United States
Environmental
Protection Agency

Science Advisory Board
Washington DC

EPA-SAB-EHC-97-010
September 1997

AN SAB REPORT: GUIDELINES FOR CANCER RISK ASSESSMENT

REVIEW OF THE OFFICE OF RESEARCH AND DEVELOPMENT'S DRAFT GUIDELINES FOR CANCER RISK ASSESSMENT BY THE ENVIRONMENTAL HEALTH COMMITTEE

September 30, 1997

EPA-SAB-EHC-97-010

Honorable Carol M. Browner
Administrator
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460

Subject: Science Advisory Board's review of the revised Guidelines for Cancer Risk Assessment

Dear Ms. Browner:

In 1983, the National Academy of Science/National Research Council (NAS/NRC) published a report entitled *Risk Assessment in the Federal Government: Managing the Process*. In that report, the NRC recommended that Federal regulatory agencies establish "inference guidelines" to insure consistency and technical quality in risk assessments, and to ensure that the risk assessment process was maintained as a scientific process separate from risk management. The Agency first issued cancer guidelines in 1986. In April 1996, the Agency released for public comment new proposed Guidelines, incorporating significant changes in the approach assessing risk. Following public comment, the Science Advisory Board's (SAB) Environmental Health Committee (EHC) was asked to review the proposed revisions, and subsequently met on February 13-14, 1997, in Washington, DC.

In terms of a "global overview." the Committee regards the proposed Guidelines as constituting a significant step forward in the "state-of-the-art" for carcinogen risk assessment. The EHC particularly commends the Agency for addressing the controversial aspects of the new Guidelines with a frank, unbiased approach to all points of view. The new Guidelines, when compared with the 1986 version, will cause risk assessors to place greater emphasis (than do the current Guidelines) on the utilization of all the available scientific information in characterizing cancer risks.

The above comments not withstanding, the EHC noted several areas in which the proposed Guidelines would benefit from clarification or revision. These areas are:

a) The Committee endorses the Guidelines' emphasis on the use of narrative discussion describing the weight of evidence, but found problems the proposed categorization, particularly in the use of multiple terms (e.g., categories/descriptors/subdescriptors). Given the complexities involved, the Committee could not come to a consensus as to how this problem should be addressed.

b) The majority of the Committee supported the Agency's handling of the issues raised when departure from the defaults are contemplated. A sizeable minority of the Committee believed that the burden of proof should rest on showing that the defaults are implausible, and also that the Guidelines should be revised to include explicit and specific criteria for judging the validity of hypotheses invoked to depart from defaults.

c) The proposed Guidelines state that: the new default procedures are "public health conservative," but that they have been revised, when appropriate, to reflect the changes in the state-of-the-art since 1986. The EHC clearly understands that the primary goal of EPA actions is public health protection and recommends that clarification be added to the final Guidelines to alleviate potential concerns by readers or users of the Guidelines, while at the same time promoting the use of good science for decision making.

d) The EHC generally endorsed the Guidelines' mode of action proposals, but suggested that the Guidelines contain specific criteria for judging that the data on mode of action are valid and adequate. This should include a discussion of mode of action that reflects the lack of a clear distinction between direct DNA damage and other mechanisms with respect to the low dose response relationship.

e) The proposed Guidelines provide for more flexibility in risk assessment and provide means for incorporating other types of biological data and information on mode of action into dose response assessment. However, some EHC Members noted that, in some areas, there may not be enough guidance, given the wide usage of the Guidelines, and in other areas additional clarification is needed. Although sympathetic to the Agency's

reluctance to develop additional criteria, the EHC believes that further guidance should be provided.

f) The EHC understands the theoretical desirability of the proposed biologically based dose response model. However, the proposed Guidelines may be misleading in this respect because no such model presently exists in a complete form, nor is it clear how to develop such in the foreseeable future. The Agency's definition of a biologically based model thus seems unnecessarily narrow. The Agency should define more clearly what is meant by a biologically based dose response model and to give guidance as to when such a model is preferred over the default linear or non-linear approaches. The enclosed report, in section 3.3.2, proposes a possible definition for the Agency's consideration.

g) To mitigate against unnecessary inconsistencies and confusion in the application of the Guidelines to determine the point of departure, the EHC suggests that further guidance be given in the Guidelines on this approach.

h) There is no explicit statement in the proposal that statistical significance should be a basic requirement for determining causality. This lack of an explicit statement has been interpreted as misleading and implying that there is a hidden intent to eliminate statistical significance as a consideration in assessing causality. Adding appropriate and specific language concerning statistical significance should rectify this problem.

i) The developing infant and child should be recognized as a population subgroup that is particularly sensitive to the carcinogenicity of a number of agents. Risk assessments should explicitly account for the differential susceptibility of the young. Other differences in susceptibility should also be taken into account when data permit.

j) *Vis-a-vis* assessing tumors, the Committee endorses the procedure adopted by the National Toxicology program of combining some closely related tumor types for statistical analyses, but otherwise conducting separate tests for different tumor types. The EHC also endorses the Agency's Guidelines that give additional significance to the finding of uncommon tumor types, multiple sites, tumors by more than one route of administration, etc. (see pp. 55-56 of the Guidelines). Lastly, because of

data on the relationship of body weight and some tumor types, the EHC encourages the Agency to consider the role of diet in a given tumor response. Consideration of this issue should be addressed when evaluating a given tumor data set.

We appreciate having been given the opportunity to address these issues, and look forward to receiving your response to our comments.

Sincerely,

/signed/
Dr. Genevieve Matanoski, Chair
Science Advisory Board

/signed/
Dr. Emil Pfitzer, Chair
Environmental Health Committee

Enclosure

NOTICE

This report has been written as a part of the activities of the Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters relating to problems facing the Agency. This report has not been reviewed for approval by the Agency and, therefore, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

ERRATA

Copies of this report distributed prior to 11/5/97 contain an error on page ii, in paragraph (h) of the Abstract.

The erroneous version stated:

h) there should be an explicit statement that statistical significance should be a basic requirement for determining causality.

The correct statement, incorporated in this document is:

g) there should be an explicit statement in the Guidelines that statistical significance should be considered in assessing causality.

ABSTRACT

EPA first issued cancer guidelines in 1986. In April 1996, new draft Guidelines were release. The Science Advisory Board's (SAB) Environmental Health Committee (EHC) was asked to review the proposed revisions, and met on February 13-14, 1997, in Washington, DC.

The Committee found the proposed Guidelines to be significant step forward in carcinogen risk assessment, and particularly commended the Agency for addressing the controversial aspects of the new Guidelines with a frank, unbiased approach. The new Guidelines will cause risk assessors to place greater emphasis on the utilization of all the available scientific information in characterizing cancer risks. However, the also Committee identified areas in which improvement was possible. These are:

a) problems with the proposed categorization (The Committee could not come to a consensus as to how this problem should be addressed.)

b) most Members supported the Agency's handling of the issues raised when departure from the defaults are contemplated. A sizeable minority of the Committee believed that the burden of proof should rest on showing that the defaults are implausible.

c) the Guidelines should be clarified to alleviate potential concerns by users about the impacts of changes on the degree to which the document remains "public health conservative."

d) the EHC generally endorsed the Guidelines' mode of action proposals, but suggested that the Guidelines contain specific criteria for judging that the data on mode of action are valid and adequate.

e) some Members noted that there may not be enough guidance on incorporating other types of biological data and information on mode of action into dose response assessment.

f) the Agency's definition of a biologically based model seems unnecessarily narrow.

g) the EHC suggests that further guidance be given in the Guidelines on determine the point of departure.

h) there should be an explicit statement in the Guidelines that statistical significance should be considered in assessing causality.

I) The developing infant/child should be recognized as a population subgroup that is particularly sensitive to the carcinogenicity of a number of agents.

j) The Committee endorses the procedure adopted by the National Toxicology program of combining some closely related tumor types for statistical analyses, but otherwise conducting separate tests for different tumor types.

KEYWORDS: cancer; risk assessment; guidelines; subpopuations; default values; mode of action; cancer mechanisms.

U.S. ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD
ENVIRONMENTAL HEALTH COMMITTEE

February 13-14, 1997

CHAIR
Dr. Emil Pfitzer, Research Institute for Fragrance Materials, Inc.,Hackensack, NJ

MEMBERS
Dr. Charles Capen (SAP), Department of Veterinary Biosciences, The Ohio State University, Columbus, OH

Dr. Adolfo Correa, The Johns Hopkins University, School of Hygiene and Public Health, Baltimore, MD

Dr. Kenny S. Crump, ICF Kaiser, Ruston, LA

Dr. Ernest E. McConnell (SAP), Raleigh, NC

Dr. Frederica Perera, School of Public Health, Columbia University, NYC, NY

Dr. Henry C. Pitot, McArdle Laboratory for Cancer Research, Madison, WI

Dr. James A. Swenberg, University of North Carolina, Chapel Hill, NC

Dr. Mark J. Utell, University of Rochester Medical Center, Rochester, NY

Dr. Lauren Zeise, California Environmental Protection Agency, Berkeley, CA

CONSULTANT
Dr. Roger O. McClellan, Chemical Industry Institute of Toxicology, Research Triangle Park, NC

SCIENCE ADVISORY BOARD STAFF
Mr. Samuel Rondberg, Designated Federal Official, US EPA, Science Advisory Board (1400), 401 M Street, SW, Washington, D.C. 20460

Ms. Mary L. Winston, Staff Secretary US EPA, Science Advisory Board (1400) 401 M Street, SW, Washington, DC 20460

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The Environmental Health Committee (EHC) met on February 13-14, 1997, in Washington, DC to review the EPA's Proposed Guidelines for Cancer Risk Assessment. The Committee, in general, endorsed the proposal, and regard it as constituting a significant step forward in carcinogen risk assessment. In particular, the new Guidelines, when implemented, will cause risk assessors to place greater emphasis (than do the current Guidelines) on the utilization of all the available scientific information in characterizing cancer risks.

As with all such reviews of a highly complex enterprise, the EHC identified areas wherein clarifications and/or revisions are advised. These areas, discussed in detail in section 3 of this report, are outlined below:

a) The Committee endorses the Guidelines' emphasis on the use of narrative discussion describing the weight of evidence, but found problems with the resultant categorizations, particularly in the use of multiple terms such as categories/descriptors/subdescriptors). The Committee could not come to a consensus as to how this problem should be addressed.

b) The majority of the Committee supported the Agency's handling of the issues raised when departure from the defaults are contemplated. A sizeable minority of the Committee believed that the burden of proof should rest on showing that the defaults are implausible, and also that the Guidelines should be revised to include explicit and specific criteria for judging the validity of hypotheses invoked to depart from defaults.

c) The proposed Guidelines state that: the proposed default procedures are "public health conservative," but that they have been revised, when appropriate, to reflect the changes in the state-of-the-art since 1986. The EHC clearly understands that the primary goal of EPA actions is public health protection and recommends that clarifications be included in the final Guidelines to alleviate potential concerns by readers or users of the Guidelines, while at the same time promoting the use of good science for decision making.

d) The EHC generally endorsed the Guidelines' mode of action proposals, but suggested that the Guidelines contain specific criteria for judging that

the data on mode of action are valid and adequate. This should include a discussion of mode of action that reflects the lack of a clear distinction between direct DNA damage and other mechanisms with respect to the low dose response relationship.

e) The proposed Guidelines provide for increased flexibility in risk assessment and provide means for incorporating other types of biological data and information on mode of action into dose response assessment. Although sympathetic to the Agency's reluctance to develop additional criteria, the EHC believes that further guidance on determing the mode of action should be provided.

f) The EHC understands the theoretical desirability of the proposed biologically based dose response model. However, the Agency's definition of a biologically based model seems to be unnecessarily narrow. The Agency should define more clearly what is meant by a biologically based dose response model and give guidance as to when such a model is preferred over the default linear or non-linear approaches.

g) To mitigate against unnecessary inconsistencies and confusion in the application of the Guidelines to determine the point of departure, the EHC suggests that further guidance be given in the Guidelines on this subject.

h) There should be an explicit statement in the Guidelines that statistical significance should be considered in assessing causality.

i) The developing infant and child should be recognized as a subgroup which is particularly sensitive to the carcinogenicity of a number of agents. Risk assessments should account explicitly for the differential susceptibility of the young. Other differences in susceptibility should also be taken into account when data permit.

j) *Viis-a-vis* assessing tumors, the Committee endorses the procedure adopted by the National Toxicology program of combining some closely related tumor types for statistical analyses, but otherwise conducting separate tests for different tumor types. The EHC also endorses the Agency's Guidelines that give additional significance to the finding of uncommon tumor types, multiple sites, tumors by more than one route of administration, etc. (see pp. 55-56 of the Guidelines). Lastly, because of

data on the relationship of body weight and some tumor types, the EHC encourages the Agency to consider the role of diet in a given tumor response.

.2 BACKGROUND AND CHARGE

2.1 Background

In 1983, the National Academy of Science/National Research Council (NAS/NRC) published a report entitled *Risk Assessment in the Federal Government: Managing the Process*. In that report, the NRC recommended that Federal regulatory agencies establish “inference guidelines” to insure consistency and technical quality in risk assessments, and to ensure that the risk assessment process was maintained as a scientific process separate from risk management. The Agency first issued cancer guidelines in 1986. In April 1996, the Agency released for public comment new proposed Guidelines, incorporating significant changes in the approach assessing risk. Following public comment, the Science Advisory Board’s (SAB) Environmental Health Committee (EHC) was asked to review the proposed revisions. The EHC subsequently met on February 13-14, 1997, in Washington, DC.

2.2 Charge

In addressing theCharge, the Committee took special note of the topics highlighted during the public comment process. In addition, under Clean Air Act Amendments of 1990, section 112, the SAB is charged to review the responses the Agency makes to recommendations of the National Academy report *Science and Judgment in Risk Assessment* with respect to cancer guidelines issues. (The EPA responses are in section 1.3 and Appendix B of the proposed Guidelines).

The detailed Charge follows:

a) Hazard Characterization: Descriptors and Narratives

In its 1986 guidelines, the EPA adopted a classification system designating agents in groups"A" through "E", Group A for "known" human carcinogens through Group E for agents with evidence of noncarcinogenicity. In its 1996 proposal, the Agency has proposed to express its conclusions through a narrative with standard descriptors.

The proposal has discussed certain limitations in the Agency’s 1986 carcinogen classification system, and has sought to move from the letter and number classification to increase the information provided to risk managers. To do so, the Agency has proposed incorporating a limited number of descriptor phrases into a weight of

evidence narrative summary of human, animal, and other key data. The proposal includes categories of descriptors: *known/likely*, *cannot be determined*, and *not likely*, with sub-descriptors. Among other issues about descriptors, the proposed Guidelines call specific attention to the descriptor/sub-descriptor term "cannot-be-determined, suggestive evidence," and seek input whether this should be made into a separate descriptor category.

Several public comments received on the weight of evidence issues reveal difficulty in getting past the category level to examine the use of the descriptors within narratives. For example, some commenters believed that "known/likely" was a single descriptor and urged separation of the terms in much the same way as the previous "A" group was held to a higher standard than the "B" groups. Even though we used the terms separately in example narratives, the categories seem to be paramount to some. Since our intent is to discard simple categorization, we need to revisit the presentation.

To clarify these points, the Risk Assessment Forum's technical panel is considering revisions to reemphasize narrative descriptors, while de-emphasizing categorization.

The panel is considering a simple presentation of the descriptors without categorizations, but with the explanation that the descriptors *known, as if known, likely,* and *not likely* are used when the weight of evidence is sufficient to support a conclusion about human carcinogenic potential. Other descriptors including *suggestive, conflicting, inadequate data, no data* would be used when the weight of evidence is not sufficient for this conclusion to be made. Thus, we might present the "suggestive" descriptor as "The weight of evidence is suggestive of carcinogenicity, but not sufficient for a confident conclusion as to human carcinogenic potential."

The original question of having a fourth, "suggestive" category may be moot if we delete the categories in favor of simply presenting the descriptors. The technical panel believes that some positive indications of carcinogenicity, in practice, fall short of making a weight of evidence to support a conclusion regarding human hazard. Some of the data bases we would formerly have called "possible" human carcinogens (group C) are of this kind. (Example: Narrative #2 in Appendix A of the proposal) However, some commenters are of the strong opinion that any indication of carcinogenicity, even a very limited animal tumor response alone, should be the basis of citing an agent as potentially carcinogenic to humans. ***The panel would like to have the Committee's views on these issues.***

b) Information requirements necessary to depart from defaults

The 1986 guidelines presented default assumptions to be used to continue assessment when gaps in data or knowledge were encountered. These were consistent with the 1985 Office of Science and Technology (OSTP) review of the science. There was little discussion of data needs for departing from defaults. The 1996 proposal provides a framework of defaults, still consistent with the OSTP review. The proposal addresses the individual, major defaults and gives very general guidance on their use and departing from them within the body of the guidance as well as in response to NRC recommendations (section 1.3).

Many respondents requested additional guidance on information requirements considered by the Agency to be sufficient to reject default assumptions, particularly for rejection of the linear default assumption in favor of a non-linear mode-of-action (MoA). Concerns were expressed over: the amount of information necessary to reject the linear default, standards of proof (plausible conservatism versus maximum use of scientific information), administrative processes by which defaults could be over-ridden, and the extent to which Agency internal peer review and determinations should be supplemented by external peer review, either through the SAB or scientific meetings and consensus.

The proposal discusses the fact that different kinds of gap filling defaults are inherently different in the amount of empirical data needed to replace them, e.g., moving from a scaling factor to toxicokinetic analysis of dose versus determining that an animal tumor response is not relevant to humans. The proposal describes two general criteria: 1) that the scientific principle underlying the use of the data be generally accepted and 2) that empirical data on the agent in question support the conclusion made.

The Committee is requested to provide views on issues of departing from defaults generally, and on mode of action (MoA) conclusions in particular, see below. Further criteria on the MoA point were discussed in a 1995 International Life Sciences Institute report titled “Low Dose Extrapolation of Cancer Risk”.

Mode-of-action (MoA) determinations

Public comments correctly focused on the new (MoA) guidance as a critical change from the Agency’s 1986 policy. The MoA guidance allows for rejection of the linear default in favor of nonlinear or combined linear/nonlinear risk assessments. A

variety of comments were elicited on the extent to which a detailed mechanistic understanding of the chemical-specific cancer process is necessary in order to make a mode-of-action determination. The spectrum of potential policy options reflected in comments can be summarized as follows:

1) *Detailed mechanistic knowledge*: Comments favoring this option highlighted the need to have a detailed knowledge of all the steps to cancer formation before departing from the linear default assumption. Related to this is the concern that nonlinear increments to a background load, such as in multiple exposure situations, may increase risk in a linear fashion having exceeded the threshold level for that mechanistic pathway.

2) *Weight-of-Evidence*: This option allows for rejection of the linear default if there is confidence that the MoA is not through direct DNA damage, combined with sufficient evidence supporting a different, nonlinear, MoA. This is the preferred standard in the proposed Guidelines.

3) *Primacy of DNA-reactivity*: A number of comments favored the dichotomous approach that if a substance is DNA-reactive it should be presumed to follow a linear MoA , and if not DNA-reactive then the default should be to a nonlinear approach. This is a position commonly espoused by a number of European countries.

Related to this issue is the option to provide both a linear and nonlinear estimate of risk, the situations in which such a dual assessment should occur, and how discrepant estimates of risk should be treated. The proposed guideline supports use of both: 1) when responses in different sites appear, with substantial supporting evidence, to have different modes of action, 2) in order to decouple consideration of the contribution of two modes of action that operate at high versus low dose

c) Dose response assessment: defining a point of departure

The 1986 guidelines provided a basic rationale for linear dose response assessment that considered the relationship to be a function of reaction of agents with DNA in a stochastic manner. While the guidelines said that other approaches might be used if information supported them, the practice of linear assessment has been predominant. The 1996 proposal bases dose response assessment on mode of action information. The process would begin with analysis of observed data, followed by a explicit decision about how to approach analysis of the range of extrapolation to low doses, based on mode of action data.

The proposal would employ a "point of departure" to mark the beginning of extrapolation. Moreover, the response data are proposed to include data on precursor effects in appropriate cases, in addition to, or in lieu of, tumor data. To maintain consistency with noncancer risk assessment, the proposal suggested using the lower 95% confidence limit on dose associated with 10% extra risk as the point of departure. Comment was requested on alternatives, including using the central estimate, $ED_{10.}$

As requested in the proposal, respondents focused on issues about determining the point of departure, including:

Comments on the use of tumor data:

1) *The choice of the point of departure, e.g., LED_{10}, ED_{10}, ED_{01}.* The Agency proposed the LED_{10} point, and asked for comment on several alternatives. Respondents ventured a variety of preferences for lower limit and central estimates and for standard versus case by case choice of the response level 1.0% to 50%. In a recent peer consultation workshop on benchmark dose estimation for noncancer risk assessment, there seemed to be a general move toward support for using the central estimate. This is consistent with a previous workshop report on a draft of the proposed cancer risk assessment guidelines.

2) *The procedure by which this point of departure should be determined.* In the absence of biologically-based models, possible options include: choosing the 10% extra risk level, as proposed by the Agency, determining the lowest point in the data that can be statistically separated from the cancer rate variability in the control group, selecting the lowest significant data result, or by using logistic regression models.

3) *The data loss inherent in choosing a single point.* Choosing a single point negates the use of much of the additional bioassay or epidemiological data at higher exposures, particularly information on the observed slope of the cancer dose-response curve.

Comments on the use of precursor data:

1) *Use of precursor data*: The proposed Guidelines allow for three alternative uses of precursor data: 1) in lieu of tumor data if the precursor is found to be a better measure of risk, 2) to provide information about likely dose response behavior in the range of extrapolation, or 3) to link

the dose response for the precursor with that for the tumor response. Respondents were universal in their concern that precursor data be clearly linked to the cancer of concern, and that adequate peer review be included in this process. It also was noted by certain comments that application of the same MoE approach to precursor data as to tumor data would act to reduce the margin below that resulting from use of tumor data. This reduction in the margin could be problematic regarding risk presentation to the public and potential research efforts, a situation that would subtract from the potential gains from confirming nonlinearity through additional research.

d) Margin of exposure (MoE) analysis

The 1986 guidelines had no approach for nonlinear dose response assessment. In the 1996 proposal, a margin-of-exposure (MoE) analysis would be used in such circumstances. This is because, without sufficient data to define model parameters, there is not a way to choose a nonlinear model. If there were sufficient data then one would have enough basis for the preferred approach of using a biologically based or case specific model. Public comments revealed a number of concerns about MoE analysis, as follows:

1) *Support to risk managers in implementing the Guidelines*. One concern is utility to risk managers. For the MoE approach, the proposal attempted to strike a balance between over-prescribing the risk managers' conclusions versus leaving them adrift with insufficient risk assessment guidance. Currently, the proposed Guidelines anticipate providing supporting analysis of factors for consideration by risk managers in determining an MoE that is adequate to protect public safety, but do not go so far as to prescribe what an adequate margin of exposure should be. It is felt that the assessment can address the question of how risk is likely to decline with exposure, and issues of inter- and intra species variation in susceptibility, but the issue of adequacy of a margin is in the hands of the risk manager.

2) Some comments suggest that the risk assessment provide an opinion as to "adequate margin of exposure" rather than a supporting analysis for the risk manager.

The Committee's view on this basic issue is requested. Would addition of case example for such analysis improve the guidance?

e) Human data

In the 1986 guidelines, there was little discussion of the issues of evaluating human studies. In the 1996 proposal, more discussion of analysis has been included. In 1986, criteria for "sufficient human data" as a basis for a "known" classification group were given which said that a positive finding was not a result of bias, confounding, or likely to be due to chance. In the proposal, bias and confounding are addressed and the Bradford-Hill criteria for judging chance are incorporated as suggested in a workshop held by the EPA (1989). For dose response assessment, the guidance relies on case-by-case judgment, providing no methodologies as data sets are so variable. Two areas of comment stand out, among the others.

1) *Dose response guidance.* The need for more guidance on dose response assessment was felt by some commenters, particularly with respect to the use of the range of observation/range of extrapolation approach and locating a point of departure.

2) *Statistical significance.* One commenter, followed by some others, expressed belief that by not retaining the 1986 list of criteria for "sufficient" human data as it was, in particular, the criterion that chance is unlikely to account for a result, the Agency has hidden an intent to eliminate statistical significance as a consideration. The commenter apparently equates chance solely with statistical significance, not with the several Bradford-Hill criteria.

Advice would be appreciated on the issue whether the approaches for dose response assessment of human data should be separated entirely from the assessment approaches for other data, leaving human data to a case-by-case treatment. Alternatively, should the two-step analysis of observed data and extrapolation be retained for all data, with generous exceptions for human studies that are not amenable to modeling in the observed range?

On the point of statistical significance, since the proposal had no intention of discarding statistical significance in judging human data, suggested clarifications of the existing discussions of the Bradford-Hill criteria (section 2.2.1.3.) and weighing evidence from human studies (section 2.6.1.) would be welcome.

f) NRC recommendations for tumor data analysis

The 1986 guidelines discussed two methods for deciding how to measure response. One was to model a tumor response at a particular site. The other was to count animals with tumor types that were individually statistically significant. The proposed Guidelines (section 3.2) provide a number of options for measuring response, with the overall aim of choosing a measure or measures of response that best represent the biology.

Several comments highlighted the proposal's failure to include among these options the explicit NRC recommendation (NRC 1994, p. 241) that EPA should first estimate cancer potencies for each tumor type and then sum tumor type-specific potencies to reach an overall cancer potency estimate. The NRC-recommended approach treats multiple tumor sites as independent events, whose effects should be added separately. The Agency's proposed approach is otherwise in that, if summing the total animal response is considered to be the best representation of response, the approach would be to sum animals with cancer rather than their site-specific tumor potencies. ***There are several factors to consider and a Committee perspective is requested.***

1) ***What is independence? Independence of tumor expression at different sites in an animal as tested by probability? Independence of mode of action for different tumor types? Which is the better basis for considering human risk potential?***

2) ***Often one observes an increase in a high background tumor (e.g. mouse liver tumors) as well as an increase in an historically low background tumor (both statistically significant). If the dose responses for these are summed, the latter will be swamped by the former.***

3) ***What are the implications of different approaches with respect to assuming site concordance?***

4) ***Whichever summing approach might be selected, ought it be a last resort, a standard approach, or be accompanied by separate analyses and discussions of different tumor types?***

g) Susceptibility Factor for Human Variability

The NRC (1994; p.219) report elaborates on human variability, recommending that "EPA could choose to incorporate into its cancer risk assessment for individual risk

a ‘default susceptibility factor’ greater than the implicit factor of 1 that results from treating all humans as identical.” NRC (1994) also recommends continued research to determine an appropriate default value for human variability. The proposed Guidelines do not incorporate a specific susceptibility factor in the linear extrapolation process. The proposal’s position is that sufficient conservatism is incorporated through the assumption that humans have susceptibility similar to that of the of rodents in the bioassay, or persons observed in epidemiological studies, combined with the inherent conservatism of the linear extrapolation process. The Agency also notes that the uncertainty resulting from human variability has been incorporated in the MoE approach to nonlinear risk assessments. The NRC (1994) recommendation that additional research be undertaken is being implemented; the Agency is funding ongoing in house and grantee research on this issue. ***Several commenters suggest that the guidelines incorporate an uncertainty factor across the board. What the factor should be for linear extrapolation is unclear. A Committee perspective is requested***.

3. DETAILED FINDINGS

Although there are far more positive, than negative, comments that could be made about the Guidelines, this report will be limited primarily to those areas (noted in detail in the response to the specific Charge issues below) wherein the Committee believes that revisions are required before the proposed Guidelines are finalized; these specific issues are discussed in detail below.

3.1 Hazard Characterization: Descriptors/Narratives

In its 1986 Cancer Risk Assessment Guidelines, the Agency adopted a carcinogen classification system assigning agents to groups "A" through "E" (Group A for "known" human carcinogens through Group E for agents with evidence of non-carcinogenicity). In its current (1996) draft revision, the Agency has proposed assessing conclusions about an agent through the use of a narrative with standard descriptors. The proposed Guidelines moved from the letter and number classification with the declared intent of increasing the information provided to risk managers. This revised classification system incorporates a limited number of descriptor phrases into a weight of evidence narrative summary of human, animal and other key data. The proposed Guidelines as published by the Agency provide three categories with 13 descriptors/subdescriptors as follows:

“known/likely”

known, in humans from epidemiology and/or experimental evidence demonstrating causality

as if known, based on plausible causality and strong experimental evidence

likely, due to tumors resulting from modes of action that are relevant to human carcinogenicity

likely, (high end of weight-of-evidence)

likely, (low end of weight-of-evidence)

“cannot be determined”

cannot be determined, but suggestive

cannot be determined, due to conflicting data

cannot be determined, due to inadequate data

cannot be determined, due to no data

"not likely"

not likely, because of available animal studies
not likely, because animal data shown to be not relevant to humans
not likely, because of dose or route dependence
not likely, because of extensive human experience.

The EHC believes that the proposed Guidelines head in the right direction by emphasizing the development of a narrative discussion describing the weight of evidence and strongly endorses this approach. Users of the Guidelines should be encouraged to support conclusions drawn by presenting in the narrative a full discussion of the weight-of-evidence for carcinogenic hazard, including all critical considerations made in reaching the conclusion. In describing the weight of evidence, the proposed Guidelines emphasize the importance of considering the relevance of the route of exposure and the sufficiency of the evidence from animal studies, and the EHC finds that this is appropriate and important.

The use of the multiple terms, "categories/descriptors/subdescriptors," caused occasional semantic confusion. **The bottom-line is that the EHC could not come to a consensus as to which, and how many, terms should be used.** Some Members found categories to be very important, and a few members indicated preference for the categories used in the International Agency for Research on Cancer () classifications that would harmonize the U.S. system with the rest of the world. It was noted that principles and procedures for IARC classifications include criteria for accepting mechanistic data and the 'strength-of-evidence' approach (see Appendix A for an extract of the relevant IARC document). Other Members felt that categories should be eliminated, and that allowing the narrative to stand with selections made from the 13 subdescriptors listed above provided the rich vocabulary that should be used to describe what is known or not known with regard to carcinogenicity. Members supporting this view found that the IARC emphasis on hazard identification limited the applicability of its criteria in US EPA risk assessments (see Appendix B for the IARC Preamble statement). There was also support for the eight descriptors proposed by Ashby, *et al.*, which is basically an expanded IARC classification.

Although a majority of the Committee believes that some headings for the narratives would be important, it was agreed that three is probably too few and thirteen probably too many. If the Agency deems it necessary to use broad categories to group subdescriptors, then the "known/likely" category should be subdivided to prevent misunderstanding about the intent to use them as separate terms, and 'suggestive' should not be lumped with 'cannot be determined.' There was

also some expressed dislike for the term “likely” and for the possible inappropriateness of the term “as if known.” Another expressed concern was that the descriptor “not likely, because animal data shown to be not relevant to humans” was illogical or ambiguous unless reworded. A possible change would be to make it similar to one of the above descriptors, e.g.,”not likely, due to tumors resulting only from modes of action that are **not** relevant to human carcinogenicity.”

The Committee recognizes that the ultimate choices of narrative descriptions will not satisfy all users of the Guidelines, and hopes that the comments provided will help EPA in making its choices.

3.2 Information Requirements to Depart from Defaults and to Use Mode of Action (MoA) Determinations

The 1986 Guidelines presented default assumptions to be used to continue assessment when gaps in data or knowledge were encountered. These were consistent with the 1985 Office of Science and Technology (OSTP) review of the science. There was little discussion of data needs for departing from defaults. The 1996 proposal provides a framework of defaults, still consistent with the OSTP review. The proposal addresses the individual, major defaults and gives very general guidance on their use and departures from them within the body of the guidance as well as in response to NRC recommendations (section 1.3).

Many respondents to the proposed Guidelines requested additional guidance on information requirements considered by the Agency to be sufficient to reject default assumptions, particularly for rejection of the linear default assumption in favor of a non-linear mode-of-action (MoA). The proposal discusses the fact that different kinds of “gap filling” defaults are inherently different in the amount of empirical data needed to replace them, e.g., moving from a scaling factor to toxicokinetic analysis of dose versus determining that an animal tumor response is not relevant to humans.

The proposal describes two general criteria: a) that the scientific principle underlying the use of the data be generally accepted; and b) that empirical data on the agent in question support the conclusion made.

Public comments to the proposed Guidelines correctly focused on the new (MoA) guidance as a critical change from the Agency’s 1986 policy. The MoA guidance allows for rejection of the linear default in favor of nonlinear or combined linear/nonlinear risk assessments. A variety of comments were elicited on the extent to

which a detailed mechanistic understanding of the chemical-specific cancer process is necessary in order to make a mode-of-action determination.

3.2.1 Issues for departing from defaults

The proposed Guidelines do not provide a list of formal decision criteria for rejecting defaults. Rather, they state that different amounts of data may be required for different situations, *i.e.*, screening assessments versus full assessments. **The EHC agreed that different types of assessments may have different data requirements and levels of analysis, and that scientific peer review is important in evaluating the adequacy of data.** As stated in the proposed Guidelines:

> *"If data are present, their evaluation may reveal inadequacies that also lead to use of the default. If data support a plausible alternative to the default, but no more strongly than they support the default, both the default and its alternative are carried through the assessment and characterized for the risk manager. If the data support an alternative to the default as the more reasonable judgment, the data are used.*"

The majority of the Committee supported the Agency's approach and opined that it is likely that many of the new risk assessments will present both linear and nonlinear assessments.

A sizeable minority of the Committee believed that the burden of proof should rest on showing that the defaults are implausible, and also that the Guidelines should be revised to include explicit and specific criteria for judging the validity of hypotheses invoked to depart from defaults. The Members holding this minority position also found it ill-advised to provide multiple assessments and recommended as a matter of science policy that the Guidelines include the following:

a) an explicit statement that, given the goal of public health protection, where there is uncertainty regarding the mode of action of a chemical, the Agency will adopt the default assumptions that the agent in question is capable of acting linearly at low dose and that positive animal data are relevant to humans; and

b) an explicit statement that the burden of proof rests on showing the defaults (low dose linearity and relevance to humans) are not plausible (see

section 3.4 of the IARC document extract in Appendix A) Some EHC Members felt that the statement should emphasize the point that the basis for departing from defaults was that the alternative was the most likely or most plausible choice, rather thandemonstrating that the default is not plausible; these Members felt that the proposed approach implied absolute knowledge or proof.

Applying a weight-of-evidence approach to risk assessments may, in some instances, help reduce the uncertainty associated with the use of defaults.

The new approach includes data on mode of action, toxicokinetics, cell proliferation, species differences and interindividual differences in susceptibility. **This approach is consistent with the recommendations of the 1994 NRC report. The EHC strongly endorses this change. The proposed Guidelines clearly state that the preferred method for developing the hazard characterization and risk assessment for a chemical relies on a strong scientific database.** It is also recognized that, for many chemicals, the assessor will not have a detailed understanding of their mode of action and related material. It is clear that regulatory decisions must still be made in the absence of such data. A linear default model will be employed in such cases that will not differ greatly from the approach outlined in the 1986 Guidelines. **The EHC agrees with the use of defaults in the absence of scientific data and does not object to the simple extrapolation procedure outlined in the proposed Guidelines.**

The proposed Guidelines indicate that the Agency has carefully thought about and decided against adopting formal criteria for judging when an analysis is sufficiently plausible and persuasive to warrant deviation from defaults (section 1.3.1 of the proposed Guidelines). Nonetheless two general criteria are given:

> *"...that the underlying scientific principle has been generally accepted within the scientific community and that supportive experiments are available that test the application of the principle to the agent under review."*

These criteria are helpful, but given the wide use of the Guidelines by the various programs within the Agency, more detailed guidance appears to be needed. There were some expressions of dissatisfaction with the term "default." A suggested alternative was to use "inference guidelines" (a term suggested in the National Research Council "Red Book" in 1983).

3.2.2 Precautionary views about risk assessment under the proposed Guidelines

The proposed Guidelines provide thoughtful background discussion on the variety of complex issues on which policy positions are taken. The rationale provided is well reasoned and transparent. One area of discussion that deserves more attention is in the characterization of the degree to which individual policy positions/defaults and the Guidelines as a whole can be characterized as public health conservative.

The proposed Guidelines state that:

> "*Generally these defaults remain public health conservative, but in some instances they have been modified to reflect the evolution of knowledge since 1986.*"

This statement leaves one with the general impression that those defaults that have not changed are conservative and those that have changed now have greater scientific justification. The general message appears to be that the overall process is still public health conservative. **Because this overall impression is likely to be translated to work products developed under these Guidelines and will thus affect the public's and risk manager's view of the risk, it deserves greater scrutiny. During the EHC meeting it was stated that it is possible that the process does not always result in evaluations that are public health conservative and examples of default as well as non-default procedures that might not be public health conservative were raised. The EHC clearly understands that the primary goal of EPA actions is public health protection and recommends adding clarifications that will alleviate potential concerns by readers or users of the Guidelines, while at the same time promoting the use of good science for decision making.**

3.2.3 Mode-of-action (MoA) Determinations

At the EHC public review meeting, there was considerable discussion regarding MoA determinations. **Most Members of the Committee agreed with the approach outlined in the proposed Guidelines, a finding that absolute proof of a MoA is neither attainable nor necessary. Rather, an alternative should displace a default when it is generally accepted in peer review as the most reasonable judgment. The EHC recommended that the Guidelines include the following:**

a) specific criteria for judging that the data on mode of action are valid and adequate to support the alternative approach (see the IARC document exract in Appendix A, pages A-3 and A-4)

b) a discussion of mode of action that reflects the lack of a clear distinction between direct DNA damage and other mechanisms with respect to the low dose response relationship. Most Members felt that the Guidelines should include some specific examples of chemicals that are accepted as threshold carcinogens.

3.3 Dose Response Assessment: Defining a Point of Departure

The 1986 Guidelines provided a basic rationale for linear dose response assessment. This rationale considered the relationship to be a function of reaction of agents with DNA in a stochastic manner. While the Guidelines said that other approaches might be used if information supported them, the practice of linear assessment has been predominant. The 1996 proposal bases dose response assessment on mode of action information. The process would begin with analysis of observed data, followed by a explicit decision about how to approach analysis of the range of extrapolation to low doses, based on mode of action data.

The proposal would employ a "point of departure" to mark the beginning of extrapolation. Moreover, the response data are proposed to include data on precursor effects in appropriate cases, in addition to, or in lieu of, tumor data. To maintain consistency with noncancer risk assessment, the proposal suggested using the lower 95% confidence limit on dose associated with 10% extra risk as the point of departure.

3.3.1 Dose-response Assessment

The Agency is to be commended for its efforts to develop a clear, well-reasoned framework for dose response assessment. The proposed Guidelines provide for more flexibility in risk assessment and provide means for incorporating other types of biological data and information on mode of action into dose response assessment However, some Members noted that, in some areas, there may not be enough guidance, given the wide usage of the Guidelines, and in other areas additional clarification is needed

Although sympathetic to the Agency's reluctance to develop additional criteria, the EHC believes that further guidance should be provided. In the case of

dose response evaluation, the Agency is encouraged to develop criteria which lead to judgments which consider the following: the reliability of predictions outside the observable range; problematic data sets such as those with poor fits and extreme curvature; the comprehensiveness of the supporting data; the degree to which alternative hypotheses have been considered and investigated; and the adequacy of the scientific peer review, considering such issues as the breadth of the review and the degree to which the peer review group can be considered to represent the scientific community at large. To aid its deliberation, the Agency is referred to the 1996 NRC report *Understanding Risk* and the 1994 NRC report *Science and Judgment in Risk Assessment.*

Unlike other parts of the Guidelines, no examples are provided in the dose response section. However, a number of the approaches proposed in the dose response section are fairly novel; consequently we have little experience on how they will work in practice. The Agency is encouraged to provide examples in the dose response section to illustrate important features of the dose response methodology, including:

a) the decision to use a biologically based model for quantitative risk assessment

b) the decision to assume that a chemical has a (low-dose) non-linear mode of action

c) extrapolation to low doses using the linear mode of action approach

d) calculation of an LED_{10} and ED_{10}

e) use of precursor data -- analyses based on non-cancer responses

Generally a number of different models will provide an equivalent fit to the data. However, some unrestricted models may give answers that are biologically unrealistic in some cases and should be avoided. **Since the issues regarding selection of a model for use in the observable range are largely generic, the EHC recommends that the Agency select a default procedure for routine use in such calculations.** A requirement of biological feasibility leads to a dose response model that, at low doses, is linear or sublinear, and does not assume an extreme superlinear shape (*e.g.*, one that predicts that the tumor response jumps abruptly from the background value to a much higher value at an infinitesimally small dose). It also calls for a model that is monotonically increasing in dose. Appropriately constrained versions of the

polynomial, Weibull and logistic models all incorporate these features and have been routinely used in the past for curve fitting. Consequently they are recommended to the Agency for consideration.

In cases where it is important to take into account early mortality in treated animals, time dependent models should be considered. Additional guidance should be provided to address cases where the dose response appears superlinear. In such cases the analyst should be encouraged to explore the basis for the non-linearity (*e.g.*, saturating pharmacokinetic processes, differences in intercurrent mortality among treatment groups) and if possible take that into account in the modeling exercise. Finally under the Weibull model, unrealistic values of effect doses and confidence bounds can result in cases of extreme sublinearity in which the shape parameter is large and precautionary guidance should be given in this regard if the Weibull is provided as a default model. The use of the unrestricted model can also lead to unrealistic values when data are superlinear.

With respect to the inferences made in performing dose response evaluations, the assumption that all humans have the same sensitivity is the same as that made in the 1986 Guidelines. As noted in the proposed Guidelines, human sensitivities can vary considerably. The Guidelines also indicate that linear extrapolation is a public health conservative assumption, but several examples suggest that in the absence of information on pharmacokinetics and time to tumor the assumption may not be correct. Other assumptions or procedures that are “public health neutral” or “not conservative” include: dose averaging (see above); the treatment of data from standard two year bioassay protocols as representative of full life exposure (exposures before weaning and in old age do not occur); the interspecies correction; presuming human tumor sites are the same as animals in cases where PBPK analyses have been performed.

The EHC supports the Agency’s efforts, reflected in these and other guidelines documents, to harmonize approaches to dose response evaluations for cancer and non-cancer endpoints, and encourages the Agency to do more.

3.3.2 Biologically Based Dose-response Models

Overall, the Guidelines provide reasonable approaches for the use of biologically based models but are unclear in some critical areas, *e.g.*, definitions of such models, description of the "state of the art" in development of biologically based dose response models, and characterization of the approach for applying biologically based dose response models. The document defines these models as ones whose

parameters are calculated independently of curve fitting of data. The EHC understands the theoretical desirability of a biologically based dose response model and applauds the encouragement to develop such approaches. **However, the proposed Guidelines may be misleading in this respect because no such model presently exists in a complete form, nor is it clear how to develop such in the foreseeable future. The Agency's definition of a biologically based model thus seems unnecessarily narrow.** One can conceive of models in which non-tumor data are used to capture the dose-response shape in the low dose range, and fitting the model to the tumor data is appropriately used to scale the dose response. On the other hand, just because a model developed independently of the tumor data predicts tumor responses at the experimental doses is no indication that it will predict responses in the low dose range. In fact, such a model would not necessarily even include a means for predicting low-dose behavior. The document also is not clear on what a "case-specific" model means, and how such a model differs from a biologically based model.

The EHC believes that it is important for the Agency to define more clearly what is meant by a biologically based dose response model and to give guidance as to when such a model is preferred over the default linear or non-linear approaches. Members of the Committee have developed the following definition for the Agency's consideration:

> *The dose response that relates the probability of tumor formation as a function of exposure to a xenobiotic is often expressed as the composition of two processes: the toxicokinetic process that describes the concentration of the active agent at the target site as a function of exposure to the parent compound, and the toxicodynamic process that expresses the probability of tumor formation as a function of the concentration of the active agent at the target site. A biologically based dose response model is one that uses either dose response data for one or more intermediate steps in the toxicodynamic process or a biological understanding of the toxicodynamic process to develop a dose response model for predicting tumor response as a function of either exposure or concentration of the active agent in the target tissue. Thus a biologically based dose response model of cancer incorporates biological information on the toxicodynamic process, and it may, in addition, incorporate such information on the toxicokinetic process.*

If a definition such as this is adopted it appears that it may be unnecessary to distinguish between a biologically based dose response model and a case-specific model.

The EHC also believes it would be helpful for the Agency to describe the "state of the art" in development of biologically based dose response models. If possible, the Agency should provide an example of an application of a biologically based dose response model that would be relied upon for low dose extrapolation. If no such model can presently be identified, a statement to that effect would also be helpful.

Given the general lack of experience in developing and applying biologically based dose response models as well as the large amount of time and resources that may be required in each instance, it seems unwise to hold up the use of such a model as “the overriding preferred approach” for dose response evaluation (see section 1.3.2.5 of the draft proposed Guidelines). Instead, the Agency may wish to indicate that the goal or ideal in cancer risk estimation is to predict risk reliably at levels of human exposure based on a clear understanding of the biology, even though this has proven to be very difficult to achieve. Nonetheless, when reliable predictions are developed the Agency will use them.

3.3.3 Defining a point of departure

The calculation of an LED_{10} for defining a point of departure involves fitting a statistical dose response model to data in the observable range. This calculation requires selecting a dose response model, selecting a risk value (*e.g.*, 10% or 1%), and calculating statistical confidence intervals. **To mitigate against unnecessary inconsistencies and confusion in the application of the Guidelines the EHC suggests that further guidance be given in the Guidelines on this approach.** The calculation of an LED_{10} is equivalent to the benchmark calculation that is used by the Agency with non-carcinogens. **The EHC sees little scientific justification for separate approaches to LED_{10} and benchmark calculations, and therefore recommends that the approaches used for non-carcinogens and (low dose) non-linear carcinogens (*e.g.*, choice of dose-response model, choice of point of departure, and use of statistical confidence bounds) be harmonized to the extent possible.** There may be scientific justification for the use of different adjustment and uncertainty factors for cancer and non-cancer. Non-cancer factors should not be routinely used for cancer without clear scientific justification for so doing.

In general, it is recommended that a single risk level, e.g., ED_{10}, be selected as the point of departure when the (low dose) non-linear approach is applied. Selection of a common value will serve to standardize the calculation and facilitate comparisons across chemicals. Use of different values could complicate an MoE evaluation since then a risk manager would somehow have to adjust for a different

point of departure in different situations. In some situations (*e.g.*, large experiments), it may be preferable to use a non-standard value. In those situations where it is preferable to have a non-standard point of departure, the evaluation should provide clear guidance to the risk manager regarding the evaluation of an MoE.

The selection of the value of 10% extra risk proposed by the Agency as the usual point of departure was supported by several members of the Committee. The point of departure is supposed to represent the lower end of the range of risks that can be measured with some reliability. A cancer bioassay typically involves 50 animals per group. Ten percent of 50 is five animals. Five responders out of 50 is only marginally statistically different from zero responders out of 50 and with a higher background a 10% increase in extra risk would be even less significant.

The consensus of the Committee was that both point estimates and statistical bounds can be useful in different circumstances, and recommended that the Agency routinely calculate and present the point estimate of the ED_{10} and the corresponding upper and lower 95% statistical bounds. It may be appropriate to emphasize point estimates in activities that involve ranking agents as to their carcinogenic hazard. On the other hand, it may be appropriate to emphasize lower statistical bounds in activities designed to develop an appropriate human exposure value, since such activities require accounting for various types of uncertainties and a lower statistical bound on the ED_{10} (LED_{10}) is a scientifically-based approach for accounting for uncertainty in the true value of the ED_{10}.

Other considerations impact this discussion. The special problem of our lack of understanding of how to address early-in-life and late-in-life exposures should feature more prominently in the section on dose averaging when appropriate. Although the approach of averaging cumulative dose over lifetime is used, biological models presented over the past 50 years as well as empirical studies (*e.g.*, DEN, radiation, saccharin) suggest that age at exposure is important and more should be done to address it. This is clearly an area where research is needed.

Little guidance is given in the proposed Guidelines on how precursor data should be used in dose response evaluation. As understood by the Committee, it was considered useful for qualitative assessments, but there was no consensus on the potential for quantitative assessment with precursor data. **The EHC consensus was that considerably more thought and work was needed in this area, with some additional guidance andsome examples provided in the document.**

Use of the terms “linear” and “non-linear” to distinguish the two procedures for dose response evaluation can lead to some misunderstanding and confusion. For example, many well-studied agents associated with a DNA reactive mode of action are observed to have non-linear dose response relationships (*e.g.*, vinyl chloride and diethylnitrosamine). Non-linear tumor dose response relationships due to increased mortality in high dose groups and dose dependent pharmacokinetics are frequently observed. **Suggested alternative terminology to address this problem is “linear at low doses” and “non-linear at low doses”; “low dose linear” and “low dose non-linear”; or “non-threshold” and “presumed threshold-like.”**

3.4 Margin of Exposure (MoE) Analysis

This term (MoE) seems to have evolved from the terms “margin of safety” and “margin of error” where circumstances or exposures in everyday life are compared to those observed experimentally to have no effect. Regarding the use of the “margin of exposure” term in the proposed Guidelines, an exposure estimate is compared to a dose associated with an undesirable outcome (*i.e.*, excess 10% cancer incidence) and not one for which lack of effect is presumed. Adoption of alternative terminology that would leave less room for confusion and misinterpretation by a risk manager and the public is recommended by the EHC.

At times during the SAB public meeting there was confusion as to the proper definition for MoE. Given the Committee’s problems with the concept, it was felt that this topic needed the addition of case studies to illustrate its use by practitioners in the field. There was no consensus on the advice to be given to a risk assessor on the magnitude of an appropriate MoE. Also, the EHC, as discussed with EPA staff in attendance at the public review meeting, recommends that the Guidelines incorporate additional information on how the Agency would use peer review in the process of deciding on alternative approaches.

3.5 Use of Human Data

In the 1986 Guidelines, there was little discussion of the issues of evaluating human studies. In the 1996 proposal, more discussion of analysis has been included. In 1986, criteria for "sufficient human data" as a basis for a "known" classification group were given which said that a positive finding was not a result of bias, confounding, or likely to be due to chance. In the proposed Guidelines, bias and confounding are

addressed and the Bradford Hill criteria for judging chance are incorporated as suggested in a workshop held by the Agency (1989). For dose response assessment, the guidance relies on case-by-case judgment, providing no methodologies, as data sets are so variable. **The EHC endorses the approach and general concepts for incorporating human data into risk assessments**.

3.5.1 Assessing adequacy of epidemiological studies

The EHC finds that the section on epidemiological studies is both thoughtful and cautious, given all the complexities of observational studies. Observational studies, such as cohort and case-control studies, represent the majority of studies on cancer etiology in humans. In these studies, the investigator has no control on the assignment of subjects to exposure groups, as is the case with clinical trials. Furthermore, assessment of exposure to the agent of interest is subject to measurement error, and the particular metric of exposure that may influence the risk of disease is often unknown. Some of these methodological issues may result in spurious associations, but others, particularly those relating to exposure, often result in a substantial attenuation of the magnitude of a real association and a marked decrease in the precision of the risk estimate. Therefore, in evaluating a particular association in an observational study, several methodological issues, including possible selection bias, confounding, and exposure misclassification need to be considered. **It was noted that it will often be important to have access to raw data for the best utilization of human data for risk assessment, and that the Agency should obtain and utilize these data whenever possible**. To further this goal, the contractual agreement for all studies receiving financial support from the Agency should be framed so that, within the concerns of privacy for study participants, the raw data developed in the study will be available to the Agency.

3.5.2 Criteria for Causality

There have been some criticisms (during the period for public comment) of the manner in which statistical significance has been treated in Sections 2.2.1.2 (Criteria for Assessing Adequacy of Epidemiologic Studies, pp. 44-48), 2.2.1.3 (Criteria for Causality, pp. 48-49), and 2.6.1 (Weighing Evidence from Human Studies, pp. 70-77) of the proposed Guidelines for Carcinogen Risk Assessment. These relate to the fact that although the draft document makes reference to the need to consider statistical significance in increasing confidence in a conclusion of causality (sections on Statistical Considerations and Weighing the Evidence from Human Studies), there is no explicit statement that statistical significance should be a basic requirement for

determining causality. **The lack of such explicit wording has been interpreted as misleading and implying that there is a hidden intent to eliminate statistical significance as a consideration in assessing causality. Adding appropriate and specific language concerning statistical significance should rectify this problem.**

It is important, in evaluating the overall evidence for an effect from epidemiological studies, to rule out chance as a possible explanation for an observed association. The degree to which chance should account for an observed association can be evaluated by calculating the statistical significance of the association. In terms of the Bradford Hill criteria for causality, statistical significance is one means for evaluating the strength of an association. Although the current version of the Guidelines appears to appropriately describe the role of statistical significance in interpreting epidemiological studies, the EHC believes, to avoid misinterpretation, the section should be revised to reflect the need to consider the statistical significance of an association. However, in an observational study, factors other than chance may affect the statistical significance of an association. These factors include selection, confounding and exposure misclassification. In some instances these factors may cause a spurious association to be statistically significant. In other instances these factors may cause a causal association not to achieve statistical significance. Also, in a data set in which numerous exploratory analyses may have been conducted, interpretation of these tests should account for the multiplicity of tests. **Because of these difficulties with the interpretation of observational studies, the EHC believes that a weight of evidence approach to evaluating causality is appropriate.**

Although the Guidelines refer to the Bradford Hill "criteria," the EHC believes that the word "guidelines" may be more appropriate. A recent text on epidemiological principles summarizes the position that most epidemiologists take with regard to causal guidelines, and that the EHC supports:

"Although causal guidelines are often referred to as criteria, this term does not seem entirely appropriate. Although it may be a desirable goal to place causal inferences on a firm quantitative and structural foundation, at present we generally do not have the information needed for doing so. The preceding list [i.e., Bradford Hill guidelines for judging whether an association is causal] should therefore be considered only guidelines which can be of most value when coupled with reasoned judgment in making decisions about causation." (Epidemiology, Leon Gordis, W. B. Saunders, 1996: 181). **Accordingly, the EHC recommends changing the word "criteria" in Sections 2.2.1.3 (Criteria for Causality), 2.2.1.4 (Assessment of Evidence of**

Carcinogenicity from Human Data), and 2.6.1 (Weight of Evidence Analysis) to "guidelines."

3.6 NRC Recommendations for Tumor Data Analysis

Independence of multiple tumor sites may be viewed in terms of biological independence or statistical independence. Independent biological action of multiple tumor sites is supported by the known mechanisms of tumor causation (mode of action). It is generally accepted that the molecular events (genetic damage) that are responsible for a given type of tumor cannot be assumed to be the same for other tumor types, even in the same animal and even though the offending chemical is the same. If they were the same, one would expect to see all tissue types qualitatively at risk or not at risk in tandem. This is not supportable by experimental evidence since a given chemical usually causes tumors in only a few tissue sites. Another reason for treating different tumor types as biologically independent events is that the dose to one tissue may be different from another based on its blood supply, fat content, etc. Furthermore, the production of metabolites and the detoxification abilities varies from tissue to tissue. As a result the mechanism of carcinogenesis in different tissues within the same animals may differ, i.e., as a promoting agent in one tissue and a complete carcinogen in another. On the other hand, statistical independence of tumor sites means that the presence of a tumor in one organ does not affect the probability of there being a tumor in a different organ. This is generally not the case. Thus, tumors in different sites are generally biologically independent, based on how the EHC is using this term, but may be statistically dependent.

Whether it is appropriate to combine different tumor types for analyses and, if so, how, depends upon whether one is interested in conducting a statistical test for the presence of a carcinogenic effect or in estimating the potency of a chemical that has been identified as a carcinogen. **In either case, combining all tumor types is not appropriate**. **In the former case, the procedure adopted by the National Toxicology program of combining some closely related tumor types for statistical analyses, but otherwise conducting separate tests for different tumor types is endorsed by the EHC.** In the latter case, estimating separate potencies for different tumor types and summing the potencies to estimate an overall carcinogenic potency tacitly assumes that different tumors are statistically independent and therefore is not recommended. An even less satisfactory approach would be to sum upper confidence intervals on carcinogenic potency. Calculating a separate potency for each tumor type is preferred to summing potencies. However this approach may underestimate the overall carcinogenic potency of a chemical that causes tumors at multiple sites. To

account for this, the Agency should consider modeling the response in which an animal is defined as a responder if it has one of the tumor types associated with exposure. This approach does not require that tumors be either statistically or biologically independent or dependent. A similar combining of responses was used recently by the Agency in its risk assessment for methylmercury. Although the combined endpoints were related to neurological deficit rather than cancer, the same principle applies to both. The Guidelines should be flexible enough to permit either of these two latter approaches (calculating potencies for individual tumor types or defining an animal as a responder if it has any tumor type associated with exposure) in individual cases, based on what is dictated by the available data.

As a general principle, chemicals that cause tumors at more than one site and in multiple species, other things being equal, are most likely more hazardous than those that produce tumors at a single site in only one species. This should be reflected in cancer risk assessments by the Agency. The EHC endorses the Agency's Guidelines that give additional significance to the finding of uncommon tumor types, multiple sites, tumors by more than one route of administration, etc. (see pp. 55-56 of the Guidelines).

3.6.1 Other Considerations

The use of historical controls should be considered in the sense of "reproducibility" of the bioassay. To use data from only a single bioassay in isolation of other control data, especially when the data are on controls conducted under the same conditions in the same strain of animals in a reasonably similar time frame are available, is not appropriate. Although concurrent controls should be given primary weight in the evaluation of the results of a bioassay, historical control data may also provide valuable insight into whether the tumor response in concurrent controls was atypical for that strain of animals. However, before concluding, based on historical control information, that the response in concurrent controls was atypical, one should carefully consider not only the average response in historical controls but also how variable that response is from study to study and whether time trends are present, as well as body weight differences between the two control groups. Other factors may also need to be considered, such as whether animals were caged in the same way in historical controls as in concurrent controls.

It is now well established that body weight is often correlated with tumor response for several tumor types. Therefore, the EHC also encourages the

Agency to consider the role of diet in a given tumor response. Consideration of this issue should be addressed when evaluating a given tumor data set.

3.7 Susceptibility Factors for Human Variability

The proposed Guidelines do not incorporate a specific susceptibility factor in the linear extrapolation process. The proposal's position is that sufficient conservatism is incorporated through the assumption that humans have susceptibility similar to that of the rodents in the bioassay, or persons observed in epidemiological studies, combined with the inherent conservatism of the linear extrapolation process. The Agency also notes that the uncertainty resulting from human variability has been incorporated in the MoE approach to nonlinear risk assessments.

There continues to be a rapid growth in our knowledge of variability in both susceptibility of humans to the development of cancer and the sensitivity of humans for induction of cancer by exogenous agents. As knowledge in the field increases, there will be a better understanding of the distribution of susceptibility and sensitivity traits across the population and for individuals within the population. The introduction of a "default susceptibility factor" must be closely linked with consideration of the "risk target" for both individuals and populations.

The EHC recognized that human variabilities within populations are being identified, such as genetic polymorphisms, nutritional status, ethnicity, age, stage of development, gender, preexisting diseases and ability for DNA repair. Each of these and their potential interactions may be significant susceptibility factors, or even potential resistance factors. Some of these factors have been estimated to have an 85 to 500 fold variation within a human population. When the penetrance of such a factor across a population is known, it should be taken into account. **The developing infant and child should be recognized as a subgroup which is particularly sensitive to the carcinogenicity of a number agents. Risk assessments should explicitly account for the differential susceptibility of the young. Other differences in susceptibility should also be taken into account when data permit.**

There was divided opinion with regard to the proposal that, when a specific susceptibility factor is not known, an uncertainty factor should nevertheless be used to account for human variability during the extrapolation process. Some Members agreed with the position in the proposed Guidelines that sufficient conservatism was already built into the process and opined that the use of a conservative one in a million cancer risk target for a population recognizes that there is a distribution of individual risks

within the population, including risks to individuals who may be more or less susceptible than the population average. Several members questioned that sufficient conservatism on human variability was present. **The EHC finds that it is premature to recommend a specific uncertainty factor for use during linear extrapolation. The Agency should continue its research on the uncertainty of human variability and the Guidelines should make the basis for conservatism in the process more explicit with examples or case studies. The Agency should readdress the question of an uncertainty factor for linear extrapolation at a later date.**

4. CONCLUSIONS

The Committee's overall impression is that the proposed Guidelines constitute a significant step forward in the "state-of-the-art" for carcinogen risk assessment. The EHC particularly commends the Agency for addressing the controversial aspects of the new Guidelines with a frank, unbiased approach to all points of view. The new Guidelines, when implemented, will cause risk assessors to place greater emphasis (than do the current Guidelines) on the utilization of all the available scientific information in characterizing cancer risks.

The above comments not withstanding, the EHC noted several areas in which the Guidelines would benefit from clarification or revision. These areas are:

a) The Committee endorses the Guidelines' emphasis on narrative descriptors, but found problems with its implementation, particularly in the use of multiple terms (i.e., categories; descriptors, and subdescriptors). **Given the complexities involved, the Committee could not come to a consensus as to how this problem should be addressed.** Some Members suggested using the International Agency for Cancer Research (IARC) classifications (see Appendix B); others advocated eliminating categories in favor of a narrative with selections made from the proposed thirteen subdescriptors. Still other Members proposed use of the eight descriptors proposed by Ashby *et al.*, which are related to the IARC scheme.

b) The majority of the Committee supported the Agency's handling of the issues raised when departure from the defaults are contemplated. A sizeable minority of the Committee believed that the burden of proof should rest on showing that the defaults are implausible, and also that the Guidelines should be revised to include explicit and specific criteria for judging the validity of hypotheses invoked to depart from defaults. TheMembers of this minority group also found it ill-advised to provide multiple assessments and recommended as a matter of science policy that the Guidelines include the following:

1) an explicit statement that, given the goal of public health protection, where there is uncertainty regarding the mode of action of a chemical, the Agency will adopt the default assumptions that the agent in question is capable of acting linearly at low dose and that positive animal data are relevant to humans; and

2) an explicit statement that the burden of proof rests on showing the defaults (low dose linearity and relevance to humans) are not plausible (see the IARC document extract in Appendix A). Some EHC Members felt that the statement should emphasize the point that the basis for departing from defaults was that the alternative was the most likely or most plausible choice, rather than demonstrating that the default is not plausible; these Members felt that the proposed approach implied absolute knowledge or proof.

c) The proposed Guidelines state that: the proposed default procedures are “public health conservative,” but that they have been revised, when appropriate, to reflect the changes in the state-of-the-art since 1986. During the public review meeting, it was stated that, in some cases, the process could result in an assessment that is not specifically public health conservative; during the ensuing discussion, it was noted that both default as well as non-default procedures could yield results that might not be public health conservative. The EHC clearly understands that the primary goal of EPA actions is public health protection and recommends adding clarifications that will alleviate potential concerns by readers or users of the Guidelines, while at the same time promoting the use of good science for decision making.

d) The EHC generally endorsed the Guidelines' Mode of Action proposals, but suggested that the Guidelines contain specific criteria for judging that the data on mode of action are valid and adequate. This should include a discussion of mode of action that reflects the lack of a clear distinction between direct DNA damage and other mechanisms with respect to the low dose response relationship. In fact, most Members felt that the Guidelines should

include some specific examples of chemicals that are accepted as threshold carcinogens.

e) The proposed Guidelines provide for more flexibility in risk assessment and provide means for incorporating other types of biological data and information on mode of action into dose response assessment. However, some EHC Members noted that, in some areas, there may not be enough guidance on mode of action determination and in other areas as well; given the wide usage of the Guidelines, these Members believe that additional clarification is needed. **Although sympathetic to the Agency's reluctance to develop additional criteria, the EHC recommends that further guidance should be provided.** In addition, it was noted that, unlike other parts of the Guidelines, no examples are provided in the dose response section. The Committee also noted that, often, a number of different models will provide an equivalent fit to the data. However, some unrestricted models may give answers that are biologically unrealistic in some cases and should be avoided. **Since the issues regarding selection of a model for use in the observable range are largely generic, the EHC recommends that the Agency select a default procedure for routine use in such calculations. The EHC supports the Agency's efforts, reflected in these and other guidelines documents, to harmonize approaches to dose response evaluations for cancer and non-cancer endpoints, and encourages the Agency to do more**..

f) The EHC understands the theoretical desirability of the proposed biologically based dose response model. **However, the proposed Guidelines may be misleading in this respect because no such model presently exists in a complete form, nor is it clear how to develop such in the foreseeable future. The Agency's definition of a biologically based model thus seems unnecessarily narrow. The Agency should define more clearly what is meant by a biologically based dose response model and to give guidance as to when such a model is preferred over the default linear or non-linear approaches.** This report, in section 3.3.2, proposes a possible definition for the Agency's for the Agency's consideration.

g) **To mitigate against unnecessary inconsistencies and confusion in the application of the Guidelines to determine the point of departure, the EHC suggests that further guidance be given in the Guidelines on this approach.** In addition, the EHC sees little scientific justification for separate approaches to LED_{10} and benchmark calculations, and therefore recommends that the approaches used for non-carcinogens and (low dose) non-linear carcinogens (*e.g.*, choice of dose-response model, choice of point of departure, and use of statistical confidence bounds) be harmonized to the extent possible. In general, it is recommended that a single risk level, e.g., ED_{10}, be selected as the point of departure when the (low dose) non-linear approach is applied. **The consensus of the Committee was that both point estimates and statistical bounds can be useful in different circumstances, and recommended that the Agency routinely calculate and present the point estimate of the ED_{10} and the corresponding upper and lower 95% statistical bounds.**

h) There is no explicit statement in the proposal that statistical significance should be a basic requirement for determining causality. This lack of an explicit statement has been interpreted as misleading and implying that there is a hidden intent to eliminate statistical significance as a consideration in assessing causality. **Adding appropriate and specific language concerning statistical significance should rectify this problem. Also, because of the difficulties with the interpretation of observational studies, the EHC believes that a weight of evidence approach to evaluating causality is appropriate.**

i) ***Vis-a-vis* assessing tumors, the Committee endorses the procedure adopted by the National Toxicology program of combining some closely related tumor types for statistical analyses, but otherwise conducting separate tests for different tumor types. The EHC also endorses the Agency's Guidelines that give additional significance to the finding of uncommon tumor types, multiple sites, tumors by more than one route of administration, etc.** (see pp. 55-56 of the Guidelines). Lastly, because of data on the relationship of body weight and some tumor types, the EHC encourages the Agency to consider the role of diet

in a given tumor response. Consideration of this issue should be addressed when evaluating a given tumor data set.

j) The developing infant and child should be recognized as a subgroup which is particularly sensitive to the carcinogenicity of a number of agents. Risk assessments should explicitly account for the differential susceptibility of the young. Other differences in susceptibility should also be taken into account when data permit.

APPENDIX A

Extract from *The Consensus Report of an IARC Monographs Working Group on Mechanisms of Carcinogenesis in Risk Identification* (IARC Internal Technical Report No. 91/002, 1991).

**********EXTRACT BEGINS**********

IV. PROPOSED PRINCIPLES AND PROCEDURES FOR USING information ON MECHANISMS IN EVALUATING CARCINOGENIC RISKS TO HUMANS

1. Introduction

The IARC Monographs programme is concerned with the evaluation of carcinogenic hazard posed to humans by chemical and other agents. In many cases, epidemiological data adequate to permit definitive evaluations of human carcinogenicity are not available. In these circumstances, in addition to consideration of animal bioassay data, assessment of data possibly relevant to the mechanism by which the putative carcinogen acts may be helpful in making the overall evaluation.

2. General principles

In view of the rapid and continuing development in understanding of fundamental mechanisms of carcinogenesis, it would be unhelpful to define strict rules concerning the need for epidemiological or other specific evidence to justify particular classifications. This raises the possibility that an agent might be placed in Group 1 in The absence of sufficient epidemiological evidence at least in the traditional sense. 'Sufficient' evidence of carcinogenicity in humans is in effect being redefined to encompass its usual scientific meaning. To define what constitutes such evidence in the current state of knowledge will be an important task of future Monographs working groups.

Many agents that might, in reality, cause substantial and/or widespread increases in cancer risk in human populations are not found to increase risk in conventional epidemiological research. The ability to measure exposures more sensitively, and closer to the target tissue, along with the ability to identify tumors with specific genetic markers, may greatly increase the capacity to detect cancer risks that were previously undetectable. (This is because the relative risks that are registered in studies that incorporate better, more specific, biologically based measures are higher

and /or have lower statistical variance.) These developments will result in the classification in Group 1 of an increased proportion of candidate agents.

Mechanisms may be understood at many different levels (see pp. 40-44); e.g., for a genotoxic carcinogen: at the level of metabolism, DNA damage, repair, mutational events, amino-acid changes in a proto-oncogene or tumor suppressor gene product, changes in function of the protein, the effect of the altered protein on cellular function or the stage in the carcinogenic process at which the change may be effective. Because different agents act by different mechanisms, the relevance and importance of each of these levels is dependent on which agent and which tumor site is being considered. It should always be made clear which levels are being considered and which are not. It should also be borne in mind that knowledge that an agent acts by one mechanism does not exclude the operation of others.

Conclusions about the operation of a particular mechanism should follow the 'strength-of- evidence' approach which is fundamental to the Agency's evaluations of carcinogenicity per se. The strength of the evidence for a particular mechanism should be assessed using terms such as 'weak', 'moderate' and 'strong'. The working groups are responsible for assessing the available data with regard to the relevance, reproducibility and sensitivity of the methods employed and other considerations of good laboratory and scientific practice.

The available evidence may show that similar mechanisms are acting in humans and experimental animals. Of particular concern are those situations in which the possibility is considered of species-specific activity. One concern would be raised when humans are the more affected or susceptible species. This could be evaluated on the basis of knowledge of mechanisms and the comparative relevance of a mechanism to animal and human responses. Another concern is raised when the putatively unaffected species is human beings. Certain principles should be applied before such species-specific activity can be concluded. It should be established, (I) for the tumor site in question, that the mechanism in question is the primary one in the tumorigenesis in that species; (ii) that the same or a similar mechanism does not operate in humans; and (iii) whether the agent induces other types of tumors in experimental animals. If other types of tumors are induced, then (i) and (ii) would have to be fulfilled for each of them. Qualitative differences, in which effects occurring in one species are not expected to occur in another, should be distinguished from quantitative differences (such as different rates of biotransformation), which may influence only the degree of response rather than the presence or absence of a response.

In view of the considerations summarized above, the Group recommended that a discussion of possible mechanisms be included in the monographs when appropriate data are available.

3. Use of data on mechanisms of carcinogenesis in the evaluation of carcinogenic risk to humans

The following steps should be taken to determine the contribution made by specific data on (known or presumed) mechanisms of carcinogenesis to evaluation of the carcinogenic risk to humans of a particular agent.

3.1 Summarization of the available data on mechanisms

For the agent, mixture or exposure circumstance being evaluated, the available data on mechanisms of carcinogenesis from studies in humans, animals and assorted tissue and cell test systems should be considered within the following four descriptive dimensions. These dimensions refer to the level at which the mechanism exists and not necessarily to the level of the system in which the observations were made. An agent may have effects on several or all dimensions.

The set of dimensions constitutes a framework that should facilitate, not rigidify, the description of the range of data available. For each dimension, examples are given; these are indicative only and would be subject to modification by each working group.

(a) Evidence of genotoxicity (i.e., structural changes at the level of the gene), for example,
- -structure-activity considerations
- -toxicokinetic considerations
- -adduct formation
- -mutagenicity
- -other genetic changes

(b) Evidence of effects on the expression of relevant genes (i.e., functional changes at the intracellular level), for example,
- -alteration of the structure or quantity of the product of a proto-oncogene or tumor suppressor gene
- -other effect on gene expression

(c) Evidence of relevant effects on cell behavior (i.e., morphological or behavioral changes at the cellular or tissue level), for example,
- -induction of mitogenesis, compensatory cell proliferation, preneoplasia and hyperplasia

-survival of premalignant or malignant cells
-metastatic potential

(d) Evidence of time and dose relationships of carcinogenic effects and interactions between agents, for example,

-early/late stage, as inferred from multistage models
-initiation/promotion/progression/malignant conversion, as defined in animal carcinogenicity experiments

These dimensions are not mutually exclusive. Thus, the mechanisms by which an agent acts on the expression of relevant genes would still be summarized under the second dimension even if it were known with reasonable certainty that those effects resulted from genotoxicity and would therefore also be described under the first heading.

3.2 Evaluation of the strength of the evidence for action of certain mechanisms of carcinogenesis

It is proposed that the strength of the evidence for each effect be evaluated in association with the framework proposed above. The evaluation could be made using terms such as ‘strong’, ‘moderate’ and ‘weak’ evidence for action of the agent via the mechanism or, where appropriate, evidence of lack of effect of the mechanism.

Ideally, operational guidelines for these evaluations would be worked out. Such guidelines may prove difficult to develop because of the diversity of the phenomena being observed. The provision of answers to the following general questions, however, may assist in evaluating the strength of the evidence for the action of particular mechanisms:

(1) Is the method valid, reliable and reproducible?
(2) Is the time sequence between exposure to the agent, measurement of the endpoint and cancer compatible with a cause-effect relationship?
(3) Have the agent and the endpoint been evaluated in several different species and/or qualitatively different test systems?
(4) Are there tenable alternatives to a cause-effect explanation for the association between exposure and the endpoint? Such alternatives would include chance, bias in design or conduct of the study and confounding of the exposure to the agent of interest with some exposure to other agent that can cause the endpoint.
(5) Is the association between the agent and the mechanistic endpoint consistent with tumor response and other information on the carcinogenic process?

3.3. Assessment of the relevance of the available data on mechanisms to the evaluation of the carcinogenic risk of the agent to humans.

A number of issues are pertinent to assessing the relevance of data on mechanisms to evaluation of the carcinogenic risk of the agent to humans, however strong the evidence for an effect of the agent on the operation of these mechanisms might be.

(1) It would be desirable to have evidence that the effect lies in the chain of events that link carcinogenic agents with cancer. Such evidence need not necessarily be available for human, although evidence of the relevance of

(2) The relevance of the test system to human responses must be considered.

(3) Generally, the closer the tissue, cells or cellular components of the system being studied are to the target cells in the mammon tissues in which the carcinogenic effect is thought to occur, the more relevant data from the experimental system are to evaluation of carcinogenic risk. Notwithstanding, the validity of the endpoint examined is of greater relevance.

(4) A particular mechanistic result is more relevant to evaluation of carcinogenic risks in humans if it can be shown that toxicokinetic variables in the test system are similar to those in human beings.

(5) The measured endpoint should be specific for the mechanism that is actually thought to be relevant to carcinogenesis. Thus, for example, when gene mutation is the relevant endpoint, greater weight would be given to results in which mutation had actually been observed than to those in which only less specific endpoints were attained.

3.4 Suggested guidelines

When the available data on mechanisms are thought to be relevant to evaluation of the carcinogenic risk of an agent to humans, they should be used in making the

overall evaluation, together with the combined evidence for animal and/or human carcinogenicity.

It is not possible to elaborate definitive guidelines for all possible situations in which mechanistic data may influence evaluation of carcinogens. The following scenarios are illustrative of the range of options available. First, information concerning mechanisms of action may confirm a particular level of carcinogen classification as indicated on the basis of epidemiological and/or animal carcinogenicity data. Second, for a particular agent, strong evidence for a mechanism of action that is relevant to carcinogenicity in humans could justify 'upgrading' its overall evaluation. Third, an overall evaluation of human cancer hazard on the basis of animal carcinogenicity data could be downgraded by strong evidence that the mechanism responsible for tumor growth in experimental animals is not relevant to humans. In keeping with the goal of public health, priority must be given to the demonstration that the mechanism is irrelevant to humans. Obviously, the absence of mechanistic data should not influence the evaluation of human carcinogenic hazards on the basis of epidemiological or animal tumor data.

Despite these caveats, and in addition to the above-mentioned proposal for systematic appraisal of mechanistic data (3.1, 3.2, 3.3), increased use of these data in the overall evaluation may require modification of the criteria for the IARC categories.

The Working Group considered the impact of increased understanding of mechanisms of carcinogenesis on the present guidelines for classification of carcinogens into Group1 to 4. It was noted that, as specified in the Preamble, the guidelines for categorization are not definitive, and exceptions can be made. Currently, for all agents in Group 1 there is sufficient evidence for cancer causation in humans from studies that have established a link between levels of exposure and increased cancer incidence. A Group 1 categorization could also be achieved by:

**********EXTRACT ENDS**********

APPENDIX B

(From IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans. Preamble, pages 7-8, Volume 67, 1996)

The *Monographs* represent the first step in carcinogenic risk assessment, which involves examination of all relevant information in order to assess the strength of the available evidence that certain exposures could alter the incidences of cancer in humans. The second step is quantitative risk estimation. Detailed, quantitative evaluations of epidemiological data may be made in the *Monographs*, but without extrapolation beyond the range of the data available. Quantitative extrapolation from experimental data to the human situation is not undertaken.

DISTRIBUTION LIST

Administrator
Deputy Administrator
Assistant Administrators
Deputy Assistant Administrator for Science, ORD
Director, Office of Science Policy, ORD
Director, Office of Air Quality Planning and Standards, OAR
Director, National Center for Environmental Assessment, ORD, RTP, NC
Director, Office of Air Quality Planning and Standards, OAR, RTP, NC
EPA Regional Administrators
EPA Laboratory Directors
EPA Headquarters Library
EPA Regional Libraries
EPA Laboratory Libraries
Library of Congress
National Technical Information Service
Congressional Research Service