United States Environmental Protection Agency

Science Advisory Board (1400)

EPA-SAB-EHC-95-014
May 1995

AN SAB REPORT:GUIDELINES FOR REPRODUCTIVE TOXICITY RISK ASSESSMENT

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

May 2, 1995

EPA-SAB-EHC-95-014

Honorable Carol M. Browner
Administrator
U.S. Environmental Protection Agency
401 M Street, S.W.
Washington, D.C. 20460

Subject: Science Advisory Board's review of the Draft Guidelines for Reproductive Toxicity Risk Assessment (EPA/600/AP-94/001, February, 1994)

Dear Ms. Browner:

The Guidelines for Reproductive Toxicity were originally proposed in 1988 as separate Guidelines for Assessing Male Reproductive Risk and Guidelines for Assessing Female Reproductive Risk. Following public comment, the EPA's Science Advisory Board (SAB) reviewed the proposed guidelines and in its report (EPA-SAB-EHC-89-005) recommended several changes, including the combining of the two guidelines into a single guideline for reproductive toxicity risk assessment. The current draft effects this combination; in addition, the female component was expanded substantially while retaining the original basic concepts.

At the request of the Office of Research and Development (ORD), the SAB's Environmental Health Committee met on July 19, 1994 to review the subject draft Guidelines. The review addressed a series of issues developed (through discussions between the Committee Chair, SAB Staff, and ORD Staff) constituting the formal Charge for the activity. This Charge included one major over-arching question: In general, does the document reflect current scientific knowledge relevant to reproductive toxicity risk assessment? Addressing this issue, the Committee found the overall scientific foundations of the draft Guidelines' positions to be generally sound. This finding notwithstanding, the Committee provided suggestions for improvement and some specific criticisms in its report.

The remainder of this letter summarizes each element of the Charge (the full text of which will be found in section 2.2 of the enclosed report) and the Committee's findings and recommendations.

a) Combining hazard identification and dose response evaluation

The Committee does not support the combination of the hazard identification and dose-response evaluation, preferring the four-step risk assessment paradigm proposed by the 1983 National Research Council committee on risk assessment. In addition, the Committee has suggested revisions for Table 5 of the draft Guidelines (the table provides a scheme for judging the available evidence on the reproductive toxicity of a particular agent). These revisions include redefinition of some of the categories of evidence to bring them into agreement with other portions of the draft document.

b) Gender-neutral default assumption

The Committee agrees that it is reasonable to assume that an agent which acts as a reproductive toxicant in one sex may also adversely affect reproductive function in the other sex; additional discussion to support this default assumption is suggested

c) Default assumption of a threshold for non-genotoxic agents

The Committee believes that the threshold assumption should be invoked only after an evaluation of the likely biological mechanism and mechanistic information indicates that linear responses would not be expected.

d) Endocrine disruptors

The Committee recommends that more discussion on this issue be incorporated in the Guidelines, noting the evidence that xenobiotics with estrogenic activity, as well as antiestrogens, androgenic, and antiandrogenic compounds, can adversely affect reproduction and development. The Committee agreed that exposure to such chemicals is a potentially serious public health hazard.

e) Need for multiple negative reproductive toxicity studies to adjudge a toxicant as "unlikely to pose a hazard."

The Committee felt that the Guidelines accurately reflected the

underlying science -- it is not possible to state with confidence that an agent is unlikely to constitute a hazard until one has ruled out the possibility that a lack of response is due to an idiosyncratic insensitivity of the species tested, or that all aspects of reproduction have been comprehensively and sensitively assessed. The Committee also felt that the burden of proof should lie with showing a lack of hazard, consistent with public health protection.

f) Susceptible populations

Given the increasing evidence that individuals and populations vary in sensitivity to environmental toxicants, the Committee recommends that EPA substantially expand the coverage of this topic in the draft document. The Committee also recommends that the Guidelines require that relevant information on differential risks to subsets of the population be incorporated into risk assessments for environmental toxicants when possible.

g) Complex mixtures and exposures

The Committee accepts the basic substance of the draft document via-a vis the discussion of these subjects. but, suggests the addition of discussion of several exposure assessment issues to strengthen the Guidelines. The Agency should develop an overall strategy to evaluate exposures to mixtures, exposures to multiple single agents, and exposures to the same agent via multiple pathways.

We appreciate the opportunity to review this document, and look forward to your response to the issues we have raised.

Genevieve M. Matanoski

Dr. Genevieve Matanoski, Chair
Science Advisory Board

Dr. Frederica Perera, Chair
Environmental Health Committee

ENCLOSURE

Distribution List

Administrator
Deputy Administrator
Assistant Administrators
Deputy Assistant Administrator for Pesticides and Toxic Substances
Deputy Assistant Administrator for Research and Development
Deputy Assistant Administrator for Water
EPA Regional Administrators
EPA Laboratory Directors
EPA Headquarters Library
EPA Regional Libraries
EPA Laboratory Libraries
Staff Director, Scientific Advisory Panel
Library of Congress
National Technical Information Service

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.

SCIENCE ADVISORY BOARD
ENVIRONMENTAL HEALTH COMMITTEE

APRIL 6, 1994

CHAIR
Dr. Frederica Perera, School of Public Health, Columbia University, New York City, NY

MEMBERS
Dr. William B. Bunn, Mobile Administrative Services, Company, Inc, Princeton, NJ

Dr. Kenny S. Crump, Clement International Corp., Ruston, LA

Dr. Rogene F. Henderson, Chemistry & Biochemical Toxicology Group, Inhalation Toxicology Research Institute, Albuquerque, NM

Dr. Richard Jackson[1], National Center for Environmental Health, Center for Disease Control, Atlanta, GA

Dr. Donald R. Mattison, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

Dr. Richard Monson, Harvard School of Public Health, Boston, MA

Dr. Martha J. Radike, Dept. of Environmental Health, University of Cincinnati, Cincinnati, OH

Dr. Ellen K. Silbergeld[2], University of Maryland, Dept. of Epidemiology, Baltimore, MD

CONSULTANTS
Dr. George P. Daston, Miami Valley Laboratories, The Proctor and Gamble Co., Ross, OH

Dr. Elaine Faustman, Dept. of Environmental Health, University of Washington, Seattle, WA

Dr. David P. Rall, Washington, D.C.

---

[1] At the time of the meeting, with the California State Department of Public Health

[2] Member, Science Advisory Board Executive Committee

Dr. Knut Ringen, Center for Worker's Rights. Washington, DC

Dr. John T. Wilson, Shreveport, LA

Dr. Lauren Zeise, Reproductive and Cancer Hazard Assessment Section, California Environmental Protection Agency, Berkeley, CA

FEDERAL EXPERT
Dr. George W. Lucier, National Institute of Environmental Health Sciences, Research Triangle Park, NC

DESIGNATED FEDERAL OFFICIAL
Mr. Samuel Rondberg, Environmental Health Committee, Science Advisory Board (1400F), U.S. Environmental Protection Agency. Washington, D.C. 20460

STAFF SECRETARY
Ms. Mary L. Winston, Environmental Protection Agency, Science Advisory Board (1400F), Washington, D.C. 20460

ABSTRACT

The Draft Guidelines for Reproductive Toxicity were reviewed by the Science Advisory Board's Environmental Health Committee on July 19, 1994. The Committee found the overall scientific foundations of the draft Guidelines' positions generally sound, but provided suggestions for improvement and some specific criticisms. The following text summarizes the Committee's findings:

a) The Committee does not support the combination of the hazard identification and dose-response evaluation, preferring the four-step risk assessment paradigm presented by the 1983 National Research Council committee (NRC, 1983) on risk assessment.

b) The Committee agrees that it is reasonable to assume that, in the absence of contraindicating information, an agent which acts as a reproductive toxicant in one sex may also adversely affect reproductive function in the other sex. However, discussion to support this default assumption is incomplete and should be developed more fully.

c) The Committee believes that a threshold should be assumed as a default only after an evaluation of other possibilities. Use of the threshold assumption should occur only after an evaluation of the likely biological mechanism and mechanistic information indicates that linear responses would not be expected.

d) The Committee recommends that more discussion on the issue of endocrine disruptors be incorporated in the Guidelines. The Committee agreed that exposure to such chemicals is a potentially serious public health hazard

e) The Committee felt that the Guidelines accurately reflected the underlying science on the need for multiple negative reproductive toxicity studies to adjudge a toxicant as "unlikely to pose a hazard."

f) Given the increasing evidence that individuals and populations vary in sensitivity to environmental toxicants, the Committee recommends that EPA substantially expand the coverage of susceptible sub-populations in the draft document.

h) The Committee accepts the basic substance of the draft document via-a vis the discussion of complex mixtures and exposures, but, suggests the addition of discussion of several exposure assessment issues to strengthen the Guidelines.

KEYWORDS: Reproductive toxicity; environmental toxicants; guidelines; risk assessment; endocrine disruptors; exposure thresholds

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY 1

2. INTRODUCTION 4
2.1 Background 4
2.2 Charge 4

3. DETAILED FINDINGS 7
3.1 Combining of Hazard ID/Dose Response Evaluation 7
3.1.1 Reliance upon scientific judgement 10
3.1.2 Discussion of statistical evaluation 10
3.1.3 Suggested revisions to Table 5 11
3.2 Gender-neutral default assumption 11
3.3 Threshold default assumption 13
3.4 Endocrine disrupters 14
3.5 Need for multiple negative studies 15
3.6 Susceptible populations 16
3.7 Risks from complex mixtures and exposures 16
3.7.1 Specific guidance on exposure assessments for reproductive toxicants 17
3.8 Scientific underpinnings of the guidelines 18

4. CONCLUSIONS 22

1. EXECUTIVE SUMMARY

The Guidelines for Reproductive Toxicity were originally proposed in the Federal Register in 1988 as separate Guidelines for Assessing Male Reproductive Risk and Guidelines for Assessing Female Reproductive Risk. Following public comment, the EPA's Science Advisory Board reviewed the proposed guidelines and recommended several changes, including the combining of the two guidelines into a single guideline for reproductive toxicity risk assessment (SAB, 1989). A draft document effecting this combination; and expanding the female component was reviewed by the Science Advisory Board's Environmental Health Committee on July 19, 1994.

The Committee found the overall scientific foundations of the draft Guidelines' positions generally sound, but provided suggestions for improvement and some specific criticisms. The following discussion summarizes the specific elements of the Charge for this review (see section 2.2 for the full Charge) and the Committee's findings.

a) Combining hazard identification and dose response evaluation

The Committee does not support the combination of the hazard identification and dose-response evaluation, preferring the four-step risk assessment paradigm presented by the 1983 National Research Council committee (NRC, 1983) on risk assessment. The Committee bases this position on three considerations:

1) The NRC paradigm is not restricted to non-threshold responses but is equally applicable to both threshold and non-threshold responses.

2) Consistency in risk assessment and communication of ideas will be fostered by continuing adherence to the paradigm unless there are compelling reasons for a departure.

3) Preserving a distinction between the two steps may enhance our knowledge of reproductive toxicity.

In addition, the Committee has suggested revisions for Table 5 of the draft Guidelines (the table provides a scheme for judging the available evidence on the reproductive toxicity of a particular agent). These revisions include a change in title to make it more descriptive, and redefinition of some of the categories of evidence to bring it into agreement with other portions of the draft document.

b) Gender-neutral default assumption

The Committee agrees that it is reasonable to assume that, in the absence of contraindicating information, an agent which acts as a reproductive toxicant in one sex may also adversely affect reproductive function in the other sex. However, discussion to support this default assumption in the main section of the document is incomplete and should be developed more fully. Also, a more detailed presentation on contraindicating information which would obviate the need for using the this default assumption is needed.

c) Default assumption of a threshold for non-genotoxic agents

The Committee believes that a threshold should be assumed as a default only after an evaluation of other possibilities. Although some of the many mechanisms by which toxicants can exert their effects may indicate threshold behavior, others do not. For example, recent studies in several laboratories have demonstrated that the shape of the dose response curve cannot be predicted solely on the knowledge that a response is receptor-mediated. Since many chemicals exert their reproductive and developmental effects by mimicking or blocking hormone action, selection of a threshold default assumption without other mechanistic or biological information could be inappropriate. Consequently, use of the threshold assumption should occur only after an evaluation of the likely biological mechanism and mechanistic information indicates that linear responses would not be expected.

d) Endocrine disruptors

The Committee recommends that more discussion on this issue be incorporated in the Guidelines, noting the evidence that xenobiotics with estrogenic activity, as well as antiestrogens, androgenic, and anti-androgenic compounds, can adversely affect reproduction and development. The Committee agreed that exposure to such chemicals is a potentially serious public health hazard. The Committee also recommends that:

1) The revised guideline document include a list of estrogen-sensitive reproductive endpoints (as identified by EPA staff at

the review meeting).

2) The Agency consider the use of risk assessment procedures that are mechanism-specific (when data permit) for assessing agents that have been identified as acting via a hormone receptor-mediated mechanism.

3) That measures of decreased sperm concentration/count be considered as a basis for regulatory action.

e) Need for multiple negative reproductive toxicity studies to adjudge a toxicant as "unlikely to pose a hazard."

The Committee felt that the Guidelines accurately reflected the underlying science. It is not possible to state with confidence that an agent is unlikely to constitute a hazard until one has ruled out the possibility that a lack of response is due to an idiosyncratic insensitivity of the species tested, or to the failure to assess comprehensively, using sensitive methods, all aspects of reproduction. The Committee also felt that the burden of proof should lie with showing a lack of hazard, consistent with public health protection. Consequently, the Committee recommends that Table 5 of the Guidelines explicitly state that data from a second species is necessary to classify an agent as being unlikely to pose a hazard.

f) Susceptible populations

Given the increasing evidence that individuals and populations vary in sensitivity to environmental toxicants, the Committee recommends that EPA substantially expand the coverage of this topic in the draft document. The Committee also recommends that the Guidelines require that relevant information on differential risks to subsets of the population be incorporated into risk assessments for environmental toxicants when possible.

h) Complex mixtures and exposures

The Committee accepts the basic substance of the draft document via-a vis the discussion of these subjects, but, suggests the addition of discussion of

several exposure assessment issues to strengthen the Guidelines. The Agency should develop an overall strategy to evaluate exposures to mixtures, exposures to multiple single agents, and exposures to the same agent via multiple pathways. In addition, exposures to multiple chemicals with a common mechanism of action should be discussed.

2. INTRODUCTION

2.1 Background

The Guidelines for Reproductive Toxicity were originally proposed in the Federal Register in 1988 as separate Guidelines for Assessing male Reproductive Risk and Guidelines for Assessing Female Reproductive Risk. Following public comment, the EPA's Science Advisory Board reviewed the proposed guidelines and recommended several changes, including the combining of the two guidelines into a single guideline for reproductive toxicity risk assessment (SAB, 1989). The current draft effects this combination; in addition, the female component was expanded substantially while retaining the original basic concepts.

Given the amount of reworking and updating required in these guidelines, and the time elapsed since the original proposals, EPA decided to solicit review and comments from interested parties before finalizing these guidelines for publication. In addition to the peer review, a *Federal Register* notice was published on March 4, 1994 announcing the availability of these guidelines for general comment.

In general, the reaction to these proposed Guidelines has been favorable from both the peer reviewers and the public commentors. A few substantive issues were raised in the peer reviews and public comments, most of which are included in the issues listed below.

2.2 Charge

Following discussions between the Chair, SAB Staff, and EPA Staff, the Environmental Health Committee identified the following issues on which tofocus its review:

a) Is it appropriate to combine the hazard identification and dose-response evaluation to reflect more accurately the process used for non-cancer health effects? *(Although this is a change from the original NAS paradigm, it is an approach that EPA has been working with for some time. This organization was used in the Guidelines for Developmental Toxicity Risk Assessment (1991)).*

b) A default assumption that supports a "gender-neutral" approach to risk

assessment for reproductive toxicity (see Section I). *(This assumption has been included to deal with the frequent situation in which sufficient data are available on only one sex that demonstrate reproductive toxicity; in this case, it is assumed that the agent may also adversely affect reproductive function in the other sex unless sufficient mechanistic evidence is available to negate the assumption.)*

c) The default assumption of a threshold for non-genotoxic agents of reproductive toxicity (see Section I). *(The argument against this assumption is that a background level of impairment already exists, and that toxic effects can add to that impairment at any level of exposure, especially for agents that also are endogenous or that act to mimic or compete with endogenous agents.)*

c) Adequate coverage of issues involving endocrine disruptors and development (see Section III).

With the recent prominence of these issues, is there adequate emphasis and guidance on hormonally mediated developmental effects and on use of data showing hormonal activity? Also, on a related point, is it appropriate to base regulatory action on decrements in sperm measures?

d) The requirement for more than one negative reproduction studies to judge that an agent is "unlikely to pose a hazard" for reproductive toxicity.

Is this requirement excessive, given the cost of conducting multi-generation reproduction studies? Currently, in the RfD/RfC process, no additional uncertainty factor is applied if only one acceptable multi-generation reproduction test is available showing no effect.

Are the criteria for evaluating the adequacy (design, power, etc.) of studies sufficiently discussed? As a scientific and/or policy matter, should a single valid positive study suffice to judge that an agent is "likely to pose a hazard," whereas a negative finding should require confirmation?

e) Is there adequate treatment of susceptible populations and individuals?

f) Are the risks from complex mixtures and multiple exposures (including agents that may act by a similar mechanism) adequately considered?

g) In general, does the document reflect current scientific knowledge relevant to reproductive toxicity risk assessment?

3. DETAILED FINDINGS

3.1 Combining of Hazard ID/Dose Response Evaluation

The draft Guidelines for Reproductive Toxicity Risk Assessment modify the risk assessment paradigm[3] proposed by the 1983 National Research Council (NRC) Committee on this topic (NRC, 1983) by combining the steps of hazard identification and dose-response assessment. The outcome of this combined step is described in the Guidelines document (page 6) as a "characterization of the health-related data as sufficient or insufficient to proceed with a quantitative risk assessment." The reason given for this approach is that whereas the NRC paradigm was developed for assessing carcinogens, EPA believes the paradigm is not applicable for assessing a threshold response, which the Agency is assuming generally holds true for reproductive effects. EPA postulates that reproductive effects are not expected to occur below some threshold level of exposure. Consequently, in the policy as drafted by EPA, determination of whether or not an agent poses a reproductive hazard (i.e., the hazard assessment) depends upon the threshold value and the level and pattern of human exposure. Specifically, the document (page 6) states: "A hazard is defined in terms of the range of effective doses, routes of exposure, timing and duration of exposure and other relevant factors." The threshold value would be estimated in the dose-response assessment step, and exposure would be evaluated in the exposure assessment step.

The Committee does not support the combination of the hazard identification and dose-response evaluation steps. The Committee wishes to make the following points and suggestions with respect to this issue:

a) The NRC paradigm is not restricted to non-threshold responses but is equally applicable to both threshold and non-threshold responses. Moreover, as discussed in section 3.3 below, the Committee questioned whether a threshold should be assumed for all reproductive toxicants.

b) Preserving a distinction between the two steps may enhance our knowledge of reproductive toxicity by allowing the identification of potential

[3] The NRC paradigm consists of four steps: a) Hazard identification: The determination of whether a particular chemical is or is not causally linked to particular health effects; b) Dose-response assessment: The determination of the relation between the magnitude of exposure and the probability of occurrence of the health effects in question; c) Exposure assessment: The determination of the extent of human exposure before or after application of regulatory controls; and d) Risk characterization: The description of the nature and often the magnitude of human risk, including attendant uncertainty.

hazards in the absence of sufficient data to conduct a dose-response analysis. Exposures that fall into this category should become high priority subjects for further testing and data collection, to support dose-response analysis.

c) The Committee feels that, since the NRC paradigm is widely applied and accepted, consistency in risk assessment and communication of ideas will be fostered by continuing adherence to the four step risk assessment paradigm unless there are compelling reasons for a departure. This paradigm has been widely used by agencies and groups involved in risk assessment since it was promulgated over ten years ago. The value and relevance of the paradigm was recently reaffirmed by the review of risk assessment practices conducted by the NRC in response to directives incorporated in the 1990 Clean Air Act Amendments (NRC, 1994; EPA 1994).

As noted, the NRC risk assessment paradigm has been widely used and has become a standard for conduct of risk assessments. Its use has helped to foster understanding and communication in risk assessment. The present draft of the reproductive Guidelines could be made more understandable by adopting functional separation of hazard identification and dose-response assessment. The Committee noted that there are several places in the guidelines where such functional separation would help to avoid potential confusion. For example, since the title to Table 5 (CATEGORIZATION OF THE HEALTH-RELATED DATA BASE HAZARD IDENTIFICATION/DOSE RESPONSE EVALUATION) involves hazard identification/dose-response evaluation, a reader familiar with the NRC paradigm would assume that this table involves elements of both hazard identification and dose-response assessment. However, the categorization does not actually involve the dose-response evaluation step. The categorization defined by Table 5 (additional comments on Table 5 follow below) is in reality a part of hazard identification and should be used to assess whether the dose-response step should be undertaken. If the guidelines were slightly reorganized to reflect this point of view then the discussion at the bottom of page 87 would follow much more logically and naturally. What is now being referred to as "completing a hazard identification/dose-response evaluation" becomes simply the dose-response step.

The hazard identification step, as defined in the NRC "Red Book" is basically an evaluation of the available data to determine if the data support continuing with the dose response evaluation and subsequent steps in the risk assessment process. For

potential reproductive toxicants, hazard assessment involves an evaluation of the evidence that an agent is a reproductive toxicant (at any dose in any species), and whether the data are sufficient to support a dose-response assessment. Deciding to proceed with a dose-response assessment should not be interpreted as implying necessarily that humans are at risk from any particular exposure level.

In the hazard evaluation step, the evaluation of data should focus upon the consistency of available information as to the nature of any reproductive effect observed, and its relevance to human health. This type of evaluation is essentially qualitative in nature, and may include evaluation of mechanism of action, the animal model(s) used, type of effect observed, the pattern of dose-response relationships, and overall consistency of available data. This step is distinct from dose-response assessment, in which data (when available) are utilized to evaluate in quantitative terms the relationship between dose or exposure and severity or probability of effect.

The Committee notes that there are some instances in which the data available are sufficient to identify a hazard, but insufficient to provide a credible basis for a dose-response analysis. That is, information may be of a qualitative, but not quantitative in nature. Qualitative information may be very important: as in, for instance, the Minimata episode, where quantitation of the risk of methyl mercury for intrauterine development was undertaken long after the identification of the epidemic, and the exact dose-response for human methyl mercury toxicity remains in some dispute (e.g., at the recent international symposium on mercury, Little Rock, November, 1994).[4] There may well be instances where qualitative information may be appropriately used to support a precautionary approach, even in the absence of quantitative data sufficient to support a full dose-response analysis. On the other hand, if hazard identification is based on a reproductive response in animals at a dose that is sufficiently high to disrupt the animal's general physical well-being, a precautionary approach may not be warranted. If we insist that the two stages should be combined, then we run the risk of throwing out information that identifies a hazard but does not support a dose response evaluation.

.The Committee notes that only minimal revision of the document would be required to separate hazard identification from dose response assessment, Sections III.A, III.B, III.C, III.E, and III.G discuss hazard identification, whereas Sections III.F and III.H discuss dose-response assessment.

[4] The Committee notes that the Minimata incident involved developmental toxicity, not reproductive effects, but believes that the lessons learned are relevant to the point concerning identification of hazard.

Pharmacokinetic considerations (III.D) are relevant to all steps of risk assessment, are most closely associated with dose assessment and should be included in depth where data are available (e.g., on changes in compartments during pregnancy). They could be addressed in a separate section or included with hazard assessment or dose-response evaluation as appropriate. In cases where pharmacokinetic data are included in the hazard assessment, it is important to determine that the differences in pharmacokinetics are qualitatively absolute, rather than quantitatively relative (e.g., absence of a critical metabolic pathway, rather than the differences in metabolic rate).

3.1.1 Reliance upon scientific judgement

Specific guidance is not provided for many of the critical decisions in risk assessment, and the document stresses use of "scientific judgement" in these decisions. Although scientific judgement is a critical element in all risk assessment, it is important for the Guidelines to define those elements of judgement sufficiently to enable all users of risk assessments to understand and evaluate the process and its relationship to explicitly stated scientific principles. This tendency is particularly apparent in the critical decision regarding determination of the No Observed Adverse Effects Level (NOAEL). The Agency should review each place in the Guidelines where scientific judgement is called for to determine whether more explicit guidance, with robust underlying scientific support, can be provided.

3.1.2 Discussion of statistical evaluation

Whereas detailed discussion is provided for various endpoints and experimental protocols, very little guidance is presented regarding the appropriate statistical tests for assessing reproductive toxicity. Statistical evaluation is critically important in establishing a NOAEL. The problems encountered in determining NOAELs (with statistical robustness and reliability) should be discussed, as should the need to provide information about the uncertainty surrounding any NOAEL or Lowest Observed Adverse Effects Level (LOAEL) (Gaylor, 1989). In particular, the discussion on evaluation of dose-response trends (page 23) could benefit from discussing specifically the use of statistical tests of trend, including the No Statistical Significance of Trend (NOSTASOT) procedure (Tukey *et al.*, 1985), used previously by EPA in specific reference dose/reference concentration (RfD/RfC) analyses for determining a NOAEL. Such procedures will generally have greater statistical power than pair-wise comparisons.

Also, for balance, the discussion (page 23) of non-significant trends or associations that may be biologically real should discuss also the likelihood that such

trends can also occur by chance. It could also emphasize that the likelihood that such a trend is a result of exposure is enhanced if similar trends are observed in a number of endpoints. Other statistical approaches should also be discussed.

The Agency appropriately emphasizes evaluation of statistical power in the evaluation of negative studies (page 67). However, formal statistical power calculations do not involve the actual data in a study and are difficult also to interpret. *A posteriori* power considerations can be better addressed through the use of confidence intervals for the effect rather than through formal power calculations. The following language is suggested as a replacement for the paragraph on page 67:

> *It is important to carefully evaluate results of a negative study, including the power of the study, and to compare the degree of concordance or discordance between that study and other studies (including careful analysis of comparability in the details, such as strain or species used, timing, reproductive status, similarity of adverse endpoints, etc). A power calculation does not reflect the observed outcomes of a study. Therefore, instead of making a formal power calculation, it may be more important to evaluate the ranges of outcomes consistent with various studies by calculating statistical confidence limits for the effects found in different studies. Studies with lower power will tend to provide wider confidence intervals. If the confidence intervals from a negative study and a positive study overlap, then there may be no conflict between the results of the two studies.*

3.1.3 Suggested revisions to Table 5

Table 5 (page 89) of the draft Guidelines document provides a scheme for judging the available evidence on the reproductive toxicity of a particular agent. Two broad categories ("Sufficient" and "Insufficient") are defined within the Table and data from all available studies are evaluated and used to judge whether available evidence allows a hazard assessment for reproductive toxicity. As mentioned earlier, the contents of Table 5 are not in full agreement with the text of the Guidelines document. As a result, the Committee has developed a suggested revision for this Table, including a change in title to make it more descriptive. The proposed revision follows on the next page.[5]

3.2 Gender-neutral default assumption

[5] The proposed revisions make the Table fairly terse; consequently the Agency may wish to add an explanatory statement or Concordance to the Guidelines . The Concordance can be longer and can be changed periodically as new scientific information becomes available or as regulatory guidance needs change

It is reasonable to assume, as stated in the draft document's Overview (p. 11), that "In the absence of information to the contrary, ... a chemical that acts as a reproductive toxicant in one sex may also adversely affect reproductive function in the other sex." The

TABLE 5 (REVISED)
CATEGORIZATION OF HEALTH-RELATED HAZARD IDENTIFICATION EVIDENCE FOR REPRODUCTIVE TOXICITY

SUFFICIENT EVIDENCE FOR TOXICITY

The Sufficient Evidence for Toxicity category includes data that collectively provide enough information to judge that a reproductive hazard exists. This category includes human and animal evidence.

Sufficient Human Evidence: This category includes data from epidemiologic studies that provide adequate evidence to judge that a causal relationship exists between exposure to an agent or mixture and human reproductive toxicity. A case series in conjunction with strong supporting evidence may also be used. In addition, mechanistic information based on animal studies that is directly to humans may be used.

Sufficient Experimental Animal Evidence/Limited Human Evidence: This category includes sufficient data from experimental animal studies and/or limited data from human studies to judge that the potential for reproductive toxicity to humans exists. The minimum information necessary is data from one study that demonstrate an adverse reproductive effect in one test species. Alternatively, minimal information for this classification may consist of data on humans for which a causal interpretation may be credible, but for which chance, bias, or confounding cannot be ruled out with reasonable confidence.

INSUFFICIENT EVIDENCE FOR TOXICITY

The Insufficient Evidence for Toxicity category includes evidence for which there are inadequate data in animals or in humans upon which to base a judgment of causality.

EVIDENCE SUGGESTING LACK OF TOXICITY

Because lack of toxicity can never be demonstrated with certainty, exposures in this category must always be judged as tentative. More than one negative study is required to support the designation of lack of toxicity

Suggestive Human Data: Data are available from multiple epidemiological studies for which there is quantitative information on exposure, credible evidence for lack of bias and confounding, sufficient information on most reproductive outcomes, and adequate sample size to lead to a narrow confidence interval around a rate ratio of 1.0 for all important reproductive outcomes.

Suggestive Animal Data: Data are available from at least two experimental animal studies for which there is quantitative information on exposure, credible evidence for lack of error, sufficient information on most reproductive outcomes, and adequate sample size to lead to a narrow confidence interval around a rate ratio of 1.0 for all important reproductive outcomes.

level of discussion of this and the other default assumptions in the Overview section is appropriate; however, discussion to support this assumption in the main section of the document (e.g., in the Dose Response/Hazard Identification section (III)) is incomplete and should be developed more fully. In addition, clearer guidance on how default assumptions should be addressed in a risk characterization (Section V) is needed.

A fuller discussion of "information to the contrary" which would obviate the need for making this default assumption is also needed. In this regard, the EPA may wish to consider the two examples about which the Committee expressed concern at the review meeting:

a) Gender differences in target organs or in tissue and temporal metabolic profiles related to potential adverse endpoints
b) Gender differences in susceptibility to xenobiotic-modified expression of isozymes (e.g., P450 and the resultant impact on the steady state metabolism of steroids important for gender organization of tissue and maintenance of reproductive status).

3.3 Threshold default assumption

.The Committee believes that use of an assumed threshold should occur only as a default after an evaluation of other possibilities. Reproductive and developmental toxicants can exert their effects by a wide variety of mechanisms. While some of these mechanisms may indicate threshold behavior; i.e., cytolethality-induced effects, other mechanisms do not imply a threshold. For example, recent studies in several laboratories have demonstrated that the shape of the dose response curve cannot be predicted solely on the knowledge that a response is receptor-mediated (Lucier *et a*l., 1993; Sewall and Lucier, 1994). Since many chemicals exert their reproductive and developmental effects by mimicking or blocking hormone action (e.g. dioxin, and environmental estrogens) selection of a threshold default assumption without other elements of mechanistic or biological information could be inappropriate.

In addition, the level of the threshold or No Observed Effects Level (NOEL) is dependent on the experimental design. Because the number of animals used in experimental studies is small (typically 10-50 per group) simply increasing the sample size will increase the chance of finding statistical significance between groups. For example, if a specific dose is given to a group of ten animals, three may have an effect compared to one out of ten in the controls. This would likely not be statistically significant, but if 100 animals were used and 30% of the animals exhibited an effect, the result would be significant. In this case the NOEL would become a Lowest Observed Effects Level (LOEL). The point is that the safety factor approach is inherently flawed, not that more animals should necessarily be used in experimental studies. The

"Benchmark Dose" method has been recommended to the Agency by the SAB (SAB 1990; 1993) as a possible alternative to the NOAEL/LOAEL approach .

Lastly, a recent study that analyzed dose response patterns for genotoxic and non-genotoxic carcinogens studied by the National Toxicology Program (315 chemicals) has implications for the threshold assumption for reproductive and developmental toxicants (Hoel and Portier, 1994). This analysis reported that non-genotoxic carcinogens were somewhat more likely to exhibit linear behavior under the test conditions than the genotoxic carcinogens. This finding was surprising and it implies that since some non-genotoxic carcinogens can act by similar mechanisms as reproductive and developmental toxicants (receptors, signal transduction pathways, enhanced mitogenesis, enzyme activation or inhibition), use of threshold assumption may not be appropriate.

There was considerable discussion of this issue by the Committee, and there was agreement that this issue was complex and our understanding of it still evolving; consequently methods used now may be inappropriate later. The Committee agreed that:

a) Although many reproductive and developmental toxicants will exhibit threshold behavior, some reproductive and developmental toxicants may not. Examples might include certain xenobiotic hormone-blocking or hormone-mimicking agents, or other chemicals that are adding to an existing molecular lesion (e.g., oxidative DNA damage) caused by background exposure to endogenous or exogenous compounds.

b) Use of the threshold assumption should occur only after an evaluation of the likely biological mechanism and available data provides evidence that linear responses would not be expected. In other words, a threshold for a reproductive or developmental toxicant should not be automatically assumed.

3.4 Endocrine disrupters

The Committee recommends that more discussion on this issue be incorporated in the Guidelines. It noted the evidence indicating that interference with endocrine targets, particularly by xenobiotics with estrogenic activity, as well as antiestrogens, androgenic, and antiandrogenic compounds, can adversely affect reproduction and development. Examples include the synthetic hormone diethylstilbestrol which has been documented to be a human developmental toxicant and the insecticide chlordecone which has been identified as a human reproductive toxicant (Williams and Uphouse, 1991). A number of other hormonal agents have been shown to be reproductive and/or developmental toxicants in animal species and are potential human hazards. The

Committee agreed that this is one of the known mechanisms of reproductive and developmental toxicity, and
that exposure to chemicals with hormonal activity is a potentially important public health hazard.

EPA Staff provided a summary of estrogen-sensitive reproductive endpoints that are measured in a number of the reproductive toxicity screening assays described in the draft Guidelines. The Committee recommended that the revised guideline document include a list of these endpoints. EPA Staff also indicated that the Agency intends to insert information regarding the effects of estrogens on the development of the male reproductive system into the section of the Guidelines on male endpoints, and on *in vitro* methods for detecting hormonal activity (The draft Guidelines already contain information on developmental effects in females -- pp. 55, 57). The Committee supports these changes.

In addition to the above, the Committee had two recommendations:

a) Where adequate data are available, the Agency should consider the use of risk assessment procedures that are mechanism-specific for assessing agents that have been identified as acting via a hormone receptor-mediated mechanism.

b) The Committee supports consideration of the use of measures of decreased sperm concentration/count as a basis for regulatory action. This position is based on the observation that the distribution of human sperm counts appears to include the minimum value necessary to ensure fertility; for individuals with sperm counts near this minimum, any measurable decrement in sperm count would adversely affect their fertility. In addition, this endpoint is one of the few biomarkers of effect available for reproductive toxicity; indications of change in this monitored endpoint should be closely evaluated.

3.5 Need for multiple negative studies

Some public comments submitted on the draft Guidelines questioned the Agency position that more than one negative study was required to constitute evidence suggesting a lack of hazard, and suggested that it is possible to categorize a substance as likely to pose a reproductive hazard based on a single positive study. These commentors felt that the expense of a comprehensive reproductive toxicity assessment is sufficiently high that it is unlikely that most companies would be willing to sponsor more than one such study on a compound that has not produced reproductive effects in a single well-conducted study.

Although the Committee was sensitive to this position, it was felt that the Guidelines, as stated, more accurately reflect the underlying science. It is not possible to state with confidence that an agent is likely to be without hazard until one has ruled out the possibility that a lack of response is due to an idiosyncratic insensitivity of the species tested, or to the failure to assess comprehensively, using sensitive methods, all aspects of reproduction. The Committee noted that in developmental toxicology, a lack of effect in a single species is also considered to be insufficient to ensure lack of human hazard; therefore, it would be difficult to classify with confidence an agent as being unlikely to have reproductive effect based on a single study. The Committee also felt that the burden of proof should lie with showing a lack of hazard, consistent with public health protection It was recommended by the Committee that Table 5 (CATEGORIZATION OF THE HEALTH-RELATED DATA BASE HAZARD IDENTIFICATION/ EVALUATION) of the Guidelines explicitly state that data from a second species are necessary to classify an agent as being unlikely to pose a hazard.

It was noted that data from a second -species generated as part of other study designs, such as sub-chronic studies, can increase the confidence in negative results; however, the magnitude of this confidence will depend on the endpoints measured. It was also noted that the Agency does not consider a lack of multiple studies to be a data gap of sufficient importance to warrant the imposition of an additional uncertainty factor. The Committee agrees with these practices.

3.6 Susceptible populations

The Committee recommends that EPA substantially expand the relevant discussion in the draft guidelines document to summarize the available data on individual and population sensitivity. There is increasing evidence that individuals and populations vary with respect to risk from environmental toxicants (NRC, 1993; 1994; Perera *et al.*, 1991). Inter-individual variability can result from variability in exposure (e.g., differences in patterns, timing, and intensity of exposure), as well as from host susceptibility (e.g., genetic, acquired, and developmental factors that may lead to heightened biologic response to exposure). Examples of the latter include polymorphisms in genes controlling xenobiotic metabolism and DNA repair, preexisting impairment or disease, nutritional deficits and stage of development. Normal physiological variation can also affect dose. Examples include mouthbreathing affecting lung deposition, as well as pregnancy, lactation, and the menopause being major events in women's life histories that can affect uptake, retention, and deposition of toxic chemicals stored in fat and bone (Mattison, Blann, and Malek, 1991).

The Committee also recommends that the Guidelines explicitly require that, when available, relevant information on differential risks to subsets of the population be

incorporated into risk assessments for environmental toxicants. Both exposure-related and biologic factors should be considered. One approach would be to present the range or distribution of risks across the population including children, women and minorities where data indicate differential exposure/susceptibility.

3.7 Risks from complex mixtures and exposures

The Committee accepts the basic substance of the draft document *via-a-vis* the discussion of these subjects. There are, however, several areas where we believe additional discussion of exposure assessment issues would strengthen the Guidelines. In general, the Committee recommends that the Agency develop an overall strategy to evaluate exposures to mixtures, exposures to multiple single agents, and exposures to the same agent via multiple pathways for reproductive toxicity endpoints. In addition, exposures to multiple chemicals with a common mechanism of action should be discussed. Human populations are generally exposed to complex mixtures through the uncontrolled general environment. Animal models should be developed to evaluate similar complex exposures under a controlled laboratory environment. Some more specific comments follow.

3.7.1 Specific guidance on exposure assessments for reproductive toxicants

The purpose of the draft document's Section IV (Exposure Assessment) is to provide clear guidance on estimating human exposure for reproductive toxicants. The discussion should be more prescriptive and provide more specific guidance for exposure assessments on issues such as patterns of exposure and reversibility of effect. A partial listing of endpoints where pattern of exposure is important to evaluate, and an indication of the significant parameters to be considered in the evaluation (e.g., age of individuals; differential exposures; peak versus average exposure) would be particularly useful. Guidance should also be given on other exposure assessment issues that arise when doing a risk assessment for a suspect reproductive toxicant. For example, the document should address ways of dealing with the frequent case where the exposure data are insufficient to assess adequately risks for specific reproductive endpoints (e.g., adding uncertainty/adjustment factors). It should also discuss aspects of the exposure assessment requiring emphasis in the risk characterization (such as inability to characterize exposures of populations that may be particularly susceptible to the reproductive endpoint of concern).

Much of the section is devoted to a discussion of exposure issues in interpreting reproductive toxicity studies (e.g., top paragraph p. 97; second & third paragraph p. 98; p. 100) and in addressing endpoints covered in the developmental toxicity guidelines (e.g., p. 101). We suggest that these items should be moved and incorporated as appropriate into earlier sections (e.g., Section III).

Questions arose during the Committee discussion of "margin of exposure." Margin of exposure (MOE) is given as "the ratio of the NOAEL from the more appropriate or sensitive species to the estimated human exposure level from all potential sources." Given that the NOAEL has not been adjusted to account for interspecies differences in pharmacokinetics or study sample size (e.g., number of animals in the selected NOAEL dose group) and that susceptibilities within the human population can be quite variable (and generally cannot be assessed), the MOE could be misleading in certain circumstances. As a simple example, consider a chronic exposure where the mouse NOAEL (in daily mg/kg) is 10 times the human dose; i.e., the MOE is 10. If the agent is active in parent form and metabolically eliminated, however, the effective dose to the human may be roughly equivalent to the mouse NOAEL and, in terms of pharmacokinetic or "effective" dose, the MOE would be roughly one. If there is significant variability in susceptibility or in patterns of exposure among people (and pattern is important), the "effective" MOE for some individuals could potentially be significantly less than one. We suggest that the MOE discussion should be modified to address these concerns.

3.8 Scientific underpinnings of the guidelines

The Committee especially endorses the separation of reproductive risk from developmental risks. There is ample scientific foundation to suggest that, for physiological, pharmacological, toxicological as well as chronological reasons, the methodological approaches for testing chemicals in animals and extrapolating those data to humans should be managed in a substantially different way for reproductive and developmental endpoints. However, it is also clear that many of the endpoints overlap (i.e., could be considered manifestations of either reproductive, developmental or mutagenic toxicity). It would strengthen this document to indicate how overlapping endpoints should be managed in characterizing risk, if they are considered manifestations of reproductive or developmental toxicity, and which risk assessment approaches might be considered appropriate depending upon the interpretation of the mechanism of toxicity producing the endpoint. In addition, some reproductive endpoints may overlap with endpoints evaluated for mutagenicity, and the guidelines should suggest alternative ways of interpreting and utilizing those endpoints depending on evidence of the structure and mechanism of action of the chemical.

With respect to the specific default assumptions, the EHC believes that it is reasonable to assume that an adverse reproductive effect observed in experimental animals represents presumptive evidence of a similar potential for humans. In addition, we support (as being reasonable positions) the default assumptions of conservation of the site and mechanism of action of xenobiotics across species and of a threshold for the dose response relationship (in many, but not all instances, as discussed above). Also, in the absence of more detailed data on site and mechanism of action, the Committee accepts the use of the most sensitive species as the most appropriate for estimating human risk. The Committee noted, however, the case of human fertility where certain

parameters of the biological process in certain human populations have little reserve capacity. For example, it has been proposed that any decrease in human sperm count is associated with a considerable decrease in male fertility; this stands in contrast to observations in laboratory rodents (Meistrich, 1984 *et seq*.).

With the caveats expressed above, we agree that, in the absence of detailed data to the contrary, it is reasonable to assume that a chemical acting in one sex may also adversely effect reproductive function in the other sex. However, we believe that testing in both sexes should be encouraged.

The above notwithstanding, some comments on the use of defaults *per se*, are also in order. The Committee appreciates the many problems faced by the Agency in assessing reproductive risk due to the limits in the state of scientific knowledge. Because of these limits, the proposed Guidelines must rely extensively on the use of defaults as

discussed above. We have two general suggestions concerning defaults, which are intended to improve both the current document and future revisions.

a) The current revision should provide additional detail on the appropriate use of default assumptions. The Committee encourages EPA to provide general criteria for deciding when to deviate from defaults, as well as examples of the proper use of defaults.

b) The Agency should develop a specific research agenda to reduce in the future the reliance on default assumptions in assessing reproductive risks.

We also recommend that more details on the mechanics of risk assessment for reproductive endpoints be included. We would encourage the agency to use this document as an opportunity to lay out some reference dose, and benchmark dose approaches for reproductive endpoints. Given the attention that the agency has directed to this issue, this would be the appropriate place to summarize these methods for consideration and discussion.

The Committee appreciates the considerable effort that has gone into combining both male and female reproductive endpoints and believes that it has added substantially to the quality and utility of the document. Although this does not mean that reproductive health effects in males or females (in the absence of a partner or child) are not important, one of the important components of the merging of the male and female reproductive risk assessment guidelines is that the unit of the couple becomes an important focus. This is also another reason why we suggest that the risk assessment strategies or methods need to be laid out in more detail. We are not aware of any other endpoint which has a couple-based biological basis; this makes risk assessment for

reproductive endpoints unique.

Despite the initial statement that reproduction is going to be considered separately from development, the document confuses them throughout the text and in the references. We would encourage the authors to go back through the document and make certain that the discussion and references are relevant to reproductive endpoints. For example, many of the references refer not to reproductive, but to developmental studies; however the text in the draft suggests they refer to reproductive endpoints.

The document should incorporate discussion of “windows of vulnerability” both during spermatogenesis and follicular development, ovulation, and corpus luteum formation. Different biological processes can result in differential susceptibility to environmental toxicants.

One of the notable gaps in the document is a lack of information on the validation of reproductive toxicity testing paradigms. We suggest that greater attention should be given by the authors to testing methods or approaches that have been validated and also

to better identifying where the lack of validation hinders or impairs the reproductive risk assessment process.

In discussion of additional test protocols (section III A7), the authors spend some time talking about dominant lethal assays but do not include discussion of other dominant assays. This is unfortunate because these other tests may be equally informative.

In the section describing sexual behavior (pg. 33), the authors assert that "Most human information comes from clinical reports in which the detection of exposure-effect associations are unlikely." The Committee disagrees with this, given that there is a large literature on effects on sexual behavior of a range of drugs. This literature is important information, and the authors are missing an opportunity to draw attention both to that literature as well as inferences from it for sexual behavior risk assessment. In addition, that literature allows the exploration of similarity of effect across species.

In the section that discusses the use of sperm evaluations (III B3d), the authors comment on a series of longitudinal study designs which have improved sensitivity. One of the complications of those longitudinal study designs, however, is that there is a substantial correlation structure within the data and none of the longitudinal study designs have used statistical techniques to account for the correlation structure in the data. Some attention to that limitation in statistical techniques in this section would be appropriate.

The issue of correlation of the data is also important in the sections in which the authors of the report describe approaches for collecting and using data from human populations. Some attention needs to be given to recommendations for handling the correlation structure in human data. This complication again points out the rationale for specific guidance for risk assessment approaches for reproductive endpoints which cannot be handled well in a generic risk assessment guideline document.

On Page 72, the authors also comment that "obtaining specimens with a high level of participation in the work force has been difficult." .This points to the need to develop more sensitive approaches to educating potential study populations, and the Guidelines should offer suggestions about tactics that improve the cooperation of study populations.

Lastly, we wish to point out some small but significant issues of terminology and the use of references. In the document, the authors describe couple-mediated approaches. Given that most of this material is directed toward the analysis of data derived from animals, it may be appropriate to consider changing "couple" to "breeding pair" when discussing animal data, but retaining "couple" when discussing human data or effects in humans. Another point of terminology is the inappropriate or confusing use of rates and indices. The authors need to make sure that they are using those terms correctly and consistently throughout the document. This is also relevant to the use of the terms “fertility” and “fecundity.” The authors would be well advised to adopt the demographer’s definition of those terms, with fecundity being the ability of a male or a female to reproduce and fertility being the actual production of live offspring. Those definitions should be used consistently throughout the document.

Throughout the document, there are places where references were omitted or are out of date. In the discussion on Page 31 of length of gestation, the authors suggest that as the length of gestation increases, birth weight may be higher. This is an example of a statement which needs to be referenced. Lengthened gestation can result in either lower or higher birth weights than term deliveries, and decreased birth weight puts the infant at greater risk of adverse outcome during or following parturition (Kassis *et al.,* 1991; McLean *et al.,* 1991; Goldenberg, *et al.,* 1989; and Eden *et al.,* 1987).

Finally, as a general comment, we were disappointed that no academic societies were asked to comment on the guidelines - given the number of societies that exist in reproductive biology and risk assessment. It would have been appropriate to give academic and clinical societies an opportunity to review and comment on the guidelines.

4. CONCLUSIONS

The Committee found the overall scientific foundations of the draft document's positions generally sound. Suggestions for improving the documents underpinnings include incorporating discussions of possible "windows of vulnerability" during the reproductive cycle or during gestation, focusing on the use of couples-based outcomes as a measure of reproductive toxicity, and providing better descriptions of categorical studies supporting the various default assumptions.

The following text reiterates each specific element of the Charge for this review (see section 2.2) and summarizes the Committee's findings.

a) Combining hazard identification and dose response evaluation

The Committee does not support the combination of the hazard identification and dose-response evaluation. In addition to the major points made below, it is inconsistent with the four-step risk assessment paradigm presented by the 1983 National Research Council committee (NRC, 1983) on risk assessment. Since it was established over ten years ago, this paradigm has been widely used by agencies and groups involved in risk assessment. The value and relevance of the paradigm was recently reaffirmed by the review of risk assessment practices conducted by the NRC in response to the Clean Air Act Amendments (NRC, 1994; EPA, 1994).

The Committee wishes to make the following points with respect to this issue:

1) The NRC paradigm is not restricted to non-threshold responses but is equally applicable to both threshold and non-threshold responses.

2) Consistency in risk assessment and communication of ideas will be fostered by continuing adherence to the paradigm unless there are compelling reasons for a departure.

3) Preserving a distinction between the two steps may enhance our knowledge of reproductive toxicity.

b) Gender-neutral default assumption

The Committee agrees that it is reasonable to assume that, in the absence of contraindicating information, an agent which acts as a reproductive toxicant in one sex may also adversely affect reproductive function in the other sex. However, discussion to support this default assumption in the main section of the document (e.g., in the Dose Response/Hazard Identification section (III) is incomplete and should be developed more fully.

Also, a more detailed presentation on contraindicating information which would obviate the need for using the this default assumption is also needed. In this regard, the EPA may wish to consider the two examples about which the Committee expressed concern at the review meeting:

1) Gender differences in target organs or in tissue and temporal metabolic profiles related to potential adverse endpoints

2) Gender differences in susceptibility to xenobiotic-modified expression of isozymes

c) Default assumption of a threshold for non-genotoxic agents

The Committee believes that use of an assumed threshold should occur only as a default after an evaluation of other possibilities. Reproductive and developmental toxicants can exert their effects by a wide variety of mechanisms. Although some of these mechanisms may indicate threshold behavior, others do not. Since many chemicals exert their reproductive and developmental effects by mimicking or blocking hormone action (e.g. dioxin, and environmental estrogens) selection of a threshold default assumption without other pieces of mechanistic or biological information could be inappropriate. Consequently, use of the threshold assumption should occur only after an evaluation of the likely biological mechanism and mechanistic information indicates that linear responses would not be expected. In other words, a threshold for a reproductive or developmental toxicant should not be automatically assumed.

d) Endocrine disruptors

The Committee recommends that more discussion on this issue be included in the Guidelines. It noted the evidence indicating that interference with

endocrine targets, particularly by xenobiotics with estrogenic activity, as well as antiestrogens, androgenic, and antiandrogenic compounds, can adversely affect reproduction and development. The Committee agreed that exposure to such chemicals is a potentially serious public health hazard.

The Committee also recommends that the revised guideline document include a list of estrogen-sensitive reproductive endpoints identified by EPA staff at the review meeting. EPA Staff also indicated that the Agency will add information on the effects of estrogens on the development of the male reproductive system into the section of the Guidelines on male endpoints, an improvement the Committee supports.

The Committee also recommends that the Agency consider the use of risk assessment procedures that are mechanism-specific (when data permit) for assessing agents that have been identified as acting via a hormone receptor-mediated mechanism; and that measures of decreased sperm concentration/count as be considered as a basis for regulatory action.

e) Need for multiple negative reproductive toxicity studies to adjudge a toxicant as "unlikely to pose a hazard."

Although the Committee considered carefully some of the public comments on this issue which disagreed with the position taken, it was felt that the Guidelines more accurately reflect the underlying science. It is not possible to state with confidence that an agent is likely to be without hazard until one has ruled out the possibility that a lack of response is due to an idiosyncratic insensitivity of the species tested, or that all aspects of reproduction have been comprehensively and sensitively assessed. The Committee also felt that the burden of proof should lie with showing a lack of hazard, consistent with public health protection. Consequently, the Committee recommends that Table 5 (CATEGORIZATION OF THE HEALTH-RELATED DATA BASE HAZARD EVIDENCE IDENTIFICATION FOR REPRODUCTIVE TOXICITY) of the Guidelines explicitly state that data from a second species was necessary to classify an agent as being unlikely to pose a hazard.

f) Susceptible populations

There is increasing evidence that individuals and populations vary with respect to risk from environmental toxicants, and the Committee recommends that EPA substantially expand the relevant discussion in the draft Guidelines document to summarize the available data on individual and population sensitivity. The Committee also recommends that the Guidelines require that relevant information on differential risks to subsets

of the population be incorporated into risk assessments for environmental toxicants when possible.

h) Complex mixtures and exposures

The Committee accepts the basic substance of the draft document via-a vis the discussion of these subjects. However, additional discussion of several exposure assessment issues would strengthen the Guidelines. The Committee recommends that the Agency develop an overall strategy to evaluate exposures to mixtures, exposures to multiple single agents, and exposures to the same agent via multiple pathways. In addition, exposures to multiple chemicals with a common mechanism of action should be discussed. Finally, the body of this report (section 3.7.1) provides some specific suggestions in improving exposure assessment for reproductive toxicants.

REFERENCES

Eden, R., Seifert, L., Winegar, A., and W. Spellacy. 1987. Perinatal characteristics of uncomplicated postdate pregnancies. *Obs. and Gyn.* 69(3 Pt 1):296-299.

EPA. 1994. Response from the Science Policy Council of the US Environmental Protection Agency to "Science and Judgment in Risk Assessment," A Report by the National Research Council (NRC). Memorandum from Robert M. Sussman, Chair SPC, May 31, 1994.

Gaylor, D. 1989. Quantitative risk analysis for quantal reproductive and developmental effects. *Environ. Health* 79:243-246.

Goldenberg, R., Davis., R., Cutter, G., Brumfield, C., and J. Foster. Prematurity, postdates, and growth retardation: the influence of use of ultrasonography on reported gestational age. *Am. J. of Obs. and Gyn.* 160(2):462-470.

Hoel, D.G., and C. Portier. 1994. Nonlinearity of dose-response functions for carcinogenicity. Environ. Health Perspect. 102(Suppl 1):109-133

Kassis, A., Mazor, M., Leiberman, J., and V. Insler. 1991. Management of post-date pregnancy: a case control study. *Israel Journal of Med. Science* 27(2):82-86.

Lucier, G. W., Portier, C. J., and M. A. Gallo. 1993. Receptor mechanisms and dose-response models for the effects of dioxins. Environ. Health Perspect. 101(l):36-44

Mattison, D., Blann, E., and A. Malek. 1991. Physiological alterations during pregnancy: impact on toxicokinetics. *Fund. and Appl. Toxicol.* 16:215-218

McLean, F., Boyd, M., Usher, R., and M. Kramer. 1991. Postterm infants: too big or too small? *Am. J. of Obs. and Gyn.* 164(2):619-624

Meistrich, M. L., 1989a. Calculation of the Incidence of Infertility in Human Populations from Sperm Measures Using the Two-Distribution Model. *Sperm Measures and Reproductive Success: Institute for Health Policy Analysis, Forum on science, Health and Environmental Risk Assessment*, pp. 275-290.

Meistrich, M. L., 1989b. Interspecies Comparison and Quantitative Extrapolation of Toxicity to the Human male Reproductive System. In: Toxicology of the Male and Female Reproductive Systems, Ed. P. K Working, Hemisphere Pub. Co., pp. 303-321.

Meistrich, M. L and C. C. Brown. 1993. Estimation of the Increased Risk of Human lnfertility from Alterations In Semen Characteristics. *Fertility and Sterility*, Vol 40, No. 2 (August).

Meistrich, M. L. and M. E. A. B. Van Beek. 1990. Radiation Sensitivity of the Human Testis. *Advances in Radiation Biology*, Vol, 14, pp. 227-268

Meistrich, M. L., Finch, M. and C. C. Lu. 1984. Strain Differences in the Response of Mouse Testicular Stem Cells to Fractionated Radiation. *Radiation Research* 97, pp. 478-487.

Perera, F., et al. 1991. Markers in Risk Assessment for Environmental Carcinogens. EMP, pp. 247-254.

NRC. 1994. Science and judgement in risk assessment, National Research Council Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Studies and Toxicology, National Academy Press, Washington DC.

NRC. 1993. Pesticides in the diets of infants and children. National Research Council Committee on Pesticides in the Diets of Infants and Children, Board of Agriculture and Board on Environmental Studies and Toxicology, Committee on Life Sciences, National Academy Press, Washington DC.

NRC. 1983. Risk Assessment in the Federal Government: Managing the Process. National Research Council, National Academy Press, Washington, DC.

SAB. 1993. Environmental Health Committee (EHC), U.S. EPA Science Advisory Board. *Review of the Office of Solid Waste and Emergency Responses Draft Risk Assessment Guidance for Superfund Human Health Evaluation* Manual. EPA-SAB-EHC-93-007.

SAB. 1990. Environmental Health Committee (EHC), U.S. EPA Science Advisory Board. *Use of Uncertainty and Modifying Factors in Establishing Reference Dose Levels*. EPA-SAB-EHC-90-005

SAB. 1989. Environmental Health Committee, (EHC) U.S. EPA Science Advisory Board. *Male and Female Reproductive Guidelines.* EPA-SAB-EHC-89-005.

Sewall, C. H., and G. W. Lucier. 1994. Dose-response and risk assessment considerations for receptor-mediated effects: case study with a TCDD hepatic tumor promotion model. In: Receptor-Mediated Biological Processes: Implications for Evaluating Carcinogenesis. Wiley-Liss, pp. 155-171.

Tukey J., Ciminera J., and B. Heyse. 1985. Testing the statistical certainty of a response to increasing doses of a drug. Biometrics 41:295-301.

Williams, J., and L. Uphouse. 1991. Vaginal cyclicity, sexual receptivity, and eating

behavior of the female rat following treatment with chlordecone. *Rerpro. Tox*. 5:65-71