Methods for Estimating Risk of Chemical Injury: Human and Non-human Biota and Ecosystems
Edited by V. B. Vouk, G. C. Butler, D. G. Hoel and D. B. Peakall

Quantitative Approaches in Use in the United States to Assess Cancer Risk

Elizabeth L. Anderson

ABSTRACT

Scientists have long used conventional toxicological methods to establish 'safe levels of exposure' for chemicals considered as having thresholds for some biological effects. These same methods cannot be used to establish safe levels of exposure for chemicals considered as having no thresholds for some effects, such as carcinogenesis. Therefore, Federal regulatory agencies in the United States use risk estimation methods to provide information for public health policy decisions concerning risks associated with exposure to chemicals considered as having no threshold effects. Acceptable exposure and risk levels are decided by policy makers who generally consider estimates of risks together with social and economic benefits derived from the use of chemicals.

This chapter discusses the development of risk estimation approaches by Federal regulatory agencies in the United States and identifies the mathematical models currently used for risk estimation. The uncertainties and limitations of these methods have led some scientists to question the utility of quantitative risk estimation. The US Environmental Protection Agency (EPA) experience, summarized in this chapter, can provide a realistic basis for evaluating the reasons for and against using such methods. Finally, shortcomings in current risk assessment methods and their use in policy decisions are explored, and areas for possible improvement, given current scientific knowledge, are identified.

1 INTRODUCTION

Conventional toxicological methods have long been available to define 'safe levels of exposure' to agents causing adverse effects which have identifiable thresholds (NAS, 1977). More recently, risk estimation methods have been developed to provide quantitative information for agents, most notably for potential carcinogens, where safe levels of exposure cannot be identified by conventional methods (EPA, 1976; Albert *et al.*, 1977; IRLG, 1979). All major Federal regulatory agencies in the United States have used information obtained

by quantitative risk estimation to establish protective public health policies. Applications include agents such as ionizing radiation, air pollutants, some food contaminants and pesticides, some water contaminants; and worker's protection, consumer protection, hazardous waste disposal, and the clean-up of uncontrolled waste sites.

In this context, risk assessment is defined as a two-step process. The first step involves a qualitative evaluation of all biomedical data available to answer the question of how likely an agent is to be a human carcinogen. The likelihood is expressed in terms of adequacy of biomedical evidence. The second step in the risk assessment process involves fitting some curve to the dose–response data and coupling it with information about population exposures to answer the following question: on the assumption the agent is a human carcinogen, what is the magnitude of health impact of current and projected exposures? These estimates are generally expressed in terms of increased individual lifetime risk in subpopulations and the number of annual cancer cases as an index of nationwide impact. There are large uncertainties associated with this extrapolation process; therefore, these estimates of risks must be used with caution.

To provide quantitative estimates of risk from low levels of exposure generally found in the ambient environment, most often one must extrapolate from high doses in the observed range, usually involving animal bioassay studies, to much lower exposures involving human populations. Although a variety of mathematical models for extrapolation have been presented in the literature, the model most often used by US regulatory agencies has been a linear nonthreshold model or a similar model having a linear nonthreshold component in the low-dose region of the dose–response curve, to estimate the upper bound of risk (Chand and Hoel, 1974; Crump, 1981a; Krewski and Van Ryzin, 1981). When adequate human data are available, they are used in preference to animal data for quantitative risk estimation. For human data, the best fit to the dose–response data is used to extrapolate from high doses to low doses (EPA, 1980b; CAG, 1982). Negative epidemiology data are used to set upper bounds on risks.

The EPA has extensive experience with the use of quantitative risk estimation as a basis for making public health policy decisions. Since other Federal agencies in the United States use similar approaches, the EPA experience is presented in some detail in this paper to illustrate quantitative approaches used in the United States. The Appendix provides a detailed discussion of the quantitative risk estimation models being used by the EPA, as previously published (EPA, 1980b).

2 DEVELOPMENT OF QUANTITATIVE RISK ASSESSMENT APPROACHES IN THE UNITED STATES

The EPA was established by executive order in December 1970. Soon afterwards, a series of actions commenced which involved the Agency in the evaluation of carcinogenesis data and the translation of these evaluations into public policy. A

brief history provides the perspective for the current EPA policy, which involves an internal process for qualitative and quantitative cancer risk assessment.

Between December 1970 and mid-1975, the EPA moved to suspend and cancel most uses of several major pesticides: DDT, aldrin and dieldrin, and chlordane and heptachlor. At the time it took these actions, the Agency lacked internal procedures for assessing the risks associated with the use of these pesticides. Instead, much of the information that focused on these potential risks came from sometimes conflicting evaluations that had been conducted by scientists outside the EPA; these evaluations were mostly qualitative. In large part, the full scientific evaluation occurred during the subsequent administrative hearings in testimony by expert witnesses called by the EPA and the registrants. In short, much of the scientific information was assimilated as a part of an adversative process and had to be summarized in legal briefs at the conclusion of the hearings.

In summarizing the testimony of their expert witnesses in several litigations, the attorneys for the EPA set forth certain summary statements which, in the legal motion, were referred to as cancer principles (Respondents brief, 1972; Respondents motion, 1975). This triggered a widespread perception that these summary statements represented the Agency's cancer policy. Because of this perception, these so-called 'cancer principles' received broad and general criticism by the scientific community, a substantial part of the private sector, and the Congress (e.g., see *The Lancet*, 1976). The major thrust of the criticism was not so much that these statements were incorrect, as it was that such a complex field as carcinogen assessment cannot be adequately covered in summary statements of this nature.

More specifically, there was a widespread concern that the Agency would simply regard all agents associated with the induction of cancer in animals as equally likely to be potential human carcinogens; treat all such agents as if they had the same potency; and regulate exposures, in the absence of a threshold or information about degree of risk, toward zero risk, so far as possible. To some this meant a Delaney approach (zero risk) for literally hundreds of environmental agents even though the EPA Administrator, in his decisions on DDT, aldrin/dieldrin, and chlordane/heptachlor, did not adopt a zero risk position, but rather attempted to qualitatively balance risks and benefits for each use.

The impracticality of aiming toward zero risk on a broad scale for a large number of economically important agents is apparent. Also, when one reviews the authorities the EPA inherited in a series of laws passed in the 1970s that deal with the control of environmental pollutants, including carcinogens, it is apparent that some basis for setting priorities is needed. The EPA authority covers eight areas: air pollution (Clean Air Act), pesticides (Federal Insecticide, Fungicide, and Rodenticide Act, FIFRA), pollution of water bodies (Federal Water Pollution Control Act), drinking water (Safe Drinking Water Act), toxic substances (Toxic Substances Control Act), hazardous wastes (Resource

Conservation and Recovery Act), uncontrolled waste sites (Comprehensive Environmental Response, Compensation, and Liability Act, usually referred to as 'Superfund') (Government Institutes, 1983), and ionizing radiation (under several authorities). In addition, four other major US regulatory agencies are also charged with regulating carcinogens under different authorities. These include: the Food and Drug Administration (FDA) (under the Food, Drug, and Cosmetic Act), the Consumer Product Safety Commission (CPSC) (under the Consumer Product Safety Act), the Occupational Safety and Health Administration (OSHA) (under the Occupational Safety and Health Act), and the US Department of Agriculture (USDA), which has some responsibilities for regulating food safety. Considering the large number of chemicals to which people are exposed, many of which have shown carcinogenic activity in laboratory animal tests, some approach was called for to determine the magnitude of the risks, as a basis for setting priorities and balancing risks against social and economic benefits.

Between 1976 and 1980, in order to provide information regarding the degree of risk associated with different levels of exposure, several US Federal regulatory agencies adopted the use of risk assessment in making health policy decisions. The EPA published guidelines for carcinogen risk assessment and established a senior health committee in the Agency to assess cancer risk (EPA, 1976; Albert *et al.*, 1977). The scientific basis for the EPA approach was consistent with recommendations from the advisory group to the National Cancer Institute, which published their report about the same time (NCABSEC, 1977).

The FDA also used risk assessment approaches in a series of decisions involving food contaminants, drugs, and cosmetics, although the FDA did not adopt guidelines for risk assessment. One example is found in an application where the FDA retracted a regulation relying on the log-probit model to establish contaminant limits for carcinogenic agents in animal foodstuffs and replaced it with a proposed regulation using the linear nonthreshold model (FDA, 1973, 1977, 1979). The FDA also used the linear model as a basis for deciding allowable limits of aflatoxin in peanut products and for permitting the use of the suspected carcinogen, lead acetate, in hair dyes (FDA, 1978, 1980). *The New York Times* carried an editorial entitled 'A Carcinogen Passes' to point out the reasonableness of an approach that recognized the insignificance of very low risk levels (*The New York Times*, 1980b).

OSHA adopted a cancer policy that mentioned possible limited use of quantitative risk assessment (OSHA, 1980). Following the Supreme Court decision on the OSHA benzene standard, however, OSHA now seems legally bound to provide at least some quantitative analysis to estimate improvements (i.e., reduced risk) in worker health associated with more stringent standards (U.S. Supreme Court, 1980). A *New York Times* editorial entitled 'So It's A Carcinogen, But How Bad?' endorsed the utility of quantitative measures of risk in deciding the extent of risk associated with potential carcinogens to set public

policy (*The New York Times*, 1980a). This media coverage, albeit limited, indicates at least the initiation of public awareness and interest in the extent of cancer risk.

Major uncertainties are inherent in the quantitative risk assessment process, for example, uncertainties associated with high- to low-dose extrapolation and extrapolation from animal to man; and only rarely is information available concerning synergistic interactions or risks in particularly susceptible groups. These limitations have led some to oppose the use of risk quantitation in the policy process. For example, an article following the Supreme Court decision on benzene cited limitations in quantitative risk assessment as a reason for opposing the use of quantitative assessment (Doniger, 1980). More recently, Weinhouse voiced similar concerns in his presidential address before the American Association for Cancer Research (Weinhouse, 1982). Nevertheless, EPA, given its regulatory responsibilities, has felt it imperative to use quantitative assessment despite admitted problems. The Agency's experience, summarized in this paper, can provide a realistic basis for evaluating the reasons for and against using quantitative risk estimation.

Seeing the need for a common approach, major US regulatory agencies joined in writing a single document to address the issues involving the identification of carcinogens and estimation of risk (IRLG, 1979). This document emphasized the importance of carefully evaluating all the positive and negative biomedical evidence for carcinogenicity and presenting the strength of this evidence clearly, whether or not quantitative estimates of risk are also presented. The document also discussed available extrapolation models for estimating cancer risk and suggested the use of the linear nonthreshold model, when only one model is selected. An effort is currently underway, chaired by the President's Office of Science and Technology Policy, to provide an updated Federal document on these issues.

3 QUANTITATIVE APPROACHES IN USE IN THE UNITED STATES WITH SPECIFIC REFERENCE TO APPLICATIONS IN THE US ENVIRONMENTAL PROTECTION AGENCY

3.1 Historical Perspective

In 1976 the EPA became the first agency to adopt guidelines for scientific evaluation of cancer risks and, further, to state that gains in public health (i.e., reductions in risks) would be balanced against social and economic concerns in making regulatory decisions, to the extent permitted by the enabling legislation (previously, risks and benefits had been balanced in all decisions involving pesticide uses, as required under FIFRA). The most experience to date in using risk assessment in the regulatory process has been in the area of carcinogenesis,

although the EPA has proposed guidelines for applying risk assessment approaches to other biological effects, such as mutagenicity. This work on mutagenicity and other effects continues (EPA, 1980a,c).

In 1976, the EPA adopted a two-step approach to risk assessment (EPA, 1976; Albert *et al.*, 1977). Risk assessment was defined as a process to answer two questions:

(1) how likely is an agent to be a human carcinogen?; and
(2) on the assumption that it is, what is the magnitude of public health impact given current and projected exposures?

Since only rarely do we know that an agent is indeed a human carcinogen, the first step involves an evaluation of all the relevant biomedical data to determine the weight of evidence that an agent might be a human carcinogen. The second step involves the quantification of risk and public health impact in terms of rough estimates of number of excess cancers for current exposures as well as estimated exposures for various regulatory options.

To answer the first question regarding likely carcinogenicity, the biomedical evidence may range from the strongest evidence, based on human data backed up by animal bioassay results, to substantial evidence provided by positive results from animal bioassay tests, to suggestive or supporting evidence provided by positive results from short-term, *in vitro* tests. The exact nature and extent of evidence, however, cannot be simply assigned to broad categories, such as the ones mentioned above, but rather each case must be judged individually by experts. These judgments are based on an evaluation of the relevant biomedical data, both negative and positive, to determine the likely carcinogenic potential. In the EPA, these evaluations have always been summarized in sections which discuss the weight of evidence for carcinogenicity. In addition, at one time, the Agency assigned the positive evidence to one of three broad categories:

(1) strongest evidence—positive epidemiologic results, backed up by animal data;
(2) substantial evidence—the broad range of positive results from animal bioassays; and
(3) suggestive evidence—positive short-term test results or borderline animal or human results.

Because these labels were being applied without full appreciation of detailed evaluation, the EPA dropped the labels, and since then it has relied primarily on summaries to present the data in terms of the nature of responses, the quality and extent of data, evaluation of both positive and negative responses, and other relevant factors, with the understanding that the regulators would take this weight of evidence into account when making regulatory decisions. To date the Agency has not adopted criteria for stratifying the weight of evidence.

The second step, that of providing quantitative estimates of public health

impact, brackets risks between an upper bound and a lower bound approaching zero. The upper bounds of risk are expressed both in terms of the individual increased cancer risks in exposed subpopulations (i.e., increased risks of, for example, one chance in a thousand or 1×10^{-3}) and the nationwide impact in terms of annual increase of cancer cases. This second quantitative step is intended to give the regulators a feeling for the potency of the suspected carcinogen and some quantitative information regarding public health impact. Uncertainties associated with the high- to low-dose risk extrapolation and extrapolation from animal to man, make it impossible to determine the risk more precisely. Nevertheless, since the potency of carcinogens can cover a range of 50 million or more, it seems important to make some attempt to take this disparity into account in making public policy decisions. The upper bound is calculated using reasonably conservative exposure estimates and the linear nonthreshold model at low doses. Since the dose–response curve at low doses is unlikely to be concave downward, the linear, nonthreshold dose–response curve is regarded by most scientists in the United States as usually setting a plausible upper bound on risks, that is, the risks are not likely to be higher. The plausibility of upper-bound estimates derived from the linear nonthreshold model is based on the correlation between carcinogenicity and mutagenicity; the nonthreshold dose–response curve for mutagenicity (in most cases); the quantal nature of DNA interactions; and the linear dose–response curves suggested by some epidemiology data, for example, on aflatoxin, radiation and cigarette smoking (see Appendix). However, the linear model could be unduly conservative if an agent exhibits either a concave curve or a threshold at low doses (Brown, 1980). In the absence of information to define mechanisms of action at low doses and interspecies correlation, extrapolation from high doses to low doses can define the risk only within rough bounds. Generally, no attempt is being made to further determine the risk within the broad bounds defined at the upper bound by the linear nonthreshold model and at the lower bound as approaching zero. This recognition that the lower bound may approach zero or be indistinguishable from zero is due to uncertainties associated with mechanisms of carcinogenesis and with the extrapolation from high doses to low doses, and from animal to man. Most often there is no biological justification to support the choice of any one model to estimate the actual risk or make point estimates of risks. Where data are available at low doses, the model that best fits the data should be used. A variety of models, however, can be used to fit the data in the observed range and to give low-dose point estimates of risk that most often will fall within these bounds, which is consistent with the above approach. Even if point estimates of risk cannot be calculated, all available information should be used to discuss the plausibility of the upper bound to establish point estimates of risk, if possible. It should be clear from this discussion that the linear nonthreshold model has been used by the EPA to place plausible upper bounds on risk, and not to establish point estimates.

Before November 1980, the CAG used the one-hit model to estimate upper-bound cancer risk from responses in animal bioassays. In response to public comments on the proposed Water Quality Criteria for Suspected Carcinogens, the EPA changed from the one-hit model to the linearized multistage model (see Appendix) to estimate the upper bound of cancer risk (EPA, 1979a,b,c, 1980b). Where human data are available, the curve best fitting the data in the observed range is selected and then extrapolated to low dose using the linear nonthreshold model. The appendix to this paper presents a detailed description of the quantitative risk assessment models used by the EPA (EPA, 1980b). A comparison of estimates using the one-hit model and the linearized multistage model shows a close agreement except in cases of steeply rising dose–response curves. When dose–response curves rise steeply, the new multistage model will provide a slope lower by as much as five times for the comparisons made (EPA, 1980b; Crump, 1981b).

Upper-bound risk estimates have inherent uncertainties and must be used with caution. However imprecise, these quantification approaches represent the best scientific tools currently available to estimate risk. The alternative is to provide no quantitative risk information to the policy process, which generally means that the level of health protection will be decided by definitions of feasibility, best available technology, and so on, all of which have considerable uncertainties and may lead to underprotective health policies or requirements to reduce trivial risks at incommensurately high costs.

The EPA risk assessment approach was certainly experimental at the time it was adopted. In practice, it has provided a conceptual basis for balancing risks against social and economic benefits. Also, quantitative risk estimation has provided an alternative to the approach aiming toward zero risks or exposures, where actual acceptable levels must be defined solely in terms of feasibility for a large number of agents introduced into the environment, and important social and economic reasons. The following section of this document presents some of these examples.

3.2 Applications of Quantitative Risk Assessment to Public Policy Decisions

Practical considerations of public health issues propelled scientists to explore methods for expressing the degree of risk associated with various exposure levels for those health effects where no 'safe level of exposure' can be identified. In most cases, risk estimation methods have been developed to answer questions of a compelling and practical nature. The following discussion provides examples of the use of risk estimates to provide information for a range of policy issues facing the EPA.

As discussed earlier, the qualitative component of risk assessment is expressed in terms of the weight of the biomedical evidence that a chemical may be a

potential human carcinogen. On the assumption that the chemical is a human carcinogen, the quantitative estimates aim at determining a plausible upper bound of risk, recognizing that the risk could approach zero if the chemical has a threshold or a concave dose–response curve.

Quantitative risk assessment, together with qualitative assessment of the biomedical evidence, has been used by the EPA for the five distinct purposes when considering public policy:

(1) to set priorities,
(2) to review residual risk after application of best available technology to see if anything more needs to be done,
(3) to balance risks against benefits,
(4) to set standards and target levels of risk, and
(5) to provide information regarding the urgency of situations where subpopulations are inadvertently exposed to toxic agents, for example, populations near uncontrolled waste sites.

Several examples of these applications are discussed below.

Under provisions of the Clean Air Act, the EPA must 'list' hazardous air pollutants and regulate sources as necessary. In order to set priorities for reviewing hundreds of agents that may be potential air pollutants, the EPA Office of Air Programs identified three groups of chemicals suspected of being present in the ambient air at levels of concern (Tables 1 and 2). Highest priority for health evaluation was given to group I, then group II, and finally group III. These priorities reflected judgments in the air office regarding those chemicals which, based on preliminary information about likely exposure and possible toxicity, might present the greatest hazard to humans from air pollution. The Carcinogen Assessment Group (CAG), one of the health subgroups in the Office of Health and Environmental Assessment, provided a qualitative weight-of-evidence statement and an index of potency expressed as an upper-bound unit risk estimate (Table 3). The unit risk estimate is the increased individual lifetime risk for a 70 kg individual breathing air containing 1 $\mu g/m^3$ of the chemical for a 70-year life-span. Notice that the potency index ranges a millionfold and that chemicals having the strongest biomedical evidence for carcinogenicity based on responses in humans may have relatively low potencies; for example, vinyl chloride with a unit risk of 10^{-6}, and benzene with a unit risk of 10^{-6}. These examples contradicts the notion that strong evidence of carcinogenicity must mean high potency as well.

In the absence of information regarding potency, regulators are inclined to regulate known human carcinogens more severely than animal carcinogens, even though some human carcinogens appear to be relatively much less potent than some chemicals whose carcinogenesis has only been demonstrated in animal studies. The weight of evidence for carcinogenicity, the unit risk estimates as a measure of potency and information concerning exposure levels have provided a

Table 1 Chemicals proposed by the Environmental Protection Agency Office of Air Programs for Unit Risk Assessment*

Group I	Group II
Acrylonitrile	Beryllium
Carbon tetrachloride	Cresols (*ortho*, *meta*, and *para*)
Chloroform	Formaldehyde
Ethylene dibromide	Maleic anhydride
Ethylene dichloride	Manganese
Ethylene oxide	Methyl chloroform
Nitrosamines (4)	Methylene chloride
Perchloroethylene	Nickel
Trichloroethylene	Nitrobenzene
Vinylidene chloride	Toluene
	Xylenes (*ortho*, *meta*, and *para*)

* Unit risk is excess lifetime risk associated with breathing 1 $\mu g/m^3$ of the chemical over a 70-year life-span for a 70 kg person.

Table 2 Chemicals proposed by the Environmental Protection Agency Office of Air Programs for Unit Risk Assessment*

Group III	
Acetaldehyde	Dioxane
Acetylene tetrachloride	Epichlorohydrin
Acrolein	Hexachlorocyclopentadiene
Allyl chloride	Methyl iodine
Benzyl chloride	Naphthylamine (1- and 2-)
Chloromethyl ether (BCME)	2-Nitropropane
Chlorobenzene	Phenol
Chloromethylmethyl ether	Phosgene
Chloroprene	Polychlorinated biphenyls (PCBs)
Dichlorobenzene (*ortho* and *para*)	Propylene oxide

* Unit risk is excess lifetime risk associated with breathing 1 $\mu g/m^3$ of the chemical over a 70-year life-span for a 70 kg person.

basis for selecting the most hazardous air pollutants for further study.

After an agent has been listed as a hazardous air pollutant, the EPA must decide which sources to regulate first, and indeed whether or not regulation is warranted. Table 4 presents a comparison of data for different source categories contributing arsenic to the ambient air. In presenting upper-bound risk to subpopulations and related upper-bound nationwide impacts, it is always necessary to rely on estimates of exposures which also have great uncertainties. Uncertainties associated with exposure estimates must always be included in

Table 3 Upper-bound unit calculations for suspected carcinogenic air pollutants*†‡

Chemical	Upper-bound unit risk estimates
Acrylonitrile	1.5×10^{-5}
Allyl chloride	5.5×10^{-8}
Arsenic	3.4×10^{-3}
Benzene	7.4×10^{-6}
Berylium	2.7×10^{-1}
Diethylnitrosamine (DEN)	1.6×10^{-2}
Dimethylnitrosamine (DMN)	5.4×10^{-3}
Dioxin¶ (2,3,7,8-tetrachloro-)	1
Ethylene dibromide	6.3×10^{-5}
Ethylene dichloride	4.2×10^{-6}
Ethylene oxide	1.8×10^{-4}
Formaldehyde	6.1×10^{-6}
Manganese	3.5×10^{-4}
Nickel	1.8×10^{-3}
N-nitroso-*N*-ethylurea (NEU)	1.2×10^{-2}
N-nitroso-*N*-methylurea (NRU)	6.7×10^{-1}
Perchloroethylene	7.6×10^{-6}
Trichloroethylene	2.7×10^{-6}
Vinyl chloride	4.1×10^{-6}
Vinylidene chloride	3.0×10^{-5}

* From US Environmental Protection Agency, Carcinogen Assessment Group Reports (EPA 1976–1981). These calculations are periodically revised as new data become available.

† The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various exrapolation and mathematical calculations.

‡ Unit risk is excess lifetime risk associated with breathing 1 $\mu g/m^3$ of the chemical over a 70-year life-span for a 70 kg person.

¶ The potency of dioxin is about 6000 times greater than for DEN; therefore, for lifetime exposure to 1 $\mu g/m^3$, the upper-bound unit risk is 100% chance of cancer occurrence. The upper-bound estimate of the potency (slope) for dioxin is $91/(\mu g/m^3)$.

the exposure assessment and taken into account in using risk assessment information.

In another example, risk assessment was used to compare residual risk, after application of best available technology to control ambient levels of vinyl chloride, with risk associated with other potential hazardous air pollutants that had not yet been fully reviewed (see Tables 5 and 6). The risk assessment information indicated that reductions in risk had been considerable for vinyl chloride and that the remaining risk was low relative to risks associated with other air pollutants that had not yet been completely studied for their hazardous

Table 4 Upper-limit lifetime cancer risk for arsenic exposures*†

Source	Number exposed in highest two groups‡	Highest two exposure levels ($\times 10^{-4}$ mg/kg/day)¶	Associated lifetime upper-bound cancer risk	Upper-bound estimates/cases per year
Copper smelters	43 800	2.7–1.5	$2.4–1.3 \times 10^{-3}$	1.5–0.821
Lead smelters	3400	0.69–0.27	$6.0–2.4 \times 10^{-4}$	0.029–0.017
Zinc smelters	37 000	0.69–0.27	$6.0–2.4 \times 10^{-4}$	0.32–0.13
Cotton gins	32	15.4–6.9	$13.4–6.0 \times 10^{-3}$	0.0061–0.0027
Pesticide manufacturing	1480	0.026–0.014	$2.2–1.2 \times 10^{-5}$	0.0004–0.00025
Glass manufacturing	11 580	0.69–0.014	$6.0–2.4 \times 10^{-4}$	0.099–0.040

* From Table 6 of the US Environmental Protection Agency, Carcinogen Assessment Group's Final Risk Assessment on Arsenic, May 2, 1980, National Technical Information Service, PB 81-206013 (EPA, 1980).

† The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various extrapolation and mathematical calculations.

‡ Population exposed to ambient levels of arsenic from the sources listed.

¶ For example, the highest exposure level for copper smelters is 2.7×10^{-4} mg/kg/day.

potential (see individual risks and nationwide impacts for arsenic and benzene, Tables 5 and 6). The Office of Air Quality Programs allocated Agency resources to consider other air pollutants and not to further reduce risks associated with vinyl chloride emissions. To date, vinyl chloride has not been further regulated.

Many decisions involving the balancing of risks and benefits under EPA's pesticide registration authorities have relied on risk assessment. Table 7 presents the quantitative risk estimates associated with three examples for which registration decisions have been made. In the case of chlorobenzilate, a pesticide used on citrus fruit, the weight of evidence for carcinogenic potential is based only on responses in the liver of both male and female mice; studies in rats were negative. On the assumption that chlorobenzilate is a human carcinogen, quantitative risk estimates indicate that risk associated with exposure of the general population is relatively low, of the order of one chance in a million, and the annual cancer rate on a nationwide basis is also relatively low. However, the risk to applicators of the pesticide was higher by two orders of magnitude. Since the pesticide act (FIFRA) requires the balancing of risks and benefits, the risk in applicators was evaluated assuming that no substitute was available for chlorobenzilate application on citrus. The EPA decided that the risks did not outweigh the benefits and therefore retained the registration of chlorobenzilate

Table 5 Comparison of upper-bound risks associated with ambient exposure to carcinogenic air pollutants*

In chemical†	Upper-bound lifetime probability of cancer death due to maximum exposure near stationary sources‡	Total number exposed‡§	Total number of cancer deaths/year at the upper-bound in US due to chemical in air‡
Arsenic	$2.4–1.3 \times 10^{-3}$	44 000	1.5–0.8
Benzene	$2.4–1.0 \times 10^{-4}$	55 000	0.1
Coke ovens	6.0×10^{-3}	1800	0.2
Vinyl chloride¶			
Before Regulation	3.8×10^{-3}	34 000	1.9
After Regulation	1.9×10^{-4}	34 000	0.1

* From the US Environmental Protection Agency, Carcinogen Assessment Group Reports (EPA, 1976–1981). These estimates may change as additional data become available.

† All risks are before regulations unless otherwise indicated.

‡ The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various extrapolation and mathematical calculations.

¶ If risks were based on the incidence of mammary tumors in the animal bioassay studies, the results would be four times higher.

§ Population exposed to ambient levels of chemicals listed. Exposure is from stationary air sources.

Table 6 Comparison of upper-bound risks associated with ambient exposure to carcinogenic air pollutants*

In chemical†	Upper-bound lifetime probability of cancer death due to average exposure near stationary sources‡	Total number exposed‡§	Total number of cancer deaths/year at the upper-bound in US due to chemical in air‡
Arsenic	4.4×10^{-5}	25 million	15.6
Benzene	2.5×10^{-5}	220 million	77.8
Coke ovens	7.0×10^{-4}	15 million	149.5
Vinyl chloride¶			
Before regulation	2.0×10^{-4}	5 million	20.0
After regulation	1.0×10^{-5}	5 million	1.0

* From the US Environmental Protection Agency, Carcinogen Assessment Group Reports (EPA, 1976–1981). These estimates may change as additional data become available.

† All risks are before regulations unless otherwise indicated.

‡ The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various extrapolation and mathematical calculations.

¶ If risks were based on the incidence of mammary tumors in the animal bioassay studies, the results would be four times higher.

§ Population exposed to ambient levels of chemicals listed. Exposure is from stationary air sources.

for use on citrus. The Agency added labelling requirements to further protect the applicators.

The next case in Table 7 involved the application of risk assessment to registration of a new pesticide Amitraz (BAAM) for use on pears and apples. Only one carcinogenesis bioassay had been performed, and it provided very weak evidence of carcinogenic activity. Nevertheless, on the assumption that BAAM is a carcinogen, calculated upper-bound risk estimates indicated relatively low estimated risk for the US population, of the order of one in a million. Balancing risks against benefits, the EPA made a decision:

(1) to permit a 3-year temporary registration of BAAM for use on pears but not apples, because substitutes were not available for pears but were available for apples; and
(2) to require submission of more definitive data before granting a permanent registration for any use.

In the final example in Table 7, risk assessment was used to balance risks and benefits for registered uses of chlordane/heptachlor. The biomedical evidence for these chemicals is reasonably strong based on liver carcinomas observed in a series of rodent bioassays. Further, these chemicals bioaccumulate and most humans carry a body burden of these chemicals in adipose tissue. Application of

Table 7 Upper-bound risk estimates for population exposure to suspected carcinogenic pesticides*

Pesticide	Population exposed	Upper-bound lifetime probability of cancer death due to exposure†	Number of expected cancer deaths/year at the Upper-bound‡
Chlorobenzilate	220 million—citrus consumption	2.1×10^{-6}	6.7
	—citrus applicators§	4.2×10^{-4} to 1.4×10^{-3}	—
Amitraz (BAAM)	220 million—apple consumption	2.5×10^{-6}	8
	220 million—pear consumption	2.0×10^{-6}	6
	1429 applicators—spraying apples	1.2×10^{-4}	0.002
	1545 applicators—spraying pears	6.4×10^{-5}	0.001
	1674 applicators—spraying pears	1.3×10^{-4}	0.003
Chlordane/heptachlor	220 million	1.6×10^{-4}‡	500†
		5.0×10^{-5}¶	150‡

* US Environmental Protection Agency, Carcinogen Assessment Group Reports (EPA, 1976–1981). These estimates may change as additional data become available.

† The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various extrapolation and mathematical calculations.

‡ Based on total tumors.

¶ Based on large carcinomas.

§ The total number of applicators was not included in the study.

quantitative risk assessment indicated risks at least one order of magnitude higher than the previous two cases presented in Table 7; considerable potential nationwide impact was also projected. The decision in this case was to cancel most uses of chlordane/heptachlor with the exception of underground applications for termite control.

Table 8 presents projected risks associated with the resumed manufacture of nitrilotriacetic acid (NTA) in the United States. This risk assessment was done because the manufacturer asked EPA for guidance as to whether or not the EPA would regulate NTA if the manufacture was resumed (the manufacture of NTA had been voluntarily suspended in the early 1970s because of early indications in animal bioassays that NTA might be a carcinogen). NTA is used in detergents to replace phosphates that contribute to eutrophication of water bodies. These risk estimates, presented in Table 8, are based on monitoring data from Canada, where NTA has been in continuous use for several years. With the exception of private wells, where only 21 samples had been analyzed, cancer risks estimated from the Canadian data indicated a low projected US risk calculated as the upper bound. Although questions were raised about the adequacy of the Canadian exposure data to estimate US exposures, the decision not to regulate the resumed manufacture of NTA cited these relatively low-risk estimates as the reason.

The final application of risk assessment is presented in Table 9, illustrating the use of quantitative risk assessment to set target levels of risk (EPA, 1980b). In this application, the EPA was obligated to recommend nationwide water quality criteria for a large number of chemicals, including suspected carcinogens. The statute under which these criteria were issued, the Federal Water Pollution Control Act, required that numerical limits be published by the Agency to protect the public health; no provisions are included in this section of the statute to incorporate social and economic concerns in setting water quality criteria. Since thresholds could not be established for suspected carcinogens, quantitative risk assessment was used to recommend concentrations of contaminants associated with risk in the range from 10^{-7} to 10^{-5} as the upper bound. These concentrations were calculated by assuming consumption of two liters of water per day and a standard intake of fish, 6.5 g per day of edible portion, to include bioaccumulation. The slope, presented in the second column of Table 9, is calculated using the linearized multistage model (Appendix). In the proposed criteria, the linear model was used to calculate the concentrations associated with a risk of 10^{-5}. In response to public comments, the Agency reviewed alternative models and decided to adopt the linearized multistage model in order to make full use of all the data points. The slope and the concentrations (in parentheses) in Table 9, were calculated using the one-hit model (EPA, 1979a,b,c). Obviously, the weight of the biomedical evidence varies enormously for the chemicals presented in Table 9, and this information should not be ignored in applying these target concentrations to local situations where the regulatory process of permitting discharges actually takes place.

Table 8 Upper bound projected lifetime cancer risk based on one-hit model from NTA exposure-response*†

Type of exposure	Number exposed	Exposure level‡ mg/kg/day	Associated cancer risk at the upper bound	Cancer cases/year at the upper bound
Public drinking water	220 million	8.0×10^{-5}	3.7×10^{-7}	1.2
(Range)		7.0×10^{-4}	3.2×10^{-6}	10.1
(Mean)		4.0×10^{-5}	1.9×10^{-7}	0.6
Private wells (Max)	66 million	up to 0.086 (insufficient data)	3.9×10^{-4} (insufficient data)	377 (insufficient data)
General				
Consumers				
Laundry	125 million	2.1×10^{-4}	1.0×10^{-6}	1.8
Dishwashing	125 million	2.1×10^{-4}	1.0×10^{-6}	1.8
Residue on unrinsed dishes	2 million	0.014	6.4×10^{-5}	1.8
Workers/	100	1.4×10^{-3}	6.4×10^{-6}	—
Manufacture		7.1×10^{-3}	3.3×10^{-5}	—
Formulations	1750	4.8×10^{-3}	2.2×10^{-5}	< 0.001
		4.8×10^{-2}	2.2×10^{-4}	

* From US Environmental Protection Agency, Office of Toxic Substances Report, Washington, DC (unpublished).

† The significant figures presented do not indicate precision or accuracy, but rather they are included to make it easier to trace the derivation of these numbers through the various extrapolation and mathematical calculations.

‡ Projected US exposures based on Canadian monitoring data.

Table 9 Guidance for water quality criteria: upper bound calculations with a lower bound approaching zero*

Chemical	Upper bound slope B_H $(mg/kg/day)^{-1}$†	Water concentrations corresponding to a risk level of 10^{-5} ($\mu g/l$) at the upper bound†§
Acrylonitrile	0.55 (2.045)	0.579 (0.084)
Aldrin	11.44 (6.349)	7.4×10^{-4} (4.6×10^{-5})
Arsenic¶	14.00	0.022
Asbestos	—	300 000 (fibers/l) (0.05)
Benzene¶	0.052	6.60
Benzidine¶	234.13	1.2×10^{-3}
Beryllium	8.84 (3.431)	0.037 (0.087)
Carbon tetrachloride	0.083 (0.091)	4.0 (2.6)
Chloroform	0.18 (0.15)	1.89 (2.1)
Chlordane	1.61 (5.36)	4.60×10^{-3} (1.20×10^{-3})
Chloroalkyl ethers		
BCME	9299.76 (13 600)	3.80×10^{-5} (2.0×10^{-5})
BCEE	1.14 (0.685)	0.30 (0.415)
Chlorinated benzenes		
HCB	1.68 (2.48)	7.2×10^{-3} (1.25×10^{-3})
Chlorinated ethanes		
1,2-di-	0.037 (0.048)	9.42 (7.0)
1,1,2-tri-	0.057 (0.123)	6.01 (2.7)
1,1,2,2-tetra	0.20 (0.164)	1.71 (1.8)
Hexa-	0.014 (0.015)	19.2 (5.9)
Dichlorobenzidine	1.69 (1.90)	0.10 (0.016)
DDT	8.42 (18.06)	2.4×10^{-4} (5.3×10^{-4})
Dichloroethylenes		
1,1-dichloroethylene	1.04 (0.25)	0.33 (1.3)
Dieldrin	30.37 (183.6)	7.1×10^{-4} (4.4×10^{-5})
Dinitrotoluene	0.31 (0.42)	1.11 (0.79)
Dioxins		
2,3,7,8-Tetrachlorodioxin	4.25×10^{5} (1.39×10^{4})	2.1×10^{-9} (4.6×10^{-7})
Diphenylhydrazine	0.77 (0.715)	0.40 (0.40)
Halomethanes	Same as chloroform	
Heptachlor	3.37 (30.31)	2.8×10^{-3} (2.3×10^{-4})
Hexachlorobutadiene	0.078 (0.049)	4.47 (0.77)
Hexachlorocyclohexane		
technical grade	4.75 (2.0)	0.052 (0.021)
alpha isomer	11.12 (2.66)	0.022 (0.016)
beta isomer	1.84 (1.513)	0.13 (0.028)
gamma isomer	1.33 (0.78)	0.185 (0.054)
Nitrosamines		
DMNA	25.88 (13.4)	1.4×10^{-2} (2.6×10^{-2})
DENA	43.48 (38.2)	8.0×10^{-3} (9.2×10^{-3})
DBNA	5.43 (26.86)	0.064 (0.013)

Table 9 (*Contd.*)

Chemical	Upper bound slope B_H $(mg/kg/day)^{-1}$†	Water concentrations corresponding to a risk level of 10^{-5} ($\mu g/l$) at the upper bound†§
N-*N*-P	2.13 (3.93)	0.160 (0.11)
PAH	11.53 (28.02)	2.8×10^{-2} (9.7×10^{-3})
PCBs	4.43 (3.25)	7.9×10^{-4} (3×10^{-4})
Tetrachloroethylene	0.039 (0.084)	8.0 (2.0)
Trichloroethylene	0.013 (0.012)	26.9 (21)
Toxaphene	1.08 (4.42)	7.1×10^{-3} (4.7×10^{-4})
Vinyl chloride¶	0.017	20

* *Federal Register* **45** (November 28) 1980, 79318–79379. This Water Quality Criteria guidance may be revised as new data become available.

† The parenthetical values were calculated using the one-hit model as originally proposed. In response to public comment, these final calculations are derived from the linearized multistage model.

§ Assuming a lifetime daily consumption of 2 liters of water and 0.0065 kg fish. Note that a daily consumption of 0.0187 kg fish was assumed in the original calculation, and some of the bioconcentration factors used in the new calculations are different from original calculations as proposed.

¶ Slope determined from epidemiological data.

These examples illustrate the application of quantitative risk assessment in a variety of practical circumstances to provide information on risk as a basis for making public health policy decisions in the United States. Between 1976 and 1980, the linear nonthreshold model at the lower end of the dose–response curve was applied to hundreds of cases, some of which are presented above, to assist policymakers in deciding how much cost, in social and economic terms, should be expended to reduce the risk to some reasonably low level. Although these policy decisions did not identify any acceptable level of risk, because each decision reflected feasibility in some measure, most often decisions fell close to an increased individual risk of about 10^{-6} as the upper bound. There were some circumstances where this level of risk was not achievable, for example, in setting haloform standards for drinking water (EPA, 1978). Decisions to accept risks higher than 10^{-5} were generally justified on grounds of social and economic tradeoffs. Risks lower than 10^{-5} generally went unregulated, as exemplified by the NTA decision, current acceptance of the residual vinyl chloride risk, and the acceptance of risk in the chlorobenzilate decision. Exceptions include the voluntary cancellation of safrole as a dog repellent (risk 10^{-7}) because of low benefits; required reduction of nitrosamine contamination in treflan (risk 10^{-7}) (EPA, 1982); and recommendation of water quality criteria associated with risks ranging from 10^{-7} to 10^{-5} (EPA, 1980b). In a large number of risk assessments

on different chemicals performed during this period, probably as high as 80–90% of the cases studied, the upper-bound risks fell into a relatively low risk range of $< 10^{-5}$. Uncertainties in exposure estimates and other uncertainties inherent in the extrapolation process need to be taken into account on a case-by-case basis. Despite these deficiencies, the use of upper-bound estimates to identify those cases where the risk may be so low, even at the upper bound, as to fall in a low priority category for regulatory consideration, as was deemed the case with NTA, has provided some basis for regulators to focus attention on the most compelling public health problems.

4 REFINING RISK ASSESSMENT APPROACHES: FUTURE TRENDS

Carcinogen risk assessment has provided the scientific basis for a range of policy decisions by Federal agencies responsible for the protection of public health in the United States. Quantitative risk estimates, generally expressed as upper-bound estimates using a linear nonthreshold model, coupled with the qualitative evaluation of the weight of the biomedical evidence, have provided policymakers with rough estimates of risk which have served well as a basis for setting priorities, balancing risks and benefits, and establishing the urgency associated with public health problems for subpopulations inadvertently exposed to suspected carcinogens, such as at uncontrolled waste sites. Quantitative upper-bound risk estimates have also been used to establish public health standards for suspected carcinogens. These standards may be unduly conservative for agents that have a concave dose–response curve or indicate a threshold at low doses. Because of this problem, quantitative risk assessment methods, in particular those that have been used to date in the United States for setting standards, are under review to see what, if anything, can be done to provide improved quantitative guidance for standard setting purposes where the upper bound may not be plausible.

In addition, the EPA experience over the last decade points out misunderstandings that seem to occur frequently when cancer risk assessments are applied in policy considerations. First, where quantitative estimates have been provided, there has been a tendency to use these risk estimates independently of the weight of biomedical evidence and then to treat all suspected carcinogens as if they are human carcinogens. An additional problem is the use of upper-bound estimates as actual estimates of risk, decoupled from important statements regarding uncertainties. Such misunderstandings can lead to errors in policy judgments.

Finally, as more information becomes available as a basis for establishing mutagenic potential for suspected carcinogens, there is an increasing interest in finding the best way to incorporate this information into guidance to establish protective policies for these substances.

Because of these and related issues, thinking in the EPA is currently focused on:

(1) methods for stratifying the weight of the biomedical evidence (in the qualitative assessment) to make it harder to ignore this information when using quantitative risk assessment;
(2) criteria for judging the weight of evidence for mutagenicity so that mutagenic potential can be more clearly expressed;
(3) possible approaches for making greater use of information about the mutagenic potential of suspected carcinogens, as well as information on reversibility, pharmacokinetics, etc., in the risk assessment process; and
(4) possible approaches for establishing quantitative guidance for chemicals where the upper-bound estimate may not be plausible.

While the outcome of this effort is uncertain, most likely the risk assessment approaches that have been used to date will continue to be used, but with some refinement. These will most likely involve stratification of the weight of evidence for carcinogenicity, greater use of information about mutagenic potential and other relevant information as it becomes available, and additional quantitative guidance, in addition to the upper-bound estimates, where the upper-bound estimates are less plausible.

ACKNOWLEDGEMENT

The entire Carcinogen Assessment Group contributed to this paper through their assessment activities from 1976 to 1982. Selected examples of their work, referenced herein, provided a data base for significant parts of the discussion that relate to the US Environmental Protection Agency. The views expressed in this article, however, are those of the author and not necessarily those of the Carcinogen Assessment Group or the US Environmental Protection Agency. Members of the Carcinogen Assessment Group are Roy E. Albert, Chairman, Robert McGaughy, Larry Anderson, Steven Bayard, David Bayliss, Chao Chen, Margaret Chu, Herman Gibb, Bernard Haberman, Charalingayya Hiremath, Dharm Singh, and Todd Thorslund.

5 REFERENCES

Albert, R. E., Train, R. E., and Anderson, E. L. (1977). Rationale developed by the Environmental Protection Agency for the assessment of carcinogenic risk. *J. natn. Cancer Inst.*, **58,** 1537–1541.

Brown, C. C. (1980). Learning about toxicity in humans from studies on animals. *ChemTech*, **13,** 350–358.

CAG (1982). *Carcinogen Assessment of Coke-Oven Emissions*. Carcinogen Assessment Group, US Environmental Protection Agency. EPA-600/6-82-003. January. Draft. Environmental Protection Agency, Washington, DC: 202 pages.

Chand, N., and Hoel, D. G. (1974). A comparison of models for determining safe levels of

environmental agents. In Proschan, F., and Serfling, R. J. (Eds.) *Reliability and Biometry, Statistical Analysis of Lifelength*, pp. 681–700. Society of Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Crump, K. S. (1981a). An improved procedure for low-dose carcinogenic risk assessment from animal data. *J. envir. Path. Tox.*, **5**, 675–684.

Crump, K. S. (1981b). Statistical aspects of linear extrapolations. In *Health Risk Analysis*, pp. 381–392. Proceedings of the Third Life Sciences Symposium. Franklin Institute Press, Philadelphia, Pennsylvania.

Doniger, D. D. (1980). Defeat in benzene exposure case no death knell for OSHA standards. *The National Law Journal* (September 15), 26–27.

EPA (Environmental Protection Agency) (1976). Interim procedures and guidelines for health risks and economic impact assessments of suspected carcinogens. *Federal Register* **41** (May 25), 21402.

EPA (1976–1981). Environmental Protection Agency internal reports from the Carcinogen Assessment Group. Available through Director of Carcinogen Assessment, Environmental Protection Agency, Washington, DC.

EPA (Environmental Protection Agency) (1978). Proposed interim drinking water standards for haloforms. *Federal Register* **43** (February 9), 5756–5780.

EPA (Environmental Protection Agency) (1979a). Water quality criteria; requests for comments (Proposed). *Federal Register* **44** (March 15), 15926–15981.

EPA (Environmental Protection Agency) (1979b). Water quality criteria: availability (Proposed). *Federal Register* **44** (July 25), 43660–43697.

EPA (Environmental Protection Agency) (1979c). Water quality criteria: availability (Proposed). *Federal Register* **44** (October 1), 56628–56656.

EPA (Environmental Protection Agency) (1980a). Mutagenicity risk assessments; proposed guidelines. *Federal Register* **45** (November 13), 74984–74988.

EPA (Environmental Protection Agency) (1980b). Water quality criteria documents; availability (Final). *Federal Register* **45** (November 28), 79316–79379.

EPA (Environmental Protection Agency) (1980c). Assessment of risks to human reproduction and to the development of the human conceptus from exposure to environmental substances. *Proceedings of EPA Symposium on Reproductive Development*, NTIS DE-82-007897.

EPA (Environmental Protection Agency) (1982). Determination concluding the rebuttable presumption against registration of trifluralin. *Federal Register* **47** (August 4), 33777.

FDA (Food and Drug Administration) (1973). Compounds used in food-producing animals. Procedures for determining acceptability of assay methods used for assuring the products of such animals (Proposal). *Federal Register* **38** (July 29), 19226.

FDA (Food and Drug Administration) (1977). Criteria and procedures for evaluating assays for carcinogenic residues (Final Order). *Federal Register* **42** (February 22), 10412.

FDA (Food and Drug Administration) (1978). Aflatoxins in shelled peanuts and peanut products used as human foods. Proposed tolerance: reopening of comment period. *Federal Register* **43** (March 3), 8808.

FDA (Food and Drug Administration) (1979). Criteria and procedures for evaluating assays for carcinogenic residues (Proposal). *Federal Register* **44** (March 20), 17070.

FDA (Food and Drug Administration) (1980). Lead acetate; listing as a color additive in cosmetics that color the hair on the scalp. *Federal Register* **45** (October 3), 72112.

Government Institutes (1983). *Environmental Statutes*, 1983 edn. Government Institutes, Inc., Rockville, Maryland: 681 pages.

IRLG (Interagency Regulatory Liaison Group) (1979). Scientific base for identification of potential carcinogens and estimation of risks. *J. natn. Cancer Inst.*, **63**, 243–268.

Krewski, D., and Van Ryzin, J. (1981). Dose response models for quantal response to toxicity data. In Scargo, D. Dawson, Rao, J. N. K., and Salch, E. (Eds.) *Statistics and Related Topics*, pp. 201–231. North-Holland, New York.

Lancet (1976). Seventeen principles about cancer, or something. Editorial. *Lancet*, **i**, 571–573.

NAS (1977). *Drinking Water and Health*, pp. 24–54. National Academy of Sciences, Washington, DC.

NCABSEC (National Cancer Advisory Board Subcommittee on Environmental Carcinogens) (1977). General criteria for assessing the evidence for carcinogenicity of chemical substances: a report. *J. natn. Cancer Inst.*, **58,** 401–465.

The New York Times (1980a). So it's a carcinogen, but how bad? (Wednesday, July 16). Editorial. SA, p. 28, Col. 1–2.

The New York Times (1980b). A Carcinogen Passes. (Sunday, November 9). Editorial. S4, p. 18E, Col. 1–2.

OSHA (Occupational Safety and Health Administration) (1980). Identification, classification and regulation of potential occupational carcinogens. *Federal Register* **45** (January 22), 5002–5035. Books 1 and 2.

Respondents brief in support of proposed findings, conclusions, and order at 63-64. In Re: Stevens Industries Inc. *et al.* (Consolidated DDT hearings April 5, 1972). Available from the Office of the General Counsel, Environmental Protection Agency, Washington, DC.

Respondents motion to determine whether or not the registration of mirex should be cancelled or amended. Attachment A. (September 9, 1975). Available from the General Counsel, Environmental Protection Agency, Washington, DC.

US Supreme Court (1980). Industrial Union Department, *AFL-CIO v. American Petroleum Institute*. 448 US 607.

Weinhouse, S. (1982). Prometheus and Pandora—cancer research on our diamond anniversary: Presidential address. *Cancer Res.*, **42,** 3471–3474.

APPENDIX

DESCRIPTION OF THE QUANTITATIVE RISK EXTRAPOLATION MODELS USED BY THE US ENVIRONMENTAL PROTECTION AGENCY*

1 CHOICE OF MODEL

There is no really solid scientific basis for any mathematical extrapolation model relating carcinogen exposure to cancer risk at the extremely low levels of exposure that must be dealt with in evaluating environmental hazards. For practical reasons, such low levels of risk cannot be measured directly using either animal experiments or epidemiological studies. We must, therefore, depend on our current understanding of the mechanisms of carcinogenesis for guidance as to which risk model to use. At the present time, the dominant view of the carcinogenic process involves the concept that most agents that cause cancer also cause irreversible damage to DNA. This position is reflected by the fact that a very large proportion of agents that cause cancer are also mutagenic. There is reason to

* Adapted from Water Quality Criteria Documents; Availability, *Federal Register*, **45,** No. 231, Friday, November 28, 1980, pp. 79350–79353.

expect that the quantal type of biological response characteristic of mutagenesis is associated with a linear nonthreshold dose–response relationship. Indeed, there is substantial evidence (from mutagenesis studies with both ionizing radiation and a wide variety of chemicals) that this type of dose–response model is the appropriate one to use. This is particularly true at the lower end of the dose–response curve; at higher doses, there can be an upward curvature, probably reflecting the effects of multistage processes on the mutagenic response. The linear nonthreshold dose–response relationship is also consistent with the relatively few epidemiological studies of cancer responses to specific agents that contain enough information to make the evaluation possible [for example, radiation-induced leukemia, breast and thyroid cancer (Court-Brown and Doll, 1957; Lewis, 1957, Myrden and Quinlan, 1974; Hempelmann *et al.*, 1975); skin cancer induced by arsenic in drinking water (Tseng *et al.*, 1968); and liver cancer induced by aflatoxin in the diet (Linsell and Peers, 1977)]. There is also some evidence from animal experiments that is consistent with the linear nonthreshold hypothesis (for instance, the initiation stage of the two-stage carcinogenesis model in rat liver and mouse skin).

Because it has the best, albeit limited, scientific basis of any of the current mathematical extrapolation models, the linear nonthreshold model has been adopted as the primary basis for extrapolation to low levels of the dose–response relationship. The risk assessments made with this model should be regarded as conservative, representing the most plausible upper limit for the risk; i.e., the true risk is not likely to be higher than the estimate but it could be smaller.

2 THE MULTISTAGE MODEL

The mathematical formulation chosen to describe the linear nonthreshold dose–response relationship at low doses is the modified multistage model developed by Crump (1980). This model employs enough arbitrary constants to be able to fit almost any monotonically increasing dose–response data, and it incorporates a procedure for estimating the largest possible linear slope (in the 95% confidence limit sense) at low extrapolated doses that is consistent with the data at all dose levels of the experiment. For this reason, it may be called a 'linearized' multistage model.

2.1 Procedure for Low-dose Extrapolation Based on Animal Carcinogenicity Data

2.1.1 Description of the Extrapolation Model

Let $P(d)$ represent the lifetime risk (probability) of cancer at dose d. The multistage model has the form

$$P(d) = 1 - \exp[-(q_0 + q_1 d + q_2 d^2 + \ldots + q_k d^k)]$$

where

$$q_i > 0, \text{ and } i = 0, 1, 2, \ldots k$$

Similarly,

$$A(d) = 1 - \exp\left[-(q_1 d + q_2 d^2 + \ldots + q_k d^k)\right]$$

where

$$A(d) = \frac{P(d) - P(0)}{1 - P(0)}$$

is the extra risk over background rate at dose d.

The point estimate of the coefficients q_i ($i = 0, 1, 2, \ldots, k$), and consequently the extra risk function $A(d)$, at any given dose, d, is calculated by maximizing the likelihood function of the data.

The point estimate and the 95% upper confidence limit of the extra risk $A(d)$ are calculated by using the computer program GLOBAL79 developed by Crump and Watson (1979). Upper 95% confidence limits on the extra risk and lower 95% confidence limits on the dose producing a given risk are determined from a 95% upper confidence limit, q_1^*, on parameter q_1. Whenever $q_1 \neq 0$, at low doses the extra risk $A(d)$ has approximately the form $A(d) = q_1 \times d$. Therefore, $q_1^* \times d$ is a 95% upper confidence limit on the extra risk and R/q_1^* is a 95% lower confidence limit on the dose producing an extra risk R. Let L_0 be the maximum value of the log-likelihood function. The upper limit, q_1^*, is calculated by increasing q_1 to a value of q_1^*, such that when the log-likelihood is remaximized subject to this fixed value, q_1^*, for the linear coefficient, the resulting maximum value of the log-likelihood L_1 satisfies the equation

$$2(L_0 - L_1) = 2.70554$$

where 2.70554 is the cumulative 90% point of the chi-square distribution with one degree of freedom, which corresponds to a 95% upper limit (one-sided). This approach of computing the upper confidence limit for the extra risk $A(d)$ is a modification of the Crump *et al.* (1977) model. The upper confidence limit for the extra risk calculated at low doses is always linear. This is conceptually consistent with the linear nonthreshold concept discussed earlier. The slope, q_1^*, is taken as an upper bound of the potency of the chemical in inducing cancer at low doses.

In fitting the dose–response model, the number of terms in the polynomial, $g(d)$, is chosen equal to $(h-1)$, where h is the number of dose groups in the experiment including the control group.

Whenever the multistage model does not fit the data sufficiently, data at the

highest dose are deleted and the model is refitted to the rest of the data. This is continued until an acceptable fit to the data is obtained. To determine whether or not a fit is acceptable, the chi-square statistic

$$\chi^2 = \sum_{i=1}^{h} \frac{(\chi_i - N_i P_i)^2}{N_i P_i (1 - P_i)}$$

is calculated, where N_i is the number of animals in the ith dose group, χ_i is the number of animals in the ith dose group with a tumour response, P_i is the probability of a response in the ith dose group estimated by fitting the multistage model to the data, and h is the number of remaining groups. The fit is determined to be unacceptable whenever chi-square (χ^2) is larger than the cumulative 99% point of the chi-square distribution with f degrees of freedom, where f equals the number of dose groups minus the number of non-zero multistage coefficients.

2.1.2 Selection and Form of Data used to Estimate Parameters in the Extrapolation Model

For some chemicals, several studies in different animal species, strains and sexes, each conducted at several doses and different routes of exposure, are available. A choice must be made as to which of the data sets from several studies are to be used in the model. It is also necessary to correct for metabolism differences between species and for differences in absorption via different routes of administration. The procedures listed below, used in evaluating these data, are consistent with the estimate of a maximum likely risk.

(1) The tumor incidence data are separated according to organ sites or tumor types. The set of data (i.e., dose and tumor incidence) used in the model is the set where the incidence is statistically significantly higher than the control for at least one test dose level and/or where the tumor incidence rate shows a statistically significant trend with respect to dose level. The data set that gives the highest estimate of lifetime carcinogenic risk, q, is selected in most cases. However, efforts are made to exclude data sets that produce spuriously high-risk estimates because of a small number of animals, that is, if two sets of data show a similar dose–response relationship and one has a very small sample size, the set of data which has the larger sample size is selected for calculating the carcinogenic potency.

(2) If there are two or more data sets of comparable size that are identical with respect to species, strain, sex and tumor sites, the geometric mean of q, estimated from each of these data sets, is used for risk assessment. The geometric mean of numbers $A_1, A_2, \ldots, A_m$ is defined as $(A_1 \times A_2 \times \ldots \times A_m)^{1/m}$.

(3) If sufficient data exist for two or more significant tumor sites in the same

study, the number of animals with at least one of the specific tumor sites under consideration is used as incidence data in the model.

(4) Following the suggestion of Mantel and Schneiderman (1975), we assume that mg/surface area/day is an equivalent dose between species. Since, to a close approximation, the surface area is proportional to the 2/3rds power of the weight as would be the case for a perfect sphere, the exposure in mg per 2/3rds power of the body weight/day is similarly considered to be an equivalent exposure. In an animal experiment, this equivalent dose is computed in the following manner. If L_e = duration of experiment; l_e = duration of exposure; m = average dose per day in mg during administration of the agent (i.e., during l_e); W = average weight of the experimental animal; then, the lifetime average exposure is

$$d = \frac{l_e \times m}{L_e \times W^{2/3}}$$

Often exposures are not given in units of mg/day, and it becomes necessary to convert the given exposure into mg/day. For example, in most feeding studies, exposure is expressed as ppm in the diet. In this case the exposure (mg/day) is derived by

$$m = \text{ppm} \times F \times r$$

where ppm is parts per million of the carcinogenic agent in the diet, F is the weight of the food consumed per day in kg and r is the absorption fraction.

In the absence of any data to the contrary, r is assumed to be one. For a uniform diet the weight of the food consumed is proportional to the calories required, which in turn is proportional to the surface area of the 2/3rds power of the weight, so that

$$m \propto \text{ppm} \times W^{2/3} \times r$$

$$\frac{M}{rW^{2/3}} \propto \text{ppm}$$

As a result, ppm in the diet is often assumed to be an equivalent exposure between species. However, we feel that this is not justified since the calories/kg of food are significantly different in the diet of man *vs* laboratory animals, primarily due to moisture content differences. Instead, we use an empirically derived food factor, $f = F/W$, which is the fraction of a species body weight that is consumed per day as food. We use the rates given as follows.

Species	W	f
Man	70	0.028
Rat	0.35	0.05
Mice	0.03	0.13

Thus, when the exposure is given as a certain dietary concentration in ppm, the exposure in mg/$W^{2/3}$ is

$$\frac{m}{r \times W^{2/3}} = \frac{\text{ppm} \times F}{W^{2/3}} = \frac{\text{ppm} \times f \times W}{W^{2/3}} = \text{ppm} \times f \times W^{1/3}$$

When exposure is given in terms of mg/kg/day $= m/Wr = s$, the conversion is simply

$$\frac{m}{rW^{2/3}} = s \times W^{1/3}$$

When exposure is by inhalation, the calculation of dose can be considered for two cases where:

(a) the carcinogenic agent is either a completely water-soluble gas or an aerosol and is absorbed in proportion to the amount of air inhaled, and
(b) where the carcinogen is a poorly water-soluble gas which reaches an equilibrium between the air breathed and the body compartments.

After equilibrium is reached, the rate of absorption of these agents is expected to be proportional to the metabolic rate, which in turn is proportional to the rate of oxygen consumption, which in turn is a function of surface area.

Case 1 Agents that are in the form of particulate matter or virtually completely absorbed gases, such as SO_2, can reasonably be expected to be absorbed proportionally to the breathing rate. In this case the exposure in mg/day may be expressed as

$$m = I \times v \times r$$

where I is inhalation rate per day in m^3, v is mg/m^3 of the agent in air, and r is the fraction absorbed.

The inhalation rates, I, for various species can be calculated from the observation that 25 g mice breathe 34.5 liters/day and 113 g rats breathe 105 liters/day (Altman and Dittmer, 1974). For mice and rats of other weights, W

(expressed in kg), the surface area proportionality can be used to determine breathing rates (in m^3/day) as follows:

$$\text{for mice, } I = 0.0345\,(W/0.025)^{2/3}\,m^3/\text{day}$$
$$\text{for rats, } I = 0.105\,(W/0.113)^{2/3} m^3/\text{day}$$

For humans, the value of 20 m^3/day is adopted as a standard breathing rate (ICRP, 1977).

The equivalent exposure in mg/$W^{2/3}$ for these agents can be derived from the air intake data in a way analogous to the food intake data. The empirical factors for the air intake per kg per day, $i = I/W$, based upon the previously stated relationships, are as follows.

Species	W	$i = I/W$
Man	70	0.29
Rat	0.35	0.64
Mice	0.03	1.3

Therefore, for particulates or completely absorbed gases, the equivalent exposure in mg/$W^{2/3}$ is

$$\frac{m}{W^{2/3}} = \frac{Ivr}{W^{2/3}} = \frac{iWvr}{W^{2/3}} = iW^{1/3}vr$$

In the absence of experimental information or a sound theoretical argument to the contrary, the fraction absorbed, r, is assumed to be the same for all species.

Case 2 The dose in mg/day of partially soluble vapors is proportional to the O_2 consumption which in turn is proportional to $W^{2/3}$ and to the solubility of gas in body fluids, which can be expressed as an absorption coefficient, r, for the gas. Therefore, when expressing the O_2 consumption as $O_2 = kW^{2/3}$, where k is a constant independent of species, it follows that

$$m = kW^{2/3} \times v \times r$$
$$d = \frac{\dot{m}}{W^{2/3}} = kvr$$

As with Case 1, in the absence of experimental information or a sound theoretical argument to the contrary, the absorbed fraction, r, is assumed to be the same for all species. Therefore, for these substances a certain concentration in ppm or $\mu g/m^3$ in experimental animals is equivalent to the

same concentration in humans. This is supported by the observation that the minimum alveolar concentration necessary to produce a given 'stage' of anesthesia is similar in man and animals (Dripps *et al.*, 1977). When the animals were exposed by the oral route, and human exposure is by inhalation (or vice versa), the assumption is made, unless there is pharmacokinetic evidence to the contrary, that absorption is equal by either exposure route.

(5) If the duration of the experiment, L_e, is less than the natural life-span of the test animal, L, the slope, q_1^*, or more generally the exponent, $g(d)$, is increased by multiplying by a factor $(L/L_e)^3$. We assume that if the average dose, d, is continued, the age-specific rate of cancer will continue to increase as a constant function of the background rate. The age-specific rates for humans increase at least by the second power of the age and often by a considerably higher power, as demonstrated by Doll (1971). Thus, we would expect the cumulative tumor rate to increase by at least the third power of age. Using this fact, we assume that the slope, q_1^*, or more generally the exponent, $g(d)$, would also increase by at least the third power of age. As a result, if the slope, q_1^* [or $g(d)$] is calculated at age L_e, we would expect that if the experiment had been continued for the full life-span, L, at the given average exposure, the slope, q_1^* [or $g(d)$], would have been increased by at least a factor of $(L/L_e)^3$.

This adjustment is conceptually consistent with the proportional hazard model proposed by Cox (1972) and the time-to-tumor model considered by Crump and Watson, (1979) where the probability of cancer by age t and at dose d is given by

$$P(d,t) = 1 - \exp[-f(t) \times g(d)]$$

3 CALCULATION OF CARCINOGENIC POTENCY BASED ON HUMAN DATA

If human epidemiological studies and sufficiently valid exposure information are available for the compound, they are always used in some way. If they show a carcinogenic effect, the data are analyzed to give an estimate of the linear dependence of cancer rates on lifetime average dose, which is equivalent to the factor q_1^*. If they show no carcinogenic effect when positive animal evidence is available, then it is assumed that a risk does exist but it is smaller than could have been observed in the epidemiological study, and an upper limit of the cancer incidence is calculated assuming hypothetically that the true incidence is just below the level of detection in the cohort studied, which is determined largely by the cohort size. Whenever possible, human data are used in preference to animal bioassay data.

In human studies, the response is measured in terms of the relative risk of the exposed cohort of individuals compared to the control group. In the analysis of

this data, it is assumed that the excess risk, or relative risk minus one, $R(X) - 1$, is proportional to the lifetime average exposure, X, and that it is the same for all ages. It follows that the carcinogenic potency is equal to $[R(X) - 1]/X$ multiplied by the lifetime risk at that site in the general population. Except for an unusually well-documented human study, the confidence limit for the excess risk is not calculated, due to the difficulty in accounting for the uncertainty inherent in the data (exposure and cancer response).

4 APPENDIX REFERENCES

Altman, P. L., and Dittmer, D. S. (Eds.). (1974). *Biology Data Book*, 2nd edn., Vol. III. Federation of American Societies for Experimental Biology, Bethesda, Maryland: 2123 pages.

Court-Brown, W. M., and Doll, R. (1957). *Leukemia and Aplastic Anemia in Patients Irradiated for Ankylosing Spondylitis*. Special Report Series No. 295, p. 1. Medical Research Council, Her Majesty's Stationery Office, London.

Cox, C. R. (1972). Regression model and life tables. *Jl R. statist. Soc. B.*, **34,** 187–220.

Crump, K. S. (1980). An improved procedure for low-dose carcinogenic risk assessment from animal data. *J. envir. Path. Tox.*, **5,** 675–684.

Crump, K. S., Guess, H. A., and Deal, L. L. (1977). Confidence intervals and test of hypotheses concerning dose–response relations inferred from animal carcinogenicity data. *Biometrics*, **33,** 437–451.

Crump, K. S., and Watson, W. W. (1979). GLOBAL79. A FORTRAN program to extrapolate dichotomous animal carcinogenicity data to low-dose. *National Institute of Environmental Health Sciences Contract No.* 1-*ES*-2123. National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina.

Doll, R. (1971). Weibull distribution of cancer: implications for models of carcinogenesis. *Jl R. statist. Soc. A.*, **13,** 133–166.

Dripps, R. O., Eckenhoff, J. E., and Vandam, L. D. (1977). *Introduction to Anesthesia, the Principles of Safe Practice*. 5th edn., pp. 121–123. W.B. Saunders Company, Philadelphia.

EPA (1980). U.S. Environmental Protection Agency. *Carcinogen Assessment Group's Final Risk Assessment on Arsenic*, May 2, 1980. PB 81-206013. National Technical Information Service, Springfield, Virginia.

Hempelmann, L. H., Hall, W. J., Phillips, M., Cooper, R. A., and Ames, W. R. (1975). Neoplasms in persons treated with X-rays in infancy: fourth survey in 20 Years. *J. natn. Cancer Inst.*, **55,** 519.

ICRP (International Commission on Radiological Protection) (1977). *Recommendation of the International Commission on Radiological Protection*. ICRP Publication 26. Pergammon Press, Oxford, New York, Frankfurt, v + 53 pages.

Lewis, E. B. (1957). Leukemia and ionizing radiation. *Science*, **125,** 965.

Linsell, C. A., and Peers, F. G. (1977). Field studies on liver cancer. In Hiatt, H. H., Watson, J. D., and Winsten, J. A. (Eds.) *Origins of Human Cancer, Book A. Incidence of Cancer in Humans*, pp. 549–556. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mantel, N., and Schneiderman, M. A. (1975). Estimating "safe levels", a hazardous undertaking. *Cancer Res.*, **35,** 1379–1386.

Myrden, J. A., and Quinlan, J. J. (1974). Breast carcinoma following multiple fluoroscopies

with pneumothorax treatment of pulmonary tuberculosis. *Ann. R. Coll. Physicians Surg. Can.*, **7**, 45.

Tseng, W. P., Chu, H. M., How, S. W., Fong, J. M., Lin, C. S., and Yeh, S. (1968). Prevalence of skin cancer in an endemic area of chronic arsenicism in Taiwan. *J. natn. Cancer Inst.*, **40**, 453–463.