EPA/600/FR-89/001

Monday
March 6, 1989

federal register

Part VII

# Environmental Protection Agency

**Proposed Amendments to the Guidelines for the Health Assessment of Suspect Developmental Toxicants; Request for Comments; Notice**

## ENVIRONMENTAL PROTECTION AGENCY

[FRL-3532-9]

### Proposed Amendments to the Guidelines for the Health Assessment of Suspect Developmental Toxicants

**AGENCY:** U.S. Environmental Protection Agency.

**ACTION:** Request for comments on the Proposed Amendments to the Guidelines for the Health Assessment of Suspect Developmental Toxicants.

**SUMMARY:** The U.S. Environmental Protection Agency (EPA) is today proposing amendments to the Guidelines for the Health Assessment of Suspect Developmental Toxicants that were issued on September 24, 1986 (51 FR 34028–34040) (hereafter "current guidelines").

These proposed amendments are intended to expand Agency guidance on the analysis of developmental toxicity data in accordance with appropriate scientific standards and with the policies and procedures established in the statutes administered by the EPA. The proposed amendments were developed as part of an interoffice guidelines development program under the auspices of the Agency's Risk Assessment Forum. The proposed amendments are based, in part, on recommendations developed in scientific workshops.

The public is invited to comment and public comments will be considered in final Agency decisions on amending the current guidelines. Commentors are asked to focus on several special issues, particularly, (1) a proposed new weight-of-evidence scheme and its use, and (2) the advantages and disadvantages of using this scheme only for hazard identification versus using it in conjunction with dose-response and exposure assessment information. Also, comments are invited on the use of the special term "reference dose for developmental toxicity ($RfD_{DT}$)." The term $RfD_{DT}$ is used to distinguish the time-limited reference dose for exposure during development from the reference dose (RfD), which generally refers to chronic exposure situations.

The proposed amendments are individually identified and explained in the Supplementary Information section of this notice. The full text of the proposed guidelines is published in the following section. As used in this notice, the term "proposed guidelines" refers to the current guidelines as modified by the proposed amendments. The request for comment applies only to the proposed amendments, but EPA will also consider any important new scientific information bearing on the proposed guidelines as a whole.

EPA's Science Advisory Board (SAB) also will review the proposed amendments at a meeting to be announced in a future FEDERAL REGISTER. Agency staff will prepare summaries of the public and SAB comments, analyses of major issues presented by commentors, and Agency responses to those comments. Appropriate comments will be incorporated, and the amended guidelines will be submitted to the Risk Assessment Forum and the Risk Assessment Council for review. The Risk Assessment Council will consider comments from the public, the SAB, and the Risk Assessment Forum in its recommendations to the EPA Administrator.

**DATE:** Public comments must be postmarked by June 5, 1989.

**ADDRESS:** Comments may be mailed or delivered to: Dr. Carole A. Kimmel, Reproductive and Developmental Toxicology Branch, Human Health Assessment Group, Office of Health and Environmental Assessment (RD–689), U.S. Environmental Protection Agency, 401 M Street SW., Washington, DC 20460.

**FOR FURTHER INFORMATION CONTACT:** Dr. Carole A. Kimmel, Telephone: 202–382–7331.

**Inspection and Copies** This notice, references, supporting documents, and other relevant materials are available for inspection and copying at the Public Information Reference Unit, (202) 382–5926, EPA Headquarters Library, 401 M Street, SW., Washington, DC, between the hours of 8:00 a.m. and 4:30 p.m.

**SUPPLEMENTARY INFORMATION:** In 1984–85, the Agency proposed risk assessment guidelines for carcinogenicity, exposure assessment, mutagenicity, developmental toxicity (49 FR 46294–46331), and chemical mixtures (50 FR 1170–1176). Following extensive scientific and public review, final guidelines were issued on September 24, 1986 (51 FR 33992–34054). Each of the guidelines set forth principles and procedures to guide EPA scientists in the conduct of Agency risk assessments, to help promote high scientific quality and Agency-wide consistency, and to inform Agency decision makers and the public about these scientific procedures.

In publishing this guidance, EPA emphasized that one purpose of its risk assessment guidelines was to "encourage research and analysis that will lead to new risk assessment methods and data," which in turn would be used to revise and improve the guidelines, and better guide Agency risk assessors. Thus, each of the 1986 risk assessment guidelines was developed and published with the understanding that risk assessment is an evolving science and that continued study could lead to changes.

As expected, Agency experience with the current Guidelines for the Health Assessment of Suspect Developmental Toxicants suggests that additional or alternate approaches should be considered for certain aspects of these guidelines. Proposals to amend the current guidelines were considered soon after their publication in September 1986 because of new reviews or re-evaluations that focused on some of the issues identified for research in the guidelines. These included several workshops and symposia cited in the Introduction to the current guidelines. In addition, much experience has been gained in using these guidelines and in instructing others in their use. Based on this experience, the proposed amendments are designed to clarify certain aspects of the current guidelines, and the terminology has been updated to be consistent with that used in other Agency guidance.

As outlined below, some of the changes involve substantive revisions to the current guidelines, while others simply clarify or reorganize current provisions. The remainder of the notice publishes the full text of the proposed guidelines, that is, the current guidelines as modified by the proposed amendments.

### Overview of Proposed Amendments

The major proposed amendments include stronger statements concerning guidance on evaluating maternal and developmental toxicity based on EPA's 1987 workshop on this topic, particularly about the inter-relationship between these end points (see Reference 3 in Section VII of the proposed guidelines). A major innovation for the proposed guidelines is a weight-of-evidence scheme for developmental toxicants (Section III.D) which was developed in a 1987 EPA workshop by experts from within and outside the Agency.

Lesser changes in the proposed guidelines include a change in the title from "Guidelines for the Health Assessment of Suspect Developmental Toxicants" to "Proposed Guidelines for Developmental Toxicity Risk Assessment." In addition, three other sections have been revised: the Human Studies section (Section III.B) was reoriented more towards risk assessment than study design, and the Dose-Response and Risk

Characterization sections (Sections IV and VI) were reorganized so that information on the NOAEL/uncertainty factor approach and low-dose extrapolation are contained in the Dose-Response section (Section IV), and the margin of exposure (MOE) approach is contained in the Risk Characterization section (Section VI).

One other proposed change is the introduction of the term $RfD_{DT}$ for the reference dose for developmental toxicity derived from dividing the NOAEL by an uncertainty factor. This is to distinguish the developmental toxicity reference dose ($RfD_{DT}$), which is based on a short-term exposure as occurs in most developmental toxicity studies, from the RfD, which the Agency derives based on a chronic or sometimes a subchronic exposure scenario. These and other proposed changes are discussed further by section.

**Section I. Introduction**

This section gives the general background information on developmental toxicity risk assessment and the magnitude of the potential for developmental toxicity problems in the general population. In the current guidelines, EPA provides the general basis for the use of data from animal studies in estimating human risk, but does not describe the assumptions generally made in this process.

The primary proposed amendment in this section is a statement of the basic assumptions made in the risk assessment process for developmental toxicity, e.g., an agent that produces an adverse developmental effect in experimental animal studies is assumed to pose a potential hazard to humans, and all four possible manifestations of developmental toxicity (i.e., death, structural abnormality, growth alteration, functional deficit) are of concern for risk assessment. The assumption of a threshold is stated, although this assumption is currently being discussed in the literature, as indicated in the proposed amendments. These assumptions help to more clearly identify the basis for the Agency's approach to risk assessment described in the proposed guidelines. In addition, some background information and references have been revised.

**Section II. Definitions and Terminology**

This section sets forth the definitions of particular terms that are widely used in the field of developmental toxicology. These include special terms such as "developmental toxicity," "altered growth," "malformations," and "variations."

The only proposed amendment in this section is the deletion of the terms "embryotoxicity" and "fetotoxicity." Because ambiguities in these terms have led to confusion and misuse, they are not used in the proposed guidelines. Thus, use of the term "developmental toxicity," which is a broader term, is encouraged and ambiguities are eliminated.

**Section III. Hazard Identification of Developmental Toxicants**

This section describes the study designs used in animal studies and the evaluation and interpretation of end points. In the current guidelines the title of this section includes the term "qualitative assessment." Also, this section recommends that other EPA risk assessment guidelines be used when carcinogenic or mutagenic effects from developmental exposures are of concern.

The proposed heading for this section no longer includes the term "qualitative assessment," since hazard identification for developmental toxicity also includes some evaluation of the dose-response nature of an effect. This change is proposed because the distinction in the current guidelines between qualitative and quantitative assessment has proved to be unsatisfactory and is not made in actual practice when using the guidelines to assess developmental toxicity data.

The discussion of potential carcinogenic effects following development exposure is proposed to be expanded somewhat, as are the statements on potential mutational events. These changes would emphasize the importance of considering potential carcinogenic and mutagenic effects resulting from developmental exposures. More extensive information on conducting risk assessments for these types of effects is provided in the Guidelines for Carcinogen Risk Assessment (51 FR 33992) or the Guidelines for Mutagenicity Risk Assessment (51 FR 34006).

*A. Laboratory Animal Studies of Developmental Toxicity: End Points and Their Interpretation*

This section provides general information on the protocols typically used to assess developmental toxicity.

There are no proposed amendments to this section.

A.1. End Points of Maternal Toxicity.

This section describes the types of maternal end points evaluated in developmental toxicity studies and provides guidance for the hazard assessment.

The proposed amendments to this section include the addition of support from adverse histopathology findings to the use of alterations in organ weights as a sign of maternal toxicity. This change would indicate more clearly the basis fof the use of maternal organ changes as signs of maternal toxicity.

A.2. End Points of Developmental Toxicity.

This section describes the types of developmental end points evaluated in developmental toxicity studies and provides guidance for the hazard assessment.

There are no proposed amendments to this section.

A.3. Functional Developmental Toxicology.

This section provides information on the state-of-the-art in the evaluation of functional effects resulting from developmental exposures. Developmental neurotoxicity is briefly reviewed, along with other areas of functional evaluation. Since the publication of the current guidelines in 1986, specific testing in this area has been proposed or required by the Agency for certain agents.

The proposed amendments to this section reflect the current regulatory status for developmental neurotoxicity testing in the Agency. The Office of Toxic Substances (OTS) recently proposed developmental neurotoxicity testing guidelines and finalized at least one test rule requiring such testing (see Reference 28 in Section VII of the proposed guidelines). In addition, the Science Advisory Panel for the Office of Pesticide Programs (OPP) has approved the development of testing guidelines for developmental neurotoxicity. The proposed amendments note these activities and identify the proposed bases for OPP and OTS requirements for such testing.

A.4. Overall Evaluation of Maternal and Developmental Toxicity.

This section discusses the relationship of maternal and developmental toxicity and the evaluation of developmental toxicity data in the presence of maternal toxicity. In the current guidelines, the statement is made that developmental effects at maternally toxic doses should not be discounted as being secondary to maternal toxicity.

A stronger statement is proposed in this section concerning the finding of developmental toxicity in the presence of maternal toxicity, i.e., when adverse developmental effects are produced only at maternally toxic doses, they are still considered to represent developmental toxicity and should not be discounted as being secondary to maternal toxicity.

Also, it is proposed that information be added on the importance of evaluating both maternal and developmental toxicity for the final characterization of risk as suggested by participants at the EPA-sponsored workshop on "The Evaluation of Maternal and Developmental Toxicity." This would indicate that maternal toxicity (even in the absence of developmental toxicity) is an important end point to evaluate in the context of all available toxicity data.

A.5. Short-Term Testing in Developmental Toxicity.

This section summarizes in vivo and in vitro approaches to short-term testing for developmental toxicity. In the current guidelines, the Chernoff/Kavlock assay is described, but more recent work, including a NIOSH-sponsored conference on this testing procedure, has appeared in the literature

The proposed amendment would update the section to include recent information on the Chernoff/Kavlock assay, in particular, that from the NIOSH-sponsored workshop on "Evaluation of the Chernoff/Kavlock Test for Developmental Toxicity."

A.6. Statistical Considerations.

This section describes approaches to the statistical evaluation of data from animal developmental toxicity studies and includes important issues of study design that affect interpretation of data.

There are no proposed amendments to this section.

*B. Human Studies*

This section describes the evaluation of human data for developmental toxic effects. In the current guidelines, this section discusses important considerations of study design and evaluation, but does not provide much guidance to the risk assessor on the relative importance of various types of human data.

The proposed amendments would reorganize and modify this section to give more specific information concerning the use of human data in risk assessment (e.g., greatest weight should be given to carefully designed epidemiologic studies with more precise measures of exposure; studies with a low probability of biased data should carry more weight in a risk assessment). These revisions would make this section consistent with similar sections in the Proposed Guidelines for Assessing Male Reproductive Risk and Female Reproductive Risk.

*C. Other Considerations*

This section discusses the importance of pharmacokinetic data and structure-activity considerations, if available, in the risk assessment of developmental toxicants.

There are no proposed amendments to this section.

**D. Weight-of-Evidence Determination**

This section describes the important considerations in determining the relative weight of various kinds of experimental and/or human evidence in estimating the risk of development toxicity in humans. In the current guidelines, various factors are listed as being important, but there is no systematic procedure for categorizing the level of confidence in the available data.

A weight-of-evidence scheme is proposed that defines three levels of confidence for data used to identify developmental hazards and to assess the risk of human developmental toxicity. The language used in the scheme is intentionally broad to allow for scientific judgment in classifying data using this scheme, and classification of agents using this scheme would require experience with developmental toxicity data. The intent of the discussion is that the scheme would not be used in isolation, but would be the first step that must be combined with information on dose-response and exposure for the final characterization of risk.

**IV. Dose-Response Assessment**

This section describes the evaluation of the dose-response data from developmental toxicity studies In the current guidelines, certain terminology (e.g., NOEL, LOEL) is used in a way that is no longer consistent with its usage in other Agency guidance In addition, certain topics (e.g., the margin of safety, now termed the margin of exposure) that are discussed as dose-response issues in the current guidelines are treated as risk characterization issues in other Agency guidance.

The proposed heading for this section no longer includes the term "quantitative assessment," since a sharp separation between qualitative and quantitative assessment in the current guidelines is not made in practice. Dose-Response Assessment is Section IV.A. in the current guidelines

The proposed amendments to this section incorporate terminology (e.g., NOAEL, LOAEL, RfD) that would make the proposed guidelines consistent with other Agency guidance The section discusses the identification of the NOAEL/LOAEL, the factors used in establishing the appropriate uncertainty factor, and the calculation of the $RfD_{DT}$. These proposed changes would also be consistent with the way in which chronic RfDs are calculated However, in the proposed guidelines, the term $RfD_{DT}$, based on short-term exposure, is introduced to distinguish it from the general RfD. An updated discussion of the status of mathematical approaches for dose-response modeling and low-dose extrapolation for developmental toxicity is also included.

**V. Exposure Assessment**

This section describes the issues of concern for developmental toxicity in the estimation of the human exposure levels. In the current guidelines, this section includes information related to human exposure-effect relationships that is actually more closely related to determining dose-effect relationship in humans.

The proposed amendments to this section, Section IV.B. in the current guidelines, include transferring some guidance from the section on determining human exposure-effect relationships to Section IV (Dose-Response Assessment) since this discussion is more involved with dose-response assessment in humans. The remaining information in this section focuses primarily on the special considerations concerning exposure assessment for developmental toxicity Another proposed change in this section would more clearly indicate that since a single exposure at the critical time in development is sufficient to produce an adverse developmental effect, the human exposure estimate used to calculate the margin of exposure is usually based on a single dose that is not adjusted for duration of exposure, and the number of exposures is not considered important unless there is evidence for a cumulative effect

**VI. Risk Characterization**

This section describes the summarization of all the toxicology and exposure data in the final stage of the risk assessment process. In the current guidelines, this section also includes a discussion of mathematical approaches to quantitative risk assessment.

The proposed amendments to the risk characterization section, Section IV C in the current guidelines, include a discussion of the Margin of Exposure approach. The discussion of dose-response models and risk extrapolation procedures has been moved to Section IV, Dose-Response Assessment in the proposed guidelines

**VII. References**

This section includes a full list of references for the proposed guidelines and is Section V in the current

guidelines. Appropriate reference changes and additions have been made to conform to the proposed amendments.

Date: February 23, 1989.

John A. Moore,

*Chairman, Risk Assessment Council*

Contents



## Proposed Guidelines for Developmental Toxicity Risk Assessment

### *I. Introduction*

### A. General

These Proposed Guidelines for Developmental Toxicity Risk Assessment (hereafter Guidelines) describe the procedures that the U.S. Environmental Protection Agency (EPA) will follow in evaluating potential developmental toxicity associated with human exposure to environmental toxicants. The Agency has sponsored or participated in several conferences that addressed issues related to such evaluations and that provided some of the scientific basis for these risk assessment Guidelines (1–6). The Agency's authority to regulate substances that have the potential to interfere adversely with human development is derived from a number of statutes that are implemented through multiple offices within the EPA. The procedures described herein are intended to promote consistency across program offices within the Agency in the assessment of developmental toxic effects.

The developmental toxicity assessments prepared pursuant to these Guidelines will be used with the requirements and constraints of the applicable statutes to arrive at regulatory decisions concerning developmental toxicity. These Guidelines provide a general format for analyzing and organizing the available data for conducting risk assessments. The Agency previously has issued testing guidelines (7, 8) that provide protocols designed to determine the potential of a test substance to induce structural and/or other adverse effects in the developing conceptus. These risk assessment Guidelines do not change any statutory or regulatory prescribed standards for the type of data necessary for regulatory action, but rather provide guidance for the interpretation of studies that follow the testing guidelines, and in addition, provide limited information for the interpretation of other studies (e.g., epidemiologic data, functional developmental toxicity studies, and short-term tests) that are not routinely required, but may be encountered when reviewing data on particular agents.

The National Research Council (9) has defined risk assessment as being comprised of some or all of the following components: hazard identification, dose-response assessment, exposure assessment, and risk characterization. In general, the process of assessing the risk of human developmental toxicity may be adapted to this format. However, the components of this format should not be considered in isolation. Instead, an appreciation of the potential for risk and the consequences of exposure can come only from consideration of the integration of all four components. Each component contributes to the final assessment of risk.

Hazard identification involves the evaluation of all available experimental animal and human data to determine if an agent is likely to cause developmental toxicity. In considering developmental toxicity, these Guidelines will address not only structural abnormalities, but also fetal and neonatal death, growth alteration, and functional abnormalities that may result from developmental exposure to environmental agents.

The dose-response assessment defines the relationship of the dose of an agent to the occurrence of developmentally toxic effects. According to the National Research Council (9), this component would usually include extrapolation from high to low doses and from experimental animals to humans. Since at present there are no mathematical extrapolation models that are generally accepted for developmental toxicity, uncertainty factors are applied to the no observed adverse effect level (NOAEL) to derive a reference dose for developmental toxicity ($RfD_{DT}$). The $RfD_{DT}$ is based on a short duration of exposure as is typically used in developmental toxicity studies in experimental animals. The use of the term $RfD_{DT}$ distinguishes it from the reference dose (RfD) which refers to chronic exposure situations (10). This approach is discussed further in these Guidelines (Section IV). Potential mathematical models are being evaluated by the Agency for application to data in this area (5).

The exposure assessment identifies populations exposed to an agent, describes their composition and size, and presents the types, magnitudes, frequencies, and durations of exposure to the agent. The exposure assessment provides an estimate of human exposure levels from all potential sources.

In risk characterization, the exposure assessment and the hazard identification and dose-response assessment are combined to estimate some measure of the risk of developmental toxicity. Here the NOAEL and the estimated human exposure levels may be compared to provide a margin of exposure (MOE). As part of risk characterization, a summary of the strengths and weaknesses in each component of the risk assessment are presented along with major assumptions, scientific judgments, and, to the extent possible, qualitative and quantitative estimates of the uncertainties. The weight-of-evidence determination should always be presented in conjunction with information on dose-response and, if available, the human exposure estimate.

Risk assessment is just one component of the regulatory process and defines the adverse health consequences of exposure to a toxic agent. The other component, risk management, combines the risk assessment with the directives of the enabling regulatory legislation, together with socioeconomic, technical, political, and other considerations, to reach a decision as to whether to control future exposure to the suspected toxic agent and, if so, the level of control. The acceptability of the uncertainty factor or the margin of exposure and risk management decisions, but the scientific bases for establishing these values are discussed here.

B. Background

The background incidence of developmental defects in the human population is quite large For example, Hertig (11) estimated that approximately 50% of human conceptuses fail to reach term; Wilcox (12), using biochemical techniques for detecting pregnancy as early as 9 days postconception, observed that 35% of pregnancies ended in an embryonic or fetal loss. Approximately 3% of newborn children are found to have one or more significant congenital malformations at birth, and by the end of the first postnatal year, about 3% more are found to have serious developmental defects (13). Of these, it is estimated that 20% are of known genetic transmission, 10% are attributable to known environmental factors, and the remainder result from unknown causes (14). Also, approximately 7.4% of children are reduced in weight at birth (i.e., below 2500 g) (15).

Close to one-half of the children in hospital wards are there because of prenatally acquired malformations (16). The Centers for Disease Control recently evaluated the enormity of the problem of developmental disabilities in the United States. Among all races, congenital anomalies, sudden infant death syndrome, and prematurity combined account for more than 50% of infant mortality in the United States (17). In addition, among the leading causes of estimated years of potential life lost (YPLL) before the age of 65, congenital anomalies ranks fifth, prematurity ranks sixth, and sudden infant death syndrome ranks seventh (18) The YPLL estimates may actually underestimate the public health impact of congenital anomalies because statistics on the following may not be represented (19): (1) Anomalies in infants who die shortly after birth may not be diagnosed and death may not be attributed to congenital anomalies; (2) YPLL estimates are based only on live births and therefore do not take into account the number of fetuses with anomalies that were spontaneously aborted or infants that were stillborn; (3) with prenatal diagnoses of chromosomal abnormalities and neural tube defects, pregnancies may be terminated and thus these statistics are not represented in the YPLL estimates.

Exposure to agents affecting development can result in any one or more of four possible manifestations (death, structural abnormality, growth alteration, and/or functional deficit) Therefore, assessment efforts should encompass a wide array of adverse developmental end points, such as spontaneous abortions, stillbirths, malformations, early postnatal mortality, reduced birth weight, and other adverse functional or physical changes that are manifested postnatally

Numerous agents have been shown to be developmental toxicants in animal test systems (18). Several of them have also been shown to be the cause of adverse developmental effects in humans, including alcohol, aminopterin, busulfan, chlorobiphenyls, diethylstilbestrol, isotretinoin, lead, organic mercury, thalidomide, and valproic acid (13, 20, 21). Although a number of agents found to be developmental toxicants in experimental animal studies have not shown clear evidence of hazard in humans, the available human data are inadequate to determine a cause and effect relationship. Comparisons of human and experimental animal data have been made for a limited number of agents that are human developmental toxicants (22–24). In these comparisons, there was almost always qualitative concordance of effects between humans and at least one species tested, also, the minimally effective dose (MED) for the most sensitive animal species was approximately 0.5 to 100 times the human MED, not accounting for differences in the incidence of effect at the MED. Thus, there is some basis for estimating the risk of exposure to human development based on data from animal studies.

However, there are a number of unknowns in the extrapolation of data from animal studies to humans. Therefore, a number of assumptions must be made which are generally applied. These assumptions are the bases for the approaches taken to risk assessment in these Guidelines.

First, an agent that produces an adverse developmental effect in experimental animal studies is assumed to pose a potential hazard to humans following exposure during development. This assumption is based on the comparisons of data for known human developmental toxicants (22–24) In almost all cases, the experimental animal data would have predicted a developmental effect in humans.

It is assumed that all of the four manifestations of developmental toxicity (death, structural abnormalities, growth alterations, and functional deficits) are of concern. In the past, there has been a tendency to consider only malformations or malformations and death as end points of concern. From the data on agents that are known human developmental toxicants (22–24), there is usually at least one experimental species that mimics the types of effects seen in humans, but in other species tested, the type of developmental perturbation may be different Thus, the appearance of any of the four manifestations is considered indicative of an agent's potential for disrupting development and producing a developmental hazard.

It is assumed that the types of developmental effects seen in animal studies are not necessarily the same as those that may be produced in humans This assumption is made because it is impossible to determine which will be the most appropriate species in terms of predicting the specific types of effects seen in humans. The fact that every species may not react in the same way is probably due to species—specific differences in critical periods, metabolism, developmental patterns, or mechanisms of action.

It is assumed that the most sensitive species should be used to estimate human risk When data are available (e.g., pharmacokinetic, metabolic) to suggest the most appropriate species, that species will be used for extrapolation. In the absence of such data, the most sensitive species is used, based on the fact that for the majority of known human developmental toxicants, humans are as sensitive or more so than the most sensitive animal species (22–24).

In general, a threshold is assumed for the dose-response curve for most developmental toxicants. This is based on the known capacity of the developing organism to compensate for or to repair a certain amount of damage at the cellular, tissue, or organ level. In addition, because of the multipotency of cells at certain stages of development, multiple insults at the molecular or cellular level may be required to produce an effect on the whole organism. There are uncertainties concerning this assumption that are being discussed currently in the literature (25, 26).

*II. Definitions and Terminology*

The Agency recognizes that there are differences in the use of terms in the field of developmental toxicology. For the purposes of these Guidelines the following definitions and terminology will be used.

Developmental toxicology. The study of adverse effects on the developing organism that may result from exposure prior to conception (either parent), during prenatal development, or postnatally to the time of sexual maturation Adverse developmental effects may be detected at any point in

the life span of the organism. The major manifestations of developmental toxicity include: (1) Death of the developing organism, (2) structural abnormality, (3) altered growth, and (4) functional deficiency.

Altered Growth. An alteration in offspring organ or body weight or size. Changes in one end point may or may not be accompanied by other signs of altered growth (e.g., changes in body weight may or may not be accompanied by changes in crown-rump length and/or skeletal ossification). Altered growth can be induced at any stage of development, may be reversible, or may result in a permanent change.

Functional Developmental Toxicology The study of alterations or delays in functional competence of the organism or organ system following exposure to an agent during critical periods of development pre- and/or postnatally.

Malformations and Variations. A malformation is usually defined as a permanent structural change that may adversely affect survival, development, or function. The term teratogenicity, which is used to describe these types of structural abnormalities, will be used in these Guidelines to refer only to structural defects. A variation is used to indicate a divergence beyond the usual range of structural constitution that may not adversely affect survival or health. Distinguishing between variations and malformations is difficult since there exists a continuum of responses from the normal to the extreme deviant. There is no generally accepted classification of malformations and variations. Other terminology that is often used, but no better defined, includes anomalies, deformations, and aberrations.

### *III Hazard Identification of Developmental Toxicants*

Developmental toxicity is expressed as one or more of a number of possible end points that may be used for evaluating the potential of an agent to cause abnormal development. The four types of effects on the conceptus that may be produced by developmental exposure to toxicants include death, structural abnormality, altered growth, and functional deficits. Of these, all four types of effects have been evaluated in human studies, but only the first three are traditionally measured in laboratory animals using the conventional developmental toxicity (also called teratogenicity or Segment II) testing protocol as well as in other study protocols, such as the multigeneration study. Although functional deficits have been shown to occur subsequent to developmental exposures in humans, such effects seldom have been evaluated in routine testing studies in experimental animals. However, functional evaluations are beginning to be examined under certain regulatory situations (27, 28).

Carcinogenic effects of developmental exposures have occurred in humans resulting from the use of diethylstilbestrol for the maintenance of pregnancy (29). Several agents have been shown to cause cancer following developmental exposures in experimental animals, and it appears from the data collected thus far that agents which are capable of causing cancer in adults may also cause transplacental or neonatal carcinogenesis (30). There is no way to predict whether adults or developing animals will be more sensitive to the carcinogenic effects of an agent. At present, testing for carcinogenesis following developmental exposure is not routinely required. However, if this type of effect is reported for an agent, it is considered appropriate to use the Guidelines for Carcinogen Risk Assessment (31) for assessing human risk. Mutational events also may occur as a result of exposure to developmental toxicants but may be difficult to discriminate from other possible mechanisms in standard studies of developmental toxicity. When mutational events are suspected from further experiments, the Guidelines for Mutagenicity Risk Assessment (32) should be consulted; however, these guidelines specifically address heritable and not somatic mutational risk.

### A. Laboratory Animal Studies of Developmental Toxicity: End Points and Their Interpretation

This section will discuss the end points examined in routinely-used protocols as well as the use of other types of studies, including functional studies and short-term tests.

The most commonly used protocol for assessing developmental toxicity in laboratory animals involves the administration of a test substance to pregnant animals (usually mice, rats, or rabbits) during the period of major organogenesis, evaluation of maternal responses throughout pregnancy, and examination of the dam and the uterine contents just prior to term (7, 8, 33–35). Other protocols may use exposures of one to a few days to investigate periods of particular sensitivity for induction of anomalies in specific organs or organ systems (36). In addition, developmental toxicity may be evaluated in studies involving exposure of one or both parents prior to conception, of the conceptus during pregnancy and over several generations, or of offspring during the late prenatal and early postnatal periods (7, 8, 27, 28, 33–35, 37). These Guidelines are intended to provide information for interpreting developmental effects related to any of these types of exposure. Since many of the end points evaluated also are related to effects on the parental reproductive systems, these Guidelines should be used in conjunction with those published on assessing male and female reproductive risk (38, 39).

Study designs should include, at a minimum, a high dose, a low dose, and one intermediate dose. The high dose should produce some maternal or adult toxicity (i.e., a level which at the least produces marginal but significantly reduced body weight, weight gain, or specific organ toxicity, and at the most produces no more than 10% mortality). The low dose should demonstrate a NOAEL for adult and offspring effects. A concurrent control group treated with the vehicle used for agent administration should be included. The route of exposure is usually oral, although data from other routes may sometimes be useful, especially if supported by pharmacokinetic information. Test animals should be selected based on considerations of species, strain, age, weight, and health status, and should be randomized to dose groups in order to reduce bias and provide a basis for performing valid statistical tests.

The next three sections discuss individual end points of maternal and developmental toxicity as measured in the conventional developmental toxicity study and the multigeneration study, and, on occasion, in postnatal studies. Other end points specifically related to reproductive toxicity are covered in the relevant risk assessment guidelines (38, 39). The fourth section deals with the integrated evaluation of all data, including the relative effects of exposure on maternal animals and their offspring, which is important in assessing the level of concern about a particular agent. It should be noted that appropriate historical control data can be helpful in the interpretation of end points of maternal and developmental toxicity.

1. *End Points of Maternal Toxicity.* A number of end points that may be observed as possible indicators of maternal toxicity are listed in Table 1. Maternal mortality is an obvious end point of toxicity; however, a number of other end points can be observed which may give an indication of the subtle effects of an agent. For example, in well-conducted studies, the fertility and gestation indices provide information on

the general fertility rate of the animal stock used and are important indicators of toxic effects to adults if treatment begins prior to mating or implantation. Changes in gestation length may indicate effects on the process of parturition.

Table 1. End Points of Maternal Toxicity

| |
|---|
| Mortality |
| Fertility Index (no. with seminal plugs or sperm/no. mated) |
| Gestation Index (no. with implants/no. with seminal plugs or sperm) |
| Gestation Length (when allowed to deliver pups) |
| Body Weight |
| Day 0 |
| During gestation |
| Sacrifice day |
| Body Weight Change |
| Throughout gestation |
| During treatment (including increments of time within treatment period) |
| Post-treatment of sacrifice |
| Corrected maternal (body weight change throughout gestation minus gravid uterine weight or litter weight at sacrifice) |
| Organ Weights (in cases of suspected specific organ toxicity and when supported by adverse histopathology findings) |
| Absolute |
| Relative to body weight |
| Food and Water Consumption (where relevant) |
| Clinical Evaluations |
| Types, incidence and duration of clinical signs |
| Enzyme markers |
| Clinical chemistries |
| Gross Necropsy and Histopathology |

Body weight and the change in body weight are viewed collectively as indicators of maternal toxicity for most species, although these end points may not be as useful in rabbits, because body weight changes in some strains of rabbits are not good indicators of pregnancy status. Body weight changes may provide more information than a daily body weight measured during treatment or during gestation. Changes in weight gain during treatment could occur that would not be reflected in the total weight change throughout gestation, because of compensatory weight gain that may occur following treatment but before sacrifice. For this reason, changes in weight gain during treatment can be examined as another indicator or maternal toxicity.

Changes in maternal body weight corrected for gravid uterine weight at sacrifice may indicate whether the effect is primarily maternal or fetal. For example, there may be a significant reduction in weight gain throughout gestation and in gravid uterine weight, but no change in corrected maternal weight gain which would generally indicate an intrauterine effect. Conversely, a change in corrected weight gain and no change in gravid uterine weight generally suggests maternal toxicity and little or not intrauterine effect. An alternate estimate of maternal weight change during gestation can be obtained by subtracting the sum of the weights of the fetuses. However, this weight does not include the uterine tissue, placental tissue, or the amniotic fluid.

Changes in other end points may also be important. For example, changes in relative and absolute organ weights may be signs of a maternal effect when an agent is suspected or causing specific organ toxicity and when such findings are supported by adverse histopathologic findings in those organs. Food and water consumption data are useful, especially if the agent is administered in the diet or drinking water. The amount ingested (total and relative to body weight) and the dose of the agent (relative to body weight) can then be calculated, and changes in food and water consumption related to treatment can be evaluated along with changes in body weight and body weight gain. Data on food and water consumption are also useful when an agent is suspected of affecting appetite, water intake, or excretory function. Clinical evaluations of toxicity may also be used as indicators of maternal toxicity. Daily clinical observations may be useful describing the profile or maternal toxicity. Enzyme markers and clinical chemistries may be useful indicators of exposure but must be interpreted carefully as to whether or not a change constitutes toxicity. Gross necropsy and histopathology data (when specified in the protocol) may aid in determining toxic dose levels. The minimum amount of information/data considered useful for evaluating maternal toxicity [as noted in the Proceedings of the Workshop on the Evaluation of Maternal and Developmental Toxicity (3)], includes: morbidity or mortality; maternal body weight and body weight gain; clinical signs of toxicity; food (and water, if dosing is via drinking water) consumption, and necropsy for gross evidence of organ toxicity. Maternal toxicity should be determined in the pregnant and/or lactating animal over an appropriate part of gestation and/or the neonatal period, and should not be assumed or extrapolated from other adult toxicity studies.

2. *End Points of Developmental Toxicity.* Because the maternal animal, and not the conceptus, is the individual treated during gestation, data generally should be calculated as incidence per litter or as number and percent of litters with particular end points. Table 2 indicates the way in which offspring and litter end points may be expressed.

Table 2. End Points of Developmental Toxicity

| |
|---|
| *Litters with implants* |
| No. implantation sites/dam |
| No. corpora lutea (CL)/dam [a] |
| Percent preimplantation loss (CL−implantations) × 100 [a]/CL |
| No. and percent live offspring [b]/litter |
| No. and percent resorptions/litter |
| No. and percent litters with resorptions |
| No. and percent late fetal deaths/litter |
| No. and percent nonlive (late fetal deaths + resorptions) implants/litter |
| No. and percent liters with nonlive implants |
| No. and percent affected (nonlive + malformed) implants/litter |
| No. and percent litters with affected implants |
| No. and percent litters with total resorptions |
| No. and percent stillbirths/litter |
| *Litters with live offspring* |
| No. and percent litters with live offspring |
| No. and percent live offspring/litter |
| Viability of offspring [c] |
| Sex ratio/litter |
| Mean offspring body weight/litter [c] |
| Mean male body weight/litter [c] |
| Mean female body weight/litter [c] |
| No. and percent externally malformed offspring/litter |
| No. and percent viscerally malformed offspring/litter |
| No. and percent skeletally malformed offspring/litter |
| No. and percent malformed offspring/litter |
| No. and percent litters with malformed offspring |
| No. and percent malformed males/litter |
| No. and percent malformed females/litter |
| No. and percent offspring with variations/litter |
| No. and percent litters having offspring with variations |
| Types and incidence of individual malformations |
| Types and incidence of individual variations |

Individual offspring and their malformations and variations (grouped according to litter and dose)
Clinical signs
Gross necropsy and histopathology

[a] Important when treatment begins prior to implantation. May be difficult to assess in mice

[b] Offspring refers both to fetuses observed prior to term or to pups following birth. The end points examined depend on the protocol used for each study

[c] Measured at selected intervals until termination of the study

When treatment begins prior to implantation, an increase in preimplantation loss could indicate an adverse effect on the fertilization process, ovum transport, uterine toxicity, the developing blastocyst, or on the process of implantation itself. If treatment begins around the time of implantation (i.e., day 6 of gestation in the mouse, rat, or rabbit), an increase in preimplantation loss probably reflects normal variability in the animals being used, but the data should be examined carefully to determine whether or not the effect is dose related. If preimplantation loss is related to dose in either case, further studies would be necessary to determine the mechanism and extent of such effects.

The number and percent of live offspring per litter, based on all litters, may include litters that have no live implants. Resorptions and late fetal deaths give some indication of when the conceptus died, and the number and percent nonlive implants per litter (post-implantation loss) is a combination of resorptions and late fetal deaths. The number and percent of litters showing an increased incidence for these end points is generally useful but may be less useful than incidence per litter because, in the former case, a litter is counted whether it has one or all resorbed, dead, or nonlive implants.

If a significant increase in postimplantation loss is found after exposure to an agent, the data may be compared not only with concurrent controls, but also with recent historical control data, since there is considerable interlitter variability in the incidence of post-implantation loss (40). If a given study control group exhibits an unusually high or low incidence of postimplantation loss compared to historical controls, then scientific judgment must be used to determine the adequacy of the study for risk assessment purposes.

The end point for affected implants (i.e., the combination of nonlive and malformed conceptuses) gives an indication of the total intrauterine response to an agent and sometimes reflects a better dose-response relationship than does the incidence of nonlive or malformed offspring taken individually. This is especially true at the high end of the dose-response curve in cases when the incidence of nonlive implants per litter is greatly increased. In such cases, the malformation rate may appear to decrease because only unaffected offspring have survived. If the incidence of prenatal death or malformation is unchanged, then the incidence of affected implants will not provide any additional dose-response information. In studies where maternal animals are allowed to deliver pups normally, the number of stillbirths per litter should also be noted.

The number of live offspring per litter, based on those litters that have one or more live offspring, may be unchanged even though the incidence of nonlive in all litters is increased. This could occur either because of an increase in the number of litters with no live offspring, or an increase in the number of implants per litter. A decrease in the number of live offspring per litter should be accompanied by an increase in the incidence of nonlive implants per litter unless the implant numbers differ among dose groups. In postnatal studies, the viability of live born offspring should be determined at selected intervals until termination of the study.

The sex ratio per litter, as well as the body weights of males and females, can be examined to determine whether or not one sex is preferentially affected by the agent. However, this is an annual occurrence.

A change in offspring body weight is a sensitive indicator of developmental toxicity, in part because it is a continuous variable. In some cases, offspring weight reduction may be the only indicator of developmental toxicity. While there is always a question remaining as to whether weight reduction is a permanent or transitory effect, little is known about the long-term consequences of short-term fetal or neonatal weight changes. Therefore, weight reduction should be used to establish the NOAEL. There are other factors that should be considered in the evaluation of fetal or neonatal weight changes. For example, in polytocous animals, fetal and neonatal weights are usually inversely correlated with litter size, and the upper end of the dose-response curve may be confounded by smaller litters and increased fetal or neonatal weight. Additionally, the average body weight of males is greater than that of females in the more commonly used laboratory animals.

Live offspring should be examined for external, visceral, and skeletal malformations. If only a portion of the litter is examined, then it is preferable that those examined be randomly selected from each litter. An increase in the incidence of malformed offspring may be indicated by a change in one or more of the following end points: the incidence of malformed offspring per litter, the number and percent of litters with malformed offspring, or the number of offspring or litters with a particular malformation that appears to increase with dose (as indicated by the incidence of individual types of malformations).

Other ways of examining the data include the incidence of external, visceral, and skeletal malformations which may indicate the general systems affected. A listing of individual offspring with their malformations and variations may give an indication of the pattern of developmental deviations. All of these methods of expressing and examining the data are valid for determining the effects of an agent on structural development. However, care must be taken to avoid counting offspring more than once in evaluating any single end point based on number or percent of offspring or litters. The incidence of individual types of malformations and variations should be examined for significant changes which may be masked if the data on all malformations and variations are pooled. Appropriate historical control data are helpful in the interpretation of malformations and variations, especially those that normally occur at a low incidence and may or may not be related to dose in an individual study. Although a dose-related increase in malformations is interpreted as an adverse developmental effect of exposure to an agent, the significance of anatomical variations is more difficult to determine, and must take into account what is known about developmental stage (e.g., with skeletal ossification) background incidence of certain variations (e.g., 12 or 13 pairs of ribs in rabbits), or other strain- or species-specific factors. However, if variations are significantly increased in a dose-related manner, these should also be evaluated as a possible indication of developmental toxicity. The Interagency Regulatory Liaison Group noted that dose-related increases in defects that may occur spontaneously are as relevant as dose-related increases in any other developmental toxicity end points (41).

3. *Functional Developmental Toxicology.* Developmental effects that

are induced by exogenous agents are not limited to death, structural abnormalities, and altered growth. Rather, it has been demonstrated in a number of instances that subtle alterations in the functional competence of an organ or a variety of organ systems may result from exposure during critical developmental periods that may occur between conception and sexual maturation. Often, these functional defects are observed at dose levels below those at which gross malformations are evident (42). Such testing has not been routinely required in the United States, but studies are beginning to be required when other information indicates the potential for adverse functional effects (27, 28). Data from postnatal studies, when available, are considered very useful for the assessment of the relative importance and severity of findings in the fetus and neonate. Often, the long-term consequences of adverse developmental outcomes noted at birth are unknown, and further data on postnatal development and function are needed to determine the full spectrum of potential developmental effects. In some cases, useful data can be derived from well-executed multigeneration studies.

Much of the early work in functional developmental toxicology was related to behavioral evaluations, and the term "behavioral teratology" became prominent in the mid 1970s. Recent advances in this area have been reviewed in several publications (43, 44). Several expert groups have focused on the functions that should be included in a behavioral testing battery (45–47), and these include: sensory systems, neuromotor development, locomotor activity, learning and memory, reactivity and/or habituation, and reproductive behavior. No testing battery has adequately addressed all of these functions, but it is important to include as many as possible. Several testing batteries have been developed and evaluated (46, 48, 49). The U.S. EPA Office of Toxic Substances (OTS) has developed a guideline for developmental neurotoxicity testing (28) that includes some evaluation of all the categories listed above except for reproductive behavior, and also includes requirements for brain weights and neuropathology. Several criteria for selecting agents for developmental neurotoxicity testing have been suggested (46), including: agents that cause central nervous system malformations, pyschoactive drugs and chemicals, adult neurotoxicants, hormonally-active agents, and chemicals that are structurally related to other developmental neurotoxicants. Data from developmental neurotoxicity studies should be evaluated in light of the data that may have triggered such testing as well as all other toxicity data available.

Less work has been done on other developing functional systems, but data have accumulated to indicate that the cardiopulmonary, immune, endocrine, digestive, and urinary systems, as well as the central nervous system are subject to alterations in functional competence (50, 51) following exposure during development. Currently, there are no standard testing procedures for these functional systems. However, when data are encountered on a chemical under review, they are considered and evaluated in the risk assessment process.

Extrapolation of functional developmental effects to humans is limited by the lack of knowledge about underlying toxicological mechanisms and their significance as is true for other end points of developmental toxicity. In comparisons made on a limited number of agents known to cause developmental neurotoxic effects in humans (52), these agents also have been shown to produce developmental neurotoxic effects in animal species. As for other end points of developmental toxicity, the assumption is made that functional effects in animal studies indicate the potential for altered development in humans. When data from functional developmental toxicity studies are encountered for particular agents, they should be evaluated and included in the risk assessment process.

Some guidance is provided here concerning important general concepts of study design and evaluation for functional developmental toxicity studies.

- Several aspects of study design are similar to those important in standard developmental toxicity studies (e.g., a dose-response approach with the highest dose producing minimal overt maternal or perinatal toxicity, number of litters large enough for adequate statistical power, randomization of animals to dose groups and test groups, litter generally considered the statistical unit, etc.).
- A replicate study design provides added confidence in the interpretation of data.
- Use of a pharmacological challenge may be valuable in evaluating function and "unmasking" effects not otherwise detectable, particularly in the case of organ systems that are endowed with a reasonable degree of functional reserve capacity.
- Use of functional tests with a moderate degree of background variability may be more sensitive to the effects of an agent than are tests with low variability that may be impossible to disrupt without being life-threatening. Butcher et al. (53) discussed this with relation to behavioral end points.
- A battery of functional tests, in contrast to a single test, usually provides a more thorough evaluation of the functioanl competence of an animal: tests conducted at several ages may provide more information about maturational changes and their persistence.
- Critical periods for the disruption of functional competence include both the prenatal and the postnatal periods to the time of sexual maturation, and the effect is likely to vary depending on the time and degree of exposure.

Although interpretation of functional data may be limited at present, it is clear that functional effects must be evaluated in light of other toxicity data, including other forms of developmental toxicity (e.g., structural abnormalities, perinatal death, and growth retardation). The level of confidence in an adverse effect may be more important than the type of change seen, and confidence may be increased by such factors as replicability of the effect either in another study of the same function or by convergence of data from tests that purport to measure similar functions. A dose-response relationship is considered an important measure of chemical effect; in the case of functional effects, both monotonic and biphasic dose-response curves are likely, and both may be appropriate depending on the function being tested. Finally, there are at least three general ways in which the data from these studies may be useful for risk assessment purposes: (1) To help elucidate the long-term consequences of fetal and neonatal findings; (2) to indicate the potential for an agent to cause functional alterations and the effective doses relative to those that produce other forms of toxicity; and (3) for existing environmental agents, to suggest organ systems to be evaluated in exposed human populations.

4. *Overall Evaluation of Maternal and Developmental Toxicity.*

As discussed previously, individual end points of maternal and developmental toxicity are evaluated in developmental toxicity studies. In order to interpret the data fully, an integrated evaluation must be performed considering all maternal and developmental end points.

Those agents that produce developmental toxicity at a dose that is

not toxic to the maternal animal are of greatest concern because the developing organism appears to be more sensitive than the adult. However. when adverse developmental effects are produced only at minimal maternally toxic doses. they are still considered to represent developmental toxicity and should not be discounted as being secondary to maternal toxicity Current information is inadequate to assume that developmental effects at maternally toxic doses result only from maternal toxicity: rather. when the lowest observed adverse effect level (LOAEL) is the same for the adult and developing organisms. it may simply indicate that both are sensitive to that dose level. Moreover. the maternal effects may be reversible while effects on the offspring may be permanent. These are important considerations for agents to which humans may be exposed at minimally toxic levels either voluntarily or in the workplace. since several agents are known to produce adverse developmental effects at minimally toxic doses in adult humans (e.g., smoking. alcohol).

Since the final risk assessment not only takes into account the potential hazard of an agent. but also the nature of the dose-response relationship. it is important that the relationship of maternal and developmental toxicity be evaluated and described Then. information from the exposure assessment is used to determine the likelihood of exposure to levels near the maternally toxic dose for each agent and the risk for developmental toxicity in humans.

If. on the other hand. maternal toxicity is seen in the absence of or at dose levels lower than those producing developmental toxicity. and if the effect level is lower than that in evaluations of other types of adult toxicity. this implies that the pregnant female is likely to be more sensitive than the nonpregnant female and the data from the pregnant female should be used to assess risk. Although the evaluation of developmental toxicity is the primary objective of standard studies within this area. maternal effects seen within the context of developmental toxicity studies should be evaluated as part of the overall toxicity profile for a given chemical.

Approaches for ranking agents according to their relative maternal and developmental toxicity have been proposed: Schardein (20) has reviewed several of these. Several approaches involve the calculation of ratios relating an adult toxic dose to a developmentally toxic dose (54–57) Such ratios may describe in a qualitative and roughly quantitative fashion the relationship of maternal (adult) and developmental toxicity However. at the U S. EPA Sponsored Workshop on the Evaluation of Maternal and Developmental Toxicity (3). there was no agreement as to the validity or utility of these approaches in other aspects of the risk assessment process. This is in part due to uncertainty about factors that can affect the ratios. For example. the number and spacing of dose levels. differences in study design (e.g., route an/or timing of exposure). and species differences in response (3. 58). can influence the maternal and developmental effects and the resulting ratios. Also. the end points used in the ratios need to be better defined to permit cross-species comparison. Until such information is available. the applicability of these approaches in risk assessment is not justified.

5. *Short-term Testing in Developmental Toxicity.* The need for short-term tests for developmental toxicity has arisen from the need to establish testing priorities for the large number of agents in or entering the environment. the interest in reducing the number of animals used for routine testing. and the expense of testing. Two approaches are considered here in terms of their contribution to the overall testing process: 1) an in vivo mammalian screen. and 2) a variety of in vitro systems. Currently. neither approach is considered as a replacement for routine in vivo development toxicity testing in experimental animals. and should not be used to make the final decision as to an agent's developmental toxicity. Rather. such tests may be useful in making preliminary evaluations of developmental toxicity. for evaluating structure-activity relationships. and for assigning priorities for further. more extensive testing. Although such short-term tests are not routinely required. data sometimes are encountered in the review of chemicals: the comments are provided here for guidance in the evaluation of such data.

*a. In vivo mamalian developmental toxicity screen.* The most widely studied in vivo short-term approach is that developed by Chernoff and Kavlock (59). This approach is based on the hypothesis that a prenatal injury. which results in altered development. will be manifested postnatally as reduced viability and/or impaired growth. When originally proposed. the test substance was administered to mice over the period of major organogenesis at a single dose level that would elicit some degree of maternal toxicity At the NIOSH Workshop on the Evaluation of the Chernoff/Kavlock Test for Developmental Toxicity (4). use of a second lower dose level was encouraged to potentially reduce the chances of false positive results. and the recording of implantation sites was recommended to provide a more precise estimate of postimplantation loss (60).

In this approach. the pups are counted and weighed shortly after birth. and again after 3–4 days. End points that are considered in the evaluation include. general maternal toxicity (including survival and weight gain). litter size. and viability. weight. and gross malformations in the offspring. Basic priority-setting categories for more extensive testing have been suggested. 1) agents that induce perinatal death should receive highest priority. 2) agents that induce perinatal weight changes should be ranked lower in priority. and 3) agents that induce no effect should receive the lowest priority (59). Another scheme that has been proposed applies a numerical ranking to the results as a means of prioritizing agents for further testing (61. 62).

The mouse was chosen originally for this test because of its low cost. but the procedure has been applied to the rat as well (63). The test will predict the potential for developmental toxicity of an agent in the species used while extrapolation of risk to other species. including humans. has the same limitations as for other testing protocols. The EPA Office of Toxic Substances has developed testing guidelines for this procedure (64). Although the testing guidelines are available. such procedures are required on a case-by-case basis. Application of this procedure in the risk assessment process within the Office of Toxic Substances has been described (65). and the experiences of a number of laboratories are detailed in the proceedings of the NIOSH workshop (1).

*b. In vitro developmental toxicity screens.* Test systems that fall under the general heading of "in vitro" developmental toxicity screens include any system that employs a test subject other than the intact pregnant mammal. Examples of such systems include: Isolated whole mammalian embryos in culture. tissue/organ culture. cell culture. and developing nonmammalian organisms. These systems have long been used to assess events associated with normal and abnormal development. but only recently have they been considered for this potential as screens in testing (66–68) Many of these systems are now being evaluated for their ability to predict the developmental toxicity of

various agents in intact mammalian systems. This validation process requires certain considerations in study design, including defined end points for toxicity and an understanding of the system's ability to handle various test agents (67, 69–71).

6. *Statistical Considerations.* In the assessment of developmental toxicity data, statistical considerations require special attention. Since the litter is generally considered the experimental unit in most developmental toxicity studies, the statistical analyses should be designed to analyze the relevant data based on incidence per litter or on the number of litters with a particular end point. The analytical procedures used and the results, as well as an indication of the variance in each end point, should be clearly indicated in the presentation of data. Analysis of variance (ANOVA) techniques, with litter nested within dose in the model, take the litter variable into account while allowing use of individual offspring data and an evaluation of both within and between litter variance as well as dose effects. Nonparametric and categorical procedures have also been widely used for binomial or incidence data. In addition, tests for dose-response trends can be applied. Although a single statistical approach has not been agreed upon, a number of factors important in the analysis of developmental toxicity data have been discussed (41, 72).

Studies that employ a replicate experimental design (e.g., two or three replicates with 10 litters per dose per replicate rather than a single experiment with 20 to 30 litters per dose group) allow for broader interpretation of study results since the variability between replicates can be accounted for using ANOVA techniques. Replication of effects due to a given agent within a study, as well as among studies or laboratories, provides added strength in the use of data for the estimation of risk.

An important factor to determine in evaluating data is the power of a study (i.e., the probability that a study will demonstrate a true effect), which is limited by the sample size used in the study, the background incidence of the end point observed, the variability in the incidence of the end point, and the analysis method. As an example, Nelson and Holson (73) have shown that the number of litters needed to detect a 5% or 10% change was dramatically lower for fetal weight (a continuous variable with low variability) than for resorptions (a binomial response with high variability). With the current recommendation in testing protocols being 20 rodents per dose group (7, 8), it is possible to detect an increased incidence of malformations in the range of 5 to 12 times above control levels, an increase of 3 to 6 times the in utero death rate, and a decrease of 0.15 to 0.25 times the fetal weight. Thus, even within the same study, the ability to detect a change in fetal weight is much greater than for the other end points measured. Consequently, for statistical reasons only, changes in fetal weight are often observable at doses below those producing other signs of developmental toxicity. Any risk assessment should present the detection sensitivity for the study design used and for the end point(s) evaluated.

Although statistical analyses are important in determining the effects of a particular agent, the biological significance of data should not be overlooked. For example, with the number of end points that can be observed in developmental toxicity studies, a few statistically significant differences may occur by chance. On the other hand, apparent trends with dose may be biologically relevant even though statistical analyses do not indicate a significant effect. This may be true especially for the incidence of malformations or in utero death where a relatively large difference is required to be statistically significant. It should be apparent from this discussion that a great deal of scientific judgment, based on experience with developmental toxicity data and with principles of experimental design and statistical analysis, may be required to adequately evaluate such data.

B. Human Studies

The category of "human studies" includes both epidemiologic studies and other reports of individual cases or clusters of events. Reports of individual cases or clusters of events may generate hypotheses of exposure-outcome associations, but require further confirmation with well-designed epidemiologic or laboratory studies. These reports of cases or clusters may give added support to associations suggested by other human or animal data, but cannot stand by themselves in risk assessments. Greatest weight should be given to carefully-designed epidemiologic studies with more precise measures of exposure, since they can best evaluate exposure-response relationships (see section IV). Epidemiologic studies in which exposure is presumed based on occupational title or residence (e.g., some case-control and all ecologic studies) may contribute data to qualitative risk assessments, but are of limited use for quantitative risk assessments because of the generally broad categorical groupings. Risk assessors should seek the assistance of professionals trained in epidemiology when conducting a detailed analysis.

1. *Examination of Clusters, Case Reports, or Case Series.* The identification of cases or clusters of adverse developmental effects is generally limited to those identified by the women involved, or clinically by their physicians. Examples of outcomes more easily identified include fetal loss in mid to late pregnancy or congenital malformations. Identification of other effects, such as embryonic loss may be difficult to separate from subfertility/infertility. Identification of such "non-events" (e.g., lack of pregnancies or children) are much harder to recognize than are developmental effects such as malformations resulting from in utero exposure. While case reports may have importance in the recognition of developmental toxicants, they may be of greatest use in suggesting topics for further investigation (74).

2. *Epidemiologic Studies.* Good epidemiologic studies provide the most relevant information for assessing human risk. As there are many different designs for epidemiologic studies, simple rules for their evalution do not exist. The following is a discussion of factors that affect the relative weight assigned a particular study in a risk assessment.

a. General design considerations. Factors that affect a study's usefulness for risk assessment include the power of the study, potential bias in data collection, control of potential risk factors, effect modifiers and confounders, and statistical factors (41, 75–80):

(1) The power of the study: The power, or ability of a study to detect a true effect, is dependent on the size of the study group, the frequency of the outcome in the general population, and the level of excess risk to be identified. In a cohort study, common outcomes, such as recognized embryo/fetal loss, require hundreds of pregnancies in order to have a high probability of detecting a modest increase in risk (e.g., 133 in both exposed and unexposed groups to detect a twofold increase; alpha $\leq 0.05$, power = 80%), while less common outcomes, such as the total of all malformations recognized at birth, require thousands of pregnancies to have the same probability (e.g., more than 1200 in both exposed and unexposed groups) (15, 75, 76, 81, 82). In case-control studies, study sizes are dependent upon the frequency of exposure within the source population

*A posteriori* determination of power of the actual study is useful in evaluating negative findings. Negative findings in a study of low power would be given considerably less weight than either a positive study, or a negative study with high power.

(2) Potential bias in data collection: Sources of bias may include selection bias and information bias (83). Selection bias may occur when an individual's willingness to participate varies with certain characteristics relating to the exposure status or health status of that individual. In addition, selection bias may operate in the identification of subjects for study. For example, for studies of very early loss, use of hospital records to identify embryonic or early fetal loss will underascertain events, because women are not always hospitalized for these outcomes. More weight would be given in a risk assessment to a study in which a more complete list of pregnancies is obtained by, for example, either interviewing the women in the study or, in a prospective study, collecting biological data (e.g., human chorionic gonadotropin measurements) of pregnancy status from study members. A second example of different levels of ascertainment of events is the use of hospital records to study congenital malformations. Hospital records contain more complete data on malformations than do birth certificates. Thus, a study using hospital records to identify congenital malformations would be given more weight in a risk assessment.

Information bias may result from misclassification of characteristics of individuals or events identified for study. Recall bias, one type of information bias, may occur when respondents with specific exposures or outcomes recall information differently than those without the exposures or outcomes. Interview bias may result when the interviewer knows a *priori* the category of exposure (for cohort studies) or outcome (for case-control studies) in which the respondent belongs. Use of highly structured questionnaires and/or "blinding" of the interviewer will reduce the likelihood of such bias. Studies with lower likelihood of such types of bias should carry more weight in a risk assessment.

When data are collected by interview or questionnaire, the appropriate respondent depends upon the type of data or study. For example, a comparison of husband-wife interviews on reproduction found the wives' responses to questions on pregnancy-related events to be considerably more complete and valid than those of the husbands (78). Studies based on interview data from the appropriate respondent (e.g., the woman when examining her pregnancy history) would carry more weight than those from proxy respondents (e.g., the man when examining his partner's pregnancy history).

Data from any source may be prone to errors or bias. Validation with an independent data source (e.g., vital or hospital records), or use of biomarkers of exposure or outcome, where possible, may indicate the presence or absence of bias and increase confidence in the results of the study. Those studies with a low probability of biased data should carry more weight (82, 84).

(3) Control of potential risk factors, effect modifiers, and confounders: Potential risk factors may include smoking, alcohol consumption, drug use, past reproductive history, and environmental and occupational exposure. Such characteristics should be examined, where appropriate, for the outcome under study, and should be controlled for in the study design and/or analysis.

The potential for characteristics of the subjects to be effect modifiers and/or confounders should also be considered. An effect modifier is a factor that produces different exposure-response relationships at different levels of the effect modifier. For example, maternal age would be an effect modifier if the risk associated with a given exposure increased with the mother's age. A confounder is associated with both the exposure and outcome, and these interrelationships could distort both the magnitude and direction of the measure of association between the exposure of interest and the outcome. For example, smoking might be a confounder in a study of the association of socioeconomic status and low birth weight, since smoking has been associated with both.

Both effect modifiers and confounders need to be controlled in the analysis to improve the estimate of the effects of exposure (85). A more in-depth discussion may be found elsewhere (83, 86). The statistical techniques used to control for these factors require careful consideration in their application and interpretation (83, 85). Studies that fail to account for these important factors should be given less weight in a risk assessment.

(4) Statistical factors. As in animal studies, pregnancies experienced by the same woman are not independent events. In animal studies, the litter is generally used as the unit of measure to deal with nonindependence of events. This approach is difficult in humans since the pregnancies are sequential, with the risk factors changing for different pregnancies (15, 41, 81, 86). If more than one pregnancy per woman is included, as is often necessary due to small study groups, the use of nonindependent observations overestimates the true size of the population at risk and artificially increases the significance level (87). Some approaches to deal with these issues have been suggested (81, 88). At this point in time, a generally accepted solution to this problem has not been developed.

b. Selection of outcomes for study. As already discussed, a number of end points can be considered in the evaluation of adverse developmental effects. However, some of the outcomes are not easily observed in humans. These include early embryonic loss and reproductive capacity of the offspring. Currently, the most feasible end points for epidemiologic studies are reproductive history studies of some pregnancy outcomes (e.g., embryo/fetal loss, birth weight, sex ratio, congenital malformations, postnatal function, and neonatal growth and survival) and measures of subfertility/infertility which in some cases might be evidence of very early embryonic loss. Factors requiring control in the design or analysis (such as other risk factors, effect modifiers, and confounders) may vary depending on the specific outcomes selected for study.

The developmental outcomes available for epidemiologic examination are limited by a number of factors, including the relative magnitude of the exposure since differing spectra of outcomes may occur at different exposure levels, the size and demographic characteristics of the population, and the ability to observe the reproductive outcome in humans. Improved methods for identifying some outcomes such as embryonic or very early fetal loss using new human chorionic gonadotropin (hCG) assays may change the spectrum of outcomes available for study (12).

Demographic characteristics of the population, such as marital status, age distribution, education, and prior reproductive history are associated with the probability of whether couples will attempt to have children. There may also be differences in the use of birth control, which would affect the number of outcomes available for study. Additionally, workers may move in and out of areas with differing levels and types of exposures, affecting the number of exposed and comparison pregnancies for study. Larger populations are usually

necessary in environmental settings, since the exposures in environmental settings are generally much lower than in occupational settings.

c. Reproductive history studies.

(1) Pregnancy outcomes: Pregnancy outcomes examined in human studies of parental exposures may include embryo/fetal loss, congenital malformations, birth weight, sex ratio at birth, and possibly postnatal survival, growth, and function. Epidemiologic studies that focus on only one type of pregnancy outcome may miss a true effect of exposure. As mentioned above, some reproductive end points can be thought of as a continuum of adverse effects; for example, a malformed stillbirth would not be included in a study of defects observed at live birth, even though the etiology could be identical (75, 89). Studies that examine multiple end points could yield more information, but the results may be difficult to interpret. Evidence of a dose-response relationship is usually an important criterion in the assessment of a toxic exposure. However, traditional dose-response relationships may not always be observed for some end points. For example, with increasing dose, a pregnancy might end in an embryo/fetal loss, rather than a live birth with malformations. A shift in the patterns of outcomes could result from differences either in level of exposure or in timing (90, 91). Therefore, a risk assessment should, when possible, attempt to look at the interrelationship of different reproductive end points and patterns of exposure.

(2) Measures of fertility: Normally, studies of subfertility/infertility would not be included in an evaluation of developmental effects. However, in humans it is difficult to identify very early embryonic loss, and to distinguish it from subfertility/infertility. Thus, studies that examine subfertility or infertility indirectly examine loss very early in the gestational period. Studies of subfertility may be thought of as the study of non-events: a couple is unable to have children within a specific time frame. Therefore, the epidemiologic measurement of reduced fertility is typically indirect, and is accomplished by comparing birth rates or time intervals between births or pregnancies. In these evaluations, the couple's joint ability to procreate is estimated. One method, the Standardized Birth Ratio (SBR, also referred to as the Standardized Fertility Ratio), compares the number of births observed to those expected based on the person-years of observation stratified by factors such as time period, age, race, marital status, parity, contraceptive use, etc. (92–94) The SBR is analogous to the Standardized Mortality Ratio (SMR), a measure frequently used in studies of occupational cohorts, and has similar limitations in interpretation (87, 95) and in usefulness for risk assessment.

Analysis of the time period between recognized pregnancies or live births has been suggested as another indirect measure of fertility (96). Because the time interval between births increases with increasing parity (97), comparisons within birth order (parity) are more appropriate. A statistical method (Cox regression) can stratify by birth or pregnancy order to help control for nonindependence of these events in the same woman.

Fertility may also be affected by alterations in sexual behavior. However, limited data are available linking toxic exposures to these alterations in humans. Moreover, such data are not easily obtained in epidemiology studies. More information on this subject is available in the Proposed Guidelines for Assessing Male Reproductive Risk (38) and the Proposed Guidelines for Assessing Female Reproductive Risk (39).

d. Community studies/surveillance programs. Epidemiologic studies may also be based upon broad populations such as a community, a nationwide probability sample, or surveillance programs (such as birth defects registries) A number of case-control studies have examined the relationship between broad classes of parental occupation in certain communities or countries, and embryo/fetal loss (98), birth defects (99–101), and childhood cancer (100, 102–104). In these reports, jobs are typically classified into broad categories based on the probability of exposure to certain classes or levels of exposure (e.g., 100). Such studies are most helpful in the identification of topics for additional study However, because of the broad groupings of types of levels of exposure, such studies are not typically useful for risk assessment of a particular agent.

Surveillance programs may also exist in occupational settings. In this case, reproductive histories and/or clinical evaluation could monitor for reproductive effects of exposures. Both could yield very useful data for risk assessment; however, a clinical evaluation program would be costly to maintain.

C. Other Considerations

1. *Pharmacokinetics.* Extrapolation of toxicity data between species can be aided considerably by the availability of data on the pharmacokinetics of a particular agent in the species tested and, when available, in humans Information on absorption, half-life, placental metabolism and transfer, comparative metabolism, and concentrations of the parent compound and metabolites in the maternal animal and conceptus may be useful in predicting risk for developmental toxicity. Such data may also be helpful in defining the dose-response curve, developing a more accurate comparison of species sensitivity, including that of humans (105, 106), determining dosimetry at target sites, and comparing pharmacokinetic profiles for various dosing regimens or routes of exposure. Pharmacokinetic studies in developmental toxicology are most useful if conducted in pregnant animals at the stage when developmental insults occur. The correlation of pharmacokinetic parameters and developmental toxicity data may be useful in determining the contribution of specific pharmacokinetic parameters to the effects observed (107).

2. *Comparisons of Molecular Structure.* Comparisons of the chemical or physical properties of an agent with those of known developmental toxicants may provide some indication of a potential for developmental toxicity Such information may be helpful in setting priorities for testing of agents or for evaluation of potential toxicity when only minimal data are available Structure/activity relationships have not been well studied in developmental toxicology, although data are available that suggest structure-activity relationships for certain classes of chemicals (e.g., glycol ethers, steroids, retinoids). Under certain circumstances (e.g., in the case of new chemicals), this is one of several procedures used to evaluate the potential for toxicity when little or no data are available.

D. Weight-of-Evidence Determination

Information from all available studies, whether indicative of potential concern or not, must be evaluated and factored into a weight-of-evidence judgment as to the likelihood that an agent may pose a risk for developmental toxicity in humans. The primary considerations are the human data (which are seldom available) and the experimental animal data. The qualitative assessment for developmental toxicity should consider quality of the data, resolving power of the studies, number and types of end points examined, relevance of route and timing of exposure, appropriateness of the dose selection, replication of effects, number of species examined, and availability of human case reports or

series, and/or epidemiologic study data. In addition, pharmacokinetic data and structure-activity considerations, as well as other factors that may affect the strength of the evidence, should be taken into account. Therefore, all data pertinent to developmental toxicity should be examined in the evaluation of a chemical's potential to cause developmental toxicity in humans, and sound scientific judgment should be exercised in interpreting the data in terms of the risk for adverse human developmental health effects.

A categorization scheme for the weight of evidence has been developed. It contains several broad categories that reflect the accumulated data base on agents and serves as an indicator of whether exposure to the substance may cause developmental toxicity in humans. It represents one important step in the evaluation of agents. However, the risk of any given exposure to an agent can only be derived from an appreciation of its intrinsic biological activity and the nature of the anticipated exposure conditions. These important aspects are developed in subsequent sections of this Guideline.

Placing an agent in a particular weight-of-evidence category such as "adequate evidence for human developmental toxicity" does not mean that it will be a developmental toxicant at every dose (because of the assumption of a threshold) or in every situation (e.g., hazard may vary significantly depending on route and timing of exposure). Thus, in the final characterization of risk, the weight-of-evidence determination should always be presented in conjunction with information on dose-response (NOAEL and/or LOAEL), and, if available, with the human exposure estimate.

The weight-of-evidence scheme (outlined in Table 3) defines three levels of confidence for data used to identify developmental hazards and to assess the risk of human developmental toxicity: definitive evidence, adequate evidence, and inadequate evidence. Within the definitive evidence and adequate evidence categories, there are subcategories for evidence indicating adverse effects and for evidence indicating no apparent effects. In both categories, the evidence required to classify an agent as demonstrating no adverse effects is greater than that required to demonstrate an adverse effect and must include evaluations of a variety of potential manifestations of developmental toxicity. Greater evidence is required because it is much more difficult both biologically and statistically to support a finding of no apparent adverse effect than one of an adverse effect. Most agents meeting current testing requirements would be expected to fall within the adequate evidence category, while many for which little or no information is available would be classified in the inadequate category. Few agents would be expected to fall into the definitive evidence category because the human data necessary to meet the criteria for this category would be difficult to obtain.

**TABLE 3. WEIGHT OF EVIDENCE SCHEME FOR DEVELOPMENTAL TOXICITY**

Definitive Evidence for
- —Human Developmental Toxicity
- —No Apparent Human Developmental Toxicity

Adequate Evidence for:
- —Potential Human Developmental Toxicity
- —No Apparent Potential Human Developmental Toxicity

Inadequate Evidence for Determining Potential Human Developmental Toxicity

Because a complex interrelationship exists among study design, statistical analysis and biological significance of the data, a great deal of scientific judgment, based on experience with developmental toxicity data and with the principles of experimental design and statistical analysis, may be required to adequately evaluate the data base. To allow for this, the language used in the scheme is intentionally broad.

*Definitive Evidence for*

—Human Developmental Toxicity

This category includes agents for which there is sufficient evidence from epidemiologic studies for the scientific community to judge that a cause and effect relationship exists. Case reports in conjunction with other supporting evidence may also be used.

—No Apparent Human Developmental Toxicity

Agents in this category have not been associated with developmental toxicity in well-executed epidemiologic studies (e.g., case control and cohort) with adequate power. A variety of potential manifestations of developmental toxicity have been studied. Supporting animal data may or may not be available.

*Adequate Evidence for*

—Potential Human Developmental Toxicity

This category includes agents for which sufficient evidence exists for them to be considered potential human developmental toxicants. The minimum evidence necessary for considering an agent a potential human developmental toxicant would include data from an appropriate, well-executed study in a single experimental animal species that demonstrates developmental toxicity, and/or strong suggestive evidence from adequate clinical/epidemiologic studies. Evidence may be modified by further data, such as studies in additional species or by other routes of exposure, and replication of the findings. Development of pharmacokinetic or mechanistic information may reduce uncertainties in extrapolation to the human. The strength of the evidence increases as it approaches the definition for definitive human developmental toxicity.

—No Apparent Potential Human Developmental Toxicity

This category includes agents with data from appropriate well-executed studies in several species (at least two) which evaluated a variety of the potential manifestations of developmental toxicity and showed no developmental effects at doses that were minimally toxic to the adult animal. In addition, there may be human data from adequate studies supportive of no adverse effects.

*Inadequate Evidence for Determining Potential Human Developmental Toxicity*

This category includes agents for which there is less than the minimum sufficient evidence necessary for assessing human risk. However, data on agents that fall into this category may be used to determine the need for additional testing or information that would then, if adequate, move the agent into the adequate evidence category.

This category includes a variety of types of information such as the lack of any data on the developmental toxicity potential of an agent, data from an appropriate well-executed study in a single species showing no developmental toxicity, data from poorly-conducted studies in animals (e.g., small numbers of animals, inappropriate dose selection, other confounding factors) or inadequate data in humans. Additionally, data on structure/activity relationships, short-term test data, pharmacokinetic data, or data on metabolic precursors of the agent of interest could be used to call for further testing but would be considered insufficient by themselves to assess human risk.

## IV. Dose-Response Assessment

When quantitative human dose-effect data are available and with sufficient range of exposure, dose-response relationships may be examined. Data on

exposure from human studies are usually qualitative, such as employment or residence histories; quantitative or dose data are frequently not available. In human studies, especially retrospective ones, linking of specific time periods and specific exposures, even on a qualitative level, may be difficult due to errors of recall or recordkeeping (where records are available). The appropriate exposure depends on the outcome(s) studied, the biologic mechanism affected by exposure, and the half-life of the exposure. The probability of misclassification of exposure status may affect the ability of a study to recognize a true effect (15, 41, 76, 108, 109).

Since data on human dose-effect relationships are rarely available, the dose-response assessment is usually based on the evaluation of tests performed in laboratory animals. Evidence for a dose-response relationship is an important criterion in the assessment of developmental toxicity, although this may be based on limited data from standard studies using three dose groups and a control group. Most human developmental toxicants that have been studied alter development at doses within a narrow range near the lowest maternally toxic dose (22). Therefore, for most chemicals, the exposure situations of concern will be those that are potentially within this range. For those few chemicals where developmental effects occur at much lower levels than maternal effects, the potential for exposing the conceptus to damaging doses is much greater. As mentioned previously (section III.A.2.), however, traditional dose-response relationships may not always be observed for some end points. For example, as the exposure level rises, embryo/fetolethal levels may be reached, resulting in an observed decrease in malformations with increasing dose (81, 90). The potential for this response pattern indicates that dose-response relationships of individual end points as well as combinations of end points (e.g., dead and malformed combined) must be carefully examined and interpreted.

Identification of a NOAEL and/or LOAEL is based on the lowest dose at which an adverse effect is detected from any adequate developmental toxicity study. Adequacy of the data to be used for determination must be judged using the weight-of-evidence approach discussed in section III.D. NOAELs and applied uncertainty factors may be used to determine a reference dose for developmental toxicity ($RfD_{DT}$) that is assumed to be below the threshold for an increase in adverse developmental effects. The $RfD_{DT}$ is based on a short duration of exposure as is typically used in developmental toxicity studies. The term $RfD_{DT}$ is used to distinguish from the RfD which refers to chronic exposure situations (10). Uncertainty factors for developmental toxicity generally include a 10-fold factor for interspecies variation and a 10-fold factor for intraspecies variation. In general, an additional uncertainty factor is not applied to account for duration of exposure. Additional factors may be applied due to a variety of uncertainties that exist in the data base. For example, the standard study design for a developmental toxicity study calls for a low dose that demonstrates a NOAEL, but there may be circumstances where a risk assessment must be based on the results of a study in which a NOAEL for developmental toxicity was not identified. Rather, the lowest dose administered caused significant effect(s) and was identified as the LOAEL. In circumstances where only a LOAEL is available, questions relative to the sensitivity of end points reported, adequacy of dose levels tested, or confidence in the LOAEL reported may require the use of an additional uncertainty factor of 10 (10). The total uncertainty factor selected is then divided into the NOAEL/LOAEL for the most sensitive end point from the most appropriate and/or sensitive mammalian species to determine the $RfD_{DT}$.

Although the Agency currently uses the NOAEL/uncertainty factor approach to establish an $RfD_{DT}$, discussions of risk extrapolation procedures have noted that improved mathematical tools are needed for developing estimates of potential human developmental risk (45, 110). Gaylor (111) suggested an approach for estimating risk that combines the use of mathematical models for low-dose estimation of risk with the application of an uncertainty factor based on a preselected level of risk. This approach is similar to approaches proposed for carcinogenesis, but does not preclude the possibility of a threshold, and may provide a more quantitative approach to estimating risk. Another approach proposed by Rai and Van Ryzin (112) and recently applied by Faustman et al. (113), uses a simple two-component developmental model in which the first component represents a dose-related risk to the litter environment and the second component expresses the risk to an individual offspring conditional upon a predisposing risk to the litter. These approaches and others have been summarized recently (5). In addition, other methods for expressing risk are being sought and will be applied, if considered appropriate.

The development of biologically-based dose-response models in developmental toxicology is limited by a number of factors, including a lack of understanding of the biological mechanisms underlying developmental toxicity, intra/interspecies differences in the types of developmental events, and the influence of maternal effects on the dose-response curve. A biological threshold is assumed for most developmental effects based on known homeostatic, compensatory, or adaptive mechanisms that must be overcome before a toxic end point is manifested, and on the rationale that the embryo is known to have some capacity for repair of damage or insult (90). In addition, most developmental deviations are probably multifactorial in nature (114). Although a threshold is assumed for developmental effects, the existence of a NOAEL in an animal study does not prove or disprove the existence or level of a true threshold; it only defines the highest level of exposure under the conditions of the study that is not associated with a significant increase in effect. The uncertainties concerning this assumption are being discussed currently in the literature (25, 26).

In conclusion, dose-response findings in developmental toxicity studies are used as part of the risk characterization. This use is dependent upon scientific judgment as to the accuracy and adequacy of the data. In addition, the slope of the dose-response curve should be considered in conjunction with a determination as to the adequacy of the exposure levels tested, the sensitivity of the end points reported, and the appropriateness of the experimental design to determine a level of confidence in the data and the resultant confidence in the LOAEL, NOAEL, and the uncertainty factors applied to obtain the $RfD_{DT}$.

## V. Exposure Assessment

In order to obtain a quantitative estimate of risk for the human population, an estimate of human exposure is required. The Guidelines for Estimating Exposures have been published separately (115) and will not be discussed in detail here. In general, the exposure assessment describes the magnitude, duration, schedule, and route of exposure. This information is developed from monitoring data and from estimates based on modeling of environmental exposures. Unique considerations for developmental

toxicity are duration and period of exposure as related to stage of development (i.e., critical periods), and the possibility that a single exposure may be sufficient to produce adverse developmental effects (i.e., repeated exposure is not a necessary prerequisite for developmental toxicity to be manifested). For these reasons, it is assumed that a single exposure at the critical time in development is sufficient to produce an adverse developmental effect. Therefore, the human exposure estimate used to calculate the margin of exposure is usually based on a single dose that is not adjusted for duration of exposure, and the number of exposures is not considered important unless there is evidence for a cumulative effect. It should be recognized also that exposure of almost any segment of the human population (i.e., fertile men and women, the conceptus, and the child up to the age of sexual maturation) may lead to risk to the developing organism.

## VI. Risk Characterization

Many uncertainties described in these Guidelines are associated with the toxicological and exposure components of risk assessments in developmental toxicology. In the past, these uncertainties have often not been readily apparent or consistently presented. The presentation of any risk assessment for developmental toxicity should be accompanied by statements concerning the weight of the evidence, dose-response relationships and assumptions underlying the estimation of the $RfD_{DT}$, esumates of human exposure, and any factors that affect the quality and precision of the assessment. The risk characterization of an agent should be based on data from the most appropriate species, or, if such information is not available, on the most sensitive species tested. It should also be based on the most sensitive indicator of toxicity, whether maternal, paternal, or developmental, and should be considered in relationship to other forms of toxicity.

In the risk characterization, the dose-response and the human exposure estimate may be combined either by comparing the $RfD_{DT}$ and the human exposure estimate or by calculating the margin of exposure (MOE). The MOE is the ratio of the NOAEL from the most appropriate or sensitive species to the estimated human exposure level from all potential sources (53). If a NOAEL is not available, a LOAEL may be used in the calculation of the MOE. In this case, the NOAEL may be estimated from the LOAEL by applying an uncerta.nty factor (10-fold) to assess the impact on the MOE (53). The MOE is presented along with a discussion of the weight of evidence, including the nature and quality of the hazard and exposure data, the number of species affected, and the dose-response information.

The $RfD_{DT}$ comparison with the human exposure estimate and the calculation of the MOE are conceptually similar but are used in different regulatory situations. The choice of approach is dependent upon several factors, including the statute involved, the situation being addressed, the data base used, and the needs of the decision maker. The $RfD_{DT}$ and/or the MOE are considered along with other risk assessment and risk management issues in making risk management decisions, but the scientific issues that must be taken into account in establishing them have been addressed here.

These Guidelines summarize the procedures that the U.S. Environmental Protection Agency will follow in evaluating the potential for agents to cause developmental toxicity. While these are the first amendments to the developmental toxicity guidelines issued in 1986, further revisions and updates will be made as advances occur in the field. Further studies that: (1) Delineate the mechanisms of developmental toxicity and pathogenesis, (2) provide comparative pharmacokinetic data, and (3) elucidate the functional modalities that may be altered by exposure to toxic agents, will aid in the interpretation of data and interspecies extrapolation. These types of studies, along with further evaluation of the relationship between maternal and developmental toxicity and the concept of a threshold, will provide for the development of improved mathematical models to more precisely assess risk.

## VII. References


1. U.S. Environmental Protection Agency. Assessment of risks to human reproduction and to development of the human conceptus from exposure to environmental substances, pp. 99–116. Available from: NTIS, Springfield, VA. DE82-007897.

2. Kimmel, G.L., K. Smith, D.M. Kochhar and R.M. Pratt. 1982. Proceedings of the consensus workshop on in vitro teratogenesis testing. Teratogenesis Carcinogen. Mutagen. 2:221–374.

3. Kimmel, G.L., C.A. Kimmel and E.Z. Francis, eds. 1987. Evaluation of maternal and developmental toxicity. Teratogenesis Carcinogen. Mutagen. 7:203–338.

4. Hardin, B.D., ed. 1987. Evaluation of the Chernoff/Kavlock test for developmental toxicity. Teratogenesis Carcinogen. Mutagen. 7:1–127.

5. Kimmel, C.A., D.G. Wellington, W. Farland, P. Ross, J.M. Manson, N. Chernoff, J.F. Young, S.G. Selevan, N. Kaplan, C. Chen, L.D. Chitlik, C.L. Siegel-Scott, G. Valoras and S. Wells. 1988. Overview of a workshop on quantitative models for developmental toxicity risk assessment. Environ. Health Perspect., in press.

6. Perlin, S.A. 1988. Using weight-of-evidence classification schemes in the assessment of non-cancer health risks. Proceedings of the Hazardous Wastes and Hazardous Materials Conference, April.

7. U.S. Environmental Protection Agency. 1982. Pesticides assessment guidelines, subdivision F. Hazard evaluation: human an domestic animals. EPA-540/9-82-025. Available from: NTIS, Springfield, VA.

8. U.S. Environmental Protection Agency. 1985. Toxic Substances Control Act test guidelines, final rules. Federal Register 50:39428–39428 and 39433–39434.

9. Committee on the Institutional Means fo the Assessment of Risks to Public Health. 1983. Risk assessment in the Federal government: managing the process. Commission on Life Sciences, National Research Council. Washington, DC. National Academy Press, pp. 17–83.

10. U.S. Environmental Protection Agency. 1987. Reference dose (RfD): description and use in health risk assessments. Appendix A. Integrated Risk Information System Documentation, vol. 1. EPA/600/8-86/032a.

11. Hertig, A.T. 1967. The overall problem in man. In: K. Benirschke, ed. Comparative aspects of reproductive failure. New York, NY: Springer-Verlag, pp. 11–41.

12. Wilcox, A.J., C.R. Weinberg, R.E. Wehmann, E.G. Armstrong, R.E. Canfield, and B.C. Nisula. 1985. Measuring early pregnancy loss: laboratory and field methods. Fertil. Steril. 44:366–374.

13. Shepard, T.H. 1986. Human teratogenicity. Adv. Pediatr. 33:225–268.

14. Wilson, J.G. 1977. Embryotoxicity of drugs in man. In: J.G. Wilson and F.C. Fraser, eds. Handbook of teratology. New York, NY: Plenum Press, pp. 309–355.

15. Selevan, S.G. 1981. Design considerations in pregnancy outcome studies of occupational populations. Scand. J. Work Environ. Health 7:76–82.

16. Shepard, T.H. 1980. Catalog of teratogenic agents. Third edition. Baltimore, MD: Johns Hopkins University Press.

17. National Center for Health Statistics. 1988. Advance report of final mortality statistics, 1986. Monthly Final Statistics Report, vol. 37 (16) Supp., DHHS Publ. No. (PHS) 88-1120.

18. Centers for Disease Control. 1988. Trends in years of potential life lost due to infant mortality and perinatal conditions: 1980–1983 and 1984–1985. Morbidity and Mortality Weekly Report 37:249–256.

19. Centers for Disease Control. 1988. Premature mortality due to congenital anomalies—United States. Morbidity and Mortality Weekly Report 37:505–506.

20. Schardein, J.L. 1983. Teratogenic risk assessment. In: H. Kalter, ed. Issues and reviews in teratology, vol. 1. New York, NY: Plenum Press, pp. 181–214.

21. Brown, N.A., and S. Fabro. 1983. The value of animal teratogenicity testing for predicting human risk. Clin. Obstet. Gynecol. 26:467–477.

22. Kimmel, C.A., J.F. Holson, C.J. Hogue and G.L. Carlo. 1984. Reliability of

experimental studies for predicting hazards to human development. NCTR Technical Report for Experiment No 6015. NCTR, Jefferson, Arkansas.

23 Nisbet, I.C.T., and N.J Karch. 1983 Chemical hazards to human reproduction. Park Ridge: Noyes Data Corp.

24. Hemminki, K., and P. Vineis. 1985 Extrapolation of the evidence on teratogenicity of chemicals between humans and experimental animals, chemicals other than drugs. Teratogenesis Carcinogen. Mutagen. 5:251–316.

25. Gaylor, D.W., D.M. Sheehan, J.F Young, and D.R. Matuson. 1988. Letter to the editor The threshold dose question in teratogenesis. Teratology 38:389–391

26. Giavinni, E. 1988. Letter to the editor Evaluation of the threshold concept in teratogenicity studies Teratology 38:393–395.

27. U.S. Environmental Protection Agency. 1986. Triethylene glycol monomethyl, monoethyl, and monobutyl ethers: proposed test rule. Federal Register 51 17883–17894.

28. U.S. Environmental Protection Agency. 1988. Feb 26. Diethylene glycol butyl ether and diethylene glycol butyl ether acetate: final test rule. Federal Register 53:5932–5953

29 Herbst, A.L., H. Ulfelder, and D C. Poskanzer. 1971. Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with appearance in young women. N Engl J. Med. 284:878.

30 Anderson, L.M., P.J Donovan, and J.M. Rice. 1985. Risk assessment for transplancental carcinogens. In: A.P. Li, ed New approaches in toxicity testing and their application in human risk assessment. New York, NY: Raven Press, pp. 179–202.

31 U.S. Environmental Protection Agency 1986. Sept. 24. Guidelines for carcinogen risk assessment. Federal Register 51(185):33992–34003

32. U.S. Environmental Protection Agency. 1986. Sept. 24. Guidelines for mutagenicity risk assessment. Federal Register 51(185):34006–34012.

33. Food and Drug Administration. 1966 Guidelines for reproduction and teratology of drugs Bureau of Drugs.

34 Food and Drug Administration. 1970 Advisory committee on protocols for safety evaluations. Panal on reproduction report on reproduction studies in the safety evaluation of food additives and pesticide residues. Toxicol. Appl Pharmacol. 16:264–296.

35 Organization for Economic Cooperation and Development (OECD) 1981. Guideline for testing of chemicals—teratogenicity.

36 Symposium on effects of radiation and other deleterious agents on embryonic development. 1954 J. Cell. Comp. Physiol 43(suppl. 1).

37. Lamb, J.C., IV. 1985. Reproductive toxicity testing: evaluating and developing new testing systems J. Amer. Coll. Toxicol. 4.163–171.

38. U.S. Environmental Protection Agency. 1988. Proposed guidelines for assessing male reproductive risk Federal Register 53.24850–24869

39 U.S Environmental Protection Agency 1988. Proposed guidelines for assessing female reproductive risk. Federal Register 53.24834–24847

40 Kimmel, C.A., and C.J Price. 1988. Conducting a study developmental toxicity studies. In D.L. Arnold et al., eds In vivo toxicity testing. New York, NY. Academic Press, in press.

41. Kimmel, C.A., G.L. Kimmel, and V. Frankos, eds 1986 Interagency Regulatory Liaison Group workshop on reproductive toxicity risk assessment Environ. Health Perspect. 66:193–221.

42. Rodier, P.M. 1978. Behavioral teratology In: J.G. Wilson and F.C. Fraser, eds Handbook of teratology, vol. 4. New York, NY: Plenum Press, pp. 397–428.

43. Riley, E.P., and C.V. Vorhees, eds. 1986. Handbook of behavioral teratology. New York, NY: Plenum Press. 522 pp

44. Kimmel, C.A. 1988. Current status of behavioral teratology—science and regulation. In: L. Goldberg, ed. CRC critical reviews in toxicology. Boca Raton, FL. CRC Press, Inc., in press

45. World Health Organization. 1984. Principles for evaluating health risks to progeny associated with exposure to chemicals during pregnancy In: Environmental Health Criteria. Geneva: World Health Organization. vol. 30

46. Buelke-Sam, J., C.A. Kimmel, and J Adams, eds. 1985. Design considerations in screening for behavioral teratogens: results of the Collaborative Behavioral Teratol. Study. Neurobehav Toxicol. Teratol 7(6):537–789

47 Food and Drug Administration 1985. Predicting neurotoxicity and behavioral dysfunction from preclinical toxicologic data Proceedings of a symposium. Sept. 30, Bethesda, MD.

48. Tanimura, T 1986. Collaborative studies on behavioral teratology in Japan Neurotoxicology 7:35–45.

49. Elsner, J., K.E. Suter, B. Ulbrich, and G Schreiner. 1986 Testing strategies in behavioral teratology IV. Review and general conclusions. Neurobehav. Toxicol. Teratol. 8:585–590.

50 Kavlock, R.M., and C.T. Grabowski, eds. 1983. Abnormal functional development of the heart, lungs, and kidneys approaches to functional teratology Prog Clin. Biol. Res., vol. 140. New York, NY: Alan R. Liss, Inc.

51. Fujii, T., and P.M. Adams. 1987. Functional teratogenesis: functional effects on the offspring after parental drug exposure Tokyo, Japan. Teikyo University Press. 238 pp.

52. Adams, J. 1986. Clinical relevance of experimental behavioral teratology Neurotoxicology 7:19–34

53. Butcher, R.E., V. Wootten, and C V. Vorhees. 1980. Standards in behavioral teratology testing: test variability and sensitivity Teratogenesis Carcinog. Mutagen. 1:49–61.

54. Johnson, E.M. 1981 Screening for teratogenic hazards: nature of the problem. Annu. Rev Pharmacol. Toxicol 21:417–429.

55 Johnson, E.M., and B.E.G. Gabel. 1983 An artificial embryo for detection of abnormal developmental biology Fundam. Appl. Toxicol. 3.243–249.

56. Fabro, S., G. Schull, and N.A. Brown. 1962. The relative teratogenic index and teratogenic potency: proposed components of developmental toxicity risk assessment. Teratogenesis Carcinogen Mutagen 2:61–76

57 Brown, N.A., and S J Freeman. 1984 Alternative tests for teratogenicity Alternatives Lab Anim. 12.7–23

58 U.S. Environmental Protection Agency 1985. Hazard evaluation division standard evaluation procedure Teratology studies Office of Pesticide Programs, Washington, DC pp. 22–23.

59 Chernoff, N., and R.J Kavlock. 1982. An in vivo teratology screen utilizing pregnant mice. J Toxicol. Environ. Health 10:541–550.

60 Kavlock, R.J., R.D. Short, Jr., and N. Chernoff. 1987. Further evaluations of an in vivo teratology screen. Teratogenesis Carcinog. Mutagen. 7:7–16.

61. Brown, J.M. 1984. Validation of an in vivo screen for the determination of embryo/fetal toxicity in mice. SRI International EPA contract no. 68–01–6079.

62. Schuler R., B. Hardin, R. Niemeyer, G Booth, K. Hazelden, V. Piccirillo, and K. Smith. 1984. Results of testing fifteen glycol ethers in a short-term, in vivo reproductive toxicity assay Environ. Health Perspect. 57:141–146.

63. Wickramaratne, G.A. de S 1987 The Chernoff-Kavlock assay: its validation and application in rats. Teratogenesis Carcinogen Mutagen. 7:73–83.

64. U.S. Environmental Protection Agency 1985. Toxic Substances Control Act Test Guidelines, final rules Federal Register 50:39428–39429

65. Francis, E.Z., and W.H Farland 1987 Application of the preliminary developmental toxicity screen for chemical hazard identification under the Toxic Substances Control Act. Teratogenesis Carcinogen. Mutagen. 7:107–117

66. Wilson, J.G. 1978. Survey of in vitro systems: their potential use in teratogenicity screening In: J.G. Wilson and F.C. Fraser, eds. Handbook of teratology, vol 4. New York, NY: Plenum Press, pp 135–153

67. Kimmel, G.L., K. Smith, D.M Kochhar, and R.M. Pratt 1982. Overview of in vitro teratogenicity testing: aspects of validation and application to screening Teratogenesis Carcinogen. Mutagen. 2:221–229.

68. Brown, N.A., and S.E. Fabro. 1982. The in vitro approach to teratogenicity testing In K. Snell, ed. Developmental toxicology. London, England: Croom-Helm, pp. 31–57

69. Kimmel, G.L. 1985. In vito tests in screening teratogens: considerations to aid the validation process. In: M Marois, ed. Prevention of physical and mental congenital defects, Part C. New York, NY. Alan R Liss, Inc., pp 259–263

70 Whitby, K.E., et al. 1987. Report of the in vitro teratology task force. Environ. Health Perspect. 72.201–248.

71. Brown, N.A. 1987. Teratogenicity testing in vitro: status of validation studies. Arch. Toxicol. Suppl. 11:106–114.

72. Haseman, J.K., and L.L. Kupper 1979 Analysis of dichotomous response data from certain toxicological experiments. Biometrics 35.281–293.

73 Nelson, C.J., and J.F. Holson 1978. Statistical analysis of teratogenic data problems and advancements J Environ Pathol Toxicol 2:187–199

74 Hogue, C.J R 1985 Developmental risks Presented at symposium on epidemiology and health risk assessment. May 14 Columbia, MD

75 Bloom. A.D 1981 Guidelines for reproductive studies in exposed human populations Guideline for studies of populations exposed to mutagenic and reproductive hazards. Report of Panel II White Plains. NY March of Dimes Birth Defects Foundation. , p. 37-110.

76 Sever. L.E., and N.A. Hessol. 1984 Overall design considerations in male and female occupational reproductive studies In: Lockey. J.E., G K. Lemasters. and W R. Keye. eds. Reproduction: the new frontier in occupational and environmental research. New York. NY: Alan R. Liss. Inc. pp 15-48

77 Hatch. M., and J. Kline. 1981. Spontaneous abortion and exposure to the herbicide 2.4.5-T a pilot study U S. Environmental Protection Agency. Washington. DC. EPA-560/6-81-006.

78. Selevan. S.G. 1980 Evaluation of data sources for occupational preganancy outcome studies (dissertation). Available from: University Microfilms. Ann Arbor. MI.

79 Axelson. O 1985 Epidemiologic methods in the study of spontaneous abortion. sources of data. methods. and sources of error In. Hemminki. K., M. Sorsa. and H. Vainio. eds. Occupational hazards and reproduction. Washington. DC. Hemisphere Pub . pp 231-238.

80 Wilcox. A J 1983 Surveillance of pregnancy loss in human populations. Am. J. Ind Med. 4 285-291.

81. Selevan. S.G 1985. Design of pregnancy outcome studies of industrial exposure In: Hemminki. K., M. Sorsa. and H Vainio. eds. Occupational hazards and reproduction. Washington. DC. Hemisphere Pub., pp 219-229

82. Stein. Z., J Kline. and P Shrout. 1985 Power in surveillance In Hemminki. K., M. Sorsa. and H Vainio. eds. Occupational hazards and reproduction. Washington. DC. Hemisphere Pub . pp 203-208.

83. Rothman. K J 1986 Modern epidemiology Boston. MA. Little. Brown and Co pp 83-94

84 Stein. A., and Hatch. M 1987 Biological markers in reproductive epidemiology: prospects and precautions. Environ. Health Perspect 74 67-75

85 Kleinbaum. D G . L.L. Kupper. and H. Morgenstern. 1982. Epidemiologic research: principle and quantitive methods. London: Lifetime Learning Publications.

86 Epidemiology Workgroup of the Interagency Regulatory Liaison Group 1981. Guidelines for documentation of epidemiologic studies Am. J. Epidemiol. 114(5) 609-613

87 McMichael. A J 1976. Standardized mortality ratios and the "healthy worker effect" Scratching beneath the surface. J Occup Med. 18 165-168.

88 Shratelli. R N Laird. and J H Ware 1984 Random-effects models for serial observations with binary responses Biometrics 40 961-971

89 Stein Z., M Susser D Warburton. J Wittes. and J Kline 1975 Spontaneous abortion as a screening device The effect of fetal surveillance on the incidence of birth defects. Am. J Epidemiol. 102:275-290.

90. Wilson. J G. 1973. Environment and birth defects. New York. NY Academic Press. pp 30-32.

91 Selevan. S G., and G K. LeMasters 1978. The dose-response fallacy in human reproductive studies of toxic exposures. J. Occup. Med. 29:451-454.

92. Wong. O., H.M.D. Utidjian. and V S. Karten. 1979 Retrospective evaluation of reproductive performance of workers exposed to ethylene dibromide J Occup. Med. 21:98-102.

93. Levine. R.J., M.J Symons. S.A. Balogh. D.M. Arndt. N.R. Kaswandik. and J W Gentile. 1980. A method for monitoring the fertility of workers: I. Method and pilot studies. J. Occup Med. 22:781-791

94. Levine. R.J., M J Symons. S.A. Balogh. T H. Milby, and M.D Whorton. 1981 A method for monitoring the fertility of workers. II. Validation of the method among workers exposed to dibromochloropropane J. Occup. Med. 23 183-188.

95. Tsai. S.P., and C.P Wen. 1986. A review of methodological issues of the standardized mortality ratio (SMR) in occupational cohort studies. Int. J. Epidemiol. 15.8-21.

96. Baird. D D., A J Wilcox. and C.R. Weinberg. 1986 Using time to pregancy to study environmental exposures Am. J Epidemiol. 124 470-480.

97 Lendon. H Human fertility the basic components 1977 Chicago. IL. The University of Chicago Press.

98. Silverman. J., J Kline. and M Hutzler 1985. Maternal employment and the chromosomal characterstics of spontaneously aborted concephons. J Occup. Med 27 427-438.

99. Hemminki. K., P Mutanen. K Luoma. and I Saloniemi 1980 Congenital malformations by the parental occupation in Finland. Int. Arch. Occup Environ. Health 46:93-98.

100. Kwa. S.-L., and L.J Fine 1980 The association between parental occupation and childhood malignancy J Occup Med 22:792-794.

101 Papier. C.M. 1985 Parental occupation and congenital malformations in a series of 35.000 births in Israel. Prog Clin. Biol Res. 163 291-294

102. Hemminki K., I Saloniemi. and T Salonen. 1981 Childhood cancer and paternal occupation in Finland J Epidemiol Commun. Health 35 11-15

103 Peters J M., S Preston-Martin and M C Yu 1981 Brain tumors in children and occupational exposure of the parents Science 213 235-237

104. Zack. M S Cannon. D Lloyd. C W Heath. Jr., J M Falletta. B. Jones. J Housworth. and S. Crowley 1980 Cancer in children of parents exposed to hydrocarbon-related industries and occupations Am J Epidemiol. 3 329-338.

105. Wilson. J G., W J Scott. E.J Ritter and R Fradkin. 1975. Comparative distribution and embryotoxicity of hydroxyures in pregnant rats and rhesus monkeys Teratology 11 169-178.

106. Wilson. J G., E.J Ritter. W J Scott and R Fradkin. 1977 Comparative distribution and embryotoxicity of acetylsalicylic acid in pregnant rats and rhesus monkeys. Toxicol. Appl Pharmacol. 41.67-78.

107 Kimmel. C.A., and J F Young. 1983. Correlating pharmacokinetics and teratogenic end points. Fundam. Appl. Toxicol 3 250-255.

108. Lemasters. G K., and S G Selevan. 1984. Use of exposure data in occupational reproductive studies. Scand. J Work Environ. Health 10:1-6.

109. Hogue. C.J.R. 1984 Reducing misclassification errors through questionnaire design. In: Lockey. J E., G K Lemasters. and W R. Keye. eds. Reproduction. the new frontier in occupational and environmental health research. New York. NY Alan R Liss. Inc., pp 81-97

110. Hogan. M.D., and D G Hoel 1982. Extrapolation to man. In A W Hayes. ed. Principles and methods of toxicology New York. NY. Raven Press. pp 711-731

111 Gaylor. D W. 1983 The use of safety factors for controlling risk J Toxicol Environ. Health 11 329-336

112. Rai. K., and J Van Ryzin 1985 A dose response model for teratological experiments involving quantal responses Biometrics 41 1-10.

113. Faustman. E.M., D G Wellington. W P. Smith. and C.A. Kimmel 1988. Characterization of a developmental toxicity dose-response model Environ Health Perspect., in press

114. Fraser. F C. 1977 Relation of animal studies to the problem in man In. J G Wilson and F C. Fraser eds Handbook of teratology. vol. 1. New York. NY Plenum Press pp 75-98.

115 U S Environmental Protection Agency 1986. Sept 24 Guidelines for estimating exposures Federal Register 51 (185) 34042-34054

[FR Doc. 89-5066 Filed 3-3-89; 8 45 am]

BILLING CODE 6560-60-M