman, both of the National Cancer Institute. The Work Group met many times over the following 18 months, sought and received advice from major science leaders in government, including Arthur Upton (director of NCI) and David Rall (director of NIEHS). The effort of the Work Group was published in the *Journal of the National Cancer Institute* in a dense and highly detailed paper entitled "Scientific Basis for Identification of Potential Carcinogens and Estimation of Risks."⁷ The contention of the Work Group was that, although the federal regulatory agencies had different legislative mandates, requiring different approaches to decision-making, the agencies could agree on common approaches to risk assessment. Although the IRLG Work Group report certainly did not settle all scientific questions and disputes, it consolidated federal agency thinking and set the stage for ensuring a systematic, relatively transparent, and consistent approach to evaluating and managing the difficult problem of carcinogens in foods and consumer products, the environment, and the workplace.

I and several other Work Group members and agency scientists and officials made public presentations in many venues regarding the IRLG effort on risk assessment. I continued in my day job at FDA, devoting much time to the public hearing the agency held on DES, trying to explain the value of risk assessment to agency scientists and policy makers, and interacting with experts in government, industry, and the academic community. I shall relate here some of the seminal events relating to the introduction into decision-making of quantitative risk assessment.

Concerns From the OSHA

Sometime during the 18 months of deliberations of the IRLG Working Group, I was invited to visit with Eula Bingham in her office at the OSHA. Dr. Bingham had come to the agency from the University of Cincinnati with a distinguished scientific

record and a strong commitment to further occupational health. The centerpiece of her program was a new proposal to regulate occupational carcinogens by establishing workplace standards based on best available control technology, also considering costs. Dr. Bingham and her staff had put an enormous effort into this proposal and had moved ahead to impose a new Permissible Exposure Limit (PEL) on benzene, based on this new approach.^{14,15}

During our meeting, in which I reported on the status of the Work Group effort, Dr. Bingham explained her deep concern about reliance on quantitative risk assessments. The approach the OSHA was proposing to manage risks from occupational carcinogens was not risk-based. Rather, a finding that a substance was a carcinogen would be sufficient to trigger technology-based controls. This approach to regulation was hazard-based and conceptually similar to the regulation I earlier described in connection with food contaminants and additives. Dr. Bingham was highly concerned that the IRLG effort, which was to put into place quantitative risk assessment methods for carcinogens, would necessitate risk-based decisions, and so would undermine the new OSHA proposal. I had no immediate response, except to say that the Work Group's effort would be considered a guideline for the conduct of risk assessment, should an agency choose to undertake such an assessment, and not a requirement that regulation be based on such an assessment. We both knew this was not an especially compelling argument, and we also knew that the FDA, and most especially the EPA, was committed to risk-based decision-making, except where explicitly prohibited by law.¹ Bingham and I could not find a clear path forward on this issue, but the IRLG effort continued, ending with the 1979 report.⁷

I recall numerous and somewhat tense calls and meetings with the EPA on this matter, and these continued until a case brought against OSHA by the American Petroleum Institute was decided by the U.S. Supreme Court in 1981. The Court ruled that the OSHA proposal was not consistent with federal law and that the agency had to demonstrate that the existing occupational exposure to a carcinogen carried a significant risk to health and that proposed exposure reductions would result in a significant reduction in risk.¹⁶ Although factors other than risk would play a role in standard setting, risk and risk reduction were essential criteria. This decision arrived well after the publication of the Work Group's guidelines but did much to ensure a permanent place for risk assessment in regulation and public health decision-making.¹⁷

The FDA and Its Critics

As I have said, the DES matter was the subject of a public hearing at the FDA, beginning in 1979, and there emerged during these hearings other views of possible limits of the risk-based approach. There was no doubt that low levels of this drug could be found in meat from treated cows, so it would seem that there should be no question that the "no residue" requirement of law was being violated. Experts engaged by the regulated industry, however, brought forth scientific arguments

that the levels of those residues carried no significant health risk and that they should be allowed. Professor Elwood Jensen of the University of Chicago, a recognized expert in hormonal carcinogenesis, testified that⁴:

There is no evidence of any fundamental difference between the hormonal action of DES and that of the endogenous hormone estradiol.

And that he knew of:

...no instance in which it is established that the cancer-enhancing effect of DES cannot be duplicated by an appropriate dose of a steroidal estrogen.

Jensen was, in effect, saying that meat naturally contains estrogen and so do the consumers of that meat, and the biological properties of the synthetic estrogen were in all respects identical to those of natural estrogens. The addition of DES contributes insignificantly to the normal, natural background levels of estrogens, and whatever it contributes is identical to the natural and safe estrogen exposures. This is in essence the “threshold” argument that is still offered by some today in relation to what have been labeled “endocrine disruptors.”⁴

A similar argument, including one based on an interesting but little examined approach to quantitative risk assessment, based on biological mechanisms, was offered by Professor Thomas Jukes of Berkeley, a prominent molecular biologist. Jukes contended that detectable residues of DES in meat posed no significant health risk. Both Jensen and Jukes, 2 superb scientists, cast significant doubts on the risk assessment approach that had been assembled by me and other government scientists.

I and others at the FDA were required to evaluate these scientific proposals, and weigh them against the counter arguments offered by other experts. We decided that the evidence in support of them was inadequate, and we also offered our own quantitative risk assessment and found that risks from detectable residues substantially exceeded those the agency considered negligible under the “Sensitivity of the Method” regulation.¹¹ Evidence that DES had properties distinct from those of natural estrogens also influenced our thinking. Commissioner Donald Kennedy, in his last official FDA act, in July 1979, signed a regulation prohibiting continued use of DES in animal production.⁴

Lessons From 2 Superb Scientists

Donald Kennedy, FDA Commissioner from 1976 to 1979, and to whom I reported during most of that period, was an outstanding scientist (neurobiology) with a superb intellect and understanding of the role of science in the public arena. He rapidly developed an understanding of the problem of carcinogens and of the essential elements of risk assessment. He provided much needed support to my own efforts and to promote the regulatory uses of risk assessment. I recall one interaction with Kennedy

that profoundly influenced my attitude and approach to advocating increased uses of risk assessment in decision-making.

I believe it occurred in the summer of 1979 and came about at the request of another highly regarded scientist, Dr. Arthur Upton, then director of the National Cancer Institute, and a giant in the area of radiation risk. I sat for an hour with Kennedy and Upton, and discussed risk assessment and policies in relation to carcinogen regulation. I sensed that Dr. Upton had discussed the IRLG effort with the 2 NCI scientists who were participating in the effort, Umberto Saffiotti and Marvin Schneiderman. I had earlier discussions with each of these scientists and knew that one of them (Saffiotti) had concerns of the same kind Eula Bingham had expressed. He also thought that quantitative risk assessment was too uncertain to be used as a guide to decisions. Dr. Upton expressed similar concerns to Kennedy and me but also recognized that if carefully conducted, described, and applied, quantitative models of risk could add much needed rigor to decision-making.

The discussion that followed was at a much-elevated level; I was mostly a listener and came away with a greatly improved understanding of science in the formulation of public policies, and the need for extreme caution in the elaboration of scientific knowledge and its limitations. The temptation to leap beyond what is truly established knowledge can be great if that leap can advance some desired policy agenda, but doing so can threaten scientific credibility and backfire. At the same time, these 2 great minds agreed, in the area of public health protection, it may be necessary, for policy reasons, to introduce certain precautionary elements into the interpretation and uses of scientific information; the goal, always, is to find the right balance in the context of the decision at hand. I have always wished I had a recording of the conversation, but I did make notes to myself about it (I’m afraid I no longer have them), and I have tried to use these as a guide in my professional life. I also know that the Kennedy–Upton discussion very much influenced the way I thought about my later role as a member of the National Academies “Red Book” study.¹ The conduct and uses of risk assessment is never to be undertaken without great care and attention to what is known and to how well it is known.

The Bingham discussion, the Jensen–Jukes commentaries on low-dose DES risks, and the Kennedy–Upton discussion provide a sense of the atmosphere surrounding quantitative risk assessment as the 1970s came to a close.

American Industrial Health Council

Perhaps the most important voice for industry during this time was that of the American Industrial Health Council (AIHC), a group founded in 1977 by several major trade associations, to deal with OSHA’s developing cancer policy. I first encountered the group as a result of its comments on the IRLG efforts. The members and advisors of AIHC included a number of scientists, and I recall many conversations and meetings with some of them during the development of the IRLG guidelines. Although some members of AIHC expressed strong opposition to the methodologies described in the IRLG report, most

avored the effort because it seemed to represent a significant step toward consistency in approach among federal regulatory agencies¹⁸:

AIHC supports the report's stated objective of ensuring that regulatory agencies evaluate carcinogenic risks consistently. We strongly urge that this initial step be followed up so that a national cancer policy is developed and conflicting policies among the regulatory agencies are minimized.

Although AIHC members expressed disappointment with the lack of opportunity for formal comment on the IRLG document, much of what AIHC proposed did influence the initiation of the Red Book effort and AIHC recommendations are discussed at some length in that report.

The Foes of Risk Assessment

It is not possible in a relatively short paper to more than briefly summarize the evolution of thought regarding the conduct of cancer risk assessment that made its way into scientific literature after it became clear in the mid-1970s that regulatory officials were interested in moving to risk-based decisions. That scientific literature and the increasing numbers of conferences and workshops it engendered began to shine bright lights on a number of scientific issues that had previously been examined in relative isolation from one another, but which, in the evolving risk assessment context, required integration.

The conduct, interpretation, and uses of epidemiological studies to identify carcinogens were the subjects of much increased attention, brought about in part by the initiation of the IARC Monograph program in 1971. Similar issues regarding cancer bioassays in animals began to appear frequently in scientific publications and programs, in part inspired by controversies such as the one, described earlier, relating to saccharin. The publication of the Ames assay in 1975 led to rapid growth in research into the utility of such assays in identifying carcinogens and their mechanisms of action. The discussions and debates devoted to these topics during the 1970s did much to improve the scientific quality and reliability of evidence relating to the identification of carcinogens and understanding how their effects were produced. Although scientific discussion of these issues and the controversies associated with them continue to these days, the increased attention paid to these sources of scientific evidence during the 1970s and early 1980s laid down one of the cornerstones of risk assessment. The other cornerstones concerned dose–risk relationships and the “low-dose” problem, and cross-species and within-species variabilities in risk. To say nothing about the issue of whether benign tumors should count!

During the period when the value of risk assessment for decisions was first given serious discussion, the scientific bases for these critical issues—low-dose extrapolations and quantification of cross-species and within-species variabilities—were relatively weak and consensus was not to be

found. The public discussions and debates regarding these matters were both lively and contentious.

I and my colleagues within the regulatory agencies nevertheless pushed ahead, primarily because the absence of any consideration of risk in decision-making (which always involved setting some limit on human exposures to avoid significant health risk) could lead in some cases to inadequate public health protection, in others, to unnecessarily restrictive limits.

The debates surrounding high-to-low dose extrapolation and variability were certainly critical to decisions regarding the preferred approaches to risk assessment, but in my experience, the most difficult issue to be resolved during this period was the question of whether quantitative risk assessment, of any kind, should play a role in decisions about carcinogen regulation. The proposed OSHA cancer policy, described earlier, and ultimately rejected by the Supreme Court, was the most visible attempt to continue hazard-based decision-making. Once convincing evidence of carcinogenicity became available for a regulatable substance (ie, its cancer hazard was established with reasonable certainty), controls, and limits on exposures to that substance would be established based on criteria such as “technical feasibility,” analytical detection limits, or complete elimination where that was technically possible. No consideration was to be given to the magnitude of risk associated with the substance.

The debate over hazard as against risk-based regulation is still with us, but it was most intense during the late 1970s perhaps because risk assessment had not yet been fully established as a regulatory approach and many did not want to see it established.

The foes of risk assessment at the time primarily emphasized the scientific uncertainties associated with extrapolations. Some focused on the high- to low-dose problem, and some on the fact that animal data were of unknown reliability for assessing human hazard or risk. Umberto Saffiotti of the NCI, mentioned earlier, wrote in 1977:

No existing method allows us to predict precisely and reliably the level of carcinogenic response in humans to chemicals which are known to be carcinogenic only from experimental studies . . .

Many similar statements can be found from other prominent NIH scientists.

The issue of dose-risk modeling was a central point of the scientific uncertainty argument. Many statistical models were being discussed as possible approaches to low-dose extrapolation, and it was apparent that, although most of these models could be used to “fit” the observed dose–risk data from cancer bioassays, they provided different estimates of risk, sometimes very large differences, at doses experienced by humans, which were typically several orders of magnitude lower than the doses at which cancer responses had been observed. Because there was no way to know which of these various models represented the truth about risk, there was no reliable way to decide which predicted risk was the true one. For some critics, huge

differences in predicted risks meant risk assessment was a useless tool for decision-making.¹

During the period leading up to the publication of the IRLG document in 1979, and in the years following that publication, I encountered these various criticisms of risk assessment many times. (And still do, although in different contexts.) I understood most of these criticisms and, in fact, thought they were true. But I also thought they missed the point. I learned to paraphrase Winston Churchill's remark about democracy being the worst form of government, except for all others. Risk assessment, with all its difficulties and uncertainties, remains superior to other approaches to public health protection. I shall elaborate on this subject when discussing the National Academies "Red Book" on risk assessment, but I shall make a few summary points here.

First, it is of course the case that no methods are available to assess risks to human health with known accuracy, from animal studies and even from observational studies in human populations that are different from the population we seek to protect. But there are ways to characterize risks that allow sufficient understanding of how well human populations will be protected under different risk management strategies.

Second, there is no situation, short of a complete prohibition on exposure, for which we can claim that any exposure is risk-free, whether that exposure relates to a carcinogen or to a chemical displaying any other form of toxicity. In fact, the use of risk assessment for carcinogens allows explicit identification of risks that are being accepted or tolerated, whereas the approach to establishing exposure limits intended to protect against all other forms of toxicity (the Lehman-Fitzhugh approach described earlier being the prototype for current approaches) provides no insight into the residual risk being accepted, and regarding how much population protection is being achieved. Certainly, thresholds exist for these forms of toxicity, but thresholds vary within populations, and quantitative approaches to evaluating risks for threshold agents could provide some characterization of how well populations are being protected (fractions having individual thresholds exceeded) at different labels of exposure.¹⁹

These and other advantages to risk-based approaches will be covered in more detail later. These advantages depend upon the availability of uniform and consistent approaches to risk assessment, and the 1979 IRLG guidelines set forth the principles for achieving these needed approaches. Improved guidance on these matters came with the Red Book and its sequelae.

The Low-Dose Risk Problem

For those having a generally favorable view of risk assessment's value, the dominant issue concerned models for moving from relatively high-dose observations of cancer risk to estimate possible risks associated with the much lower doses typically associated with exposures occurring in human populations. Other issues requiring extrapolation—from animals to humans, from one human population to another, etc—certainly arose for discussion, but received less attention

because approaches to them had been earlier developed for other types of toxic responses.

It is not surprising that many scientists frequently offered credible arguments suggesting that many carcinogens acted through mechanisms involving a threshold in the dose–risk relationship. Such agents, it was suggested, should be assessed using the standard methods used for agents exhibiting other forms of toxicity.

Two excellent publications in *Science* by Jerome Cornfield on carcinogenic risk assessment and relevant dose–response models provided a clear picture of the state-of-the-science in the mid-1970s.^{20,21} Thresholds, concluded Cornfield, could be derived from the models reviewed, but only if the carcinogenic agent were completely deactivated prior to any initiating event. There was no enthusiasm in our IRLG Working Group for promoting the threshold concept for carcinogens; several of us thought thresholds were likely for some carcinogens (saccharin was on our minds), but thought those proposing thresholds for carcinogens should have the burden of providing evidence to support such a hypothesis. The IRLG paper concluded that “any dose may induce or promote cancer.”^{7(p265)} In retrospect, I believe a better conclusion, one consistent with available dose–risk models, might have been that “any dose that reaches the target site could increase the risk of cancer development.” At what point that risk and the corresponding dose became of concern is another matter.

Much attention during this time was focused on the Armitage and Doll's 1961 formulation of the multistage model of carcinogenesis.¹³ The model assumes that cancer originates in a cell that has undergone a series of somatic mutations, taking place in finite steps. Each mutational stage is depicted as a Poisson process, in which transitions occur at rates that increase with dose in a linear fashion. Other models (including several “time-to-tumor” models) were under much discussion. In a now famous 1976 publication, Guess and Crump demonstrated how a linear hypothesis could be incorporated into the multistage model; upper confidence limits on the linear term were applied to estimate *upper bounds on low-dose risks*. Crump later demonstrated that this mode of extrapolation produced results similar to those derived from application of their one-hit model, thought by many to be most appropriate for “single-hit” carcinogenic processes.¹³

In 1980, I attended what I thought was one of the most comprehensive symposiums on health risk analysis, organized by the Oak Ridge National Laboratory (ORNL), with proceedings published in 1981.²² The 36 papers in the proceedings covered all aspects of risk assessment, and collectively provided, together with commentaries from the audience, an excellent record of the state-of-the-science in 1980; it also offered perspectives on significant policy questions.

With respect to dose–risk modeling, I found the broad overview by Bernard Altshuler and Kenny Crump's exposition of the linearized multistage model to be most useful. Richard Peto's commentary, including his deep skepticism about any form of extrapolation and his advocacy of reliance on simple, data-based potency estimates, ushered in a discussion of the

role of policy in model selection, based on the needs of regulators. (The role of policy in such selections is fully elaborated in the Red Book, discussed later.) Roy Albert, then at New York University Medical Center and still at this time a consultant to the EPA's Cancer Assessment Group, noted that "it's very difficult to recommend an amount of money to spend on remedial action on the basis of a potency estimate."

I was also attentive to Richard Peto's comment on Kenny Crump's presentation²³:

If one wants to make extrapolations to get point estimates of upper confidence limits down toward zero based on dichotomous data, I think the method you have described is obviously the method of choice.

The linearized, multistage model became the EPA's default model for cancer risk assessment, to be replaced later by the simple linear, no-threshold model proposed in 1980 by David Gaylor and Ralph Kodell. Gaylor actually commented on this "linear interpolation algorithm" during the Oak Ridge symposium. This approach yields virtually the same upper bound estimates of low-dose risk as does the linearized multistage model, but without any assumptions regarding its biological basis. Gaylor and Kodell state²⁴:

In this paper, we have modified and developed the suggestion of Mantel and Bryan. We do not extrapolate outside the experimental data range with the parametric model used to describe the results in the experimental dose range. We do not use arbitrarily "conservative" slopes to extrapolate to lower doses, but use linear interpolation to obtain an upper bound on the risk at lower dosages. The purpose of this paper is to provide a justification for such a procedure, to provide a simple widely applicable mathematical algorithm for performing low dose risk assessment from dose response data, and to examine the performance of this procedure on a variety of dose response curves for toxicological data, including but not limited to carcinogenesis.

Although I had always assumed that it was not possible to claim that accurate estimates of low-dose cancer risk could be derived, the discussions at the ORNL symposium also convinced me that upper bound estimates of low-dose risk were obtainable and useful for decisions. I also understood that many other decisions in risk assessments, including data review and evidence weighing, selection of specific data sets for risk assessments, the value and uses of mechanistic information, dealing with a range of issues related to biological variability, estimating human exposures, and analyzing and treating uncertainty, all contributed to the assessment of risk and decisions resting on these assessments. Low-dose risk modeling was far from the whole story.

Risk Analysis

In 1969, Chauncey Starr of the Electric Power Research Institute published a paper in *Science* entitled "Social Benefit vs.

Technological Risk."²⁵ This article, heavily focused on safety risks, opened up many new areas of inquiry regarding the social acceptance of risk, and the important question of ensuring public safety and health without unnecessarily losing the benefits of technological innovation; in the decade following Starr's publication, what came to be called "risk analysis" became an organized field of study. It included not only the assessment of risk but also the questions of risk perception, risk communication, public attitudes about risks of different kinds, risk-benefit and cost trade-offs, decision-making under uncertainty, and various risk management questions. The growing interest in these broader areas of risk inspired the initiation in 1981 of the journal *Risk Analysis*. While at the FDA in 1979, I entertained a visit from Robert Cumming of the ORNL, and he convinced me to get involved in that journal's development. The Society for Risk Analysis (SRA) was created in 1981, and every type of threat to human safety and health is now cultivatable territory for members of SRA and its journal.

An outstanding book published in 1976 by William Lowrance, written when he was a resident fellow at the National Academy of Sciences, and entitled "Of Acceptable Risk" was my first introduction to the wider world of risk analysis. Of particular interest to me was the clarity he brought to the questions of risk "acceptability" and its relationship to safety²⁶:

a thing is safe of its attendant risks are judged to be acceptable.

This definition raises many serious questions: (1) how do we measure risks associated with the "thing," and how good are our measurements; (2) who gets to be the judge and why should the judge be trusted; and (3) why should any risk be "acceptable"? Lowrance deals with these and similar questions with clarity and thoughtfulness. I suspect many risk analysts would today find some of Lowrance's views a bit antiquated, but I think in the whole they have survived well.

Those working in the deep forests of risk assessment should always strive to achieve greater awareness and understanding of the social and policy contexts of their work. In my own case, I have long believed that the methods long in place to derive "acceptable" exposure levels for toxic agents operating through threshold mechanisms, because they fail to provide any understanding of levels of risk associated with these exposures, offer no opportunity for decision-makers (which would today be referred to as risk managers) to evaluate and judge acceptability. Further, these "bright line" (safe/unsafe) models have no value in evaluating the magnitudes of risk reduction achieved under different risk mitigation strategies, a valuable piece of information for decision-makers.¹⁹ One may question the cancer risk assessment methodology, but it does offer quantitative measures or risks, and thus opportunity for clarity regarding risk acceptance. I recognize that most toxicologists are comfortable with the traditional approach to threshold effects, but I believe a better understanding of what is required for risk information to be truly useful might inspire thought regarding methods for quantifying risk for all forms of toxicity.²⁷ In sum, distinguishing between thresholds for individuals and

populations, which requires recognition of the fact that individual thresholds vary across populations, has been an overlooked problem in risk assessment that limits its utility in many circumstances (see the later discussion concerning the NRC's 2009 Silver Book report).

The Red Book

I left my position at the FDA in mid-1980, recognizing that quantitative risk assessment had not attracted large numbers of enthusiastic supporters at the agency. I was pleased to note that under the leadership of Alan Rulis, Ron Lorentzen, and a few others, the agency had found value in risk assessment for establishing criteria for acceptance of unavoidable residual levels of carcinogens in, for example, food-packaging materials. Meanwhile, the EPA was plowing ahead and applying quantitative risk assessments in a range of areas, including the difficult problems associated with remediation of Superfund and other hazardous waste sites. The OSHA, too, began a program of regulating workplace carcinogens based on application of risk assessment methods according to the mandate of the Supreme Court's benzene decision. One interesting feature of the OSHA's efforts concerned its decisions to tolerate risks for occupational carcinogens significantly greater than the EPA found acceptable for carcinogens in the environment.

I have not mentioned one important development that occurred while I was still employed at the FDA and that was conceived and executed by Gilbert Omenn of the President's Office of Science and Technology Policy (OSTP). Omenn was a physician and geneticist who came to a high position in the White House in 1977 and remained there until 1982. During the period in which I was much involved in the IRLG program, I met with Omenn and his staff to fill him in on our activities and the direction of our work. Omenn, I learned, was supportive of a quantitative approach to risk assessment and had underway the development of a report on the topic. Omenn wisely saw an important role for the OSTP in ensuring that federal government agencies maintained strong scientific support for their activities and that consistency across those agencies in the use of scientific information was also achieved. He realized that risk assessment was becoming central to many areas of decision-making, and he and his staff published, in 1978, an important report on the topic.

The report focused on the content of risk assessment but was not highly directive on the specific methods to be used. It emphasized the need for transparency and consistency, and also set forth the content of and distinction between the scientific assessment process and the development and implementation of regulatory approaches.²⁸ Much of the debate that arose during the early 1980s concerned what came to be called the "separation" of science and policy. Many critics of regulatory practices claimed that the science behind regulation was too often twisted to achieve predetermined policy (or "political") outcomes. The most common criticisms related to alleged biases in the selection of data and models in risk assessments, designed to achieve results that would best serve the political

interests of decision-makers. To put it bluntly, risk assessors were accused of making risks "disappear" if their existence made decisions too difficult, and of greatly overstating risks if doing so made it easier to achieve desired outcomes. The seriousness of this issue—distortion of science to achieve desired policy outcomes—during the early 1980s, when risk assessment was becoming increasingly prominent in many regulatory domains, is difficult to exaggerate.¹ As I was becoming a practitioner of risk assessment and a consultant to the EPA on some new applications of risk assessment, particularly in the area of Superfund and hazardous waste remediation decisions, I witnessed firsthand many often-bitter skirmishes relating to the undue influence of policy in risk assessment results, and the so-called "cherry-picking of data" problem.

Congressman Donald Ritter, Republican of Pennsylvania, with a doctorate in science from MIT, introduced the Risk Analysis Research and Demonstration Act of 1982. The proposed Act had, I thought, many attractive features, and I testified in its favor at a hearing in 1982. Among other things, it called for a federal study, organized by the Regulatory Affairs Office of OMB, of risk analysis and its applications across the federal government, to include recommendations for research to improve risk assessment, and various projects to demonstrate its value and applications. Ritter's bill went nowhere but was heavily discussed in the scientific and trade press and had many features which, I suggest, would be valuable to consider even today. I will note that the several scientists and advocates who testified alongside me during Ritter's hearing expressed a wide range of views, some harshly critical of efforts to enlarge the application of risk assessments in regulatory discussions. The impassioned remarks of Nicholas Ashford, then a professor in the Technology and Law Program at MIT, particularly caught my attention. Ashford forcefully argued that risk assessment, because of its highly technical nature, was simply a device for slowing down and delaying regulatory decisions. This argument has much power, and is still heard today, as I shall note in the closing section.

"Risk Assessment in the Federal Government: Managing the Process" was published in 1983 by the Committee on the Institutional Means for Assessment of Risks to Public Health of the National Research Council.¹ The Committee was chaired by Reuel Stallones ("Stoney"), an epidemiologist at the School of Public Health of the University of Texas at Houston. I was invited to serve on the Committee and I was more than eager to do so. The committee had only 14 members and included health and social scientists and policy and regulatory experts, who came from positions in academia, industry and consulting, nongovernmental agencies, and government research institutions. A few members had extensive experience in risk assessment, but most, including the Chair, did not. But those without specific experience in risk assessment brought to the Committee deep knowledge and understanding of science and the role of science in forming public policy. The Committee was well served by a superb staff, headed by Lawrence McCray. McCray was a superb social scientist who provided much needed

guidance and support to our group; he went on to become a faculty member at MIT.

It may seem odd that a Committee assigned to examine risk assessment was not overloaded with risk assessment experts. I do not intend to repeat here the major findings of the 1983 Red Book, already discussed thousands of times in the literature, but I will emphasize a matter that is not often discussed—the objectives the Committee was asked to fulfill. The Committee was formed in response to a congressional directive to:

- (i) Assess the merits of separating the analytic functions of developing risk assessments from the regulatory functions of making policy decisions.
- (ii) Consider the feasibility of designating a single institution to do risk assessments for all regulatory agencies.
- (iii) Consider the feasibility of developing risk assessment guidelines for use by all regulatory agencies.

These objectives all relate to the raging debates I have described pertaining to the alleged improper incursions of policy makers into the conduct of risk assessment or the tendencies of risk assessors to offer up risk findings they believe policy makers would prefer to have. The “solutions” to this problem envisioned in the Committee’s legislative mandate can be readily inferred from these objectives. Risk assessments might be kept “pure” by ensuring the complete insulation of risk assessors from regulatory policy makers, perhaps even by creating a “National Institute of Risk Assessment” (see (ii) above). The guideline questions indicate the value of having “blueprints” all risk assessors must follow, without interference from decision-makers. These were the questions the Committee was asked to consider. It was not asked to offer opinions on how risk assessments were to be conducted, and it did not do so. But it did make many important recommendations.

Almost all public discussions about the Red Book begin and end with a presentation of its risk assessment framework and the 4 steps of risk assessment. This was no doubt important and helped to advance the field but resolved no significant controversies. Two other sets of recommendations and their support did help to resolve controversies, but their importance in advancing the field has not been given enough attention.

First, the Committee made it clear that risk assessments could not be completed without the inclusion of certain “science policy” decisions to deal with ever-present gaps in data and, most importantly, basic knowledge. It was critical, the Committee noted, that the best possible scientific basis be developed for decisions aimed at public health protection and that risk assessments were the necessary basis. But unless uncertainties were dealt with, risk assessments could never be completed—a completely unsatisfactory outcome.

At many steps of risk assessment, including the all-important dose–risk relationship in dose ranges for which empirical data are generally unavailable, inferences beyond the available data must be made. The options available to make these inferences are to be examined using the best available methods of science, they are to be ranked, if possible,

according to their relative scientific merits, and the option to be used—across all risk assessments—is selected. The selection involves an unavoidable policy choice. It is not the kind of policy choice associated with risk management; it is rather a “science policy” choice. These “inference” options and their selection do not apply only to carcinogen risk assessment but to all assessments of toxicity risks. The Committee offered guidance on how the inference options to be used, at each step of risk assessment where they are needed, was best selected. Uncertainties are inherent in the process and are to be explicitly considered.

The second area of major importance is the Red Book’s efforts on risk assessment guidelines. Guidelines on the conduct of risk assessment were necessary to ensure consistency and uniformity in approach across all agencies. The guidelines should, according to the Red Book Committee, include approaches for identifying the data and methods to be used in risk assessment, but also the options for making inferences beyond the data and the choice of options to be used (“the defaults”). The Red Book Committee emphasized that, in specific cases, in which well-developed scientific support became available, it should be possible to deviate from the guidelines. In effect, scientific data might become available in specific cases to reduce the need for the defaults.

These 2 critical recommendations were developed to deal with the Committee’s charge regarding institutional separation of risk assessment activities from the regulatory decision-making agencies. The Committee believed that the use by all regulatory agencies of uniform guidelines which contained agreed-upon inference options and defaults would be adequate to resolve the problem of case-by-case interjections of policy biases into the risk assessment process. Separating scientists involved in risk assessment from those who had to use assessments for decisions would, the Committee thought, lead to serious inefficiencies in communication and lack of clarity regarding the specific problems to be examined and addressed. Assessors and managers should be able to engage in useful (and necessary) discussion without undue influence of the managers on the conduct of risk assessments; guidelines for risk assessment would be essential to ensure the success of this approach. No separate “risk assessment agency” should be considered.

The 3 major recommendations of the Red Book Committee, out of a total of 10, were set forth in the Red Book’s summary. The third major recommendation, about which I have written elsewhere,²⁹ was directed at the Congress and urged the development of a Board on Risk Assessment methods. The Board would track developments in risk assessment and periodically revise inference guidelines, study the usefulness of agency approaches, and identify research needs. This important recommendation, if followed, would have done much to help ensure the efficient and scientifically adequate functioning of risk assessment, both within and outside of government, on a continuing basis.

No one paid attention to this recommendation.

Aftermath

Bernard Goldstein became chief scientist at the EPA at about the time the Red Book was released. He was an enthusiastic supporter of its findings and recommendations and wanted to see recommendations followed at his agency. Goldstein also understood that risk assessment was both controversial and not well understood. I was at that time advising the EPA's Superfund office on how risk assessment might be applied to decisions about site remediation, and one day ended up in Goldstein's office, along with other senior EPA scientists. Goldstein had in mind an ambitious program directed at training stakeholders from other federal and state agencies, regulated industries, NGOs, and academics who were serving as advisers to the EPA and related agencies. The list of stakeholders Goldstein had in mind included not only experienced scientists but also scientists with little experience and policy and legal experts. The program was implemented and I recall a 2-day training exercise (I think at a hotel in or near Annapolis, sometime in 1984) for about 120 people.

I was asked to help with the training, developed an exercise based on a hypothetical carcinogen (dinitrochickenwire [DNC]), and invented data regarding its hazards, dose-risk relationships, and human exposures. Trainees worked in groups of about 8 people and were required to work through the data and various options for evaluating the data and reaching conclusions about DNC's risks. Each of the 20 or so groups was then required to present its conclusions to the entire group of trainees. The exercise was quite well received, and most people left the meeting with substantially improved understanding of the content of risk assessment, the significant impediments to achieving highly certain results, and even regarding the borders between assessment and management. (I have been involved in offering the DNC exercise to many different groups, most recently in China and Australia.) Bernard Goldstein deserves much credit for creating an environment in which risk assessment would be recognized as an important basis for regulatory and similar public health decisions.

The publication of the Red Book generated immense interest in risk assessment, and I and many others were involved for a year or more in seminars and workshops devoted to the subject. The EPA got busy developing the recommended guidelines and also incorporating them into its many regulatory programs.

While the publication of the Red Book brought much clarity and understanding and opened many pathways for its application, controversies over the conduct of risk assessment did not vanish. This is not in the least surprising, because uncertainties are ever present in risk assessment, and scientific debate regarding them is not only expected but also essential. The existence of guidelines, including the establishment of inference options and default assumptions, certainly placed some bounds on these debates, and this was necessary to allow agencies to complete assessments without endless debates over scientifically irresolvable issues.

By the time of the Red Book's publication some state agencies had undertaken efforts to develop their own policies and

implementing guidelines.³⁰ I became aware of the important role of states in chemical regulation when I was invited by James Solyst, soon after the Red Book's release, to brief the board of the National Governor's Association on its purpose and content. I also attended a meeting of state agency scientists and administrators held at Times Beach, Missouri, a site of significant dioxin contamination that had received national attention. The Red Book was a major topic of discussion at the meeting, and it was clear that some states were ready to implement its recommendations with high enthusiasm.

My most memorable encounter at the meeting was with Thomas Burke, then a senior regulatory and public health official in New Jersey. Tom has, for many years, been on the faculty of the Johns Hopkins Bloomberg School of Public Health and serves as director of the School's Risk Sciences and Public Policy Institute. At our very first meeting in Times Beach, I found Tom to be not only a quick learner but someone with significant understanding and foresight regarding the importance of risk assessment, not only to regulation, but in the broader world of public health science and policy. Tom has had major roles in many National Academies efforts, the most notable of which was his role as Chair of the Committee that produced the extraordinary 2009 report "Science and Decisions: Advancing Risk Assessment." The recommendations of this report have not received sufficient attention, but I contend that if risk assessment has a future, much of that future is to be found in this report.¹⁹

Tom served for about 2 years under the Obama administration as the EPA Science Advisor and Deputy Assistant Administrator for Research and Development. I have been a friend of Tom since the Times Beach meeting and agreed to meet him for a drink the day after the presidential election of 2016. Like most Americans we had been shocked by Trump's victory. We also knew that Tom would not remain much longer at the EPA and we thought that Trump's plan for the agency was likely to be destructive. We had our drink, but it was a gloomy evening.

We had our drink at the Willard Hotel, where Abraham Lincoln had stayed when he arrived in Washington after his election to the Presidency.

Some of the Missing Parts, Including Ed Calabrese

There are many omissions in my story, both because there were during the decade I cover many activities I had little or no hand in, and also because I had to place some limits on what I chose to relate. Readers should not, therefore, assume that what I have offered is anything like a complete history of this important time. In addition, I have omitted the contributions of many individuals, including those who were important influences on my own thinking.

I am not an adequate writer to convey the nature and intensities of the many debates and disagreements that accompanied my journey, but I must say I found most civil and good-intentioned. I did suffer a few *ad hominem* attacks, interestingly from some traditional toxicologists who thought I was both

trying to wreck their traditions and that I was inadequately educated in those traditions, but these were not at all representative.

I do want to offer a few observations about the work of Ed Calabrese, a scientist well-known to the readers of this journal and to the world's community of dose-response scientists. Ed, in a series of extraordinary review articles, has researched and documented serious concerns about the scientific basis for the assumption of linearity at low dose, first in connection with radiation-induced cancer and then in connection with its adoption, by the EPA in particular, for chemically induced cancers. Calabrese has painstakingly reviewed many publications on radiation-induced mutagenicity and found that much of it falls short of providing evidence to support a relationship between somatic mutation theories of carcinogenesis and low-dose linearity for carcinogens, whether induced by radiation or by chemicals.

I have not devoted sufficient effort to provide an adequate appraisal of Calabrese's work, but it does deserve, I suggest, far more attention from the mainstream risk analysis community involved in regulatory and public health practice. Calabrese draws from his work the conclusion that what he terms "the flawed acceptance of linearity at low dose" by a 1977 Committee of the National Academies that was providing guidance to the EPA on carcinogenic contaminants of drinking water³¹ was of signal importance in driving regulators to the low-dose linearity model. This report may have been important to the EPA, but in my experience and as suggested in this article, there were many other forces at work that moved regulators in this direction.

Calabrese is also apparently dismayed that precautionary policies have been at play in some of the defaults used by regulators. But, as I have tried to make clear in this article, such policies are inevitable when science is uncertain and decisions have to be made. I do agree with Calabrese's contention that errors in science should be corrected when they are found and that there may have been inadequate attention to some of the scientific foundations for "one-hit" and "irreversibility" and other hypotheses that moved carcinogen risk toward low-dose linearity and that "science policy" (as in the Red Book) should not be decided until what science can and cannot tell us has been made clear. But I do suggest that it would take an enormous effort to revisit, verify, or reject all of the science upon which concepts of thresholds and low-dose linearity are based, and it is completely unclear how such an effort might be accomplished (of course, this is not a problem Ed should be asked to solve). Nevertheless, Calabrese's work is powerful and reflects the determination to uncover facts that, were they found to be true, would upset decades of risk assessment work and regulation.

A Bit of Summation

Finished and Unfinished Business

What is now called the National Academies of Science, Engineering and Medicine (NASEM) has issued, since the

publication of the Red Book, a long series of reports relating to risk assessment and other elements of risk analyses.¹⁹ Most of these reports have been sponsored by the EPA, but several other agencies of the federal government have also sought advice from the NASEM. A number of these reports have been devoted to identifying approaches to improve the practice of risk assessment, and many focus on new applications. What can be found in all of these reports is a commitment to the principles and analytical frameworks first laid out in the Red Book. The 2009 report "Science and Decisions" does create a new and valuable decision-making framework, in which the conduct of risk assessment follows a "problem formulation and scoping" exercise, designed to ensure that the risk assessment will be directed at the questions important to ultimate decision-making. The committee that produced this report recognized that many programs lacked emphasis on maximizing the utility of risk assessments for ultimate decisions. This and other recent NASEM reports have also placed emphasis on ensuring that scientific uncertainties, inherent to all risk assessments, are adequately described and considered in the risk management process.³²

In addition to a wealth of guidance from committees of NASEM, there are of course studies from many other institutions, both governmental and nongovernmental. One can readily perceive that much of this guidance has penetrated the major risk assessment programs of federal and state government, not perfectly but nevertheless usefully and effectively.

In my view, however, significant issues that arose during the decade about which I have written remain unattended or inadequately addressed. Some have been the subject of expert studies; but actual practice remains largely unaffected. I note 3 such issues.

1. *Guidelines.* I have emphasized the need for continuing review and updating of risk assessment guidelines. The EPA has been, by far, the most diligent in producing guidelines, but even that agency's efforts have fallen short, and other important agencies have produced little. The difficult but essential issue of justifying default assumptions, and criteria for moving away from them in specific cases, has received insufficient attention. Without adequate guidelines, the production of risk assessments having adequate transparency is hampered. Risk assessments also become subject unnecessarily to controversies that should be resolved within guidelines, and so their timely production is impeded. As I noted earlier, the failure of government institutions to support the Red Book's recommendation regarding a standing Board on Risk Assessment Methods has been an impediment to progress.
2. *Quantification of Risk.* Cancer risks are generally subjected to quantitative assessment, but quantification of risk for all other forms of toxicity has not been significantly pursued. Instead, "bright-line: models," such as those represented by the RfD, ADI, and TDI, are the norm. These approaches provide no guidance to

decision-makers on the risks associated with agents acting through threshold mechanisms or to agents other than carcinogens acting through nonthreshold mechanisms. Methods for probabilistic assessments have been suggested by the WHO/International Program on Chemical Safety and demonstrated by others.^{27,33} The “Science and Decisions” report mentioned earlier called for similar quantification efforts, heavily dependent upon modes of toxic action. Probabilistic approaches are both more revealing regarding the risks that remain under different exposure scenarios and are useful (necessary?) for the many decisions requiring trade-offs of one kind or another. Resistance of many toward probabilistic assessments for all forms of toxicity remains strong; indeed, the matter is hardly discussed.

Risk assessments should contain assessments of risk!

3. *Risk-Based Decision-Making.* I earlier described some of the opposition to risk-based decision-making that emerged from important stakeholders during the decade when risk assessment rose to prominence. The alternative approach focuses primarily on the identified hazards associated with a chemical (the type of toxicity it can cause) and seeks to reduce exposures to substances having what are judged to be particularly serious hazardous properties (eg, carcinogens, reproductive and developmental toxicants, endocrine-disrupting substances). In many cases, total elimination of the hazardous substance is sought and in others reductions to the lowest levels thought to be reasonably achievable; in neither case is the risk posed by the substance considered.

Arguments for hazard-based decisions typically point to the uncertainties associated with risk assessments, the fact that risk assessments require much more data (eg, on dose–response and exposure) and take much more time to complete. These powerful arguments do, however, seem counter to the fact that almost all laws governing chemical regulation call for risk-based approaches. They also seem counter to the fact that hazard-based approaches can often lead to the unnecessary elimination of valuable substances without achieving significant health benefits. Introducing new chemicals to replace eliminated chemicals also raises the possibility of introducing unanticipated risks. The hazard-based approach to decisions has many advocates and is most seen in the voluntary actions of manufacturers and users of targeted substances because risk-based laws and regulation do not apply in such circumstances. Substitutions to replace chemicals in products can be perilous if not carefully done, with proper attention to risk.

Thus, the question of risk-based approaches which I encountered in the years during which risk assessment

began to achieve recognition remains in force. It seems unlikely that the risk-based requirements of our laws will be reversed, so that risk assessment will remain a priority. The arguments for the much simpler hazard-based approaches are, however, appealing in many ways and will likely influence the policies and decisions that can be taken on a voluntary basis. Although the 2 approaches do conflict on a scientific level (simplistically, one divides the world of chemicals into “the toxic” and “the non-toxic;” and the second into “the risky” and “the not so risky”), it is a conflict that may not require a resolution unless it leads to unforeseen and unacceptable consequences.

I am highly grateful for having been given the opportunity to offer my recollections and perceptions of risk assessment’s rise during the period I have covered, and I hope this subjective historical summary will somehow help support continued review and improvement of assessment and its uses.

Many of us have concerns that, at present, the federal government does not seem interested in devoting efforts to improve and foster science-based decisions. But risk assessment is now a worldwide project that should see continuing support. And interest in science-based policies is certainly not forever lost in the United States. I was extremely fortunate to be present at the creation and hope to be able to witness the continuing incorporation of these types of innovative approaches to both data development and risk assessment, much discussed now among scientists from an extraordinary range of disciplines. Maybe policy-free (default-free) risk assessment is too much to hope for, but current research tools and methods seem to point to that possibility.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note

1. The strict prohibition of the Delaney Clause continued to apply, as it does today, to any intentionally and directly added food substance that is carcinogenic.

References

1. NRC. *Risk Assessment in the Federal Government: Managing the Process*. Washington, DC: Author; 1983.
2. FDA. *Assessment of Estimated Risk Resulting from Aflatoxins in Consumer Product Peanut Products and Other Contaminants*. Rockville, MD: Author; 1979.
3. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina. *N Engl J Med*. 1971;284(16):878-881.
4. Zervos C, Rodricks JV. FDA’s ban of DES in meat production. *Am Stat*. 1982;36(3):278.

5. 21 U.S.C. Federal Food, Drugs, and Cosmetic Act. Food Additives Amendment of 1958.
6. Lehman A, Fitzhugh O. 100-fold margin of safety. *Quarterly Bulletin*. 1954;18:33-35.
7. IRLG (Interagency Regulatory Liaison Group). Scientific bases for identification of potential carcinogens and estimation of risks. *JNCI J Natl Cancer Inst*. 1979;63(1):241-268.
8. Mantel N, Bryan WR. "Safety" testing of carcinogenic agents. *JNCI J Natl Cancer Inst*. 1961;27(2):455-470.
9. Rodricks JV. Regulation of carcinogens in food. *Ann N Y Acad Sci*. 1981;363(1):29-35.
10. Hutt PB, Merrill R. Regulation of carcinogens. In: *Food and Drug Law*. Westbury, NY: The Foundation Press; 1991:863-963.
11. FDA. Chemical compounds in food-producing animals; criteria and procedures for evaluating assays for carcinogenic residues. *Fed Regist*. 1977;42:10412-10429.
12. Albert RE, Train RE, Anderson E. Rationale developed by the Environmental Protection Agency for the assessment of carcinogenic risks. *J Natl Cancer Inst*. 1977;58(5):1537-1541.
13. Crump KS, Hoel DG, Langley CH, Hod DG. Fundamental carcinogenic processes and their implications for low dose risk assessment fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer Res*. 1976;36(9 pt 1): 2973-2979.
14. OSHA (Occupational Safety and Health Administration). Proposed rule: identification, classification and regulation of potential occupational carcinogens. *Fed Regist*. 1977;42:54148.
15. OSHA (Occupational Safety and Health Administration). Final rule: identification, classification and regulation of potential occupational carcinogens. *Fed Regist*. 1980;46:5001.
16. Industrial Union Dept. v. Amer. Abu Dhabi, UAE: Petroleum Institute; 1980.
17. OSHA (Occupational Safety and Health Administration). Identification, classification and regulation of potential occupational carcinogens. 1982;47:187.
18. AIHC (American Industrial Health Council). *Recommended Alternative to OSHA's Generic Carcinogen Proposal*. Scarsdale, NY: Author; 1978.
19. NRC (National Research Council). *Science and Decisions*. Washington, DC: National Academies Press; 2009.
20. Cornfield J. Carcinogenic risk assessment. *Science*. 1977; 198(4318):693-699.
21. Cornfield J. Models for carcinogenic risk assessment. *Science*. 1978;202(4372):1107-1109.
22. Richmond CR, Walsh PJ, Copenhaver ED. Health risk analysis. In: *Proceedings of the Third Life Science Symposium*. Philadelphia, PA: Franklin Institute; 1980.
23. Peto R. See Ref.22:391.
24. Gaylor DW, Kodell RL. Linear interpolation algorithm for low dose risk assessment of toxic substances. *J Environ Pathol Toxicol*. 1980;4(5-6):305-312. <http://www.ncbi.nlm.nih.gov/pubmed/7217854>. Accessed August 17, 2018.
25. Starr C. Social benefit versus technological risk. *Science*. 1969; 165(3899):1232-1238. <http://www.ncbi.nlm.nih.gov/pubmed/5803536>. Accessed August 17, 2018.
26. Lowrance WW. *Of Acceptable Risk: Science and the Determination of Safety*. Hebden Bridge, England: Will Kaufman; 1976.
27. WHO (World Health Organization). *Harmonization Project Document 11 Guidance Document on Evaluating and Expressing Uncertainty in Hazard Characterization*. Geneva, Switzerland: World Health Organization; 2014.
28. OSTP (Office of Science and Technology Policy). *Potential Human Carcinogens: Methods of Identification and Characterization. Part 1*; USA: US Government's Office of Science and Technology Policy; 1978.
29. Rodricks JV. What happened to the Red Book's second most important recommendation? *Hum Ecol Risk Assess*. 2003;9(5): 1169-1180. doi: 10.1080/10807030390240373.
30. State of California - Health and Welfare Agency. Sacramento, CA. October, 1982. Carcinogen Identification Policy: A Statement of Science as a Basis of Policy; Section 2: Methods for Estimating Cancer Risks from Exposures to Carcinogens.
31. National Academy of Sciences. *Drinking Water and Health*. National Academy of Sciences: Washington, DC; 1977.
32. IOM (Institute of Medicine). *Environmental Decisions in the Face of Uncertainty*. Washington, DC: National Academies Press; 2013.
33. Chiu WA, Slob W. A unified probabilistic framework for dose-response assessment of human health effects. *Environ Health Perspect*. 2015;123(12):1241-1254.