G-000-1101.7

4446

GUIDELINES FOR CARCINOGEN RISK ASSESSMENT (FEDERAL REGISTER NOTICE)

09/24/86

USEPA
13
GUIDE DOC

federal register

Wednesday
September 24, 1986

4446

Part II

Environmental Protection Agency

Guidelines for Carcinogen Risk Assessment

0 1

ENVIRONMENTAL PROTECTION AGENCY

[FRL-2984-1]

Guidelines for Carcinogen Risk Assessment

AGENCY: U.S. Environmental Protection Agency (EPA).

ACTION: Final guidelines for carcinogen risk assessment.

SUMMARY: The U.S. Environmental Protection Agency is today issuing five guidelines for assessing the health risks of environmental pollutants. These are:

Guidelines for Carcinogen Risk Assessment
Guidelines for Estimating Exposures
Guidelines for Mutagenicity Risk Assessment
Guidelines for the Health Assessment of Suspect Developmental Toxicants
Guidelines for the Health Risk Assessment of Chemical Mixtures

This notice contains the Guidelines for Carcinogen Risk Assessment; the other guidelines appear elsewhere in today's **Federal Register.**

The Guidelines for Carcinogen Risk Assessment (hereafter "Guidelines") are intended to guide Agency evaluation of suspect carcinogens in line with the policies and procedures established in the statutes administered by the EPA. These Guidelines were developed as part of an interoffice guidelines development program under the auspices of the Office of Health and Environmental Assessment (OHEA) in the Agency's Office of Research and Development. They reflect Agency consideration of public and Science Advisory Board (SAB) comments on the Proposed Guidelines for Carcinogen Risk Assessment published November 23, 1984 (49 FR 46294).

This publication completes the first round of risk assessment guidelines development. These Guidelines will be revised, and new guidelines will be developed, as appropriate.

EFFECTIVE DATE: The Guidelines will be effective September 24, 1986.

FOR FURTHER INFORMATION CONTACT: Dr. Robert E. McGaughy, Carcinogen Assessment Group, Office of Health and Environmental Assessment (RD-689), U.S. Environmental Protection Agency, 401 M Street SW., Washington, DC 20460, 202-382-5898.

SUPPLEMENTARY INFORMATION: In 1983, the National Academy of Sciences (NAS) published its book entitled *Risk Assessment in the Federal Government: Managing the Process.* In that book, the NAS recommended that Federal regulatory agencies establish "inference guidelines" to ensure consistency and technical quality in risk assessments and to ensure that the risk assessment process was maintained as a scientific effort separate from risk management. A task force within EPA accepted that recommendation and requested that Agency scientists begin to develop such guidelines.

General

The guidelines published today are products of a two-year Agencywide effort, which has included many scientists from the larger scientific community. These guidelines set forth principles and procedures to guide EPA scientists in the conduct of Agency risk assessments, and to inform Agency decision makers and the public about these procedures. In particular, the guidelines emphasize that risk assessments will be conducted on a case-by-case basis, giving full consideration to all relevant scientific information. This case-by-case approach means that Agency experts review the scientific information on each agent and use the most scientifically appropriate interpretation to assess risk. The guidelines also stress that this information will be fully presented in Agency risk assessment documents, and that Agency scientists will identify the strengths and weaknesses of each assessment by describing uncertainties, assumptions, and limitations, as well as the scientific basis and rationale for each assessment.

Finally, the guidelines are formulated in part to bridge gaps in risk assessment methodology and data. By identifying these gaps and the importance of the missing information to the risk assessment process, EPA wishes to encourage research and analysis that will lead to new risk assessment methods and data.

Guidelines for Carcinogen Risk Assessment

Work on the Guidelines for Carcinogen Risk Assessment began in January 1984. Draft guidelines were developed by Agency work groups composed of expert scientists from throughout the Agency. The drafts were peer-reviewed by expert scientists in the field of carcinogenesis from universities, environmental groups, industry, labor, and other governmental agencies. They were then proposed for public comment in the **Federal Register** (49 FR 46294). On November 9, 1984, the Administrator directed that Agency offices use the proposed guidelines in performing risk assessments until final guidelines become available.

After the close of the public comment period, Agency staff prepared summaries of the comments and analyses of the major issues presented by the commentors, and proposed changes in the language of the guidelines to deal with the issues raised. These analyses were presented to review panels of the SAB on March 4 and April 22-23, 1985, and to the Executive Committee of the SAB on April 25-26, 1985. The SAB meetings were announced in the **Federal Register** as follows: February 12, 1985 (50 FR 5811) and April 4, 1985 (50 FR 13420 and 13421).

In a letter to the Administrator dated June 19, 1985, the Executive Committee generally concurred on all five of the guidelines, but recommended certain revisions, and requested that any revised guidelines be submitted to the appropriate SAB review panel chairman for review and concurrence on behalf of the Executive Committee. As described in the responses to comments (see Part B: Response to the Public and Science Advisory Board Comments), each guidelines document was revised, where appropriate, consistent with the SAB recommendations, and revised draft guidelines were submitted to the panel chairmen. Revised draft Guidelines for Carcinogen Risk Assessment were concurred on in a letter dated February 7, 1986. Copies of the letters are available at the Public Information Reference Unit, EPA Headquarters Library, as indicated elsewhere in this notice.

Following this Preamble are two parts: Part A contains the Guidelines and Part B, the Response to the Public and Science Advisory Board Comments (a summary of the major public comments, SAB comments, and Agency responses to those comments).

The Agency is continuing to study the risk assessment issues raised in the guidelines and will revise these guidelines in line with new information as appropriate.

References, supporting documents, and comments received on the proposed guidelines, as well as copies of the final guidelines, are available for inspection and copying at the Public Information Reference Unit (202-382-5926), EPA Headquarters Library, 401 M Street, SW., Washington, DC, between the hours of 8:00 a.m. and 4:30 p.m.

I certify that these Guidelines are not major rules as defined by Executive Order 12291, because they are nonbinding policy statements and have no direct effect on the regulated community. Therefore, they will have no effect on costs or prices, and they will

have no other significant adverse effects on the economy. These Guidelines were reviewed by the Office of Management and Budget under Executive Order 12291.

Dated: August 22, 1986.

Lee M. Thomas,

Administrator.

Contents

Part A: Guidelines for Carcinogen Risk Assessment

I. Introduction

II. Hazard Identification

A. Overview
B. Elements of Hazard Identification
1. Physical-Chemical Properties and Routes and Patterns of Exposure
2. Structure-Activity Relationships
3. Metabolic and Pharmacokinetic Properties
4. Toxicologic Effects
5. Short-Term Tests
6. Long-Term Animal Studies
7. Human Studies
C. Weight of Evidence
D. Guidance for Dose-Response Assessment
E. Summary and Conclusion

III. Dose-Response Assessment, Exposure Assessment, and Risk Characterization

A. Dose-Response Assessment
1. Selection of Data
2. Choice of Mathematical Extrapolation Model
3. Equivalent Exposure Units Among Species
B. Exposure Assessment
C. Risk Characterization
1. Options for Numerical Risk Estimates
2. Concurrent Exposure
3. Summary of Risk Characterization

IV. EPA Classification System for Categorizing Weight of Evidence for Carcinogenicity From Human and Animal Studies (Adapted From IARC)

A. Assessment of Weight of Evidence for Carcinogenicity From Studies in Humans
B. Assessment of Weight of Evidence for Carcinogenicity From Studies in Experimental Animals
C. Categorization of Overall Weight of Evidence for Human Carcinogenicity

V. References

Part B: Response to Public and Science Advisory Board Comments

I. Introduction
II. Office of Science and Technology Policy Report on Chemical Carcinogens
III. Inference Guidelines
IV. Evaluation of Benign Tumors
V. Transplacental and Multigenerational Animal Bioassays
VI. Maximum Tolerated Dose
VII. Mouse Liver Tumors
VIII. Weight-of-Evidence Categories
IX. Quantitative Estimates of Risk

Part A: Guidelines for Carcinogen Risk Assessment

I. Introduction

This is the first revision of the 1976 Interim Procedures and Guidelines for Health Risk Assessments of Suspected Carcinogens (U.S. EPA, 1976; Albert et al., 1977). The impetus for this revision is the need to incorporate into these Guidelines the concepts and approaches to carcinogen risk assessment that have been developed during the last ten years. The purpose of these Guidelines is to promote quality and consistency of carcinogen risk assessments within the EPA and to inform those outside the EPA about its approach to carcinogen risk assessment. These Guidelines emphasize the broad but essential aspects of risk assessment that are needed by experts in the various disciplines required (e.g., toxicology, pathology, pharmacology, and statistics) for carcinogen risk assessment. Guidance is given in general terms since the science of carcinogenesis is in a state of rapid advancement, and overly specific approaches may rapidly become obsolete.

These Guidelines describe the general framework to be followed in developing an analysis of carcinogenic risk and some salient principles to be used in evaluating the quality of data and in formulating judgments concerning the nature and magnitude of the cancer hazard from suspect carcinogens. It is the intent of these Guidelines to permit sufficient flexibility to accommodate new knowledge and new assessment methods as they emerge. It is also recognized that there is a need for new methodology that has not been addressed in this document in a number of areas, e.g., the characterization of uncertainty. As this knowledge and assessment methodology are developed, these Guidelines will be revised whenever appropriate.

A summary of the current state of knowledge in the field of carcinogenesis and a statement of broad scientific principles of carcinogen risk assessment, which was developed by the Office of Science and Technology Policy (OSTP, 1985), forms an important basis for these Guidelines; the format of these Guidelines is similar to that proposed by the National Research Council (NRC) of the National Academy of Sciences in a book entitled *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983).

These Guidelines are to be used within the policy framework already provided by applicable EPA statutes and do not alter such policies. These Guidelines provide general directions for analyzing and organizing available data. They do not imply that one kind of data or another is prerequisite for regulatory action to control, prohibit, or allow the use of a carcinogen.

Regulatory decision making involves two components: risk assessment and risk management. Risk assessment defines the adverse health consequences of exposure to toxic agents. The risk assessments will be carried out independently from considerations of the consequences of regulatory action. Risk management combines the risk assessment with the directives of regulatory legislation, together with socioeconomic, technical, political, and other considerations, to reach a decision as to whether or how much to control future exposure to the suspected toxic agents.

Risk assessment includes one or more of the following components: hazard identification, dose-response assessment, exposure assessment, and risk characterization (NRC, 1983).

Hazard identification is a qualitative risk assessment, dealing with the process of determining whether exposure to an agent has the potential to increase the incidence of cancer. For purposes of these Guidelines, both malignant and benign tumors are used in the evaluation of the carcinogenic hazard. The hazard identification component qualitatively answers the question of how likely an agent is to be a human carcinogen.

Traditionally, quantitative risk assessment has been used as an inclusive term to describe all or parts of dose-response assessment, exposure assessment, and risk characterization. Quantitative risk assessment can be a useful general term in some circumstances, but the more explicit terminology developed by the NRC (1983) is usually preferred. The dose-response assessment defines the relationship between the dose of an agent and the probability of induction of a carcinogenic effect. This component usually entails an extrapolation from the generally high doses administered to experimental animals or exposures noted in epidemiologic studies to the exposure levels expected from human contact with the agent in the environment; it also includes considerations of the validity of these extrapolations.

The exposure assessment identifies populations exposed to the agent, describes their composition and size, and presents the types, magnitudes, frequencies, and durations of exposure to the agent.

In risk characterization, the results of the exposure assessment and the dose-response assessment are combined to estimate quantitatively the carcinogenic risk. As part of risk characterization, a summary of the strengths and weaknesses in the hazard identification, dose-response assessment, exposure assessment, and the public health risk estimates are presented. Major assumptions, scientific judgments, and, to the extent possible, estimates of the uncertainties embodied in the assessment are also presented, distinguishing clearly between fact, assumption, and science policy.

The National Research Council (NRC, 1983) pointed out that there are many questions encountered in the risk assessment process that are unanswerable given current scientific knowledge. To bridge the uncertainty that exists in these areas where there is no scientific consensus, inferences must be made to ensure that progress continues in the assessment process. The OSTP (1985) reaffirmed this position, and generally left to the regulatory agencies the job of articulating these inferences. Accordingly, the Guidelines incorporate judgmental positions (science policies) based on evaluation of the presently available information and on the regulatory mission of the Agency. The Guidelines are consistent with the principles developed by the OSTP (1985), although in many instances are necessarily more specific.

II. Hazard Identification

A. Overview

The qualitative assessment or hazard identification part of risk assessment contains a review of the relevant biological and chemical information bearing on whether or not an agent may pose a carcinogenic hazard. Since chemical agents seldom occur in a pure state and are often transformed in the body, the review should include available information on contaminants, degradation products, and metabolites.

Studies are evaluated according to sound biological and statistical considerations and procedures. These have been described in several publications (Interagency Regulatory Liaison Group, 1979; OSTP, 1985; Peto et al., 1980; Mantel, 1980; Mantel and Haenszel, 1959; Interdisciplinary Panel on Carcinogenicity, 1984; National Center for Toxicological Research, 1981; National Toxicology Program, 1984; U.S. EPA, 1983a, 1983b, 1983c; Haseman, 1984). Results and conclusions concerning the agent, derived from different types of information, whether indicating positive or negative responses, are melded together into a weight-of-evidence determination. The strength of the evidence supporting a potential human carcinogenicity judgment is developed in a weight-of-evidence stratification scheme.

B. Elements of Hazard Identification

Hazard identification should include a review of the following information to the extent that it is available.

1. *Physical-Chemical Properties and Routes and Patterns of Exposure.* Parameters relevant to carcinogenesis, including physical state, physical-chemical properties, and exposure pathways in the environment should be described where possible.

2. *Structure-Activity Relationships.* This section should summarize relevant structure-activity correlations that support or argue against the prediction of potential carcinogenicity.

3. *Metabolic and Pharmacokinetic Properties.* This section should summarize relevant metabolic information. Information such as whether the agent is direct-acting or requires conversion to a reactive carcinogenic (e.g., an electrophilic) species, metabolic pathways for such conversions, macromolecular interactions, and fate (e.g., transport, storage, and excretion), as well as species differences, should be discussed and critically evaluated. Pharmacokinetic properties determine the biologically effective dose and may be relevant to hazard identification and other components of risk assessment.

4. *Toxicologic Effects.* Toxicologic effects other than carcinogenicity (e.g., suppression of the immune system, endocrine disturbances, organ damage) that are relevant to the evaluation of carcinogenicity should be summarized. Interactions with other chemicals or agents and with lifestyle factors should be discussed. Prechronic and chronic toxicity evaluations, as well as other test results, may yield information on target organ effects, pathophysiological reactions, and preneoplastic lesions that bear on the evaluation of carcinogenicity. Dose-response and time-to-response analyses of these reactions may also be helpful.

5. *Short-Term Tests.* Tests for point mutations, numerical and structural chromosome aberrations, DNA damage/repair, and *in vitro* transformation provide supportive evidence of carcinogenicity and may give information on potential carcinogenic mechanisms. A range of tests from each of the above end points helps to characterize an agent's response spectrum.

Short-term *in vivo* and *in vitro* tests that can give indication of initiation and promotion activity may also provide supportive evidence for carcinogenicity. Lack of positive results in short-term tests for genetic toxicity does not provide a basis for discounting positive results in long-term animal studies.

6. *Long-Term Animal Studies.* Criteria for the technical adequacy of animal carcinogenicity studies have been published (e.g., U.S. Food and Drug Administration, 1982; Interagency Regulatory Liaison Group, 1979; National Toxicology Program, 1984; OSTP, 1985; U.S. EPA, 1983a, 1983b, 1983c; Feron et al., 1980; Mantel, 1980) and should be used to judge the acceptability of individual studies. Transplacental and multigenerational carcinogenesis studies, in addition to more conventional long-term animal studies, can yield useful information about the carcinogenicity of agents.

It is recognized that chemicals that induce benign tumors frequently also induce malignant tumors, and that benign tumors often progress to malignant tumors (Interdisciplinary Panel on Carcinogenicity, 1984). The incidence of benign and malignant tumors will be combined when scientifically defensible (OSTP, 1985; Principle 8). For example, the Agency will, in general, consider the combination of benign and malignant tumors to be scientifically defensible unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin. If an increased incidence of benign tumors is observed in the absence of malignant tumors, in most cases the evidence will be considered as limited evidence of carcinogenicity.

The weight of evidence that an agent is potentially carcinogenic for humans increases (1) with the increase in number of tissue sites affected by the agent; (2) with the increase in number of animal species, strains, sexes, and number of experiments and doses showing a carcinogenic response; (3) with the occurrence of clear-cut dose-response relationships as well as a high level of statistical significance of the increased tumor incidence in treated compared to control groups; (4) when there is a dose-related shortening of the time-to-tumor occurrence or time to death with tumor; and (5) when there is a dose-related increase in the proportion of tumors that are malignant.

Long-term animal studies at or near the maximum tolerated dose level (MTD) are used to ensure an adequate power for the detection of carcinogenic

activity (NTP, 1984; IARC, 1982). Negative long-term animal studies at exposure levels above the MTD may not be acceptable if animal survival is so impaired that the sensitivity of the study is significantly reduced below that of a conventional chronic animal study at the MTD. The OSTP (1985; Principle 4) has stated that,

> The carcinogenic effects of agents may be influenced by non-physiological responses (such as extensive organ damage, radical disruption of hormonal function, saturation of metabolic pathways, formation of stones in the urinary tract, saturation of DNA repair with a functional loss of the system) induced in the model systems. Testing regimes inducing these responses should be evaluated for their relevance to the human response to an agent and evidence from such a study, whether positive or negative, must be carefully reviewed.

Positive studies at levels above the MTD should be carefully reviewed to ensure that the responses are not due to factors which do not operate at exposure levels below the MTD. Evidence indicating that high exposures alter tumor responses by indirect mechanisms that may be unrelated to effects at lower exposures should be dealt with on an individual basis. As noted by the OSTP (1985), "Normal metabolic activation of carcinogens may possibly also be altered and carcinogenic potential reduced as a consequence [of high-dose testing]."

Carcinogenic responses under conditions of the experiment should be reviewed carefully as they relate to the relevance of the evidence to human carcinogenic risks (e.g., the occurrence of bladder tumors in the presence of bladder stones and implantation site sarcomas). Interpretation of animal studies is aided by the review of target organ toxicity and other effects (e.g., changes in the immune and endocrine systems) that may be noted in prechronic or other toxicological studies. Time and dose-related changes in the incidence of preneoplastic lesions may also be helpful in interpreting animal studies.

Agents that are positive in long-term animal experiments and also show evidence of promoting or cocarcinogenic activity in specialized tests should be considered as complete carcinogens unless there is evidence to the contrary because it is, at present, difficult to determine whether an agent is only a promoting or cocarcinogenic agent. Agents that show positive results in special tests for initiation, promotion, or cocarcinogenicity and no indication of tumor response in well-conducted and well-designed long-term animal studies should be dealt with on an individual basis.

To evaluate carcinogenicity, the primary comparison is tumor response in dosed animals as compared with that in contemporary matched control animals. Historical control data are often valuable, however, and could be used along with concurrent control data in the evaluation of carcinogenic responses (Haseman et al., 1984). For the evaluation of rare tumors, even small tumor responses may be significant compared to historical data. The review of tumor data at sites with high spontaneous background requires special consideration (OSTP, 1985; Principle 9). For instance, a response that is significant with respect to the experimental control group may become questionable if the historical control data indicate that the experimental control group had an unusually low background incidence (NTP, 1984).

For a number of reasons, there are widely diverging scientific views (OSTP, 1985; Ward et al., 1979a, b; Tomatis, 1977; Nutrition Foundation, 1983) about the validity of mouse liver tumors as an indication of potential carcinogenicity in humans when such tumors occur in strains with high spontaneous background incidence and when they constitute the only tumor response to an agent. These Guidelines take the position that when the only tumor response is in the mouse liver and when other conditions for a classification of "sufficient" evidence in animal studies are met (e.g., replicate studies, malignancy; see section IV), the data should be considered as "sufficient" evidence of carcinogenicity. It is understood that this classification could be changed on a case-by-case basis to "limited," if warranted, when factors such as the following, are observed: an increased incidence of tumors only in the highest dose group and/or only at the end of the study; no substantial dose-related increase in the proportion of tumors that are malignant; the occurrence of tumors that are predominantly benign; no dose-related shortening of the time to the appearance of tumors; negative or inconclusive results from a spectrum of short-term tests for mutagenic activity; the occurrence of excess tumors only in a single sex.

Data from all long-term animal studies are to be considered in the evaluation of carcinogenicity. A positive carcinogenic response in one species/strain/sex is not generally negated by negative results in other species/strain/sex. Replicate negative studies that are essentially identical in all other respects to a positive study may indicate that the positive results are spurious.

Evidence for carcinogenic action should be based on the observation of statistically significant tumor responses in specific organs or tissues. Appropriate statistical analysis should be performed on data from long-term studies to help determine whether the effects are treatment-related or possibly due to chance. These should at least include a statistical test for trend, including appropriate correction for differences in survival. The weight to be given to the level of statistical significance (the p-value) and to other available pieces of information is a matter of overall scientific judgment. A statistically significant excess of tumors of all types in the aggregate, in the absence of a statistically significant increase of any individual tumor type, should be regarded as minimal evidence of carcinogenic action unless there are persuasive reasons to the contrary.

7. *Human Studies.* Epidemiologic studies provide unique information about the response of humans who have been exposed to suspect carcinogens. Descriptive epidemiologic studies are useful in generating hypotheses and providing supporting data, but can rarely be used to make a causal inference. Analytical epidemiologic studies of the case-control or cohort variety, on the other hand, are especially useful in assessing risks to exposed humans.

Criteria for the adequacy of epidemiologic studies are well recognized. They include factors such as the proper selection and characterization of exposed and control groups, the adequacy of duration and quality of follow-up, the proper identification and characterization of confounding factors and bias, the appropriate consideration of latency effects, the valid ascertainment of the causes of morbidity and death, and the ability to detect specific effects. Where it can be calculated, the statistical power to detect an appropriate outcome should be included in the assessment.

The strength of the epidemiologic evidence for carcinogenicity depends, among other things, on the type of analysis and on the magnitude and specificity of the response. The weight of evidence increases rapidly with the number of adequate studies that show comparable results on populations exposed to the same agent under different conditions.

It should be recognized that epidemiologic studies are inherently capable of detecting only comparatively large increases in the relative risk of

cancer. Negative results from such studies cannot prove the absence of carcinogenic action; however, negative results from a well-designed and well-conducted epidemiologic study that contains usable exposure data can serve to define upper limits of risk; these are useful if animal evidence indicates that the agent is potentially carcinogenic in humans.

C. Weight of Evidence

Evidence of possible carcinogenicity in humans comes primarily from two sources: long-term animal tests and epidemiologic investigations. Results from these studies are supplemented with available information from short-term tests, pharmacokinetic studies, comparative metabolism studies, structure-activity relationships, and other relevant toxicologic studies. The question of how likely an agent is to be a human carcinogen should be answered in the framework of a weight-of-evidence judgment. Judgments about the weight of evidence involve considerations of the quality and adequacy of the data and the kinds and consistency of responses induced by a suspect carcinogen. There are three major steps to characterizing the weight of evidence for carcinogenicity in humans: (1) Characterization of the evidence from human studies and from animal studies individually, (2) combination of the characterizations of these two types of data into an indication of the overall weight of evidence for human carcinogenicity, and (3) evaluation of all supporting information to determine if the overall weight of evidence should be modified.

EPA has developed a system for stratifying the weight of evidence (see section IV). This classification is not meant to be applied rigidly or mechanically. At various points in the above discussion, EPA has emphasized the need for an overall, balanced judgment of the totality of the available evidence. Particularly for well-studied substances, the scientific data base will have a complexity that cannot be captured by any classification scheme. Therefore, the hazard identification section should include a narrative summary of the strengths and weaknesses of the evidence as well as its categorization in the EPA scheme.

The EPA classification system is, in general, an adaptation of the International Agency for Research on Cancer (IARC, 1982) approach for classifying the weight of evidence for human data and animal data. The EPA classification system for the characterization of the overall weight of evidence for carcinogenicity (animal, human, and other supportive data) includes: Group A—Carcinogenic to Humans; Group B—Probably Carcinogenic to Humans; Group C—Possibly Carcinogenic to Humans; Group D—Not Classifiable as to Human Carcinogenicity; and Group E—Evidence of Non-Carcinogenicity for Humans.

The following modifications of the IARC approach have been made for classifying human and animal studies.

For human studies:

(1) The observation of a statistically significant association between an agent and life-threatening benign tumors in humans is included in the evaluation of risks to humans.

(2) A "no data available" classification is added.

(3) A "no evidence of carcinogenicity" classification is added. This classificaton indicates that no association was found between exposure and increased risk of cancer in well-conducted, well-designed, independent analytical epidemiologic studies.

For animal studies:

(1) An increased incidence of combined benign and malignant tumors will be considered to provide sufficient evidence of carcinogenicity if the other criteria defining the "sufficient" classification of evidence are met (e.g., replicate studies, malignancy; see section IV). Benign and malignant tumors will be combined when scientifically defensible.

(2) An increased incidence of benign tumors alone generally constitutes "limited" evidence of carcinogenicity.

(3) An increased incidence of neoplasms that occur with high spontaneous background incidence (e.g., mouse liver tumors and rat pituitary tumors in certain strains) generally constitutes "sufficient" evidence of carcinogenicity, but may be changed to "limited" when warranted by the specific information available on the agent.

(4) A "no data available" classification has been added.

(5) A "no evidence of carcinogenicity" classification is also added. This operational classification would include substances for which there is no increased incidence of neoplasms in at least two well-designed and well-conducted animal studies of adequate power and dose in different species.

D. Guidance for Dose-Response Assessment

The qualitative evidence for carcinogenesis should be discussed for purposes of guiding the dose-response assessment. The guidance should be given in terms of the appropriateness and limitations of specific studies as well as pharmacokinetic considerations that should be factored into the dose-response assessment. The appropriate method of extrapolation should be factored in when the experimental route of exposure differs from that occurring in humans.

Agents that are judged to be in the EPA weight-of-evidence stratification Groups A and B would be regarded as suitable for quantitative risk assessments. Agents that are judged to be in Group C will generally be regarded as suitable for quantitative risk assessment, but judgments in this regard may be made on a case-by-case basis. Agents that are judged to be in Groups D and E would not have quantitative risk assessments.

E. Summary and Conclusion

The summary should present all of the key findings in all of the sections of the qualitative assessment and the interpretive rationale that forms the basis for the conclusion. Assumptions, uncertainties in the evidence, and other factors that may affect the relevance of the evidence to humans should be discussed. The conclusion should present both the weight-of-evidence ranking and a description that brings out the more subtle aspects of the evidence that may not be evident from the ranking alone.

III. Dose-Response Assessment, Exposure Assessment, and Risk Characterization

After data concerning the carcinogenic properties of a substance have been collected, evaluated, and categorized, it is frequently desirable to estimate the likely range of excess cancer risk associated with given levels and conditions of human exposure. The first step of the analysis needed to make such estimations is the development of the likely relationship between dose and response (cancer incidence) in the region of human exposure. This information on dose-response relationships is coupled with information on the nature and magnitude of human exposure to yield an estimate of human risk. The risk-characterization step also includes an interpretation of these estimates in light of the biological, statistical, and exposure assumptions and uncertainties that have arisen throughout the process of assessing risk.

The elements of dose-response assessment are described in section III.A. Guidance on human exposure assessment is provided in another EPA

document (U.S. EPA, 1986); however, section III.B. of these Guidelines includes a brief description of the specific type of exposure information that is useful for carcinogen risk assessment. Finally, in section III.C. on risk characterization, there is a description of the manner in which risk estimates should be presented so as to be most informative.

It should be emphasized that calculation of quantitative estimates of cancer risk does not require that an agent be carcinogenic in humans. The likelihood that an agent is a human carcinogen is a function of the weight of evidence, as this has been described in the hazard identification section of these Guidelines. It is nevertheless important to present quantitative estimates, appropriately qualified and interpreted, in those circumstances in which there is a reasonable possibility, based on human and animal data, that the agent is carcinogenic in humans.

It should be emphasized in every quantitative risk estimation that the results are uncertain. Uncertainties due to experimental and epidemiologic variability as well as uncertainty in the exposure assessment can be important. There are major uncertainties in extrapolating both from animals to humans and from high to low doses. There are important species differences in uptake, metabolism, and organ distribution of carcinogens, as well as species and strain differences in target-site susceptibility. Human populations are variable with respect to genetic constitution, diet, occupational and home environment, activity patterns, and other cultural factors. Risk estimates should be presented together with the associated hazard assessment (section III.C.3.) to ensure that there is an appreciation of the weight of evidence for carcinogenicity that underlies the quantitative risk estimates.

A. Dose-Response Assessment

1. *Selection of Data.* As indicated in section II.D., guidance needs to be given by the individuals doing the qualitative assessment (toxicologists, pathologists, pharmacologists, etc.) to those doing the quantitative assessment as to the appropriate data to be used in the dose-response assessment. This is determined by the quality of the data, its relevance to human modes of exposure, and other technical details.

If available, estimates based on adequate human epidemiologic data are preferred over estimates based on animal data. If adequate exposure data exist in a well-designed and well-conducted negative epidemiologic study, it may be possible to obtain an upper-bound estimate of risk from that study. Animal-based estimates, if available, also should be presented.

In the absence of appropriate human studies, data from a species that responds most like humans should be used, if information to this effect exists. Where, for a given agent, several studies are available, which may involve different animal species, strains, and sexes at several doses and by different routes of exposure, the following approach to selecting the data sets is used: (1) The tumor incidence data are separated according to organ site and tumor type. (2) All biologically and statistically acceptable data sets are presented. (3) The range of the risk estimates is presented with due regard to biological relevance (particularly in the case of animal studies) and appropriateness of route of exposure. (4) Because it is possible that human sensitivity is as high as the most sensitive responding animal species, in the absence of evidence to the contrary, the biologically acceptable data set from long-term animal studies showing the greatest sensitivity should generally be given the greatest emphasis, again with due regard to biological and statistical considerations.

When the exposure route in the species from which the dose-response information is obtained differs from the route occurring in environmental exposures, the considerations used in making the route-to-route extrapolation must be carefully described. All assumptions should be presented along with a discussion of the uncertainties in the extrapolation. Whatever procedure is adopted in a given case, it must be consistent with the existing metabolic and pharmacokinetic information on the chemical (e.g., absorption efficiency via the gut and lung, target organ doses, and changes in placental transport throughout gestation for transplacental carcinogens).

Where two or more significantly elevated tumor sites or types are observed in the same study, extrapolations may be conducted on selected sites or types. These selections will be made on biological grounds. To obtain a total estimate of carcinogenic risk, animals with one or more tumor sites or types showing significantly elevated tumor incidence should be pooled and used for extrapolation. The pooled estimates will generally be used in preference to risk estimates based on single sites or types. Quantitative risk extrapolations will generally not be done on the basis of totals that include tumor sites without statistically significant elevations.

Benign tumors should generally be combined with malignant tumors for risk estimates unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin. The contribution of the benign tumors, however, to the total risk should be indicated.

2. *Choice of Mathematical Extrapolation Model.* Since risks at low exposure levels cannot be measured directly either by animal experiments or by epidemiologic studies, a number of mathematical models have been developed to extrapolate from high to low dose. Different extrapolation models, however, may fit the observed data reasonably well but may lead to large differences in the projected risk at low doses.

As was pointed out by OSTP (1985; Principle 26),

> No single mathematical procedure is recognized as the most appropriate for low-dose extrapolation in carcinogenesis. When relevant biological evidence on mechanism of action exists (e.g., pharmacokinetics, target organ dose), the models or procedures employed should be consistent with the evidence. When data and information are limited, however, and when much uncertainty exists regarding the mechanism of carcinogenic action, models or procedures which incorporate low-dose linearity are preferred when compatible with the limited information.

At present, mechanisms of the carcinogenesis process are largely unknown and data are generally limited. If a carcinogenic agent acts by accelerating the same carcinogenic process that leads to the background occurrence of cancer, the added effect of the carcinogen at low doses is expected to be virtually linear (Crump et al., 1976).

The Agency will review each assessment as to the evidence on carcinogenesis mechanisms and other biological or statistical evidence that indicates the suitability of a particular extrapolation model. Goodness-of-fit to the experimental observations is not an effective means of discriminating among models (OSTP, 1985). A rationale will be included to justify the use of the chosen model. In the absence of adequate information to the contrary, the linearized multistage procedure will be employed. Where appropriate, the results of using various extrapolation models may be useful for comparison with the linearized multistage procedure. When longitudinal data on tumor development are available, time-to-tumor models may be used.

It should be emphasized that the linearized multistage procedure leads to

a plausible upper limit to the risk that is consistent with some proposed mechanisms of carcinogenesis. Such an estimate, however, does not necessarily give a realistic prediction of the risk. The true value of the risk is unknown, and may be as low as zero. The range of risks, defined by the upper limit given by the chosen model and the lower limit which may be as low as zero, should be explicitly stated. An established procedure does not yet exist for making "most likely" or "best" estimates of risk within the range of uncertainty defined by the upper and lower limit estimates. If data and procedures become available, the Agency will also provide "most likely" or "best" estimates of risk. This will be most feasible when human data are available and when exposures are in the dose range of the data.

In certain cases, the linearized multistage procedure cannot be used with the observed data as, for example, when the data are nonmonotonic or flatten out at high doses. In these cases, it may be necessary to make adjustments to achieve low-dose linearity.

When pharmacokinetic or metabolism data are available, or when other substantial evidence on the mechanistic aspects of the carcinogenesis process exists, a low-dose extrapolation model other than the linearized multistage procedure might be considered more appropriate on biological grounds. When a different model is chosen, the risk assessment should clearly discuss the nature and weight of evidence that led to the choice. Considerable uncertainty will remain concerning response at low doses; therefore, in most cases an upper-limit risk estimate using the linearized multistage procedure should also be presented.

3. *Equivalent Exposure Units Among Species.* Low-dose risk estimates derived from laboratory animal data extrapolated to humans are complicated by a variety of factors that differ among species and potentially affect the response to carcinogens. Included among these factors are differences between humans and experimental test animals with respect to life span, body size, genetic variability, population homogeneity, existence of concurrent disease, pharmacokinetic effects such as metabolism and excretion patterns, and the exposure regimen.

The usual approach for making interspecies comparisons has been to use standardized scaling factors. Commonly employed standardized dosage scales include mg per kg body weight per day, ppm in the diet or water, mg per m^2 body surface area per day, and mg per kg body weight per lifetime. In the absence of comparative toxicological, physiological, metabolic, and pharmacokinetic data for a given suspect carcinogen, the Agency takes the position that the extrapolation on the basis of surface area is considered to be appropriate because certain pharmacological effects commonly scale according to surface area (Dedrick, 1973; Freireich et al., 1966; Pinkel, 1958).

B. Exposure Assessment

In order to obtain a quantitative estimate of the risk, the results of the dose-response assessment must be combined with an estimate of the exposures to which the populations of interest are likely to be subject. While the reader is referred to the Guidelines for Estimating Exposures (U.S. EPA, 1986) for specific details, it is important to convey an appreciation of the impact of the strengths and weaknesses of exposure assessment on the overall cancer risk assessment process.

At present there is no single approach to exposure assessment that is appropriate for all cases. On a case-by-case basis, appropriate methods are selected to match the data on hand and the level of sophistication required. The assumptions, approximations, and uncertainties need to be clearly stated because, in some instances, these will have a major effect on the risk assessment.

In general, the magnitude, duration, and frequency of exposure provide fundamental information for estimating the concentration of the carcinogen to which the organism is exposed. These data are generated from monitoring information, modeling results, and/or reasoned estimates. An appropriate treatment of exposure should consider the potential for exposure via ingestion, inhalation, and dermal penetration from relevant sources of exposures including multiple avenues of intake from the same source.

Special problems arise when the human exposure situation of concern suggests exposure regimens, e.g., route and dosing schedule, that are substantially different from those used in the relevant animal studies. Unless there is evidence to the contrary in a particular case, the cumulative dose received over a lifetime, expressed as average daily exposure prorated over a lifetime, is recommended as an appropriate measure of exposure to a carcinogen. That is, the assumption is made that a high dose of a carcinogen received over a short period of time is equivalent to a corresponding low-dose spread over a lifetime. This approach becomes more problematical as the exposures in question become more intense but less frequent, especially when there is evidence that the agent has shown dose-rate effects.

An attempt should be made to assess the level of uncertainty associated with the exposure assessment which is to be used in a cancer risk assessment. This measure of uncertainty should be included in the risk characterization (section III.C.) in order to provide the decision-maker with a clear understanding of the impact of this uncertainty on any final quantitative risk estimate. Subpopulations with heightened susceptibility (either because of exposure or predisposition) should, when possible, be identified.

C. Risk Characterization

Risk characterization is composed of two parts. One is a presentation of the numerical estimates of risk; the other is a framework to help judge the significance of the risk. Risk characterization includes the exposure assessment and dose-response assessment; these are used in the estimation of carcinogenic risk. It may also consist of a unit-risk estimate which can be combined elsewhere with the exposure assessment for the purposes of estimating cancer risk.

Hazard identification and dose-response assessment are covered in sections II and III.A., and a detailed discussion of exposure assessment is contained in EPA's Guidelines for Estimating Exposures (U.S. EPA, 1986). This section deals with the numerical risk estimates and the approach to summarizing risk characterization.

1. *Options for Numerical Risk Estimates.* Depending on the needs of the individual program offices, numerical estimates can be presented in one or more of the following three ways.

a. Unit Risk—Under an assumption of low-dose linearity, the unit cancer risk is the excess lifetime risk due to a continuous constant lifetime exposure of one unit of carcinogen concentration. Typical exposure units include ppm or ppb in food or water, mg/kg/day by ingestion, or ppm or $\mu g/m^3$ in air.

b. Dose Corresponding to a Given Level of Risk—This approach can be useful, particularly when using nonlinear extrapolation models where the unit risk would differ at different dose levels.

c. Individual and Population Risks—Risks may be characterized either in terms of the excess individual lifetime risks, the excess number of cancers

produced per year in the exposed population, or both.

Irrespective of the options chosen, the degree of precision and accuracy in the numerical risk estimates currently do not permit more than one significant figure to be presented.

2. *Concurrent Exposure.* In characterizing the risk due to concurrent exposure to several carcinogens, the risks are combined on the basis of additivity unless there is specific information to the contrary. Interactions of cocarcinogens, promoters, and initiators with known carcinogens should be considered on a case-by-case basis.

3. *Summary of Risk Characterization.* Whichever method of presentation is chosen, it is critical that the numerical estimates not be allowed to stand alone, separated from the various assumptions and uncertainties upon which they are based. The risk characterization should contain a discussion and interpretation of the numerical estimates that affords the risk manager some insight into the degree to which the quantitative estimates are likely to reflect the true magnitude of human risk, which generally cannot be known with the degree of quantitative accuracy reflected in the numerical estimates. The final risk estimate will be generally rounded to one significant figure and will be coupled with the EPA classification of the qualitative weight of evidence. For example, a lifetime individual risk of 2×10^{-4} resulting from exposure to a "probable human carcinogen" (Group B2) should be designated as: 2×10^{-4} [B2]. This bracketed designation of the qualitative weight of evidence should be included with all numerical risk estimates (i.e., unit risks, which are risks at a specified concentration or concentrations corresponding to a given risk). Agency statements, such as **Federal Register** notices, briefings, and action memoranda, frequently include numerical estimates of carcinogenic risk. It is recommended that whenever these numerical estimates are used, the qualitative weight-of-evidence classification should also be included.

The section on risk characterization should summarize the hazard identification, dose-response assessment, exposure assessment, and the public health risk estimates. Major assumptions, scientific judgments, and, to the extent possible, estimates of the uncertainties embodied in the assessment are presented.

IV. EPA Classification System for Categorizing Weight of Evidence for Carcinogenicity From Human and Animal Studies (Adapted From IARC)

A. Assessment of Weight of Evidence for Carcinogenicity From Studies in Humans

Evidence of carcinogenicity from human studies comes from three main sources:

1. Case reports of individual cancer patients who were exposed to the agent(s).
2. Descriptive epidemiologic studies in which the incidence of cancer in human populations was found to vary in space or time with exposure to the agent(s).
3. Analytical epidemiologic (case-control and cohort) studies in which individual exposure to the agent(s) was found to be associated with an increased risk of cancer.

Three criteria must be met before a causal association can be inferred between exposure and cancer in humans:

1. There is no identified bias that could explain the association.
2. The possibility of confounding has been considered and ruled out as explaining the association.
3. The association is unlikely to be due to chance.

In general, although a single study may be indicative of a cause-effect relationship, confidence in inferring a causal association is increased when several independent studies are concordant in showing the association, when the association is strong, when there is a dose-response relationship, or when a reduction in exposure is followed by a reduction in the incidence of cancer.

The weight of evidence for carcinogenicity [1] from studies in humans is classified as:

1. Sufficient evidence of carcinogenicity, which indicates that there is a causal relationship between the agent and human cancer.
2. Limited evidence of carcinogenicity, which indicates that a causal interpretation is credible, but that alternative explanations, such as chance, bias, or confounding, could not adequately be excluded.
3. Inadequate evidence, which indicates that one of two conditions prevailed: (a) there were few pertinent data, or (b) the available studies, while showing evidence of association, did not exclude chance, bias, or confounding and therefore a causal interpretation is not credible.
4. No data, which indicates that data are not available.
5. No evidence, which indicates that no association was found between exposure and an increased risk of cancer in well-designed and well-conducted independent analytical epidemiologic studies.

B. Assessment of Weight of Evidence for Carcinogenicity From Studies in Experimental Animals

These assessments are classified into five groups:

1. Sufficient evidence [2] of carcinogenicity, which indicates that there is an increased incidence of malignant tumors or combined malignant and benign tumors: [3] (a) in multiple species or strains; or (b) in multiple experiments (e.g., with different routes of administration or using different dose levels); or (c) to an unusual degree in a single experiment with regard to high incidence, unusual site or type of tumor, or early age at onset.

Additional evidence may be provided by data on dose-response effects, as well as information from short-term tests or on chemical structure.

2. Limited evidence of carcinogenicity, which means that the data suggest a carcinogenic effect but are limited because: (a) the studies involve a single species, strain, or experiment and do not meet criteria for sufficient evidence (see section IV. B.1.c); (b) the experiments are restricted by inadequate dosage levels, inadequate duration of exposure to the agent, inadequate period of follow-up, poor survival, too few animals, or inadequate reporting; or (c) an increase in the incidence of benign tumors only.
3. Inadequate evidence, which indicates that because of major qualitative or quantitative limitations, the studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect.
4. No data, which indicates that data are not available.
5. No evidence, which indicates that there is no increased incidence of neoplasms in at least two well-designed

[1] For purposes of public health protection, agents associated with life-threatening benign tumors in humans are included in the evaluation.

[2] An increased incidence of neoplasms that occur with high spontaneous background incidence (e.g., mouse liver tumors and rat pituitary tumors in certain strains) generally constitutes "sufficient" evidence of carcinogenicity, but may be changed to "limited" when warranted by the specific information available on the agent.

[3] Benign and malignant tumors will be combined unless the benign tumors are not considered to have the potential to progress to the associated malignancies of the same histogenic origin.

and well-conducted animal studies in different species.

The classifications "sufficient evidence" and "limited evidence" refer only to the weight of the experimental evidence that these agents are carcinogenic and not to the potency of their carcinogenic action.

C. Categorization of Overall Weight of Evidence for Human Carcinogenicity

The overall scheme for categorization of the weight of evidence of carcinogenicity of a chemical for humans uses a three-step process. (1) The weight of evidence in human studies or animal studies is summarized; (2) these lines of information are combined to yield a tentative assignment to a category (see Table 1); and (3) all relevant supportive information is evaluated to see if the designation of the overall weight of evidence needs to be modified. Relevant factors to be included along with the tumor information from human and animal studies include structure-activity relationships; short-term test findings; results of appropriate physiological, biochemical, and toxicological observations; and comparative metabolism and pharmacokinetic studies. The nature of these findings may cause one to adjust the overall categorization of the weight of evidence.

TABLE 1.—ILLUSTRATIVE CATEGORIZATION OF EVIDENCE BASED ON ANIMAL AND HUMAN DATA [1]

Human evidence	Animal evidence				
	Sufficient	Limited	Inadequate	No data	No. Evidence
Sufficient	A	A	A	A	A
Limited	B1	B1	B1	B1	B1
Inadequate	B2	C	D	D	D
No data	B2	C	D	D	E
No evidence	B2	C	D	D	E

[1] The above assignments are presented for illustrative purposes. There may be nuances in the classification of both animal and human data indicating that different categorizations than those given in the table should be assigned. Furthermore, these assignments are tentative and may be modified by ancillary evidence. In this regard all relevant information should be evaluated to determine if the designation of the overall weight of evidence needs to be modified. Relevant factors to be included along with the tumor data from human and animal studies include structure-activity relationships, short-term test findings, results of appropriate physiological, biochemical, and toxicological observations, and comparative metabolism and pharmacokinetic studies. The nature of these findings may cause an adjustment of the overall categorization of the weight of evidence.

The agents are categorized into five groups as follows:

Group A—Human Carcinogen

This group is used only when there is sufficient evidence from epidemiologic studies to support a causal association between exposure to the agents and cancer.

Group B—Probable Human Carcinogen

This group includes agents for which the weight of evidence of human carcinogenicity based on epidemiologic studies is "limited" and also includes agents for which the weight of evidence of carcinogenicity based on animal studies is "sufficient." The group is divided into two subgroups. Usually, Group B1 is reserved for agents for which there is limited evidence of carcinogenicity from epidemiologic studies. It is reasonable, for practical purposes, to regard an agent for which there is "sufficient" evidence of carcinogenicity in animals as if it presented a carcinogenic risk to humans. Therefore, agents for which there is "sufficient" evidence from animal studies and for which there is "inadequate evidence" or "no data" from epidemiologic studies would usually be categorized under Group B2.

Group C—Possible Human Carcinogen

This group is used for agents with limited evidence of carcinogenicity in animals in the absence of human data. It includes a wide variety of evidence, e.g., (a) a malignant tumor response in a single well-conducted experiment that does not meet conditions for sufficient evidence, (b) tumor responses of marginal statistical significance in studies having inadequate design or reporting, (c) benign but not malignant tumors with an agent showing no response in a variety of short-term tests for mutagenicity, and (d) responses of marginal statistical significance in a tissue known to have a high or variable background rate.

Group D—Not Classifiable as to Human Carcinogenicity

This group is generally used for agents with inadequate human and animal evidence of carcinogenicity or for which no data are available.

Group E—Evidence of Non-Carcinogenicity for Humans

This group is used for agents that show no evidence for carcinogenicity in at least two adequate animal tests in different species or in both adequate epidemiologic and animal studies.

The designation of an agent as being in Group E is based on the available evidence and should not be interpreted as a definitive conclusion that the agent will not be a carcinogen under any circumstances.

V. References

Albert, R.E., Train, R.E., and Anderson, E. 1977. Rationale developed by the Environmental Protection Agency for the assessment of carcinogenic risks. J. Natl. Cancer Inst. 58:1537–1541.

Crump, K.S., Hoel, D.G., Langley, C.H., Peto, R. 1976. Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Res. 16:2973–2979.

Dedrick, R.L. 1973. Animal Scale Up. J. Pharmacokinet. Biopharm. 1:435–461.

Feron, V.J., Grice, H.C., Griesemer, R., Peto, R., Agthe, C., Althoff, J., Arnold, D.L., Blumenthal, H., Cabral, J.R.P., Della Porta, G., Ito, N., Kimmerle, G., Kroes, R., Mohr, U., Napalkov, N.P., Odashima, S., Page, N.P., Schramm, T., Steinhoff, D., Sugar, J., Tomatis, L., Uehleke, H., and Vouk, V. 1980. Basic requirements for long-term assays for carcinogenicity. In: Long-term and short-term screening assays for carcinogens: a critical appraisal. IARC Monographs, Supplement 2. Lyon, France: International Agency for Research on Cancer, pp 21–83.

Freireich, E.J., Gehan, E.A., Rall, D.P., Schmidt, L.H., and Skipper, H.E. 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. Cancer Chemother. Rep. 50:219–244.

Haseman, J.K. 1984. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environ. Health Perspect. 58:385–392.

Haseman, J.K., Huff, J., and Boorman, G.A. 1984. Use of historical control data in carcinogenicity studies in rodents. Toxicol. Pathol. 12:126–135.

Interagency Regulatory Liaison Group (IRLG). 1979. Scientific basis for identification of potential carcinogens and estimation of risks. J. Natl. Cancer Inst. 63:245–267.

Interdisciplinary Panel on Carcinogenicity. 1984. Criteria for evidence of chemical carcinogenicity. Science 225:682–687.

International Agency for Research on Cancer (IARC). 1982. IARC Monographs on the

Evaluation of the Carcinogenic Risk of Chemicals to Humans. Supplement 4. Lyon, France: International Agency for Research on Cancer.

Mantel, N. 1980. Assessing laboratory evidence for neoplastic activity. Biometrics 36:381–399.

Mantel, N., and Haenszel, W. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. 22:719–748.

National Center for Toxicological Research (NCTR). 1981. Guidelines for statistical tests for carcinogenicity in chronic bioassays. NCTR Biometry Technical Report 81–001. Available from: National Center for Toxicological Research.

National Research Council (NRC). 1983. Risk assessment in the Federal government: managing the process. Washington, D.C.: National Academy Press.

National Toxicology Program. 1984. Report of the Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation of the National Toxicology Program. Board of Scientific Counselors. Available from: U.S. Government Printing Office, Washington, D.C. 1984–421–132:4726.

Nutrition Foundation. 1983. The relevance of mouse liver hepatoma to human carcinogenic risk: a report of the International Expert Advisory Committee to the Nutrition Foundation. Available from: Nutrition Foundation. ISBN 0–935368–37–x.

Office of Science and Technology Policy (OSTP). 1985. Chemical carcinogens: review of the science and its associated principles. Federal Register 50:10372–10442.

Peto, R., Pike, M., Day, N., Gray, R., Lee, P., Parish, S., Peto, J., Richard, S., and Wahrendorf, J. 1980. Guidelines for simple, sensitive, significant tests for carcinogenic effects in long-term animal experiments. In: Monographs on the long-term and short-term screening assays for carcinogens: a critical appraisal. IARC Monographs. Supplement 2. Lyon, France: International Agency for Research on Cancer. pp. 311–426.

Pinkel, D. 1958. The use of body surface area as a criterion of drug dosage in cancer chemotherapy. Cancer Res. 18:853–856.

Tomatis, L. 1977. The value of long-term testing for the implementation of primary prevention. In: Origins of human cancer. Hiatt, H.H., Watson, J.D., and Winstein, J.A., eds. Cold Spring Harbor Laboratory, pp. 1339–1357.

U.S. Environmental Protection Agency (U.S. EPA). 1976. Interim procedures and guidelines for health risk and economic impact assessments of suspected carcinogens. Federal Register 41:21402–21405.

U.S. Environmental Protection Agency (U.S. EPA). 1980. Water quality criteria documents; availability. Federal Register 45:79318–79379.

U.S. Environmental Protection Agency (U.S. EPA). 1983a. Good laboratory practices standards—toxicology testing. Federal Register 48:53922.

U.S. Environmental Protection Agency (U.S. EPA). 1983b. Hazard evaluations: humans and domestic animals. Subdivision F. Available from: NTIS, Springfield, VA. PB 83–153916.

U.S. Environmental Protection Agency (U.S. EPA). 1983c. Health effects test guidelines. Available from: NTIS, Springfield, VA. PB 83–232984.

U.S. Environmental Protection Agency (U.S. EPA). 1986. Guidelines for estimating exposures.

U.S. Food and Drug Administration (U.S. FDA). 1982. Toxicological principles for the safety assessment of direct food additives and color additives used in food. Available from: Bureau of Foods, U.S. Food and Drug Administration.

Ward, J.M., Griesemer, R.A., and Weisburger, E.K. 1979a. The mouse liver tumor as an endpoint in carcinogenesis tests. Toxicol. Appl. Pharmacol. 51:389–397.

Ward, J.M., Goodman, D.G., Squire, R.A., Chu, K.C., and Linhart, M.S. 1979b. Neoplastic and nonneoplastic lesions in aging (C57BL/6N x C3H/HeN)F_1 (B6C3F_1) mice. J. Natl. Cancer Inst. 63:849–854.

Part B: Response to Public and Science Advisory Board Comments

I. Introduction

This section summarizes the major issues raised during both the public comment period on the Proposed Guidelines for Carcinogen Risk Assessment published on November 23, 1984 (49 FR 46294), and also during the April 22–23, 1985, meeting of the Carcinogen Risk Assessment Guidelines Panel of the Science Advisory Board (SAB).

In order to respond to these issues the Agency modified the proposed guidelines in two stages. First, changes resulting from consideration of the public comments were made in a draft sent to the SAB review panel prior to their April meeting. Secondly, the guidelines were further modified in response to the panel's recommendations.

The Agency received 62 sets of comments during the public comment period, including 28 from corporations, 9 from professional or trade associations, and 4 from academic institutions. In general, the comments were favorable. The commentors welcomed the update of the 1976 guidelines and felt that the proposed guidelines of 1985 reflected some of the progress that has occurred in understanding the mechanisms of carcinogenesis. Many commentors, however, felt that additional changes were warranted.

The SAB concluded that the guidelines are "reasonably complete in their conceptual framework and are sound in their overall interpretation of the scientific issues" (Report by the SAB Carcinogenicity Guidelines Review Group, June 19, 1985). The SAB suggested various editorial changes and raised some issues regarding the content of the proposed guidelines, which are discussed below. Based on these recommendations, the Agency has modified the draft guidelines.

II. Office of Science and Technology Policy Report on Chemical Carcinogens

Many commentors requested that the final guidelines not be issued until after publication of the report of the Office of Technology and Science Policy (OSTP) on chemical carcinogens. They further requested that this report be incorporated into the final Guidelines for Carcinogen Risk Assessment.

The final OSTP report was published in 1985 (50 FR 10372). In its deliberations, the Agency reviewed the final OSTP report and feels that the Agency's guidelines are consistent with the principles established by the OSTP. In its review, the SAB agreed that the Agency quidelines are generally consistent with the OSTP report. To emphasize this consistency, the OSTP principles have been incorporated into the guidelines when controversial issues are discussed.

III. Inference Guidelines

Many commentors felt that the proposed guidelines did not provide a sufficient distinction between scientific fact and policy decisions. Others felt that EPA should not attempt to propose firm guidelines in the absence of scientific consensus. The SAB report also indicated the need to "distinguish recommendations based on scientific evidence from those based on science policy decisions."

The Agency agrees with the recommendation that policy, judgmental, or inferential decisions should be clearly identified. In its revision of the proposed guidelines, the Agency has included phrases (e.g., "the Agency takes the position that") to more clearly distinguish policy decisions.

The Agency also recognizes the need to establish procedures for action on important issues in the absence of complete scientific knowledge or consensus. This need was acknowledged in both the National Academy of Sciences book entitled *Risk Management in the Federal Government: Managing the Process* and the OSTP report on chemical carcinogens. As the NAS report states, "Risk assessment is an analytic process that is firmly based on scientific considerations, but it also requires judgments to be made when the available information is incomplete. These judgments inevitably draw on both scientific and policy considerations."

The judgments of the Agency have been based on current available scientific information and on the combined experience of Agency experts. These judgments, and the resulting guidance, rely on inference; however, the positions taken in these inference guidelines are felt to be reasonable and scientifically defensible. While all of the guidance is, to some degree, based on inference the guidelines have attempted to distinguish those issues that depended more on judgment. In these cases, the Agency has stated a position but has also retained flexibility to accommodate new data or specific circumstances that demonstrate that the proposed position is inaccurate. The Agency recognizes that scientific opinion will be divided on these issues.

Knowledge about carcinogens and carcinogenesis is progressing at a rapid rate. While these guidelines are considered a best effort at the present time, the Agency has attempted to incorporate flexibility into the current guidelines and also recommends that the guidelines be revised as often as warranted by advances in the field.

IV. Evaluation of Benign Tumors

Several commentors discussed the appropriate interpretation of an increased incidence of benign tumors alone or with an increased incidence of malignant tumors as part of the evaluation of the carcinogenicity of an agent. Some comments were supportive of the position in the proposed guidelines, i.e., under certain circumstances, the incidence of benign and malignant tumors would be combined, and an increased incidence of benign tumors alone would be considered an indication, albeit limited, of carcinogenic potential. Other commentors raised concerns about the criteria that would be used to decide which tumors should be combined. Only a few commentors felt that benign tumors should never be considered in evaluating carcinogenic potential.

The Agency believes that current information supports the use of benign tumors. The guidelines have been modified to incorporate the language of the OSTP report, i.e., benign tumors will be combined with malignant tumors when scientifically defensible. This position allows flexibility in evaluating the data base for each agent. The guidelines have also been modified to indicate that, whenever benign and malignant tumors have been combined, and the agent is considered a candidate for quantitative risk extrapolation, the contribution of benign tumors to the estimation of risk will be indicated.

V. Transplacental and Multigenerational Animal Bioassays

As one of its two proposals for additions to the guidelines, the SAB recommended a discussion of transplacental and multigenerational animal bioassays for carcinogenicity.

The Agency agrees that such data, when available, can provide useful information in the evaluation of a chemical's potential carcinogenicity and has stated this in the final guidelines. The Agency has also revised the guidelines to indicate that such studies may provide additional information on the metabolic and pharmacokinetic properties of the chemical. More guidance on the specific use of these studies will be considered in future revisions of these guidelines.

VI. Maximum Tolerated Dose

The proposed guidelines discussed the implications of using a maximum tolerated dose (MTD) in bioassays for carcinogenicity. Many commentors requested that EPA define MTD. The tone of the comments suggested that the commentors were concerned about the uses and interpretations of high-dose testing.

The Agency recognizes that controversy currently surrounds these issues. The appropriate text from the OSTP report has been incorporated into the final guidelines which suggests that the consequences of high-dose testing be evaluated on a case-by-case basis.

VII. Mouse Liver Tumors

A large number of commentors expressed opinions about the assessment of bioassays in which the only increase in tumor incidence was liver tumors in the mouse. Many felt that mouse liver tumors were afforded too much credence, especially given existing information that indicates that they might arise by a different mechanism, e.g., tissue damage followed by regeneration. Others felt that mouse liver tumors were but one case of a high background incidence of one particular type of tumor and that all such tumors should be treated in the same fashion.

The Agency has reviewed these comments and the OSTP principle regarding this issue. The OSTP report does not reach conclusions as to the treatment of tumors with a high spontaneous background rate, but states, as is now included in the text of the guidelines, that these data require special consideration. Although questions have been raised regarding the validity of mouse liver tumors in general, the Agency feels that mouse liver tumors cannot be ignored as an indicator of carcinogenicity. Thus, the position in the proposed guidelines has not been changed: an increased incidence of only mouse liver tumors will be regarded as "sufficient" evidence of carcinogenicity if all other criteria, e.g., replication and malignancy, are met with the understanding that this classification could be changed to "limited" if warranted. The factors that may cause this re-evaluation are indicated in the guidelines.

VIII. Weight-of-Evidence Categories

The Agency was praised by both the public and the SAB for incorporating a weight-of-evidence scheme into its evaluation of carcinogenic risk. Certain specific aspects of the scheme, however, were criticized.

1. Several commentors noted that while the text of the proposed guidelines clearly states that EPA will use all available data in its categorization of the weight of the evidence that a chemical is a carcinogen, the classification system in Part A, section IV did not indicate the manner in which EPA will use information other than data from humans and long-term animal studies in assigning a weight-of-evidence classification.

The Agency has added a discussion to Part A, section IV.C. dealing with the characterization of overall evidence for human carcinogenicity. This discussion clarifies EPA's use of supportive information to adjust, as warranted, the designation that would have been made solely on the basis of human and long-term animal studies.

2. The Agency agrees with the SAB and those commentors who felt that a simple classification of the weight of evidence, e.g., a single letter or even a descriptive title, is inadequate to describe fully the weight of evidence for each individual chemical. The final guidelines propose that a paragraph summarizing the data should accompany the numerical estimate and weight-of-evidence classification whenever possible.

3. Several commentors objected to the descriptive title E (No Evidence of Carcinogenicity for Humans) because they felt the title would be confusing to people inexperienced with the classification system. The title for Group E, No Evidence of Carcinogenicity for Humans, was thought by these commentors to suggest the absence of data. This group, however, is intended to be reserved for agents for which there exists credible data demonstrating that the agent is not carcinogenic.

Based on these comments and further discussion, the Agency has changed the

title of Group E to "Evidence of Non-Carcinogenicity for Humans."

4. Several commentors felt that the title for Group C, Possible Human Carcinogen, was not sufficiently distinctive from Group B, Probable Human Carcinogen. Other commentors felt that those agents that minimally qualified for Group C would lack sufficient data for such a label.

The Agency recognizes that Group C covers a range of chemicals and has considered whether to subdivide Group C. The consensus of the Agency's Carcinogen Risk Assessment Committee, however, is that the current groups, which are based on the IARC categories, are a reasonable stratification and should be retained at present. The structure of the groups will be reconsidered when the guidelines are reviewed in the future. The Agency also feels that the descriptive title it originally selected best conveys the meaning of the classification within the context of EPA's past and current activities.

5. Some commentors indicated a concern about the distinction between B1 and B2 on the basis of epidemiologic evidence only. This issue has been under discussion in the Agency and may be revised in future versions of the guidelines.

6. Comments were also received about the possibility of keeping the groups for animal and human data separate without reaching a combined classification. The Agency feels that a combined classification is useful; thus, the combined classification was retained in the final guidelines.

The SAB suggested that a table be added to Part A, section IV to indicate the manner in which human and animal data would be combined to obtain an overall weight-of-evidence category. The Agency realizes that a table that would present all permutations of potentially available data would be complex and possibly impossible to construct since numerous combinations of ancillary data (e.g., genetic toxicity, pharmacokinetics) could be used to raise or lower the weight-of-evidence classification. Nevertheless, the Agency decided to include a table to illustrate the most probable weight-of-evidence classification that would be assigned on the basis of standard animal and human data without consideration of the ancillary data. While it is hoped that this table will clarify the weight-of-evidence classifications, it is also important to recognize that an agent may be assigned to a final categorization different from the category which would appear appropriate from the table and still conform to the guidelines.

IX. Quantitative Estimates of Risk

The method for quantitative estimates of carcinogenic risk in the proposed guidelines received substantial comments from the public. Five issues were discussed by the Agency and have resulted in modifications of the guidelines.

1. The major criticism was the perception that EPA would use only one method for the extrapolation of carcinogenic risk and would, therefore, obtain one estimate of risk. Even commentors who concur with the procedure usually followed by EPA felt that some indication of the uncertainty of the risk estimate should be included with the risk estimate.

The Agency feels that the proposed guidelines were not intended to suggest that EPA would perform quantitative risk estimates in a rote or mechanical fashion. As indicated by the OSTP report and paraphrased in the proposed guidelines, no single mathematical procedure has been determined to be the most appropriate method for risk extrapolation. The final guidelines quote rather than paraphrase the OSTP principle. The guidelines have been revised to stress the importance of considering all available data in the risk assessment and now state, "The Agency will review each assessment as to the evidence on carcinogenic mechanisms and other biological or statistical evidence that indicates the suitability of a particular extrapolation model." Two issues are emphasized: First, the text now indicates the potential for pharmacokinetic information to contribute to the assessment of carcinogenic risk. Second, the final guidelines state that time-to-tumor risk extrapolation models may be used when longitudinal data on tumor development are available.

2. A number of commentors noted that the proposed guidelines did not indicate how the uncertainties of risk characterization would be presented. The Agency has revised the proposed guidelines to indicate that major assumptions, scientific judgments, and, to the extent possible, estimates of the uncertainties embodied in the risk assessment will be presented along with the estimation of risk.

3. The proposed guidelines stated that the appropriateness of quantifying risks for chemicals in Group C (Possible Human Carcinogen), specifically those agents that were on the boundary of Groups C and D (Not Classifiable as to Human Carcinogenicity), would be judged on a case-by-case basis. Some commentors felt that quantitative risk assessment should not be performed on any agent in Group C.

Group C includes a wide range of agents, including some for which there are positive results in one species in one good bioassay. Thus, the Agency feels that many agents in Group C will be suitable for quantitative risk assessment, but that judgments in this regard will be made on a case-by-case basis.

4. A few commentors felt that EPA intended to perform quantitative risk estimates on aggregate tumor incidence. While EPA will consider an increase in total aggregate tumors as suggestive of potential carcinogenicity, EPA does not generally intend to make quantitative estimates of carcinogenic risk based on total aggregate tumor incidence.

5. The proposed choice of body surface area as an interspecies scaling factor was criticized by several commentors who felt that body weight was also appropriate and that both methods should be used. The OSTP report recognizes that both scaling factors are in common use. The Agency feels that the choice of the body surface area scaling factor can be justified from the data on effects of drugs in various species. Thus, EPA will continue to use this scaling factor unless data on a specific agent suggest that a different scaling factor is justified. The uncertainty engendered by choice of scaling factor will be included in the summary of uncertainties associated with the assessment of risk mentioned in point 1, above.

In the second of its two proposals for additions to the proposed guidelines, the SAB suggested that a sensitivity analysis be included in EPA's quantitative estimate of a chemical's carcinogenic potency. The Agency agrees that an analysis of the assumptions and uncertainties inherent in an assessment of carcinogenic risk must be accurately portrayed. Sections of the final guidelines that deal with this issue have been strengthened to reflect the concerns of the SAB and the Agency. In particular, the last paragraph of the guidelines states that "major assumptions, scientific judgments, and, to the extent possible, estimates of the uncertainties embodied in the assessment" should be presented in the summary characterizing the risk. Since the assumptions and uncertainties will vary for each assessment, the Agency feels that a formal requirement for a particular type of sensitivity analysis would be less useful than a case-by-case evaluation of the particular assumptions and uncertainties most significant for a particular risk assessment.

[FR Doc. 86-19601 Filed 9-23-86; 8:45 am]

BILLING CODE 6560-50-M

012