EPA/630/P-02/002F
December 2002
Final Report

A REVIEW OF THE REFERENCE DOSE AND REFERENCE CONCENTRATION PROCESSES

Prepared for the
Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC

Reference Dose/Reference Concentration (RfD/RfC) Technical Panel

Bob Benson (OPRA/Region 8)
Gary Foureman (NCEA/ORD)
Lee Hofmann (PARMS/OSWER)
Carole Kimmel (NCEA/ORD)*
Gary Kimmel (NCEA/ORD)
Susan Makris (OPP/OPPTS)
Deirdre Murphy (OAQPS/OAR)
Edward Ohanian (OST/OW)
Jennifer Orme-Zavaleta (NHEERL/ORD)
Deborah Rice (NCEA/ORD)
Jennifer Seed (OPPT/OPPTS)
Hugh Tilson (NHEERL/ORD)
Vanessa Vu (SAB Staff Office, formerly OSCP/OPPTS and NCEA/ORD)

*Technical Panel Chair

Technical Advisors

Amy Mills, IRIS Director, NCEA/ORD
Bill Wood, RAF Director, NCEA/ORD

Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF ABBREVIATIONS AND ACRONYMS viii
PREFACE x
EXECUTIVE SUMMARY xiii

1. INTRODUCTION, PURPOSE, AND SCOPE 1-1

2. REVIEW OF THE CURRENT USE OF ACUTE, SHORT-TERM, AND LONGER-TERM REFERENCE VALUES 2-1
   2.1. REVIEW OF CURRENT LESS-THAN-LIFETIME REFERENCE VALUES 2-1
      2.1.1. Acute Reference Exposure (ARE) Methodology 2-1
      2.1.2. Acute Exposure Guidelines (AEGL) Program 2-3
      2.1.3. Office of Pesticide Programs (OPP) Procedures for Setting Acute and Intermediate RfDs 2-5
      2.1.4. Office of Water (OW) Health Advisories (HAs) 2-6
      2.1.5. Agency for Toxic Substances and Disease Registry (ATSDR) Minimal Risk Levels (MRLs) 2-9
   2.2. SUMMARY OF CURRENT METHODS FOR SETTING ACUTE, SHORT-TERM, AND LONGER-TERM REFERENCE VALUES 2-9
   2.3. RECOMMENDATION 2-11

3. REVIEW OF TESTING GUIDELINES WITH RESPECT TO LIFE STAGE ASSESSMENT 3-1
   3.1. EVALUATION OF CURRENT GUIDELINE TESTING PROTOCOLS 3-2
      3.1.1. Exposures and Endpoints Related to General Toxicity Testing 3-4
         3.1.1.1. Acute and Short-Term Toxicity Studies 3-6
         3.1.1.2. Subchronic and Chronic Toxicity Studies 3-9
      3.1.2. Exposures and Endpoints Related to Evaluation of Reproductive Toxicity 3-10
         3.1.2.1. Overview of Tests 3-10
         3.1.2.2. Gaps in Life Stage of Assessment 3-13
         3.1.2.3. Gaps in Assessment Endpoints 3-14
         3.1.2.4. Gaps in Duration/Latency Assessment 3-14
      3.1.3. Exposures and Endpoints Related to Evaluation of Neurotoxicity 3-15
         3.1.3.1. Overview of Tests 3-15

CONTENTS (continued)

3.1.3.2. Gaps in Life Stage of Assessment 3-17
3.1.3.3. Gaps in Assessment Endpoints 3-18
3.1.3.4. Gaps in Duration/Latency Assessment 3-18
3.1.4. Exposures and Endpoints Related to Evaluation of Immunotoxicity 3-19
3.1.4.1. Overview of Tests 3-19
3.1.4.2. Gaps in Life Stage of Assessment 3-22
3.1.4.3. Gaps in Assessment Endpoints 3-22
3.1.4.4. Gaps in Duration/Latency Assessment 3-23
3.1.5. Exposures and Endpoints Related to Evaluation of Cardiovascular Toxicity 3-23
3.1.5.1. Overview of Tests 3-23
3.1.5.2. Gaps in Life Stage of Assessment 3-23
3.1.5.3. Gaps in Assessment Endpoints 3-25
3.2. CONCLUSIONS AND RECOMMENDATIONS 3-26
3.2.1. Conclusions 3-27
3.2.2. Recommendations 3-29
3.3. OPTIONS FOR ALTERNATIVE TESTING APPROACHES 3-31
3.3.1. Alternative Acute Toxicity Testing Protocol 3-31
3.3.2. Alternative Chronic Toxicity Testing Protocols 3-33
3.3.2.1. The Expanded Chronic/Carcinogenicity Study 3-33
3.3.2.2. The Unified Screening Study 3-35

4. FRAMEWORK FOR SETTING ACUTE, SHORT-TERM, LONGER-TERM, AND CHRONIC REFERENCE VALUES 4-1
4.1. DEFINITIONS OF EXPOSURE DURATIONS FOR USE IN SETTING REFERENCE VALUES 4-2
4.2. PROPOSED CHANGES IN THE REFERENCE VALUE DEFINITIONS 4-4
4.3. CHARACTERIZATION OF THE EXTENT OF THE HEALTH-RELATED DATABASE FOR SETTING REFERENCE VALUES 4-7
4.3.1. Review of Studies 4-7
4.3.1.1. Adequacy of Studies 4-8
4.3.2. Issues to be Considered in Characterizing the Database for Risk Assessment 4-11
4.3.2.1. The Weight-of-Evidence Approach 4-11
4.3.2.2. Use of Human and Animal Data in Risk Assessment 4-12
4.3.2.3. Characterization of Effects in Potentially Susceptible Subpopulations 4-15
4.3.3. Characterization of the Extent of the Database 4-17
4.3.3.1. Extent of the Database 4-20
4.4. DERIVATION OF REFERENCE VALUES 4-21

CONTENTS (continued)

4.4.1. Sample Reference Values and Selection of Endpoints to Use as the POD for Reference Values 4-22
4.4.2. Dose Adjustment for Duration of Exposure 4-24
4.4.2.1. Duration Adjustment Procedures for Inhalation Exposures to Continuous-Exposure Scenarios 4-24
4.4.2.2. Duration Adjustment for Inhalation Developmental Toxicity Studies—A Current Exception 4-25
4.4.2.3. Duration Adjustment for Acute Reference Values—Discontinuous Scenarios of 24 Hours or Less 4-26
4.4.3. Derivation of a HEC or a HED 4-29
4.4.3.1. PBPK Models and Derivation of HEDs and HECs: Estimating Internal Dose 4-30
4.4.3.2. Default Procedures and Derivation of HECs from the RfC Methodology: Derivation and Application of DAFs 4-32
4.4.3.3. HECs and Children—A Special Case? 4-34
4.4.3.4. Deriving a HED for Oral and Dermal Exposure—Use of $BW^{3/4}$ as a Cross-Species DAF 4-35
4.4.4. Other Issues 4-37
4.4.5. Application of Uncertainty/Variability Factors 4-38
4.4.5.1. Recommendations for Application of UFs 4-40
4.4.5.2. Interspecies UF 4-42
4.4.5.3. Intraspecies UF 4-42
4.4.5.4. LOAEL-to-NOAEL UF 4-44
4.4.5.5. Database UF 4-44
4.4.5.6. Subchronic-to-Chronic-Duration UF 4-45
4.4.5.7. Modifying Factor (MF) 4-46
4.4.6. Future Directions 4-47
4.4.6.1. Chemical-Specific Adjustment Factors (CSAFs) 4-47
4.4.6.2. Probabilistic Approaches 4-48
4.4.7. Summary of Key Points from the Case Studies 4-49

5. RECOMMENDATIONS 5-1

APPENDIX A: ISSUES RAISED BY THE 10X TASK FORCE A-1

APPENDIX B: CASE STUDIES—EVALUATING AND SELECTING HEALTH ENDPOINTS FOR DERIVING REFERENCE VALUES B-1

GLOSSARY G-1

REFERENCES R-1

LIST OF TABLES

Table 2-1. Duration definitions used for various reference values 2-10

Table 2-2. Uncertainty/safety factors for various reference values 2-12

Table 3-1. Approximate age at equivalent life stages in several species 3-3

Table 3-2. Systems/endpoints evaluated by routine toxicity guideline testing protocols . . 3-5

Table 4-1. Factors for evaluation of the weight of evidence regarding the likelihood of effects in humans 4-14

Table 4-2. Factors for evaluating of evidence regarding identification and characterization of susceptible subpopulations 4-17

Table 4-3. DAFs based on $BW^{3/4}$ for various species 4-36

Table 4-4. BMDLs, ingested dose, and RfDs for various endpoints from the Faroe Islands, New Zealand, and the NRC integrative analysis 4-51

LIST OF FIGURES

Figure 3-1. Exposures and endpoints related to general toxicity evaluations 3-7

Figure 3-2. Exposures and endpoints related to reproductive evaluations 3-11

Figure 3-3. Exposures and endpoints for neurotoxicity evaluations 3-16

Figure 3-4. Exposures and endpoints for immunotoxicity evaluations 3-20

Figure 3-5. Exposures and endpoints related to cardiovascular evaluations 3-24

Figure 3-6. Alternative acute toxicity protocol 3-32

Figure 3-7. Expanded chronic/carcinogenicity study 3-34

Figure 3-8. Unified screening study 3-36

Figure 4-1. Concentration-by-duration plot showing the effect of the exponent in the $C^n \times T = k$ on extrapolation across time. 4-27

Figure 4-2. Current and proposed generalized procedures for deriving HECs or HEDs from animal exposures 4-31

LIST OF ABBREVIATIONS AND ACRONYMS

ACE II	Angiotensin converting enzyme II
ADI	Acceptable Daily Intake
AEGL	Acute exposure guideline level
ARE	Acute reference exposure
ATSDR	Agency for Toxic Substances and Disease Registry
AUC	Area under the curve
BMC	Benchmark concentration
BMCL	Benchmark concentration lower confidence limit
BMD	Benchmark dose
BMDL	Benchmark dose lower confidence level
BMR	Benchmark response
CatReg	Categorical Regression (software)
CFSAN	Center for Food Safety and Nutrition
CNS	Central nervous system
CSAF	Chemical-specific adjustment factor
DAF	Dosimetric adjustment factor
DNT	Developmental neurotoxicity
ECE-1	Endothelin-converting enzyme-1
ELISA	Enzyme-linked immunosorbent assay
FDA	Food and Drug Administration
FQPA	Food Quality Protection Act
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
GD	Gestational day
GLP	Good Laboratory Practices
HA	Health Advisory
HEC	Human equivalent concentration
HED	Human equivalent dose
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
LOAEL	Lowest-observed-adverse-effect-level
MF	Modifying factor
MOE	Margin of exposure
MRL	Minimal risk level
NAAQS	National Ambient Air Quality Standards
NCEA	National Center for Environmental Assessment
NK	Natural Killer
NOAEL	No-observed-adverse-effect-level
OAR	Office of Air and Radiation
OECD	Organisation for Economic Cooperation and Development
OPP	Office of Pesticide Programs

LIST OF ABBREVIATIONS AND ACRONYMS (continued)

OPPTS	Office of Prevention, Pesticides, and Toxic Substances
OSWER	Office of Solid Waste and Emergency Response
OW	Office of Water
P	Parental
PAD	Population adjusted dose
PBPK	Physiologically-based pharmacokinetic model
PFC	Plaque-forming cell
PND	Postnatal day
POD	Point of departure
PRA	Plasma renin activity
RDDR	Regional deposited dose ratio
RGDR	Regional gas dose ratio
RfC	Reference concentration
RfD	Reference dose
SAB	Science Advisory Board
SPF	Specific pathogen free
SRBC	Sheep Red Blood Cells
TSCA	Toxic Substance Control Act
TWA	Time-weighted average
UF	Uncertainty factor

PREFACE

The U.S. Environmental Protection Agency (EPA) Risk Assessment Forum was established to promote scientific consensus on risk assessment issues and to ensure that this consensus is incorporated into appropriate risk assessment guidance. To accomplish this, the Forum assembles experts throughout EPA in a formal process to study and report on these issues from an Agency-wide perspective. For major risk assessment activities, the Forum has established Technical Panels to conduct scientific reviews and analyses. Members are chosen to ensure that necessary technical expertise is available.

The RfD/RfC Technical Panel (hereafter the Technical Panel) was established by the Risk Assessment Forum in early 1999 in response to a request from the Agency's 10X Task Force[1] to the Science Policy Council and the Forum. In the process of developing a strategy for implementing the Food Quality Protection Act (FQPA) relative to protecting children's health and application of the 10X safety factor, the 10X Task Force produced two draft reports, one on toxicology (U.S. EPA, 1999b) and one on exposure data requirements (U.S. EPA, 1999c) that were used by the Office of Pesticide Programs to develop a policy document for implementation of the FQPA safety factor (U.S. EPA, 2002b).

The draft 10X toxicology report (U.S. EPA, 1999b) raised a number of issues that relate to the derivation of the oral reference dose (RfD) and inhalation reference concentration (RfC). Examples of these issues include the following. (1) Appropriate application of a database uncertainty factor (UF) or modifying factor for studies that are considered necessary but are absent or judged inadequate that may show children to be significantly more sensitive or susceptible than adults. Addressing this issue also implicates aspects of other UFs that relate to children's health, including the factor for inter-individual variability in humans (e.g., response of the aged vs. response of the younger adult or child), and the interspecies UF (e.g., young animals vs. young humans). (2) How to account for degree of concern for potential toxicity to children in the RfD/RfC process. Degree of concern, as used in the 10X toxicology report, refers to the characterization of the database as to the likelihood that the agent under review would have effects in humans within the context of dose, route, duration, and timing of exposure. (3) The

[1]The 10X Task Force was created by the EPA Administrator to explore the adequacy of current testing approaches for pesticides for protecting children's health and to recommend approaches for implementing the additional 10X safety factor mandated by the 1996 Food Quality Protection Act.

use of developmental toxicity data as the basis for reference values[2] of chronic duration (RfDs or RfCs) and the appropriate setting of acute, short-term, and longer-term reference values, including the application of developmental toxicity data for these shorter-duration reference values. (4) The appropriateness and/or rationale for adjusting the no-observed-adverse-effect level (NOAEL) or the benchmark dose from developmental toxicity data with inhalation exposures using a concentration-times-time (C x t) adjustment, as is done for other study types.

The Technical Panel also was asked to consider the need for additional toxicity test protocols related to children's health as recommended by the 10X Task Force, when such protocols should be required, and how the data should be interpreted for risk assessment purposes. These additional protocols include (1) collection of toxicokinetic data, both in adults and at different developmental stages; (2) direct dosing of neonates, especially when early exposure is of concern; (3) perinatal carcinogenesis studies and appropriate triggers for when they should be required; (4) developmental immunotoxicity testing and appropriate triggers; (5) advanced developmental neurotoxicity testing, in particular, cognitive testing that is more similar to that used in humans; and (6) exposure assessments that are more compatible with the dose-response assessment. (See Appendix A for more a detailed discussion of the issues raised by the 10X Task Force.)

The Science Policy Council and the Risk Assessment Forum agreed that these issues should be examined—with input from various program offices within the Agency and from the outside scientific/policy community—on a broader scale than just for pesticides. This charge was expanded by the Forum to include a more in-depth review of a number of issues related to the RfD/RfC process, in part because of several other Forum activities that were underway. These activities included development of *Framework for the Harmonization of Cancer and Noncancer Risk Assessment*, revision of *Benchmark Dose Guidance Document*, and revision of *Guidelines for Carcinogen Risk Assessment*. In addition, the RfD/RfC derivation process (Barnes and Dourson, 1998; U.S. EPA, 1994, 2002c) had not been evaluated in detail for a number of years, and several scientific issues concerning children's health, for example, neurotoxicity and immunotoxicity, have become increasingly important in risk assessment. These various but related activities have prompted the need to re-examine the RfD/RfC process and to coordinate these efforts with other related activities. In particular, it was important that efforts continue to focus on moving toward the goal of harmonization of risk assessment

[2]The term reference value is used generically here to refer to values such as the RfD, RfC, acute reference exposure (ARE), Health Advisory (HA), acute exposure guideline level (AEGL), minimal risk level (MRL), or other similar values.

approaches for all health endpoints. This document represents the review and deliberations of the RfD and RfC processes by the Risk Assessment Forum Technical Panel.

EXECUTIVE SUMMARY

This document summarizes the review and deliberations of the Risk Assessment Forum's RfD/RfC Technical Panel and its recommendations for improvements in oral reference dose/inhalation reference concentration (RfD/RfC) process as well as additional efforts that are needed. It discusses revisions to the framework for the derivation of reference values. The document is a review, not guidance, but it does make recommendations that should be considered in the implementation of changes in the current process and/or development of needed guidance.

The Technical Panel reviewed most of the issues relating to hazard characterization for developing reference values and the need for developing reference values for different durations of exposure as well as the process of deriving reference values, but it did not go into detail on the quantitative aspects of the dose-response process, which is being covered in other Forum activities. The Technical Panel views the RfD/RfC process as one that should be continually evolving as new information becomes available and new scientific and risk assessment approaches are developed. This does not mean that current RfDs or RfCs are invalid, but these new scientific issues should be included in the process of re-evaluating of current reference values.

This document reviews and discusses a number of issues and provides conclusions and recommendations that are intended to improve the RfD/RfC process. The Technical Panel has provided specific recommendations for the development of guidance in some cases and more general conclusions and recommendations in others. In the latter cases, the Technical Panel felt that development of specific recommendations was beyond the scope of its efforts or that policies needed to be further developed before specific guidance could be written to implement the recommendations. The document is divided into five chapters:

Chapter 1 provides an introduction, background, purpose, and scope for the project.

Chapter 2 reviews current approaches to developing acute, short-term, and longer-term reference values as well as the chronic reference values, the RfD and the RfC. This chapter incorporates the presentations and discussions on developing less-than-lifetime values from briefings to the Technical Panel and a colloquium held August 2, 2000, and includes discussions of the proposed Acute Reference Exposure (ARE) methodology for acute inhalation exposures, the Acute Exposure Guideline Level (AEGL) Program, the Office of Pesticide Programs' procedures for setting acute and longer-term duration RfDs, the Office of Water's Health Advisories, and the Agency for Toxic Substances and Disease Registry's Minimal Risk Levels.

On the basis of its review of the various approaches to setting acute, short-term, and longer-term reference values, the Technical Panel concurred with the recommendation of the 10X Task Force that such values should be set, where possible, and that they should be incorporated into the Integrated Risk Information System (IRIS) database. In addition, the Technical Panel recommended that this process be done in a consistent manner using standardized definitions for acute, short-term, longer-term, and chronic durations that are consistent with current practice. These values can then be used by various program offices, where applicable. A framework for deriving these additional values is presented in Chapter 4.

Chapter 3 reviews the current Office of Prevention, Pesticides and Toxic Substances' harmonized health effects testing guidelines for the purpose of determining the data available for setting various duration reference values. The intent of this review is not to suggest that additional testing be conducted for each and every chemical in order to fill in the information gaps identified for those organ systems evaluated. Nor is it suggested that alternative testing protocols that are discussed in this chapter should be conducted for every chemical or become part of current toxicology testing requirements or that these alternative protocols are the only options available. Rather, it is the goal of this document to provide a basis for the development of innovative alternative testing approaches and the use of such data in risk assessment, and to then illustrate some aspects of this concept with a few examples. In reviewing the current testing protocols, target organs/systems that are evaluated were reviewed as was the thoroughness of testing with respect to life stage assessment, endpoint assessment, route, timing and duration of exposure, and latency to response. These issues were all considered important in evaluating potentially susceptible subpopulations, including life stages. The testing guideline protocols were reviewed overall for these issues; in addition, four biological systems were evaluated in depth, two that are fairly thoroughly evaluated (the reproductive and nervous systems) and two that are evaluated to a more limited extent (the immune and cardiovascular systems). In each case, an overview of the tests for the particular system is given, as well as a more specific discussion of gaps in life stage of assessment, gaps in assessment endpoints, and gaps in duration and latency assessment.

The Technical Panel has made a number of recommendations concerning toxicity testing, including development of a strategy for approaches to toxicity testing, with guidance on how and when to use existing and newly recommended guidelines; development of guidelines or guideline study protocols that will provide more systematic information on toxicokinetics and toxicodynamics (i.e., mechanism or mode of action), including at different life stages; development of protocols for acute and short-term studies that provide more comprehensive data

for setting reference values; modification of existing guideline study protocols to provide more comprehensive coverage of life stages for both exposure and outcomes; collection of more information from less-than-lifetime exposure to evaluate latency to effect and reversibility of effect; development of guidelines or guideline study protocols to assess immunotoxicity, carcinogenicity, and cardiovascular toxicity at different life stages; and exploration of the feasibility of setting dermal reference values for direct toxicity at the portal of entry, including sensitization.

A primary goal of this review was to provide the basis for recommendations for the development of a strategy for approaches to toxicity testing and for innovative alternative testing approaches to provide data for risk assessment. The Technical Panel is suggesting that alternative strategies and guidance for testing approaches be developed that incorporate information on toxicokinetics and mode of action early in the process, thus allowing a more targeted testing approach. In addition, alternative protocols are discussed that are aimed at more efficient use of animals and resources in combined studies that would provide more extensive data on life stages, endpoints, and other factors not well characterized in current testing approaches. Recommendations are also made about research areas that should be encouraged to aid in better study design and interpretation of data for risk assessment.

Finally, an example of an alternative testing protocol for acute exposure and evaluation that incorporates the types of endpoints and evaluations optimal for setting acute reference values is discussed. Two sample alternative protocols are presented for chronic exposures and options are discussed for combining studies and evaluations to include a wider array of life stage and endpoint assessments.

Chapter 4 discusses a number of modifications to the existing framework for use in deriving reference values, both for the current chronic reference values (RfD and RfC) as well as for acute, short-term, and longer-term reference values. The approach to reference values discussed here is intended for risk assessments of any type of health effect known or assumed to be produced through a nonlinear and/or threshold mode of action (which may include U-shaped or other nonmonotonic dose-response curves as well as thresholds). Thus, the Technical Panel recommends moving away from the dichotomy between "cancer" and "noncancer." The term "noncancer" has been removed from the reference value definition, denoting the move toward defining approaches for low-dose estimation or extrapolation based on mode of action. Two case studies that illustrate many of the concepts discussed in this chapter are presented in more detail in Appendix B. The Technical Panel recommends including the acute, short-term, longer-term, and chronic reference values derived on the basis of the recommendations in this report in

IRIS after appropriate internal, external, and consensus review. Standard exposure durations are proposed, as is a definition for the reference value, including a designation for route and duration of exposure.

The Technical Panel is aware that there will be data limitations for an individual chemical that may preclude development of all four reference values, and it is aware that time and resources need to be considered when implementing these recommendations. The IRIS program has begun to implement a pilot program to test whether development of the expanded array of reference values is practical and can be accomplished without unduly delaying the completion of an IRIS file. As a part of the pilot, the IRIS program will need to identify the methods to be used in deriving these additional values.

The Technical Panel recommends that endpoint-specific reference values should not be developed, including the reference dose for developmental toxicity, RfD_{DT}. Rather, a sample reference value should be calculated for each relevant and appropriate endpoint and these should be considered in the derivation of various duration reference values. The reference values should be derived to be protective of all types of effects for a given duration of exposure.

An expanded approach to the evaluation of studies and characterization of the extent of the database as a whole is recommended; in particular, several factors are discussed that should be considered in a weight-of-evidence approach for characterizing hazard for the population as a whole as well as for potentially susceptible subpopulations. Those considerations for assessing level of concern raised by the Toxicology Working Group of the 10X Task Force have been incorporated into this approach.

In the context of this framework, the Technical Panel recommends a somewhat different approach to characterizing the extent of the database for reference values. Instead of specifying particular studies, this approach emphasizes the types of data needed (both in terms of human and animal data) for deriving reference values, and it recommends the use of a narrative description of the extent of the database rather than a single confidence ranking of high, medium, or low. To characterize the database, the Technical Panel has developed a description of a "minimal" database and a "robust" database as a way of describing the range of data that can be used for deriving a reference value, and the Panel urges the use of a great deal of scientific judgement in the process of summarizing the extent of the database, including its strengths and limitations.

The narrative approach is intended to emphasize the types of data available (both human and animal) as well as the data gaps that could improve the derivation of reference values. This approach should encourage the use of a wider range of information in deriving reference values,

taking into consideration the issues of duration, timing, and route of exposure; the types and extent of endpoint assessments (i.e., structure and function); the life stages evaluated; and the potential for latent effects and/or reversibility of effects.

Dosimetric adjustment of values for deriving a human equivalent concentration (HEC) for inhalation exposure is discussed, as is the derivation of a human equivalent dose (HED) for oral or dermal exposure. The Technical Panel recommends that duration adjustment procedures to continuous exposures based on concentration times time (C x t) be used as a default procedure for inhalation developmental toxicity studies as for other health effects from inhalation exposures. In addition, further evaluation of current dosimetric adjustments for deriving HECs should be pursued to confirm or assess the relevance for population subgroups (particularly for children).

Because of the recommendation for deriving several duration reference values, the Technical Panel recommends that the data for the point of departure (POD) be evaluated on the basis of a comparison of all relevant endpoints carried through the derivation of sample reference values, with selection of the limiting value(s) as the final step rather than on the basis of selection of a single "critical study" and "critical effect." To aid in this evaluation, the use of an exposure-response array is recommended as a visual display of all relevant and appropriate endpoints and durations of exposure in order to determine the range of numerical values for each reference value.

The Technical Panel makes a number of recommendations concerning the application of uncertainty factors (UFs) for reference value derivation. In particular, it is imperative that the IRIS documentation contain a justification for the individual factors selected for each chemical or assessment because rigid application of UFs could lead to an illogical set of reference values. Although default factors of 10 are recommended, with 3 used in place of half-power values (i.e., $10^{0.5}$) when occurring singly, the exact value of the UF chosen should depend on the quality of the studies available, the extent of the database, and scientific judgment. Sound scientific judgment should be used in the application of UFs to derive reference values that are applied to the value chosen for the POD derived from the available database (BMDL, NOAEL, or LOAEL).

The Technical Panel recommends that if there is uncertainty in more than four areas of extrapolation, it is unlikely that the database is sufficient to derive a reference value. Even when there is uncertainty in four areas, the database should be carefully evaluated to determine whether the derivation of a reference value is appropriate. In addition, the Technical Panel recommends limiting the total UF applied to a chronic reference value for any particular

chemical to 3000. This maximum of 3000 applies only to the UFs and does not include the various adjustment factors discussed in Chapter 4.

The intraspecies UF is applied to account for variations in susceptibility within the human population (interhuman variability) and the possibility (given a lack of relevant data) that the database available is not representative of the dose/exposure-response relationship in the subgroups of the human population that are most sensitive to the health hazards of the chemical being assessed. Because the RfD/RfC is defined to be applicable to "susceptible subgroups," this UF was established to account for uncertainty in that regard. In general, the Technical Panel reaffirms the importance of this UF, recommending that reduction of the intraspecies UF from a default of 10 be considered only if data are sufficiently representative of the exposure/dose-response data for the most susceptible subpopulation(s). At the other extreme, a 10-fold factor may sometimes be too small because of factors that can influence large differences in susceptibility, such as genetic polymorphisms. The Technical Panel urges the development of data to support the selection of the appropriate size of this factor, but recognizes that often there are insufficient data to support a factor other than the default.

The Technical Panel urges continued research and evaluation of the similarities and differences between the general population and susceptible subpopulations, particularly children and the elderly, in their responses to particular agents. From such evaluations, the protectiveness of the 10-fold default factor should continue to be assessed. The Technical Panel urges the development of data to support the selection of the appropriate size of this factor, but it recognizes that often there are insufficient data to support a factor other than the default. The database UF is intended to account for the potential for deriving an underprotective RfD/RfC as a result of an incomplete characterization of the chemical's toxicity. In addition to the identification of toxicity information that is lacking, review of existing data may also suggest that a lower reference value might result if additional data were available. Consequently, in deciding to apply this factor to account for deficiencies in the available data set, and in identifying its magnitude, the assessor should consider both the data lacking and the data available for particular organ systems as well as life stages. The Panel considers the purpose of the modifying factor (MF) to be sufficiently subsumed in the general database UF, and recommends that use of the MF be discontinued.

Given that there are several UFs that can be used to deal with data deficiencies as part of the current reference value process, and given that these are assumed to overlap to some extent, the Technical Panel agrees with the 10X Task Force Toxicology Working Group that the current interspecies, intraspecies, and database deficiency UFs, if appropriately applied using the

approaches recommended in this review, will be adequate in most cases to cover concerns and uncertainties regarding the potential for pre- and postnatal toxicity and the completeness of the toxicology database. In other words, an additional uncertainty factor is not needed in the RfC/RfD methodology because the currently available factors are considered sufficient to account for uncertainties in the database from which the reference values are derived (and does not exclude the possibility that these UFs may be decreased *or* increased from the default value of 10). The approach to using chemical-specific data for toxicokinetic and toxicodynamic components of the interspecies UF is part of the current RfC methodology. The Technical Panel encourages the Agency to develop its own guidance for chemical-specific adjustment factors (CSAFs) on the basis of some of the available methodologies (e.g., the International Programme on Chemical Safety [IPCS]).

Several other issues discussed by the Technical Panel were considered more appropriate for deliberation by other panels/committees, for example, further consideration of the use of BMD modeling approaches for deriving reference values; harmonization of the approaches for HEC and HED derivation for all types of health effects; further evaluation of approaches such as probabilistic analysis for characterizing variability and uncertainty in toxicity reference values; further evaluation of appropriate adjustment of doses for duration of exposure for acute toxicity data; and further evaluation of duration adjustment for short-term and longer-term reference values analogous to the subchronic-to-chronic duration UF for chronic reference values.

Chapter 5 summarizes the recommendations of the Technical Panel.

1. INTRODUCTION, PURPOSE, AND SCOPE

The RfD/RfC Technical Panel (hereafter the Technical Panel) was established by the U.S. Environmental Protection Agency's (EPA's, or the Agency's) Risk Assessment Forum in early 1999 to review the current oral reference dose (RfD) and inhalation reference concentration (RfC) processes, in particular with respect to how well children and other potentially susceptible subpopulations are protected; to consider new scientific issues that have become more important and of greater concern in risk assessment; and to raise issues that should be explored or developed further for application in the RfD/RfC process. This document summarizes the review and deliberations of the Technical Panel and its recommendations for improvements in the process as well as additional efforts that are needed. It discusses revisions to the framework for the derivation of RfDs and RfCs. The document is a review, not guidance, but it does make recommendations that should be considered in the implementation of changes in the current process and/or development of needed guidance.

Many of the recommendations made in this report are consistent with the Agency's commitment to harmonization of health risk assessment procedures, including the harmonization of approaches for noncancer and cancer endpoints, and to making efficient use of animal testing to achieve this goal. As noted in several places in the document, all such topics have not been discussed and resolved by the Agency. For instance, the differences in scaling factors used for cancer and noncancer derivations from oral exposure data are raised as an issue that has not been resolved; thus, there will likely be a need for revised or further guidance on this issue.

Although mixtures or multiple chemical exposures are not specifically discussed in this review, most of the recommendations are applicable to the approach to risk assessment of mixtures. The Agency's mixtures risk assessment guidelines should be consulted for issues specific to the evaluation of mixtures (U.S. EPA, 1986, 2000a). In addition, the Agency has recently issued the draft *Framework for Cumulative Risk Assessment* (U.S. EPA, 2002a), which deals with the issue of multiple stressors and their overall impacts on exposure-effect relationships. The risk assessment approaches discussed within this framework are likely to be the subject of further guidance as well.

The Technical Panel attempted to review most of the issues relating to hazard characterization for developing reference values, to the need for developing reference values for different durations of exposure, and to the process of deriving reference values. The Technical Panel did not go into detail on the quantitative aspects of the dose-response process, as this is being covered in other Forum activities (e.g., the benchmark dose [BMD] guidance document

and the quantitative dose-response aspects of the cancer guidelines revision process). The Technical Panel approached its review from the point of view that the RfD/RfC process has been and should be a continually evolving process. Thus, as new information becomes available and new scientific and risk assessment approaches are developed, they are incorporated into new RfDs and RfCs as these values are developed or as current RfDs and RfCs are reevaluated. This process of incorporating new science does not invalidate current RfDs or RfCs, because consideration of these new scientific issues is included in the reevaluation of current values; higher or lower values or, in some cases, no change in the current value may result.

This report provides conclusions and recommendations that are intended to improve the RfD/RfC process. The audience for this review is primarily the Integrated Risk Information System (IRIS) program, IRIS chemical managers, and other scientists within the Agency who are involved in developing the RfDs and RfCs, as well as IRIS users and the program offices within EPA that develop RfDs and RfCs or similar values (see Chapter 2), particularly resource managers who may be impacted by the potential for additional workload due to several of the recommendations. The Technical Panel has provided specific recommendations for guidance in some cases and more general conclusions and recommendations in others. In the latter cases, the Technical Panel felt that development of specific recommendations was beyond the scope of its efforts or that policies needed to be further developed before specific guidance could be written to implement the recommendations.

The methodology recommended in the RfD document is considered generally applicable to both cancer and noncancer endpoints where dose-response relationships are thought to be either nonlinear or consistent with a threshold. Although the emphasis in this document is on the calculation of RfDs and RfCs, the same processes and considerations are applicable to the margin of exposure (MOE), as discussed in the draft cancer risk assessment guidelines (U.S. EPA, 1999a).

The Technical Panel discussed a number of issues concerning a revised framework for the RfD/RfC process, with particular emphasis on the extent to which children and other potentially susceptible subpopulations are considered. The next three chapters summarize these issues, and several recommendations are made. Chapter 2 reviews current approaches to developing acute, short-term, and longer-term reference values as well as the chronic reference values, the RfD and the RfC. Chapter 3 reviews the current testing guidelines with respect to life stage assessment and discusses the gaps in life stage assessment, endpoint assessment, and assessment of duration and latency. Alternative testing protocols and strategies as options for combining studies and evaluations are discussed.

Chapter 4 provides constructive commentary on the current framework used in deriving reference values and on the need and possibilities for calculating reference values for different durations and routes of exposure. In addition, an expanded approach to evaluating studies and characterizing the extent of the database as a whole is presented and discussed, including dosimetric adjustment, the application of uncertainty factors (UFs), and derivation of sample reference values for each appropriate and relevant endpoint to aid in selecting the point of departure (POD) for deriving reference values.

The final chapter (Chapter 5) summarizes all of the recommendations of the Technical Panel. Two case studies that illustrate several of the recommended changes are also included as Appendix B.

2. REVIEW OF THE CURRENT USE OF ACUTE, SHORT-TERM, AND LONGER-TERM REFERENCE VALUES

The Technical Panel considered the recommendation of the 10X Task Force that acute, short-term, and longer-term reference values as well as chronic reference values should be set for environmental agents (see Appendix A). It is likely that the endpoints critical for setting acute, short-term, and longer-term reference values may differ from those for setting chronic RfDs and RfCs, although studies that use acute and short-term exposure conditions from which the appropriate data for many types of effects could be derived are not often available. Data on acute and short-term health effects must often be derived from observations after the first exposure in a repeated-exposure testing protocol.

Several acute and short-term values currently are set for various chemical types and media. For example, acute and chronic oral RfDs are set for pesticides, with some intermediate values set for occupational and residential pesticide exposures. Health advisories (HAs) of several durations have been developed for drinking water. In addition, the Office of Solid Waste and Emergency Response, the Office of Prevention, Pesticides, and Toxic Substances (OPPTS), and other program offices and regional offices use values derived through the interagency acute exposure guidelines (AEGL) process for emergency response planning. The National Center for Environmental Assessment (NCEA) is currently developing the acute reference exposure (ARE) methodology for acute inhalation exposures. These developments are reviewed in more detail below.

2.1. REVIEW OF CURRENT LESS-THAN-LIFETIME REFERENCE VALUES

The Technical Panel was briefed by representatives of several Agency offices on the methods currently used to set various less-than-lifetime reference values. Subsequently, on August 2, 2000, a Risk Assessment Forum colloquium was held on this topic (CDM Group, Inc., 2000). Each of the methods was presented and discussed. In addition, a recommendation by the Technical Panel to begin deriving acute, short-term, and longer-term reference values as well as chronic values and to standardize the definitions for each duration was presented and discussed. Each method presented is summarized below.

2.1.1. Acute Reference Exposure (ARE) Methodology

The ARE methodology is being developed at the request of the EPA's Office of Air and Radiation. It is intended for development of reference values for acute inhalation exposures of

24 hours or less. The criteria air pollutants are not included, because they are assessed within the National Ambient Air Quality Standards (NAAQS) setting process.[1] The ARE is defined as an inhalation exposure of 24 hours or less that is not likely to cause noncancer adverse effects. The ARE can be applied to intermittent exposures or to a continuous exposure. AREs are being developed in order to address the acute risk aspects of risk-related provisions of the hazardous air pollutant sections of the 1990 Clean Air Act Amendments. The ARE methodology is described in a 1998 EPA external review draft document (U.S. EPA, 1998a). The method builds on the procedures of the RfC methodology.

The ARE methodology includes three approaches in order to accommodate the varying types of data available for acute exposure. The first two approaches, the no-observed-adverse-effect level (NOAEL) and the benchmark concentration (BMC) are familiar. The third approach, categorical regression (CatReg), is newer. The NOAEL approach is useful for chemicals that have limited available data and for which no or limited dose-response relationships have been established. The BMC approach is suitable for analysis of studies that establish dose-response relationships. The CatReg approach requires multiple studies that report not only dose and response, but also duration; it is most applicable for data-rich chemicals. A feature of the CatReg approach is that effects data are grouped into severity categories (e.g., mild or severe to lethal) to which sophisticated regression procedures are then applied.

Adjustments for deriving ARE values of different durations (e.g., 15 minutes or 8 hours) are made differently for the CatReg approach than for the NOAEL and BMC approaches. For any approach, the preferred adjustment procedure is to use a pharmacokinetic model, if available. When the NOAEL or BMD approach is used, the default procedure is to use the multiple of concentration times time (C x t) (C^n x t = k; ten Berge et al., 1986) to extrapolate from short to long duration and to use the same concentration as obtained for long duration to extrapolate from long to short duration. When more than one duration is available, interpolation is performed. When the CatReg approach is used, the procedure involves reading the values directly from the concentration duration curve that is generated by the CatReg software. These approaches are explained more fully and illustrated in Chapter 4.

A minimal data set has not been defined for the ARE. Also, extrapolation from the oral to the inhalation route of exposure is not addressed in the ARE approach. UFs in the ARE approach include a lowest-observed-adverse-effect level- (LOAEL-) to-NOAEL UF of 10 and a

---

[1]Criteria air pollutants are those air pollutants for which NAAQS have been established under the Clean Air Act; at present, the six criteria air pollutants are particulate matter, ozone, carbon monoxide, nitrogen oxides, sulfur dioxide, and lead.

default value of 10 for interspecies and for intraspecies extrapolation. No factor is assigned for database inadequacies and study quality.

In 1998, the EPA Science Advisory Board (SAB) reviewed the ARE methodology document and made a number of comments that addressed, among other things, issues about the NOAEL and BMC approaches, the need for addressing protection of children, the dosimetry adjustment and duration extrapolation, and the CatReg approach. The SAB discussed the fact that the CatReg model, as currently set up, forces parallelism of the concentration-duration curves for the various severity categories. In addition, there were concerns about judging severity categories across various target organs and species, and there was discussion about the reliability of the confidence limits around the maximum likelihood estimate and about the appropriateness of the approach used to accommodate group versus individual data. This methodology has since (March 2001) undergone an Agency review by the Risk Assessment Forum. The principal comments from this review concerned reevaluation of whether CatReg should remain as an approach in the ARE methodology and further evaluation of the procedures for cross-species dosimetry adjustment. Revision of the ARE methodology is currently underway. In addition to revising the ARE methodology and CatReg software documents, NCEA-Research Triangle Park will develop a framework for adding AREs to the IRIS database.

2.1.2. Acute Exposure Guidelines (AEGL) Program

The primary purpose of the AEGL program is to develop guideline levels for once-in-a-lifetime short-term exposures to airborne concentrations of acutely toxic chemicals (NRC, 2000). AEGLs are needed for a wide variety of emergency planning, response, and prevention applications. AEGLs represent threshold exposure limits for the general public and are applicable to emergency exposure periods ranging from 10 minutes to 8 hours. Specific values are set for 10 minutes, 30 minutes, 1 hour, 4 hours, and 8 hours. It is believed that the recommended exposure levels are applicable to the general population, including infants and children and other individuals (e.g., asthmatics) who may be sensitive or susceptible. It is recognized that certain individuals who may be subject to unique or idiosyncratic responses could experience the effects described at concentrations below the corresponding AEGL level.

The AEGL-1, AEGL-2, and AEGL-3 levels are distinguished by varying degrees of severity of toxic effects. With increasing airborne concentrations above each AEGL level there is a progressive increase in the likelihood of occurrence and the severity of effects described for each level.

AEGL-1 is the airborne concentration of a substance above which it is predicted that the general population, including susceptible individuals, could experience notable discomfort, irritation, or certain asymptomatic, nonsensory effects. However, the effects would not be disabling and would be transient and reversible upon cessation of exposure.

AEGL-2 is the airborne concentration of a substance above which it is predicted that the general population, including susceptible individuals, could experience irreversible or other serious, long-lasting adverse health effects or an impaired ability to escape.

AEGL-3 is the airborne concentration of a substance above which it is predicted that the general population, including susceptible individuals could experience life-threatening health effects or death.

Airborne concentrations below AEGL-1 represent exposure levels that could produce mild and progressively increasing odor, taste, and sensory irritation or certain asymptomatic, nonsensory effects.

UFs are used for extrapolations. If there are no appropriate human data, an interspecies UF of 1, 3, or 10 is used. The factors considered when deciding on a specific value include (1) the species tested (type, appropriateness, and range), (2) the toxicological endpoint observed and the likely mechanism of action, (3) the range of response in the species tested, (4) the variability of response among the species tested, and (5) pharmacokinetic differences among the species tested. An intraspecies UF of 1, 3, or 10 is also used. The factors considered when assigning a specific value include (1) the toxicological endpoint observed and the likely mechanism of action, (2) the range of response among humans and subpopulations, and (3) pharmacokinetic differences among individuals. Individual factors of 3 are often used to ensure that the final values are not overly conservative.

Adjustment for duration is conducted using the equation $C^n \times t = k$. If data are available for the endpoint of concern, the value of n is derived from regression analysis. If data are not available for the endpoint of concern, the value of n is usually derived from lethality data by regression analysis and used for the other endpoints. If the study duration is greater than 1 hour, the 10-minute value is usually assigned equal to the 30-minute value. If no data are available to derive a value of n, a value of 3 is used to extrapolate to shorter durations, and a value of 1 is used to extrapolate to longer durations. As mentioned above, this procedure is further explained and illustrated in Chapter 4.

2.1.3. Office of Pesticide Programs (OPP) Procedures for Setting Acute and Intermediate RfDs

OPP developed methodologies for acute dietary as well as occupational and residential risk assessments during the process of re-registration following the 1988 revision to the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). In 1998, a guidance document, *Toxicology Endpoint Selection Process* (U.S. EPA, 1998b), was presented to the FIFRA Scientific Advisory Panel for review and comment. This document, which provided the basis for procedures that are still in place, describes toxicology endpoint selection for less-than-lifetime dietary and occupational/residential risk assessments for pesticides. It includes guidance on evaluating toxicity studies that are relevant for use, selecting appropriate endpoints for hazard identification, the process of hazard identification, the influence of dermal absorption in hazard identification, the criteria for the use of the NOAEL and the LOAEL, and the use of MOEs in risk assessments. Since this guidance was first issued, some changes have evolved, such as the replacement of the acute MOE with the acute RfD and the addition of standard consideration of short- and intermediate-term incidental nondietary ingestion exposures for toddlers.

Toxicology Endpoint Selection Process (U.S. EPA, 1998b) describes the types of studies that are most likely to provide appropriate endpoints for the various exposure durations and risk assessments that will be conducted for each pesticide. OPP can rely on the availability of a wide variety of standard guideline toxicity studies from which to select endpoints because such studies are required by regulation for any pesticide registration (40 CFR Part 158). Additionally, OPP considers other sources of toxicology data, such as studies published in the open literature, as appropriate.

For the establishment of the acute RfD, OPP uses a weight-of-evidence approach in evaluating all the available data. Three guideline studies have been found to be particularly useful by OPP: the acute neurotoxicity study, the prenatal developmental toxicity study, and the developmental neurotoxicity (DNT) study.

Acute effects from subchronic and chronic dietary studies are also used in the establishment of the acute dietary RfD. Careful scrutiny of toxicological data from early in the first week of treatment can sometimes identify effects that can be described as acute. However, for a number of reasons, this option has not often been used. These reasons include the absence of detailed toxicological observations other than morbidity and mortality checks in subchronic and chronic studies before the end of the first week of treatment (i.e., after 7 days of treatment), the nature of the dietary exposure (i.e., each daily exposure results from an extended period of nightly feeding rather than from a discrete acute dose), and the possibility that apparent adverse

effects during the first week of treatment may be related to palatability issues as the animals adjust to treated feed.

OPP does not calculate short- or intermediate-term references doses. However, risk assessments are conducted for incidental nondietary ingestion exposures to toddlers—a very specific population subgroup—that result from the use of a pesticide in and around the home or other nonoccupational sources such as schools, parks, and golf courses. The post-application risk assessment considers or accounts primarily for incidental ingestion of (1) the dry pesticide materials (granules or pellets) used to treat outdoor residential areas, (2) pesticide residues in soil that are ingested by toddlers who play in treated areas (e.g., yards, gardens, playgrounds) as a result of normal mouthing activities, and (3) pesticide residues that are transferred to the skin of toddlers playing in treated areas and are subsequently ingested as a result of hand-to-mouth transfer. These risk assessments consider short-term (1 day to 1 month) and intermediate-term (1–6 months) exposure durations. Risks are expressed as MOEs. The MOE approach is used because these exposures are considered to be nondietary in source and are based on high-end values or (when adequate site- or chemical-specific field data are unavailable) on assumptions.

OPP also conducts short-term, intermediate, and long-term (longer than 6 months) dermal and inhalation risk assessments for occupational and residential exposures. The MOE approach is also used to calculate the risk for these nondietary exposure scenarios. A difficulty that OPP often faces when conducting these risk assessments is that dermal absorption and inhalation toxicity data are often not available for food-use pesticides; in that case, appropriate assumptions are applied, and the available oral toxicity data are converted for use in dermal and inhalation risk assessment.

Toxicology Endpoint Selection Process (U.S. EPA, 1998b) does not address the use of UFs in acute dietary risk assessment. In practice, however, the same 10-fold inter- and intraspecies UFs are used in calculating the acute dietary RfD as are used for the chronic RfD. Other standard UFs may be used when appropriate (e.g., the LOAEL-to-NOAEL threefold factor). Others are not appropriate for an acute risk assessment, for example, the threefold subchronic-to-chronic factor. However, no standard set of "core" studies has been defined for acute dietary risk assessment; therefore, a database UF is not used. If appropriate endpoints and doses cannot be selected for acute dietary risk assessment from the studies in the database, then an acute RfD is not calculated.

2.1.4. Office of Water (OW) Health Advisories (HAs)

The OW HA program was initiated in 1978 to provide guidance on unregulated contaminants found in drinking water. Since then, HAs have also been developed for regulated

contaminants. HAs are derived for contaminants that are known to or are likely to occur in drinking water and that may cause adverse, noncarcinogenic health effects (Orme and Ohanian, 1991). The approach for developing HAs is based on recommendations from the National Academy of Sciences (NAS, 1977). HAs are developed for specific exposure durations (1 day, 10 days, longer-term, and lifetime) that reflect different emergency contamination situations. HAs are not legally enforceable, but they do serve as technical guidance to assist in emergency spills or contamination situations or for determining unreasonable risks to health under sections 1415 and 1416 of the Safe Drinking Water Act. They also are issued at the request of State or local governments or to fill a need for criteria, guidelines, or standards. HAs undergo scientific peer review and can function as a preliminary risk assessment, if necessary.

The following assumptions are used in setting the various HAs. The 1-day HA represents a concentration of the contaminant in drinking water that is considered protective of adverse noncancer health effects in a 10 kg child. The 10 kg child serves as the protected individual for the less-than-lifetime HAs because a child of this size is likely to receive a greater dose on a mg/kg basis. This 1-day HA can serve as a guideline for each day for up to 5 consecutive days of exposure. The 1-day HA is usually derived from experimental studies of 7 days duration or less.

The 10-day HA is considered protective of these effects in a 10 kg child for each day for up to 14 days of continuous exposure and may be based on experimental studies of 30-day duration or less.

The longer-term HA, which is based on subchronic exposure studies covering 10% of an animal's lifetime, is considered protective of an exposure period in humans of up to 7 years (i.e., 10% of an individual's lifetime). The longer-term HA is developed to protect both a 10 kg child and a 70 kg adult.

The lifetime HA is considered protective of lifetime exposures and is usually based on chronic or subchronic or other more relevant experimental data. The Lifetime HA is based on the chronic oral RfD, adjusted for a 70 kg adult drinking 2 L water per day; the value is apportioned by a relative source contribution, for example, 20% of the toxicant represented by intake of water.

HA levels are generally based on available, well-conducted studies that involve humans or animals. Data from drinking water studies are preferred; however, data from dietary or gavage studies can also be used. In the absence of oral data, studies by other routes of exposure, such as inhalation or injection, are considered. Following identification of an appropriate study to develop a HA, the NOAEL or the LOAEL is adjusted for water consumption by the protected

individual. For a child, the assumed water consumption level is 1 L/day; for an adult, 2 L/day is used.

When data are absent for setting a 1-day or a 10-day HA, OW uses scientific judgment on how to handle any given situation on the basis of the overall weight of evidence. In the absence of short-term toxicity studies, a subchronic or chronic study may be used to develop a less-than-lifetime HA. Given the pressure under which HAs need to be calculated, many assessments are based on whatever toxicological data are available and on scientific judgment. Although this may be an overly conservative approach, OW considers the error to be protective of public health.

OW applies the same factors for minimum data as those outlined in the Agency's RfD methodology. For example, in emergency situations, missing data are accounted for by applying another factor of 3 or 10. Or, for instance, where inhalation data might be applied to estimate a HA based on water consumption, a factor may be applied to account for differences in absorption. Judgments based on toxicokinetic and toxicodynamic considerations are reached through intensive consultation.

Calculation of HAs is straightforward and familiar, and in most cases the NOAEL/UF approach is used. For each of the less-than-lifetime HA values, it is assumed that all of an individual's exposure to a contaminant comes from a drinking water source. The calculation of the lifetime HA differs from that of the less-than-lifetime values in that a relative source contribution factor is included. This factor adjusts the exposure to reflect the portion that is likely to be contributed from drinking water. Unless actual exposure data are available, a default factor of 20% is used to reflect the assumed contribution to exposure from drinking water. Also, in cases where there is limited evidence suggesting a carcinogenic potential of a contaminant, an additional "policy" factor of 10 is applied in calculating the lifetime HA.

The methodology for developing HAs was reviewed by the SAB and the FIFRA Scientific Advisory Panel in 1986. Each HA that is developed undergoes external peer review and Agency review before it is released to the public. The availability of the HAs is announced in the *Federal Register* and distributed through the Safe Drinking Water Hotline and the Water Docket and by the Office of Science and Technology in OW. In addition, HAs have been published in a collection of books and are available in English, Japanese, and Italian.

2.1.5. Agency for Toxic Substances and Disease Registry (ATSDR) Minimal Risk Levels (MRLs)

The ATSDR is tasked with establishing MRLs, which are defined as

> " *... an estimate of daily human exposure to a hazardous substance that is likely to be without appreciable risk of adverse noncancer health effects over a specified route and duration of exposure.*"

MRLs are considered by ATSDR to be substance-specific estimates intended to be screening levels in identifying contaminants and potential health effects that may be of concern; they do not define clean-up or action levels. The derivation procedures for MRLs have many similarities and parallels to the derivation of RfDs and RfCs; MRLs are based on careful scientific consideration of noncancer health effects only, not on consideration of cancer effects. A list of various procedural specifics employed in the derivation of MRLs, including specific effects and the level of severity, is codified in a *Federal Register* notice (ATSDR, 1996). The definition of an MRL differs expressly from EPA's definition of an RfD or an RfC in that both route and duration are included. The current routes of concern for MRL derivation are oral and inhalation (but not dermal).

The EPA procedures and methodologies discussed above address the issue of duration through a variety of extrapolation procedures. For MRLs, however, duration is addressed by providing for the designation of MRLs in three different duration categories: acute = $\leq$14 days, intermediate = 15–364 days, and chronic = $\geq$365 days. These duration categories are absolute and apply to all species, regardless of relative life span. Thus, it is possible for a contaminant to have a total of six different MRL values: two routes by three different durations.

The use of UFs is a parallel practice in RfD/RfC and MRL derivation. The UFs used by ATSDR are intraspecies 1, 3, 10; interspecies 1, 3, 10; and LOAEL/NOAEL 3, 10. The modifying factor (MF) can include database considerations, that is, deficiencies in the data or overestimates from bioaccumulative chemicals.

2.2. SUMMARY OF CURRENT METHODS FOR SETTING ACUTE, SHORT-TERM, AND LONGER-TERM REFERENCE VALUES

In summary, several methods are used by various EPA programs for setting acute, short-term, and longer-term reference values. The definitions for each of the durations used for the methods reviewed are included in Table 2-1. Because there are some differences in these

Table 2-1. Duration definitions used for various reference values

Reference value duration	Definition
Acute	
ARE	Inhalation single continuous exposure values for durations ≤ 24 hrs (to be protective of intermittent exposures)
AEGL	10 and 30 min; 1, 4, and 8 hrs
OPP acute RfD	Maximum 1-day dietary exposure
OW 1-day HA	1 day (5 successive daily doses)
ATSDR acute MRL	≤14 days
Standardized definition[a]	24 hrs or less
Short-term	
ARE	NA
AEGL	NA
OPP short-term RfD	1 day–1 month
OW 10-day HA	10 days (7–14 successive daily doses)
ATSDR MRL	NA
Standardized definition[a]	>24 hrs up to 30 days
Longer-term	
ARE	NA
AEGL	NA
OPP intermediate RfD	1–6 months
OW longer-term HA	Approximately 10% of life span in humans (90 days to 1 year in test species)
ATSDR intermediate MRL	15–364 days
Standardized definition[a]	>30 days up to approximately 10% of the life span in humans (>30–90 days in typically used laboratory species)

[a] See Chapter 4 for further discussion of these definitions.

definitions, standardized definitions were discussed at the Risk Assessment Forum Colloquium (CDM Group, 2000), and these are shown in Table 2-1. Definitions for durations are further discussed in Chapter 4.

A comparison of the UFs applied for various reference values is shown in Table 2-2. Although there is some variation in the UFs applied, those for animal-to-human extrapolation (U_A), for within-human variability (U_H), and for LOAEL-to-NOAEL (U_L) are fairly consistent. Less consistent is the way in which database deficiencies (U_D) are taken into consideration, particularly for pesticides where the Food Quality Protection Act (FQPA) safety factor is used to account for deficiencies in the database related to children's health risks.

Duration extrapolation for each of these values was also reviewed. Some type of duration adjustment of the NOAEL or the BMD is done for the ARE and the AEGL methods, and there appears to be consistency in the use of C^n x t for extrapolating from shorter to longer exposures but in using the same value (i.e., no duration adjustment) when extrapolating from longer to shorter exposures. Duration extrapolation is not done for the OPP RfDs, the OW HAs, or the ATSDR MRLs.

2.3. RECOMMENDATION

On the basis of its review of the various approaches to setting acute, short-term, and longer-term reference values, the Technical Panel concurred with the recommendation of the 10X Task Force that acute, short-term, and longer-term reference values should be set, where possible, and that they be incorporated into the IRIS database. In addition, the Technical Panel recommended that these values be set in a consistent manner, using standardized definitions for acute, short-term, longer-term, and chronic durations that are consistent with current practice. These values can then be used by various program offices, where applicable. A scheme for deriving these additional values is presented in Chapter 4.

Table 2-2. Uncertainty/safety factors for various reference values

Reference value	UF[a]				FQPA[b]
	U_A	U_H	U_L	U_D	
ARE	1, 3, 10	1, 3, 10	1, 3, 10	ND	NA
AEGL	1, 3, 10	1, 3, 10	3[c]	ND[d]	NA
OPP acute and intermediate RfDs	10	10	3, 10	ND[e]	10±
OW HAs	1, 3, 10	1, 3, 10	1, 3, 10	case-specific	NA
ATSDR MRLs	1, 3, 10	1, 3, 10	1, 3, 10	ND[d]	NA

[a] Uncertainty factors: U_A = animal-to-human; U_H = within-human variability; U_L = LOAEL-to-NOAEL; U_D = database deficiency.
[b] Additional safety factor required under FQPA.
[c] Endpoint = lethality, not really a LOAEL-to-NOAEL adjustment in this case.
[d] Database deficiencies considered, and a factor may be included for intermediate RfDs if, for example, there is no reproduction and fertility study.
[e] Overlaps with the FQPA safety factor (see U.S. EPA, 2002b)

ND = not done
NA = not applicable

3. REVIEW OF TESTING GUIDELINES WITH RESPECT TO LIFE STAGE ASSESSMENT

As a first step in determining the data necessary for setting various duration reference values for protecting potentially susceptible subpopulations, the Technical Panel reviewed the current OPPTS Series 870 health effects testing guidelines[2] to determine what information is gathered in these studies. The intent of this review is not to suggest that additional testing be conducted for each and every chemical in order to fill in the information gaps identified for those organ systems evaluated. Nor is it suggested that the alternative testing protocols that are discussed in this chapter should be conducted for every chemical or become part of current toxicology testing requirements or that these alternative protocols are the only options available. Rather, it is the goal of this document to provide a basis for the development of innovative alternative testing approaches and the use of such data in risk assessment and to then illustrate some aspects of this concept with a few examples.

Development of a toxicology testing paradigm that is based not on rigid conformance to a list of required guideline screening studies but rather on the application of knowledge about the chemical is encouraged. Under such a paradigm, both the selection of studies that would be required as well as the design of the tests themselves could be influenced by other substantive and reliable information about the chemical. For example, the incorporation of toxicokinetic and mode-of-action data early in the development of the testing strategy for a chemical would provide particularly valuable direction for development of research protocols.

Other input could include toxicity and dose-response data from other guideline or nonguideline studies, in vitro screening assays, structure-activity relationships, studies that examine age-related sensitivity or susceptibility to chemical exposure, and information on potential or actual exposure to humans. These data could be used to inform a more targeted approach in the design of individual studies or of an overall testing strategy and might in some cases result in a reduction in the number of animals used in testing or support a position that a traditionally required toxicology test should be waived.

The purpose of the review of the current OPPTS guidelines was to understand which target organ systems are evaluated in current testing protocols and how thorough the testing protocols are with respect to life stage assessment; endpoint assessment; route, timing, and

---

[2]The guidelines are available on the OPPTS web page (http://www.epa.gov/docs/OPPTS_Harmonized/870_Health_Effects_Test _Guidelines/Series/).

duration of exposure; reversibility; and latency to response. These issues were all considered of importance in evaluating potentially susceptible subpopulations, including children. The following sections give an overview of the current testing protocols evaluated in this way and, for certain organ/functional systems, provide a more in-depth analysis as to whether and how current protocols address these issues. The organs/functional systems that were examined in greater detail included the reproductive and the nervous systems, which were selected to represent systems that are thought to be rather well-evaluated. The immune and the cardiovascular systems were selected for review because the current evaluation of these systems is limited. It should be noted that testing guidelines were not originally designed with a focus on evaluations of different life stages or different durations of exposure. Therefore, a number of gaps in life stage assessment, endpoint assessment, timing and duration of exposure, reversibility, and latency to response were noted for each organ system that is reviewed in depth.

The last section provides recommendations for alternative testing approaches that are designed to make more efficient use of animals and resources in combined studies that would provide more extensive data on life stages, endpoints, and other factors not well characterized in current testing approaches.

3.1. EVALUATION OF CURRENT GUIDELINE TESTING PROTOCOLS

The following tables and figures summarize the exposures and endpoints covered in current testing guidelines, what is covered for each organ system/endpoint measured, and the relative depth of evaluation for each system/endpoint. In addition, the life stages covered by exposures and outcomes are illustrated. The discussions that correspond to the figures give an overview of the tests that are currently available and the gaps in assessment of life stages, endpoints, timing and duration of exposure, and latency to response. Together, these analyses provide a clear picture of the testing guidelines currently available, the systems/endpoints measured, the life stages during which exposures and outcomes are measured, the timing and duration of exposures included, and the degree of detail covered for both structural and functional outcomes.

In order to make comparisons among laboratory animal species and humans in terms of life stages covered, the approximate ages that correspond to specific events or life stages (e.g., birth, weaning, puberty, etc.) in different species are shown in Table 3-1, and these events/life stages are indicated in the figures. In a few cases, no data could be found on appropriate ages corresponding to particular life stages. In particular, the ages for mature adults and older adults often were not available, and there is some controversy about what constitutes old age in today's

Table 3-1. Approximate age at equivalent life stages in several species

Rat		Mouse		Rabbit		Beagle dog		Human	
Life stage	**Age**	**Life stage**	**Age**	**Life stage**	**Age**	**Life stage**	**Age**	**Life stage**	**Age**
Embryonic	GD 0–16	Embryonic	GD 0–15	Embryonic	GD 0–19	Embryonic	GD 0–30?	Embryonic	GD 0–58
Fetal[a]	GD 16–22 (22–23 days)	Fetal	GD 15–20 (18–22 days)	Fetal	GD 19–32 (30–32 days)	Fetal	GD 30–63 (53–71 days)	Fetal	GD 58–267
Neonate[b]	PND 0–14	Neonate	PND 0–14	Neonate	PND 0–21?	Neonate	PND 0–21	Neonate	PND 0–30
Weaning[c]	PND 21	Weaning	PND 21 (19–28)	Weaning	PND 42 (42–56)	Weaning	PND 42	Infancy	PND 30–1 yr
								Toddler	2–3 yrs
Young	PND 22–35	Young	PND 21–35	Young	PND 42–?	Young	1.5–5 mos	Preschool	3–6 yrs
								Elementary school age	6–12 yrs
Puberty	PND 35–60	Puberty	PND 35–?	Puberty	3–8 mos	Puberty	5–7 mos	Adolescence	12–21 yrs
Sexual maturity	2.5–3 mos	Breeding age	1.5–2 mos	Breeding age	6–9 mos	Breeding age	12 mos	Young adult	21–40 yrs
Mature adult	5–18 mos	Mature adult		Mature adult		Mature adult		Mature adult	40–65 yrs?
Old adult	18 mos–2 yrs+	Old adult		Old adult		Old adult	~15 yrs	Old adult	>65 yrs?

[a] Range of gestation length in parentheses.
[b] Some neonatal events in rodents occur in utero in humans.
[c] Range of weaning ages in parentheses.

GD = gestation day
PND = postnatal day

population. A background paper on aging commissioned as part of this review discusses this issue to some extent (Versar Inc., 2001a). In animal studies, the use of dietary restriction has been shown to affect aging and life span to a significant extent, so the issue of what constitutes an older animal is also somewhat controversial.

3.1.1. Exposures and Endpoints Related to General Toxicity Testing

Table 3-2 provides an overview of the biological systems and other endpoints that are evaluated by routine toxicity test designs. The table includes all of the routine test designs that are available in Agency testing guidelines for evaluating toxicity and most of the test designs that focus on specific biological functions. The acute and subchronic studies are intended to give general information on the potential toxicity of an agent by screening the major organ systems, in particular, the liver, the kidney, and the gastrointestinal tract. This information can be used to determine where to look in more detail at specific organ system structure and function. The chronic studies, which are usually done in combination with a carcinogenicity study, evaluate general toxicity in all major organ systems. Several testing guidelines have been developed with the idea that certain systems should be evaluated frequently in more detail (e.g., the nervous system) or that the general toxicity studies do not provide any indication of a potential for effects (e.g., reproductive and developmental toxicity studies). More detailed information about specific aspects of guideline test designs for certain systems (e.g., life stages covered, exposure periods, outcomes measured, etc.) is included in the figures.

Table 3-2 is shaded and marked to indicate the extent of the evaluation of a particular system/endpoint within a particular test design. XXX indicates that the system/endpoint is a primary focus of the particular test design and that detailed assessment of the dose-response relationship of an exposure is carried out within some defined life stage and exposure period for major elements of the system/endpoint. XX indicates those systems/endpoints for which some histopathology or clinical measure of system function is carried out. X indicates those systems/endpoints that are assessed in some observational or gross manner. "0" indicates that the system/endpoint cannot be included, generally because of the design of the test. Blank cells indicate that the system/endpoint is not presently included but could be if the test design were altered appropriately.

It is obvious from the table that few systems/endpoints are examined in any significant detail. The systems/endpoints under the acute test designs are for the most part observational in nature. The acute inhalation toxicity with histopathology guideline (40 CFR 799.9135) was developed under the Toxic Substances Control Act for characterizing the exposure-response

Table 3-2. Systems/endpoints evaluated by routine toxicity guideline testing protocols[a]

Guideline[b]	Systems															Other endpoints		
	Lung-respira-tory	Cardio-vascular	Hema-tologic	Musculo-skeletal	Skin	Eye	Gastro-intest-inal	Kidney-urinary	Liver	Immuno-logical	Repro-ductive	Neuro-logical	Endocrino-logical	Pharmaco-kinetic-metabolic	Mutagenic	Cancer	Immediate death	Short life span
Acute, oral	X	X	X				X	X	X		X	X				0	XX	0
Acute, inhalation	XX	X	X			X	X	X	X		X	X				0	XX	0
Acute, dermal	X	X	X		XX		X	X	X		X	X				0	XX	0
Subchronic, oral	XX	XX	XX	XX	X	XX	XX	XX	XX	XX	XX	XX	XX			0	XX	X
Subchronic, inhalation	XX	XX	XX	XX	X	XX	XX	XX	XX	XX	XX	XX	XX			0	XX	X
Subchronic, dermal	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX			0	XX	X
21-day, dermal	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX			0	XX	0
Chronic, oral	XX	XX	XX	XX	X	XX	XX	XX	XX	XX	XX	XX	XX			XXX	XX	XX
Chronic, inhalation	XX	XX	XX	XX	X	XX	XX	XX	XX	XX	XX	XX	XX			XXX	XX	XX
Chronic, dermal	XX	XX	XX	XX	XXX	XX	XX	XX	XX	XX	XX	XX	XX			XXX	XX	XX
Prenatal developmental toxicity	X	XX		XX	X	XX	X	XX	X		XX	XX				0	X	0
Reproduction and fertility effects	X	X		X	X		X	XX	XX	X	XXX	X	XX				X	X
Neurotoxicity, acute												XXX					X	
Neurotoxicity, subchronic												XXX					X	
Neurotoxicity, acute-delayed												XXX					X	
Neurotoxicity, subchronic-delayed												XXX					X	
Neurotoxicity, chronic												XXX					X	X
Developmental neurotoxicity											X	XXX	X				X	
Operant behavior												XXX					X	
Peripheral nerve function												XXX					X	
Sensory evoked potential												XXX					X	
Eye irritation, primary						XX								0	0	0	X	0
Dermal irritation, primary					XX									0	0	0	X	0
Dermal, sensitization					X					X				0	0	0	X	0
Dermal, penetration					X									XX	0	0	X	0
Metabolism/pharmacokinetics														XXX	0	0	X	0
Genetic toxicity														0	XXX	0	X	0
Immunotoxicity										XXX				0	0	0	X	0

[a] X indicates that some observational or gross endpoints are included; XX indicates level X plus histopathology or some clinical measure of system function. The prenatal developmental toxicity study includes a more in-depth structural evaluation. XXX indicates the major focus of the evaluation. 0 indicates that this endpoint cannot be included as a major aspect in this protocol. A blank indicates that an aspect is not routinely included but could be.

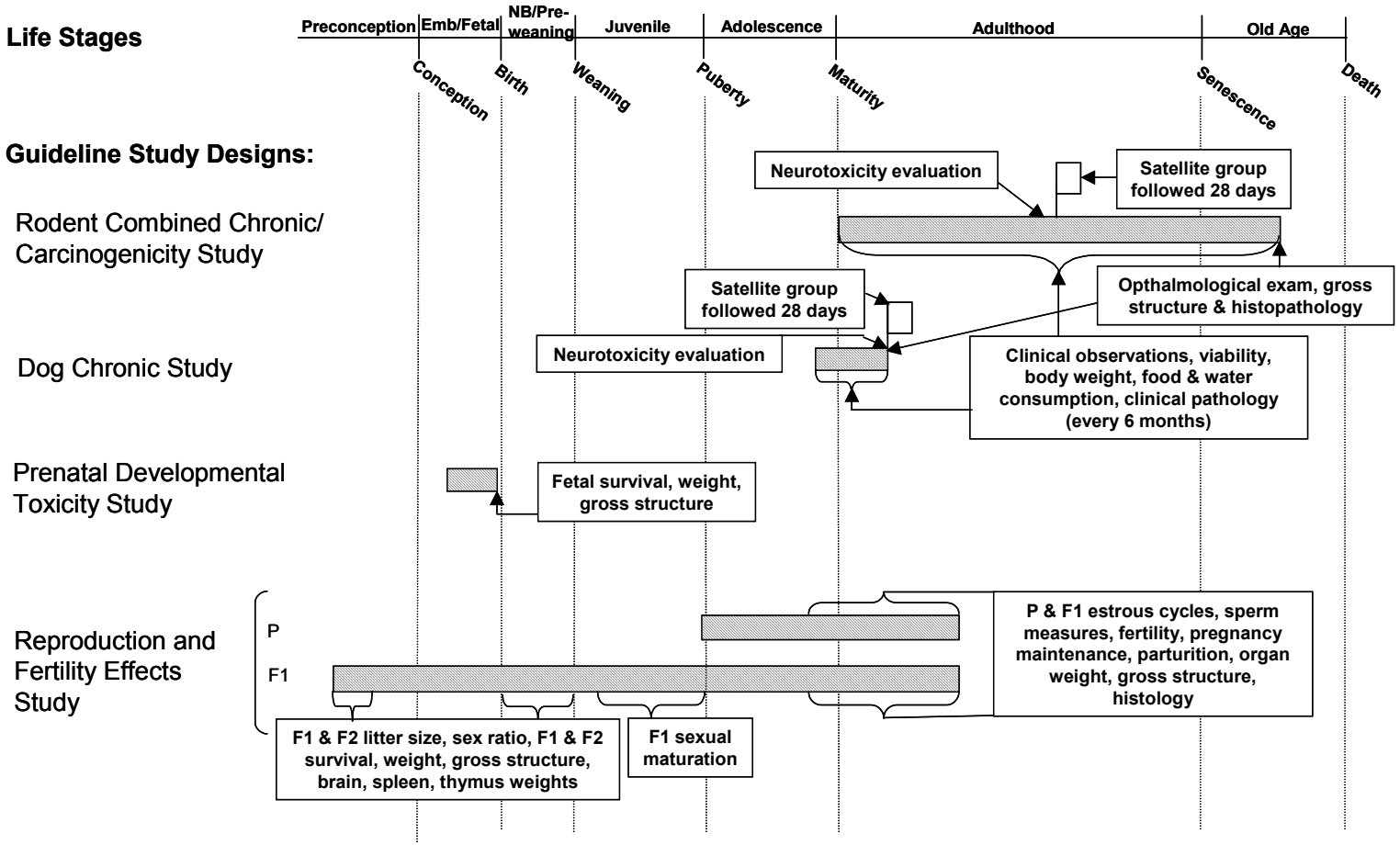
[b] A Series 870 guideline(s) exists for conducting each of the above tests.

relationship for sensitive endpoints following acute inhalation exposure and the toxicologic response following acute high exposures (see further discussion in section 3.1.1.1). Acute toxicity information is useful in establishing reference values for short-duration exposures and for establishing dose-ranges for subchronic and chronic studies. The subchronic and chronic test designs evaluate most endpoints with somewhat greater detail than do the acute test designs. Although the histopathology and/or clinical measures of system function are screening in nature, there is greater confidence that with this level of examination the dose-response relationship will be more clearly defined. Nevertheless, it should be recognized that most systems/endpoints are evaluated at a screening level, and detailed analyses of pathology and function are generally not carried out. Even in those test designs that do incorporate detailed analyses, these analyses are limited in regard to the life stages, exposure periods, and measures that are assessed.

Figure 3-1 shows the study designs that are used for general toxicity testing superimposed on a time line that indicates the life stages during which exposure occurs (hatched bars) and endpoints are measured (indicated in the boxes). The guideline studies shown represent the minimum requirement for derivation of a chronic oral RfD. Similar studies are required for the chronic inhalation RfC, with appropriate endpoints for inhalation exposure and toxicity included. In some cases, only a 90-day subchronic study is available instead of the chronic studies shown. Because the relative length of time between life stages varies among species, the placement of exposures and endpoints on the figures is not necessarily to scale. The following sections discuss the studies that address acute and short-term toxicity as well as chronic toxicity. Similar figures related to specific organ system toxicity testing are shown in subsequent sections.

3.1.1.1. *Acute and Short-Term Toxicity Studies*

3.1.1.1.1. ***Overview of tests.*** The primary purpose of the guideline acute toxicity tests (870.1100 acute oral; 870.1200 acute dermal; and 870.1300 acute inhalation) and other short-term studies (e.g., 14–28-day studies, no OPPTS guidelines available) is to identify hazards (focusing on route-specific lethality) from short-term exposure studies, provide a basis for classification and



Emb = Embryonic NB = Newborn P = Parental F1 and F2 = 1st and 2nd filial generations

Figure 3-1. Exposures and endpoints related to general toxicity evaluations. Endpoints shown are for oral exposures; endpoints specific to inhalation and dermal exposure are included for studies by those routes of exposure.

labeling, and enable the selection of exposure ranges for longer-term studies.[3] Acute guideline studies are conducted in young adult animals, with a 14-day post-exposure observation period. Other than mortality, the endpoints include cage-side observations, body weight at the end of the observation period, gross pathology changes at necropsy, and histopathological examination of organs showing evidence of gross pathology in animals surviving 24 hours or more. Two other available guideline studies cover acute exposures followed by extensive assessment of a specific organ system. The first is the acute inhalation toxicity study with histopathology (40 CFR 799.9135), which was developed for hazardous air pollutants. This study includes assessments of liver, kidney, and broncho alveolar lavage samples for several indicators of cellular damage (e.g., total protein, cell count, percent leukocytes) and a phagocytosis assay to determine macrophage activity. For the respiratory tract histopathology, detailed specifications are provided.

The second expanded study that includes observations following an acute exposure is the acute neurotoxicity study (870.6200), which was developed for the evaluation of neurotoxic chemicals and includes assessments of functional behavior and motor activity at the time of peak effect and again at 14-days post-treatment and histopathology of the central and peripheral nervous systems at 14-days post-treatment. The prenatal developmental toxicity study (870.3700) in two species (typically rats and rabbits) and the DNT study (870.6300) can also provide relevant data for acute risk assessment because maternal observations are often recorded daily and there is a presumption that effects during development may result from a single exposure.

3.1.1.1.2. ***Gaps in life stage of assessment.*** Acute/short-term testing is done only in prenatally exposed animals and in young adults. No direct information is available from any of these studies on acute or short-term exposure in postweaning young animals or aged animals.

---

[3]Alternative test protocols have been adopted by the Organization for Economic Cooperation and Development for acute toxicity testing for oral, dermal, and inhalation exposure, including the fixed-dose procedure, the acute toxic class method, and the up-and-down procedure. All are designed to minimize animal usage and provide minimal hazard and dose-response information for classification, labeling, and dose selection. In the future, EPA plans to put primary reliance on the up-and-down procedure for testing of technical grade pesticides, although the other tests may be acceptable in some circumstances, e.g., testing of pesticidal products. These studies are not designed to provide information for use in less-than-lifetime risk assessments.

3.1.1.1.3. ***Gaps in assessment endpoints.*** Data on only a limited number of toxicological endpoints are available from guideline acute toxicity (lethality) studies except in the case of the acute inhalation toxicity guideline study with histopathology and the acute neurotoxicity study. Consequently, these studies often are not suitable for use in deriving reference values unless additional data, such as those from subchronic studies (e.g., hematological, clinical, histology of more organs), are collected. Some data from animals examined at early times might be available in guideline subchronic or chronic studies. These data could augment the results from guideline acute studies.

3.1.1.1.4. ***Gaps in duration of exposure/latency to response assessment.*** There is no guideline study for short-term toxicity testing, although the prenatal developmental toxicity studies in rats and rabbits and the DNT study include repeated dosing of maternal animals for periods of less than 25 days. Because of the post-exposure observation period in acute guideline studies and in the DNT study, some information on latency to effect and reversibility of effect may be available.

3.1.1.2. ***Subchronic and Chronic Toxicity Studies***

The subchronic exposure studies (870.3100, 870.3150, 870.3200, 870.3250, 870.3465) are used for setting chronic RfDs and RfCs when a chronic study is not available. The guideline studies for chronic exposures (870.4100, 870.4200, 870.4300) (1 year in rodents, although the typical study is a combined chronic and carcinogenicity study with a 2-year exposure) provide an in-depth look at a number of organ systems, and in some cases they evaluate both structure and function (see Figure 3-1). The chronic study in nonrodents, usually dogs, involves a 12-month exposure with similar endpoints assessed as in rodents. The prenatal developmental toxicity study (870.3700) in two species (typically rats and rabbits), the DNT study (870.6300), and the reproduction and fertility effects study (870.3800), typically in rats, are also considered in setting chronic RfDs or RfCs.

3.1.1.2.1. ***Gaps in life stage of assessment.*** The subchronic and chronic studies are conducted in young adult animals, with exposure in the chronic/carcinogenicity study continuing into old age. No information is available from chronic studies in pre- or postnatal animals. Exposures in subchronic study protocols do not include pre- or postnatal development, although the reproduction and fertility effects study does provide data on subchronic exposures in animals that are exposed before birth, through prenatal and postnatal development up to mating of the F1

males and females and through pregnancy (F1 young adult females). No subchronic toxicity evaluations are conducted in aged animals. No chronic studies are conducted in pre- or postnatal animals, although aged animals are exposed and evaluated as part of the chronic study protocol.

3.1.1.2.2. ***Gaps in assessment endpoints.*** The greatest gaps appear to be the lack of routine testing for subchronic neurotoxicity in adults, immunotoxicity testing in adults, and more thorough toxicokinetics in animals at various life stages. Gaps in assessment endpoints during prenatal and postnatal development are discussed in the next section. Assessment endpoints for routine toxicity testing in old age are completely lacking, as is background information on endpoints related to the aging process itself.

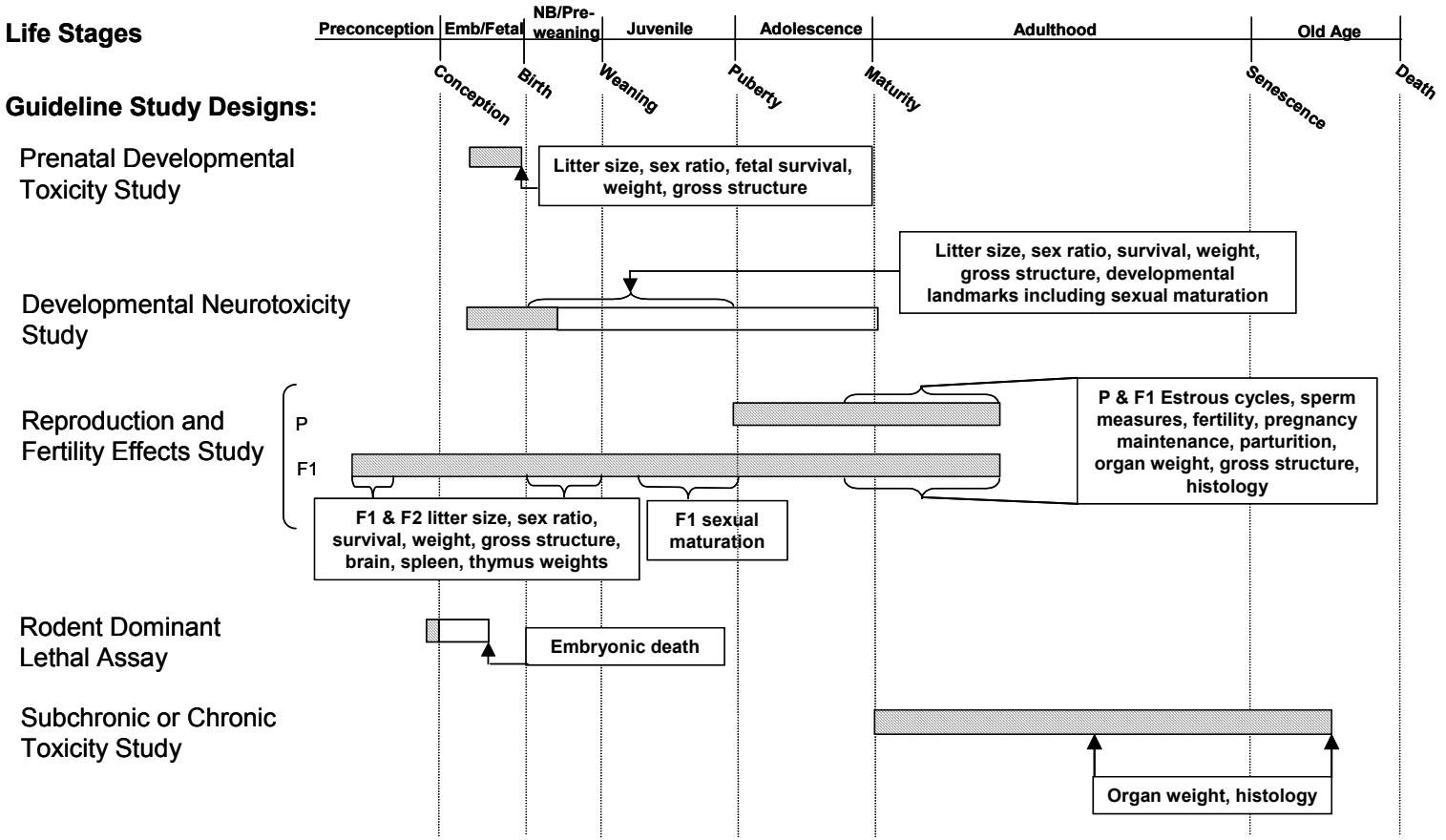
3.1.1.2.3. ***Gaps in duration/latency assessment.*** Chronic studies that include prenatal and postnatal exposure into old age are lacking. The so-called chronic study in dogs is actually a short-term study, as it does not cover at least 10% of the life span. Chronic studies that include a satellite group in which exposure is stopped after 12 months in rodents do assess latency to response for a brief period of time (28 days or more).

3.1.2. Exposures and Endpoints Related to Evaluation of Reproductive Toxicity

3.1.2.1. *Overview of Tests*

The reproductive organs are examined structurally in a number of general guideline screening studies, including the 90-day subchronic study (OPPTS 870.3100, 870.3150, 870.3250, 870.3465), chronic/carcinogenicity studies (OPPTS 870.4100, 870.4200, 870.4300), the prenatal developmental toxicity study (OPPTS 870.3700), and the reproduction and fertility effects study (OPPTS 870.3800), which is a two-generation reproduction study. In addition, extensive assessment of numerous functional aspects of the reproductive system is conducted in the reproduction and fertility effects study. Specific functional effects on the reproductive system of male animals can also be assessed in the rodent dominant lethal assay (OPPTS 870.5450). As illustrated in Figure 3-2, these studies include a variety of both structural and functional assessments of the reproductive system over a wide sampling of life stages.

In guideline subchronic and chronic/carcinogenicity studies, gross structural evaluation and general qualitative histopathology are conducted on reproductive organs and tissues. The animals in these studies are adults, but at the time of organ assessment they may be young (e.g., rats 45 days to 5 months of age from a subchronic study), mature (e.g., rats 5–18 months of age from a reproduction study), or old animals (e.g., rats 18 months to 2 years of age from a chronic study), depending on the protocol.



Emb = Embryonic NB = Newborn P = Parental F1 and F2 = 1st and 2nd filial generations

Figure 3-2. Exposures and endpoints related to reproductive evaluations.

Standard guideline prenatal developmental toxicity studies are designed to evaluate the potential effects of the test substance on the developing fetus. Observations on the reproductive capacity of the maternal animals in this study generally consist only of clinical observations (including any abnormalities of pregnancy maintenance) and gross necropsy data (including uterine). Selected fetuses are examined for gross structural changes to the internal reproductive organs. In studies that employ methods of serial sectioning in the process of soft tissue examination, a limited macroscopic evaluation of the internal structure and integrity of the reproductive organs is performed; however, the fetal tissues are not examined microscopically. Additionally, there are no assessments of organ function in this study design.

In the guideline reproduction and fertility effects study, rats are exposed to the test substance over the duration of two generations, beginning when the first generation animals are young adults of approximately 6–9 weeks of age. Daily exposure continues during all phases of development and reproductive function. Adult animals of both generations are killed as mature adults, generally prior to reaching reproductive senescence (that is, the cessation of normal reproductive function) or an age that would be considered geriatric in that species. Assessments of reproductive capability and function are conducted at least once in each generation. These assessments include direct evaluation of the age of sexual maturation, estrous cyclicity (immediately prior to mating), sperm measures (at termination), mating success, fertility and fecundity, implantation, pregnancy maintenance, gestation duration, parturition, and success of lactation (e.g., maternal nurturing and nesting behavior).

Indirect assessments of some reproductive functions are also evaluated. These observations are based on evidence of normality in a structure, function, or process that is dependent on normal functioning of the component parts, including, for example, hormonal homeostasis, ejaculation, accessory gland function, placental function, milk production, pup nursing behavior or ability, and, to some extent, reproductive senescence (although the adult animals are terminated at the end of each generation, when they are only around 6 months of age; therefore, there are no assessments conducted in older rats). Gross structural assessments of the whole animal are conducted on adult and immature animals throughout the course of the study; gross internal (organ) structural assessments are conducted on offspring that are killed at litter standardization (postnatal day [PND] 4), weaning (PND 21), and termination of each generation (mature adults). Histopathological evaluation of the reproductive organs (gonads and accessory structures) is conducted only in the mature parental adult animals that are killed at the termination of each generation. The guideline specifies a very focused pathological examination of the reproductive organs in this study.

The dominant lethal assay is not conducted for every chemical, but it may be conducted in response to a concern raised by other developmental or reproductive toxicity findings in the database. In this study, sexually mature adult males are treated with the test substance to determine whether there is an effect in the germinal tissue that does not cause dysfunction in the gamete but is lethal to the fertilized egg or developing embryo. Exposed males are mated with untreated females, and uterine contents are evaluated. Evidence of pre- and/or postimplantation loss is generally thought to be indicative of treatment-related chromosomal damage in germinal tissue.

3.1.2.2. *Gaps in Life Stage of Assessment*

Determination of gaps in the assessment of potential effects of any chemical across all life stages requires consideration of both the exposure period and the time of assessment. In the prenatal developmental toxicity study, animals are exposed from implantation through gestation. The reproductive organs are examined for gross structural changes, but no microscopic examination is conducted. There is no follow-up of the animals to determine the functional consequences of prenatal exposure. In the reproduction and fertility effects study, the F1 animals are exposed from preconception throughout prenatal and postnatal development until after mating. The reproductive organs are examined macroscopically at weaning and adulthood. The maturation of the reproductive system is assessed, as is its function. Thus, the study provides a fairly thorough assessment of structure and function following exposure during many critical periods of development. In the parental generation, the animals are exposed as young adults, and the structure and function of the reproductive organs are assessed.

The dominant lethal study, when conducted, assesses a single aspect of the function of the reproductive system for one sex, although a detailed structural assessment is not conducted. In the subchronic and chronic studies, the animals are exposed beginning as young adults, and the structure—but not the function—of the reproductive organs is assessed. Therefore, the major gaps include (1) the lack of functional assessment (particularly the age of onset of reproductive senescence) in older adult animals following adult-only exposures, and (2) the lack of structural and functional assessments in older adult animals following developmental exposures.

The onset of reproductive senescence can be marked by findings such as altered hormonal homeostasis, disruption of estrous cyclicity, diminished sperm measures (number, motility, or morphology), or gonadal atrophy. Studies in rodents have demonstrated the adverse effects of a number of agents (e.g., ionizing radiation, chemotherapeutic agents, polycyclic aromatic hydrocarbons, and agents that form epoxides, such as 1, 3-butadiene and 4-vinylcyclohexene) on reproductive senescence (reviewed by Hoyer and Sipes, 1996).

In humans, premature reproductive senescence has been associated with cigarette smoking (Jick et al., 1977). In addition to potentially diminishing fertility in individuals who are only slightly past prime reproductive age, early reproductive senescence can adversely affect the general health of the aged human. For example, hormonal alterations that are associated with early senescence have been linked to abnormalities of cardiovascular function, osteoporosis, and even a predisposition to early mortality.

3.1.2.3. *Gaps in Assessment Endpoints*

As described above, there are identifiable gaps in the endpoints that are used to assess reproductive toxicity in guideline studies. Currently, there is no assessment of functional endpoints in older animals following adult exposures, and there are no structural or functional endpoints assessed in older animals following developmental exposures, including reproductive senescence. In addition, concerns have recently been raised about the ability to detect rare malformations of the reproductive organs and abnormalities in the maturation of the reproductive system in the two-generation reproductive toxicity study. This concern relates particularly to endocrine-active chemicals. In the current guideline, three pups/sex/litter are examined macroscopically at weaning.

Questions have been raised about whether these weanlings should be retained until day 45 (females) or day 60 (males) to ensure that any later-appearing gross or functional changes are detected. This issue is currently being examined within the endocrine validation/standardization program.

3.1.2.4. *Gaps in Duration/Latency Assessment*

There are no studies that include acute or chronic exposures that can be used to assess the development of the reproductive system. As indicated above, it has been suggested that animals be retained until older ages in the two-generation study in order to assess later-appearing structural or functional changes in reproductive organs. In addition, there is no consideration of latent responses for reproductive toxicity, such as early onset of reproductive senescence, as a result of an exposure earlier in life in any of the studies that can be used to evaluate reproductive toxicity, except for a few endpoints in the DNT study.

3.1.3. Exposures and Endpoints Related to Evaluation of Neurotoxicity

3.1.3.1. *Overview of Tests*

Observation of the animals for signs of overt toxicity and routine gross pathological assessment of the nervous system is required under OPPTS acute, subchronic, and chronic study protocols (870.100–870.400 series). In rat studies, age at initiation of testing is to be 8–12 weeks under acute and subchronic testing protocols. In acute studies, cage-side observation and gross neuropathology are the only endpoints required under 870.100 (oral, dermal, or inhalation exposure). Motor activity, grip strength, and sensory reactivity and neuropathology are measured in the rodent oral study, the dermal 21–28- and 90-day subchronic studies, and the 90-day inhalation study. In rodent subchronic studies, specific assessment for neurotoxicity is performed at or near the end of the study, although observations of the animals, including those for detection of overt neurotoxicity, are made routinely throughout the study. No specific functional tests for neurotoxicity are required for nonrodent subchronic studies, although observation and neuropathology are required.

Chronic toxicity studies (oral, dermal, inhalation) are to be performed in two species (one rodent) over a 12-month period, regardless of the life span of the species. Exposure in rodents is to begin no later than 8 weeks of age. Motor activity, grip strength, and sensory reactivity are to be assessed at or near the end of the study, but no earlier than the 11th month. Clinical observation is performed weekly throughout the study and would presumably detect gross neurological abnormality. In current practice, the chronic study is often combined with the carcinogenicity test, in which dosing extends for 24 months in rats and 18 months in mice (OPPTS 870.4300). Motor activity would be performed at 11–12 months only, as in the chronic study, and not again until near the end of exposure.

The neurotoxicity screening battery (870.6200) is designed to be included in acute, subchronic, or chronic toxicity studies (Figure 3-3). The endpoints examined extend those required in the 870.100 series, although there is no guidance as to when these extended batteries would be required. The functional observation battery includes a ranking system for general reactivity, activity, and gait abnormalities, as well as forelimb and hindlimb grip strength, landing foot splay, sensorimotor reactivity to sensory stimuli, and pain reception. Motor activity and a more detailed neuropathological observation are also required in this battery. For acute studies, assessments are made before initiation of dosing, at the estimated peak of activity within 8 hours of dosing, and at 7 and 14 days post-dosing. For subchronic studies, assessments are performed pre-exposure and at 7, 8, and 13 weeks of exposure. For chronic studies, assessment is at pre-exposure and every 3 months post-exposure. There is no specific guidance regarding

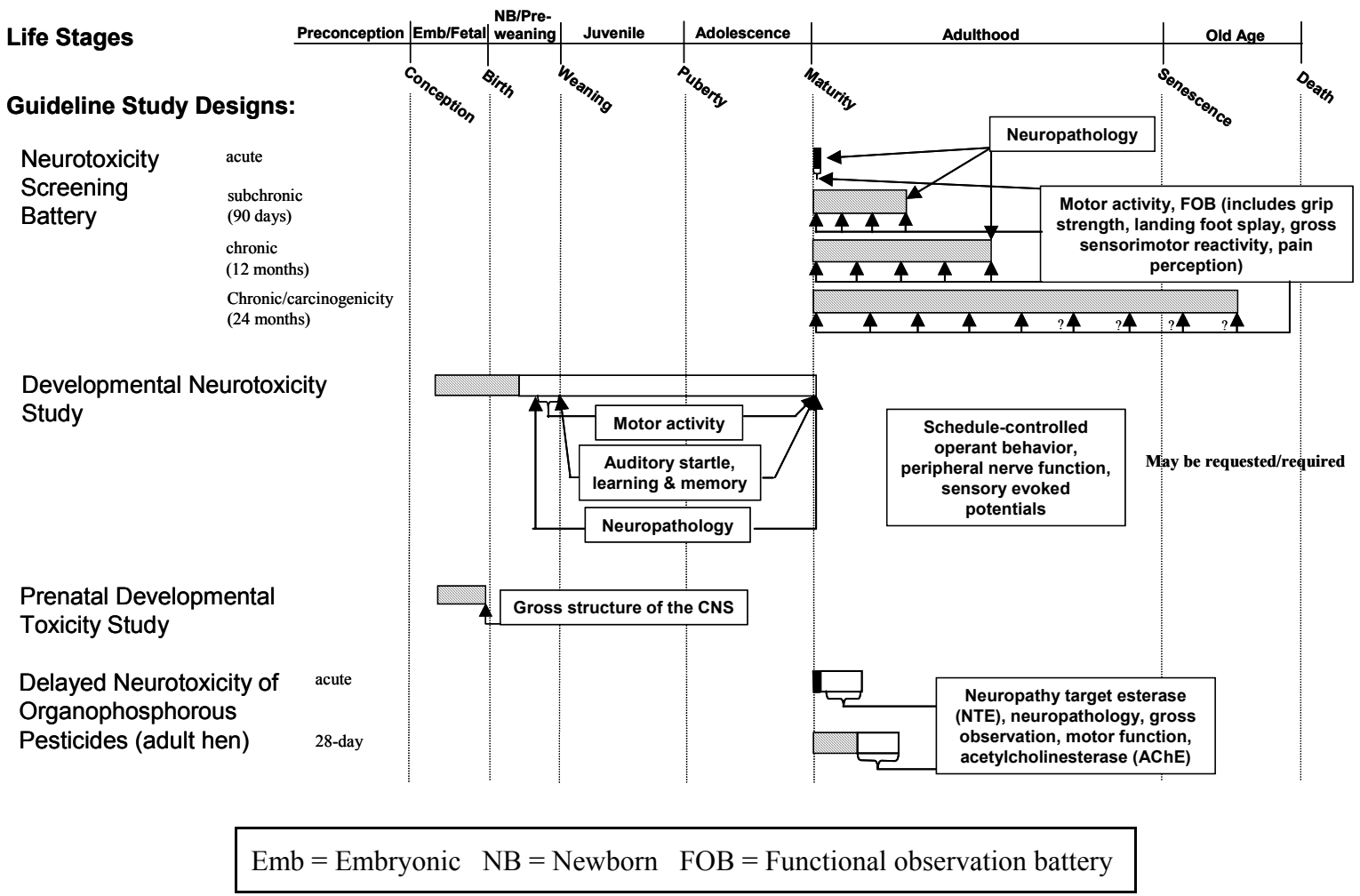


Figure 3-3. Exposures and endpoints for neurotoxicity evaluations.

the assessment schedule for the combined chronic/carcinogenicity study, but presumably the schedule required for the chronic study would be maintained.

The DNT study protocol (870.6300) currently requires dosing of the dams from gestational day (GD) 6 through PND 10, although the requirement may soon be extended to PND 21 (i.e., until weaning). Motor activity is measured at PNDs 13, 17, 21, and 60. Auditory startle is measured around weaning and at PND 60, as is a test of learning and memory, which may be the same test or different tests at the two time points. Cage-side observation of both dams and pups is required, and neuropathology in the pups is required at PND 11 and at the termination of the study (usually PND 60). The prenatal developmental toxicity study (870.3700) requires dosing of the dams on GDs 6–20 in rats and 6–29 in rabbits. Gross structural evaluation of the nervous system is evaluated as part of the fetal examinations conducted in this study.

3.1.3.2. *Gaps in Life Stage of Assessment*

One of the most significant gaps revealed by Figure 3-3 is the lack of exposure or assessment under any protocol during old age. For example, following acute exposure, assessment is for 14 days in juvenile or young adult animals. The chronic exposure protocol extends exposure into adulthood and the combined chronic/carcinogenicity protocol extends exposure up to approximately the aged period in the rat, but neurotoxicology assessments are not performed in aged animals. Thus, none of the protocols assess potential effects of chemicals on aging as a function of exposure during development. This may be important, because studies in animals have shown that developmental exposure to agents that cause neurotoxicity, such as trimethyl tin, can accelerate the onset of cognitive deficits measured later in life. Other studies with methyl mercury have documented early-onset sensory dysfunction in monkeys exposed during development. Furthermore, current testing protocols do not provide information collected at different life stages—that is, comparison of effects of exposure during infancy, adulthood, or old age. This is important, because life stage-dependent differences in pharmacokinetic, and possibly toxicodynamic, parameters could result in quantitatively or qualitatively different effects at different life stages.

Under the DNT protocol, there currently is no requirement to perform kinetic studies to ascertain either in utero or postnatal exposure. There is no mechanism to guarantee exposure postnatally (i.e., direct dosing of pups) because the compound may not be excreted into breast milk or it may be excreted only at very low concentrations. This is of particular importance, because the early postnatal period in the rodent is equivalent to a prenatal life stage in humans.

There is no long-term follow-up assessment to detect delayed neurotoxic effects, a situation that is arguably more worrisome for developmental exposure than for exposure later in life.

3.1.3.3. *Gaps in Assessment Endpoints*

The nervous system is one of the most fully assessed organ systems in the EPA/OPPTS 870 guidelines. Nonetheless, most of the endpoint assessments are designed to be screening procedures rather than sensitive assessments of nervous system function. In addition, the assessments required are different in the neurotoxicity screening battery than in the DNT study. The adult neurotoxicity screening battery does not require assessment of learning and memory or auditory startle. The lack of assessment of cognitive function in the neurotoxicity screening battery constitutes a significant omission that should be addressed.

It may also be pointed out that even in the developmental protocol, the tests that are used to assess learning and memory may be very simple, potentially revealing only relatively gross deficits. In addition, although potentially more sensitive cognitive, sensory, and motor tests are available (Figure 3-3), there is no guidance as to what would trigger a requirement for these assessments. Except for the protocol for delayed neurotoxicity for organophosphorous pesticides in the hen, there is no assessment of neurochemical endpoints. Additionally, the required neuropathological assessments may also be considered screening.

Minimal morphometric analysis, consisting of the thickness of "representative" layers in the neocortex, hippocampus, and cerebellum, is required in the DNT study. No morphometric analyses are required in the adult neurotoxicity testing protocols. Although more sophisticated tests would presumably not be performed on all agents, more sophisticated measures could be triggered by results from screening tests. It also may be advisable to require more sensitive tests in instances of particular concern, for example, adding more extensive morphometric analysis to the DNT protocol.

In summary, although the nervous system is one of the most thoroughly assessed systems in the 870 test guideline studies, it must be kept well in mind when interpreting the results that these are screening tests. Positive findings must be viewed as indicative of relatively overt toxicity, not so-called subtle effects.

3.1.3.4. *Gaps in Duration/Latency Assessment*

One of the principles in the neurotoxicity risk assessment guidelines (U.S. EPA, 1998c) is that neurotoxicity could occur after one or a few exposures, such as in the case of an organophosphate insecticide that produces a delayed neuropathy, or only after a series of repeated exposures, as in the case of acrylamide. For DNT, it is assumed that a single exposure

to a chemical during a critical period of development could result in an adverse effect on the developing nervous system. There are, however, few data that compare the effects of a single exposure to a chemical with the effects of the same chemical given multiple times during development.

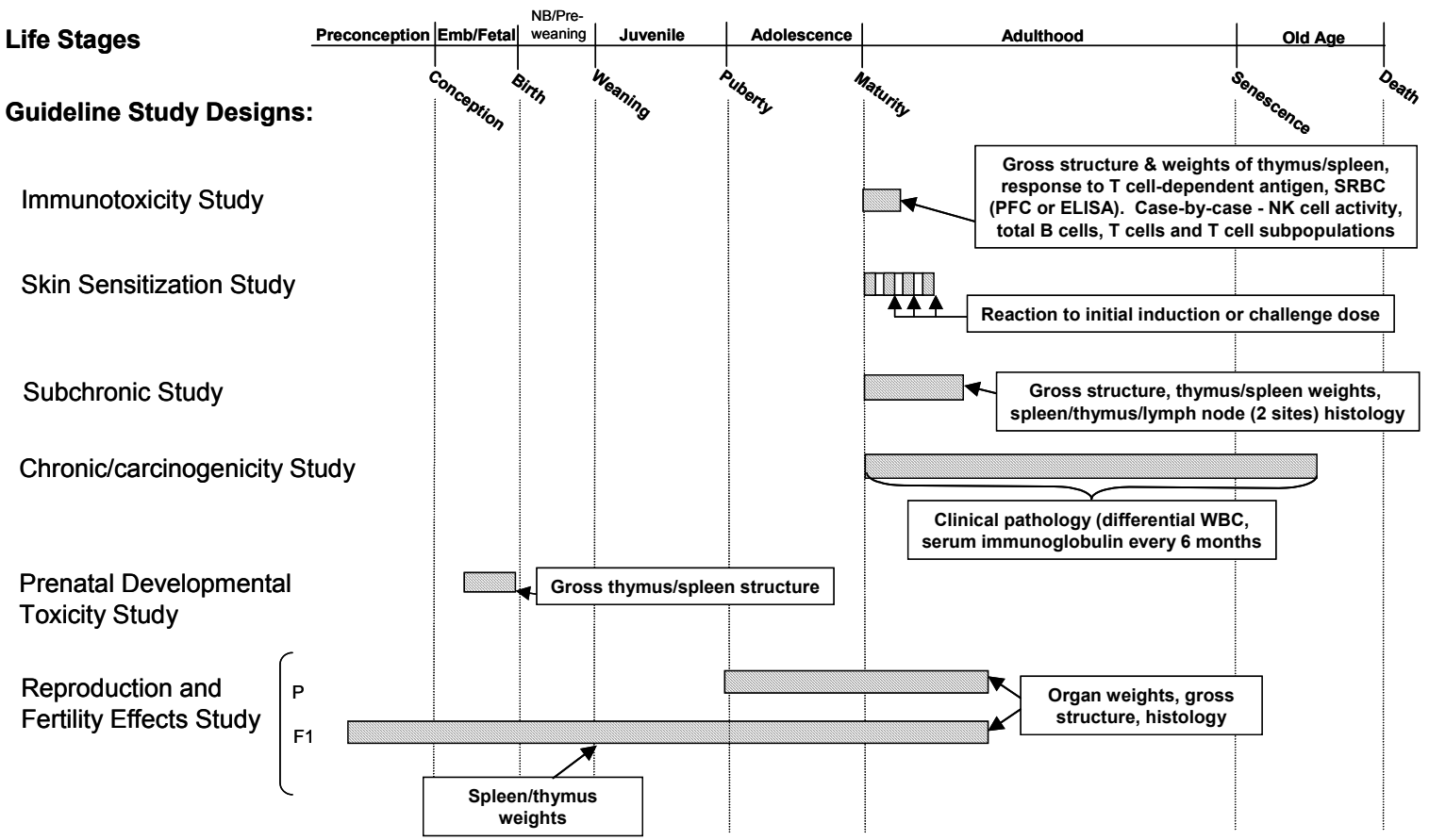
3.1.4. Exposures and Endpoints Related to Evaluation of Immunotoxicity

3.1.4.1. *Overview of Tests*

Examination of the macro- and/or microscopic structural anatomy of immune system organs and tissues is performed in a number of general guideline screening studies, including the acute inhalation toxicity with histopathology guideline (40 CFR 799.9135), the 90-day subchronic study (OPPTS 870.3100, 870.3150, 870.3250, 870.3465), the chronic/carcinogenicity studies (OPPTS 870.4100, 870.4200, 870.4300), the prenatal developmental toxicity study (OPPTS 870.3700), and the two-generation reproduction study (OPPTS 870.3800). In addition, functional assessments of the immune system are evaluated in the skin sensitization study (OPPTS 870.2600) and the immunotoxicity testing guideline (OPPTS 870.7800) (see Figure 3-4).

In the guideline immunotoxicity study, young adult rats (6–8 weeks of age) are exposed to the test substance for 28 days, at which time they are terminated. The spleen and thymus are examined macroscopically, and organ weights are recorded; a histopathological evaluation is not performed. Assessments of immune system function include an evaluation of the response to the T cell-dependent antigen, sheep red blood cells (SRBC). The SRBC antigen response assays can be conducted either by an antibody plaque-forming cell (PFC) assay or an immunoglobulin quantification by enzyme-linked immunosorbent assay (ELISA). In addition, an assessment of natural killer (NK) cell activity and/or enumeration of splenic or peripheral blood total B cells, total T cells, and T cell subpopulations may be required on a case-by-case basis.

The skin sensitization study has been generally conducted in guinea pigs as a Guinea Pig Maximization Test (GPMT) or a Buehler test. In a recent review by the FIFRA Science Advisory Panel (U.S. EPA, 2001a), it was recommended that in the future, skin sensitization methods should preferentially include the local lymph node assay (LLNA), which uses young adult mice. The skin sensitization test involves an initial intradermal (GPMT) and/or epidermal (Buehler, LLNA) exposure of the test animal to a substance, followed by a challenge exposure approximately 1 week later. In the guinea pig tests, sensitization is determined by examining the reaction to the challenge exposure and comparing this reaction with that of the initial induction



Emb = Embryonic NB = Newborn P = Parental F1= 1st filial generation SRBC = Sheep red blood cell
PFC = Placque-forming cells ELISA = Enzyme-linked immunosorbent assay NK = Natural killer
WBC = White blood cells

Figure 3-4. Exposures and endpoints for immunotoxicity evaluations.

exposure. In the LLNA, proliferation of lymphocytes is measured (as a function of in vivo radioisotope incorporation into cellular DNA) in draining lymph nodes proximal to the application site. Hence, although the GPMT and Buehler tests result in a qualitative assessment of hypersensitivity, the LLNA provides a quantitative dose-response evaluation. Histopathological evaluation of the skin is not required with any of these methods, but it may be conducted. No other immune system endpoints or organs are evaluated in this study.

In guideline subchronic and chronic/carcinogenicity studies, an evaluation of macroscopic structure and general qualitative histopathology are conducted on only a few immune system tissues. In studies that include young adult animals (e.g., rats 45 days to 5 months of age from a subchronic study), the spleen, thymus, and lymph nodes from two locations (one near to and the other distant from the site of administration) are examined; the spleen and thymus are weighed. In chronic and carcinogenicity study guidelines, there is no requirement that the thymus be examined and/or weighed. For rodents (e.g., rats or mice 18 months to 2 years of age), it is reasonable to assume that the thymus would have undergone normal age-related atrophy by study termination. However, the thymus might be present at early interim sacrifices of rodents (e.g., at 6 months or 12 months of study) during a long-term study, and it would certainly be present at study termination in a canine chronic study (at which point the dogs are young adults of only approximately 1.5 years of age).

Differential white cell counts in the circulating blood are examined at study termination in the subchronic study and at approximately 6-month intervals in long-term studies. Serum immunoglobulin levels may be measured at the same intervals. Perturbations may indicate increased immune system response to some unspecified initiator, but this information does not address the adequacy of immune system function. In the same manner, histopathological evaluation of other organ systems in the subchronic and chronic/carcinogenicity studies may identify cellular alterations that are nonspecific indicators of an effect on immune response, for example, the presence of increased numbers of macrophages in lung tissue or an increased incidence of inflammatory dermal lesions.

In the reproduction and fertility effects study in rats, a macroscopic evaluation of all organ systems is conducted in a sample of offspring at weaning and in the mature adult parental animals at the termination of each generation. Additionally, the spleen and thymus are weighed for those pups that are necropsied at weaning; these measurements are intended to provide information on the need for further evaluation of the immunotoxic potential of a chemical to the immature animal.

In the prenatal developmental toxicity study, an evaluation of the macroscopic structure of the thymus and spleen is conducted in at least half of the fetuses from each litter.

3.1.4.2. *Gaps in Life Stage of Assessment*

In the available guideline studies, assessments of organs with immune system function are conducted in fetuses following prenatal exposure, in weanling animals following pre- and postnatal exposure, and in young and/or mature adult animals at a variety of time points. With prenatal exposures and evaluation at early life stages, these assessments consist entirely of the evaluation of macroscopic changes, with no microscopic examination. Toxicokinetic data that characterize the exposure in the young (i.e., exposure of the fetus to the chemical or its metabolites via the placenta or of the neonate via breast milk) are not routinely required and are seldom available.

Some detailed structural assessment (histopathology) is conducted in mature or older adult animals. Indirect assessment of immune system function is conducted in adult animals of various ages via the evaluation of peripheral blood cells and chemistry. Direct functional assessments of the immune system are conducted only in young adult animals; generally this age group is selected for assessment because of the anticipated robustness of the immune response.

There is no guideline that examines potential perturbation of immune system function following early pre- and/or postnatal exposure (often referred to as a developmental immunotoxicity study). Comparisons of immune effects following exposure at various life stages (i.e., during in utero or postnatal development, adulthood, or old age), including data that analyze whether these effects are more severe in one age group or whether the effects are persistent, are not required. To achieve even a minimal assessment of immune system structure and function, a broad variety of studies would need to be conducted and assessed; yet there could still be relatively low confidence in the ability of the results of these combined studies to predict the outcome of age-specific insults to the immune system.

3.1.4.3. *Gaps in Assessment Endpoints*

There are identifiable gaps in the endpoints that are used to assess immunotoxicity in guideline studies. For example, for fetuses, immature animals, and old animals (rodents), assessments are composed entirely of the evaluation of macroscopic structural changes, with no histopathological or functional evaluations. In mature adult animals, thorough macroscopic and microscopic structural assessments, as well as routine hematological testing (e.g., blood cell counts), are performed; however, those assessments are generally very limited in young animals, and guideline requirements do not consider species differences. The only assessments of functional integrity of the immune system are provided by the guideline sensitization study and the 28-day immunotoxicity study. These studies are conducted only in young adult animals, and

they include only a few examples of potential immune system response (e.g., hypersensitivity, humoral immunity, or nonspecific cell-mediated immunity). In very young and very old animals, there is no direct assessment of immunological function. No assessment of autoimmune effects is conducted in any of the current guideline protocols.

3.1.4.4. *Gaps in Duration/Latency Assessment*

Latent effects on immune function that result from early lifetime exposure are not assessed; these can include effects in aged animals that result from in utero, neonatal, or young-adult exposure. Exacerbation of effects in relation to aging and response to subsequent immunological challenge are not routinely or systematically assessed to any extent. The two-generation reproduction study offers an opportunity to evaluate immunotoxic response in adulthood that resulted from prenatal or early postnatal exposure. In the chronic toxicity studies in rodents, aged animals are available for evaluation. However, in both cases there is little focus on the evaluation of the immune system. Only indirect evidence of perturbation of the immune system may be observed through macroscopic and microscopic evaluation of various organs; corollary functional assessment is not performed. Response to an immunological challenge is examined only in the hypersensitization study, and even when the results from this study are positive, no further specific assessment of the immune system is pursued.

3.1.5. Exposures and Endpoints Related to Evaluation of Cardiovascular Toxicity

3.1.5.1. *Overview of Tests*

Gross observation of the heart and major vessels augmented by conditional standard pathology is mentioned in most applicable OPPTS Series 870 health effect guidelines (Figure 3-5).

3.1.5.2. *Gaps in Life Stage of Assessment*

The period from birth to maturity is essentially without toxicological monitoring of cardiovascular endpoints for both repeated chronic and single acute-exposure regimes.

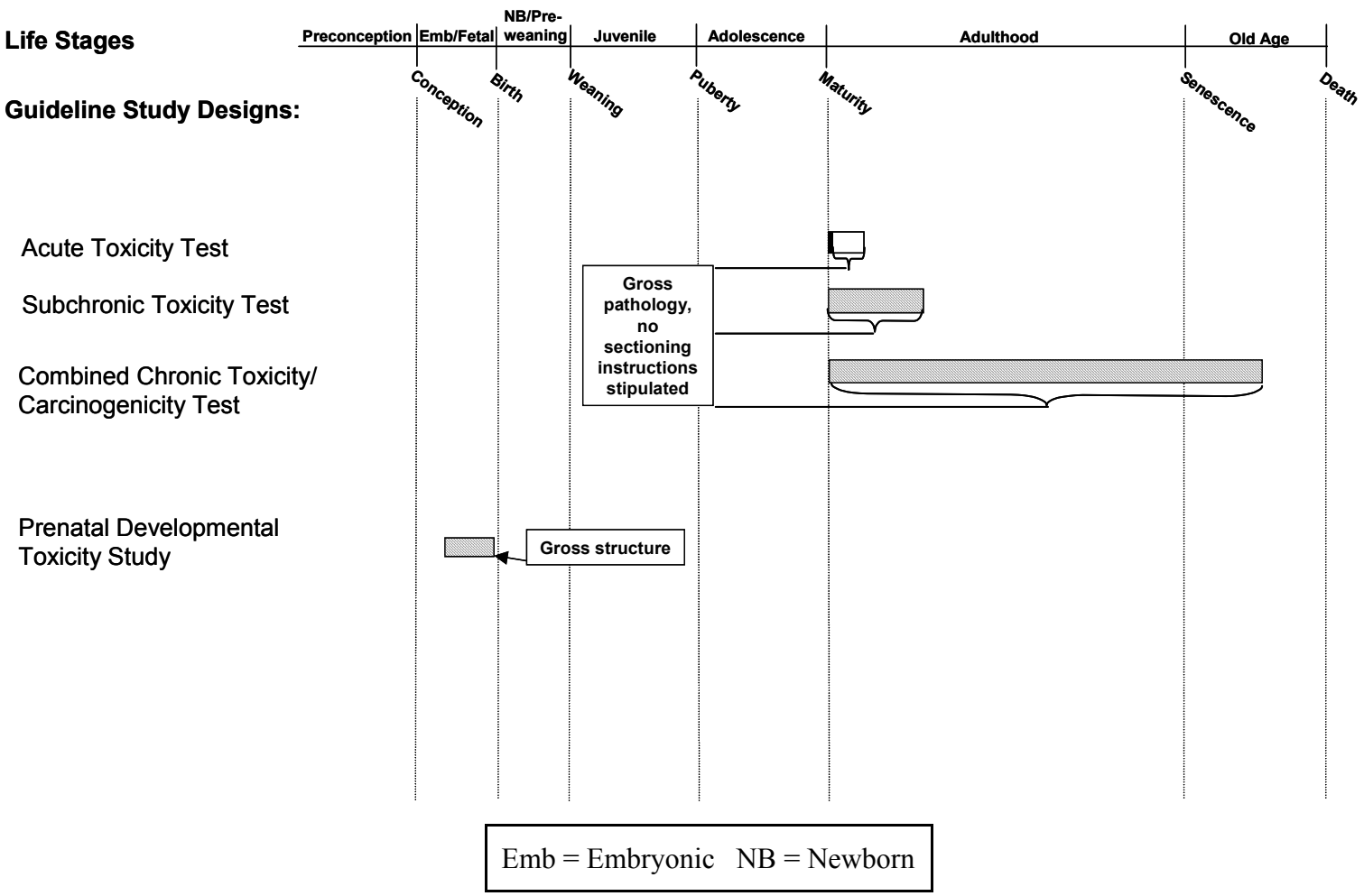


Figure 3-5. Exposures and endpoints related to cardiovascular evaluations.

3.1.5.3. *Gaps in Assessment Endpoints*

Gross observation only of the heart is provided for in most OPPTS guidelines. Functional clinical or histopathological cardiac examination is not currently part of any testing guideline. Even gross pathology could be improved and brought into line with current cardiovascular evaluation by separating, weighing, and constructing right and left ventricle-to-body weight ratios to give an evaluation of cardiac hypertrophy. Also, guidelines regarding sectioning procedures for the heart, either number or plane, could be provided.

No simple cardiac functional evaluation is currently available, including even systolic or diastolic blood pressures. It should be noted that telemetric in-dwelling echocardiograms (ultrasound examinations of the heart) can be used to detect occlusions and atherosclerosis and to detect alterations in cardiac output. Combination echocardiograms and electrocardiogram analysis can detect cardiac wave forms as well as heart rate variability in high- and low-frequency power ranges (i.e., beat-to-beat changes in heart rate ascribed to varying control by the autonomic nervous system). Heart rate variability may be critical in explaining toxicity, as was shown in recent work associating exposures to fine particulate matter with decreases in heart rate variability in elderly humans (Creason et al., 2001). Both echocardiograms and electrocardiograms can be done on rats down to 100 g, well within the size range of juvenile and adolescent rats.

Chemicals can produce degenerative and/or inflammatory changes in the peripheral blood vessels as a consequence of an excessive pharmacologic effect or by an interaction with a vascular structural or functional macromolecule. As a result of sustained arterial vasoconstriction, peripheral arterial lesions consisting of intimal proliferation and medial degenerative changes could result in gangrene. Also, chemicals can induce or enhance atheroma formation, which is characterized by endothelial damage with increased permeability, monocyte adhesion, and endothelial proliferation.

Selected representative techniques to study the peripheral vascular system consist of flow measurement techniques (Smith et al., 1994), such as electromagnetic flowmetry, pulsed Doppler flowmetry, transit time flowmetry, laser Doppler fluxmetry, and laser scanner methods. These techniques allow investigation of blood flow in vessels as large as the aorta and as small as the capillary, determination of the level of perfusion in tissues, and calculation of the derived hemodynamic variable of resistance.

The two major noninvasive techniques for determining microvascular velocity are the flying spot technique and the dual-slit technique. External ultrasound may be used to examine internal vascular dimensions. A noninvasive assessment of arterial flow in rodents and monkeys can be performed using Doppler spectrum analysis (duplex ultrasound technology) (Leopold et

al., 1997). This test detects arterial compromise in extremities, functional severity, and the hemodynamic significance of vascular lesions. In most cases, the locations in the arteries involved can be designated. Information regarding the extent and effectiveness of collateral circulation can also be gained. This testing is a valuable tool for monitoring early flow compromise secondary to chronic reoccurrence of anastomotic or distal disease.

Several blood/plasma tests for clinical assessment are in active use in cardiovascular research. In general, these are tests that may be used to document a cardiovascular accident (within 48–96 hours). Their utility for risk assessment has yet to be evaluated. Specific enzymes currently being used by the research community for these purposes include LDH-I, creatinine kinase-II, and troponin. Other enzymes useful as prognostic indicators of risk of a cardiovascular accident include angiotensin converting enzyme II, plasma renin activity, endothelin-converting enzyme-1, and catecholamines (epinephrine and norepinephrine).

3.2. CONCLUSIONS AND RECOMMENDATIONS

A review of current testing guidelines was conducted to determine the types of data available for setting reference values. The approach used was to evaluate testing guidelines from the point of view of (1) life stages covered, (2) endpoints assessed generally and for specific organ systems, (3) timing and duration of exposure, and (4) evaluation of reversibility and latency to response.

The relevance of these issues to the health evaluation of children and other potentially susceptible subpopulations should be apparent from the gaps identified in each of the above sections regarding life stage assessment, endpoints assessed, timing and duration of exposures included in guideline studies, reversibility, and latency to response. Although a number of areas of toxicity testing have been discussed, this review should not be considered exhaustive, and other health effects may be as or more important for particular chemicals than those reviewed in detail here.

Issues of particular concern for children's health that have not been discussed in great detail here are effects related to asthma and other respiratory tract toxicity. For both children and the elderly, renal and liver function can be a major factor in the disposition, metabolism, and excretion of chemicals and, therefore, their toxicity. Thus, the evaluation of toxicity and the interpretation of data in terms of its completeness will always require scientific judgment about whether or not adequate data have been collected on effects of importance at the appropriate life stages and timing and duration of exposure, for example, for a given agent.

Effects seen at the termination of a chronic study may be due to cumulative damage from a continued repeated chemical insult, but they could also be a latent response from an earlier

single or short-term multiple exposure. Thus, latent effects might be revealed in chronic studies, but it would not be clear whether they were the result of acute/short-term exposure or the chronic exposure. Specific information on the latency of a response would follow only from a clearer understanding of the mechanism of the effect and from actual "stop exposure" protocols (e.g., the satellite studies depicted in Figure 3-1) or from shorter-term exposures with follow-up over a much longer period of time. It thus follows that any chemical database that does not have exposure-response studies of lifetime duration or any specific exposure-latency protocols would not cover the possibility of latent effects.

Effects that persist throughout a designated post-exposure period may be considered irreversible; those that do not are reversible. For chronic lifetime exposures, designation of an effect as irreversible or reversible is academic, as exposure is presumed to be lifetime (i.e., there is no post-exposure period). For shorter-term values (e.g., acute, short-term) where an appreciable period of time post-exposure is anticipated, designation of an effect as reversible or irreversible becomes more relevant. Derivation of a reference value based on shorter-term exposure guideline protocols would have to fully consider the aspect of reversibility in interpretation of the data. It is important to understand the difference between an endpoint that is truly reversible and one that is related to or is a precursor of other adverse effects. For example, low birth weight may be "reversible" through catch-up growth postnatally, but it also may be related to developmental delays or other health outcomes that result from prenatal growth reduction/retardation.

3.2.1. Conclusions

From this review, the Technical Panel reached the following major conclusions:

1. There are a number of gaps in life stages covered in current guideline testing protocols, particularly in terms of the exposure periods included. In particular, there is minimal evaluation of aged animals, especially after exposures that include early development.

2. There are a number of gaps in the evaluation of endpoints included for certain systems; for example, the evaluations of the cardiovascular and immune systems in various guideline studies were reviewed as examples of systems that are minimally covered. Other systems, for example, the reproductive and nervous systems, are evaluated in more detail, but even in these systems there are gaps that need to be

considered; notably, functional evaluations are not always included or integrated with structural evaluations of particular systems.

3. Acute and short-term exposure studies are either not available or include only gross effects, so that the data needed to derive acute and short-term reference values are often not available.

4. Latency to response and reversibility are only rarely evaluated directly. These types of effects could have a major impact on hazard characterization, especially in designing acute and short-term test guideline protocols and ultimately on the risk management options that can be used for intervention or prevention.

5. Although not more specifically discussed, it is clear that there is a lack of information on toxicokinetics. The available data are generally limited to studies that are conducted in young adult animals, but there are no guideline protocols for toxicokinetic evaluations during development or in older age related to exposures and outcomes.

6. The underlying assumption that the internal dose of the active form of an agent to the target site is the relevant measure of dose clearly underscores toxicokinetics as an essential tool that must be used in both hazard identification and dose-response evaluations. This should not only continue to be a central and critical area of exploration, it should be an area of direct application to assessment activities to address various issues, including but not limited to (a) design of studies, (b) delivery to the fetus/neonate, (c) dose scaling, (d) toxicokinetic and toxicodynamic considerations, and (e) route extrapolation.

A white paper on pharmacokinetics commissioned by the Technical Panel (Versar Inc., 2001b) is meant to serve as a technical resource for the application of toxicokinetics to these and other issues addressed throughout this document. Another white paper on aging (Versar Inc., 2001a) also addresses issues of changing pharmacokinetics during this life stage.

7. Portal-of-entry effects (i.e., respiratory, gastrointestinal, dermal) are acknowledged as being important in the effects of chemicals, and they may preclude systemic toxicity as being sentinel. Chronic oral RfDs and inhalation RfCs have been developed for a

number of agents, but rarely have dermal RfDs been derived. In some cases, oral RfDs and oral cancer potency factors have been used to assess systemic toxicity from dermal exposures. However, the dermal route of exposure can result in different patterns of distribution, metabolism, and excretion than those that occur from the oral route. Dermal contact with a chemical may also result in direct dermal toxicity, such as allergic contact dermatitis, urticaria reactions, chemical irritation, and skin cancer.

The dose-response relationship for the portal-of-entry effects in skin is likely to be independent of any associated systemic toxicity exhibited by a particular chemical. Therefore, there is a long-term need for the development of dermal RfDs that consider both the systemic toxicity effects and the portal-of-entry effects of individual chemicals. In addition, there is a need for data on the dermal uptake of chemicals from soil, water, and air, including information about specific chemical forms and bioavailability from different soil types that contribute to variations in uptake. Different exposure duration RfDs, such as acute chemical injury to the skin, need to be developed.

3.2.2. Recommendations

On the basis of the review of the guideline toxicity studies, the Technical Panel makes the following recommendations (in no particular order) regarding the development of testing procedures and guidance for their use. In identifying the need for development of specific protocols, the Technical Panel is not recommending that these tests be used for every chemical or in all circumstances, as pointed out at the beginning of the chapter.

- *Develop a strategy for alternative approaches to toxicity testing, with guidance on how and when to use existing and newly recommended guidelines.* Information on all aspects of a chemical should be considered in the strategy for testing, including chemical-physical characteristics, intended use, and toxicokinetic and toxicodynamic (mode of action) data, to allow a more efficient and targeted testing approach. In addition, the strategy should consider life stages in evaluating exposures and outcomes, as well as other sensitive subpopulations.

- *Develop guidelines or guideline study protocols that will provide more systematic information on toxicokinetics and toxicodynamics (i.e., mechanism or mode of action), including at different life stages.* Such studies could provide information that would be relevant to susceptible subpopulations, including life stages (i.e., inform the

selection of the intraspecies UF). Such studies also could provide information on species differences (i.e., inform the selection of the interspecies UF). Finally, such studies can provide information to conduct route-to-route extrapolations and reduce the number of route-specific tests required to derive a reference value.

- *Develop protocols for acute and short-term studies that provide more comprehensive data for setting reference values (see Section 3.3).* The existing protocols for acute studies (except for the acute inhalation protocol with histopathologic evaluation) generally collect data only on what could be called frank effects, which may not be protective of more subtle effects.

- *Modify existing guideline study protocols to provide more comprehensive coverage of life stages for both exposure and outcomes (see Section 3.3)*. Existing guideline studies do not include, for example, the evaluation of toxic effects that may occur in old age from prenatal or early postnatal exposure (including carcinogenesis) or premature aging from exposure earlier in life.

- *Collect more information from less-than-lifetime exposure to evaluate latency to effect and to evaluate reversibility of effect.* Existing guideline studies, with the exception of the acute tests and some developmental toxicity studies, expose animals up to the time of testing. Some form of "stop exposure" studies would provide useful information that could increase or decrease the level of concern for an observed toxic event.

- *Develop guidelines or guideline study protocols to assess immunotoxicity, carcinogenicity, and cardiovascular toxicity at different life stages.* Immunotoxicity and cardiovascular toxicity are presently looked at only in a cursory manner. Carcinogenicity is currently evaluated only after chronic exposure to adult animals. There is a need to integrate functional measurements into evaluations of these and other systems.
- *Explore the feasibility of setting dermal reference values for direct toxicity at the portal of entry, including sensitization.* Reference values have been derived for lesions in the gastrointestinal and respiratory tracts from direct exposure. The lack of procedures for dealing with similar effects on the skin is a glaring omission.

3.3. OPTIONS FOR ALTERNATIVE TESTING APPROACHES

The Technical Panel explored alternative testing protocols for acute toxicity testing as well as alternative protocols for subchronic/chronic toxicity testing. These are offered here as alternatives that may be used, depending on the agent being tested or the type of reference values needed.

3.3.1. Alternative Acute Toxicity Testing Protocol

The current EPA test guidelines for acute toxicity focus on the determination of an LD50 in adult test species. A gross necropsy is conducted on the animals, and histologic evaluation of target organs may or may not be conducted. Therefore, very limited information is obtained from the current protocol that would be useful for determining an acute reference value. However, a number of alternative study designs are available that would provide information for consideration in establishing the acute reference value (Gad and Chengelis, 1998).

One basic study design is shown in Figure 3-6. In this protocol, a control group and a minimum of three dose groups with 10 animals/sex/group are used. The animals are dosed once on day 1 and followed for 2 weeks. Clinical signs of toxicity are recorded daily, food consumption and body weights are recorded on days 1–4, 8, and 14. There is an interim sacrifice of 5 animals/sex/group at 3 days after dosing and a final sacrifice of the remaining animals at 2 weeks after dosing. At both sacrifices, hematological and clinical chemistry analyses are conducted, as is a urinalysis. The animals are necropsied, organ weights are recorded, and the organs are examined histologically.

Because the purpose of this study design is to provide hazard and dose-response information rather than determination of an LD50, the dose levels should be chosen accordingly. This study would initially be conducted on adult animals. As information is obtained from other toxicology and/or toxicokinetic studies, it may be necessary to conduct the study with animals at different life stages and to include other endpoints.

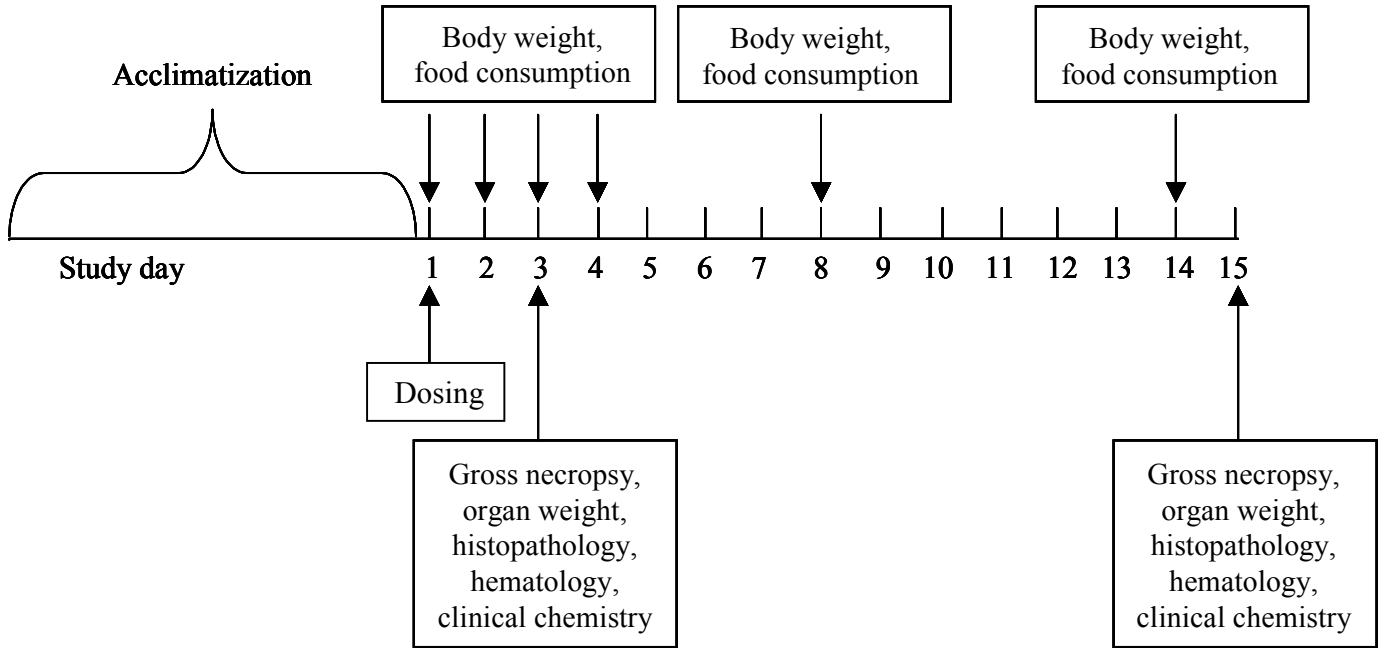


Figure 3-6. Alternative acute toxicity protocol.

3.3.2. Alternative Chronic Toxicity Testing Protocols

As stated, a review of currently available EPA guideline toxicology studies (OPPTS 870 Series) demonstrates that there is no single protocol that addresses continuous exposure through all life stages of any test species. To address this issue, two possible alternative study designs were considered: the "expanded chronic/carcinogenicity study" and the "unified screening study." These are described in some detail below and are illustrated in accompanying figures. The intent of this discussion is to demonstrate the advantages (and disadvantages) of exploring nontraditional testing paradigms; however, such discussion does not constitute a recommendation for implementation. For many chemicals, the existence of adequate (by Agency standards) stand-alone studies would preclude the need for further testing, with or without expanded or combined protocols such as those described below. In any case, any proposal to use alternative study designs in a regulatory setting should be thoroughly discussed by Agency and registrant scientists prior to study initiation.

3.3.2.1. *The Expanded Chronic/Carcinogenicity Study*

An example of a study design that would incorporate lifetime (in utero through old age) exposure is the expanded chronic/carcinogenicity study (shown in Figure 3-7), which could serve as a replacement for a standard guideline chronic/carcinogenicity study in rats. In this expanded study, female rats are assigned to treatment groups, mated, and treated with test substance throughout gestation and lactation. When pups are weaned on PND 21, they are assigned individual animal numbers and maintained within their established treatment group. Prenatal and early postnatal exposure to the test substance in this study is similar to that required for the in utero carcinogenicity study that is used to evaluate food additive chemicals for regulation by the Food and Drug Administration's Center for Food Safety and Nutrition.

The difference here is that the study duration is extended to a period of 3 years (vs. a typical chronic duration of 2 years for rats), with interim sacrifices scheduled at yearly intervals. The total number of animals used in this expanded study is greater than for a standard guideline chronic/carcinogenicity study because of the additional interim sacrifice; for each annual segment, the sacrifice of 25 rats/sex/group is required. To reduce this number, the study could be conducted with fewer animals per segment (e.g., 20/sex/group), or only two sacrifices could be scheduled (e.g., at 1.5 and 3 years). Of course, such actions will either reduce the power of the evaluation for tumor data or will eliminate examination of an important life phase.

Parameters typical of a guideline chronic/carcinogenicity study are examined in this expanded study (e.g., mortality, clinical observations, body weight, food consumption, clinical

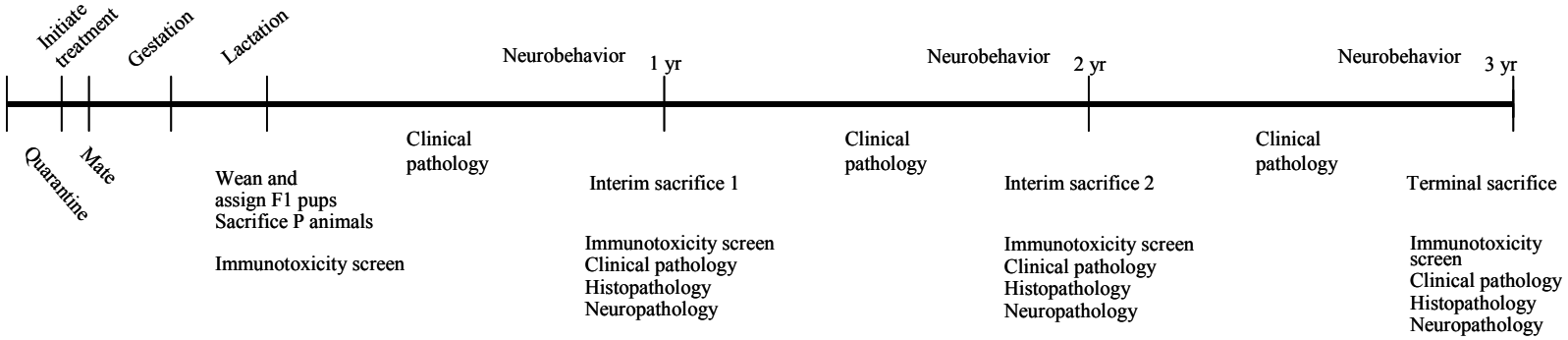


Figure 3-7. Expanded chronic/carcinogenicity study.

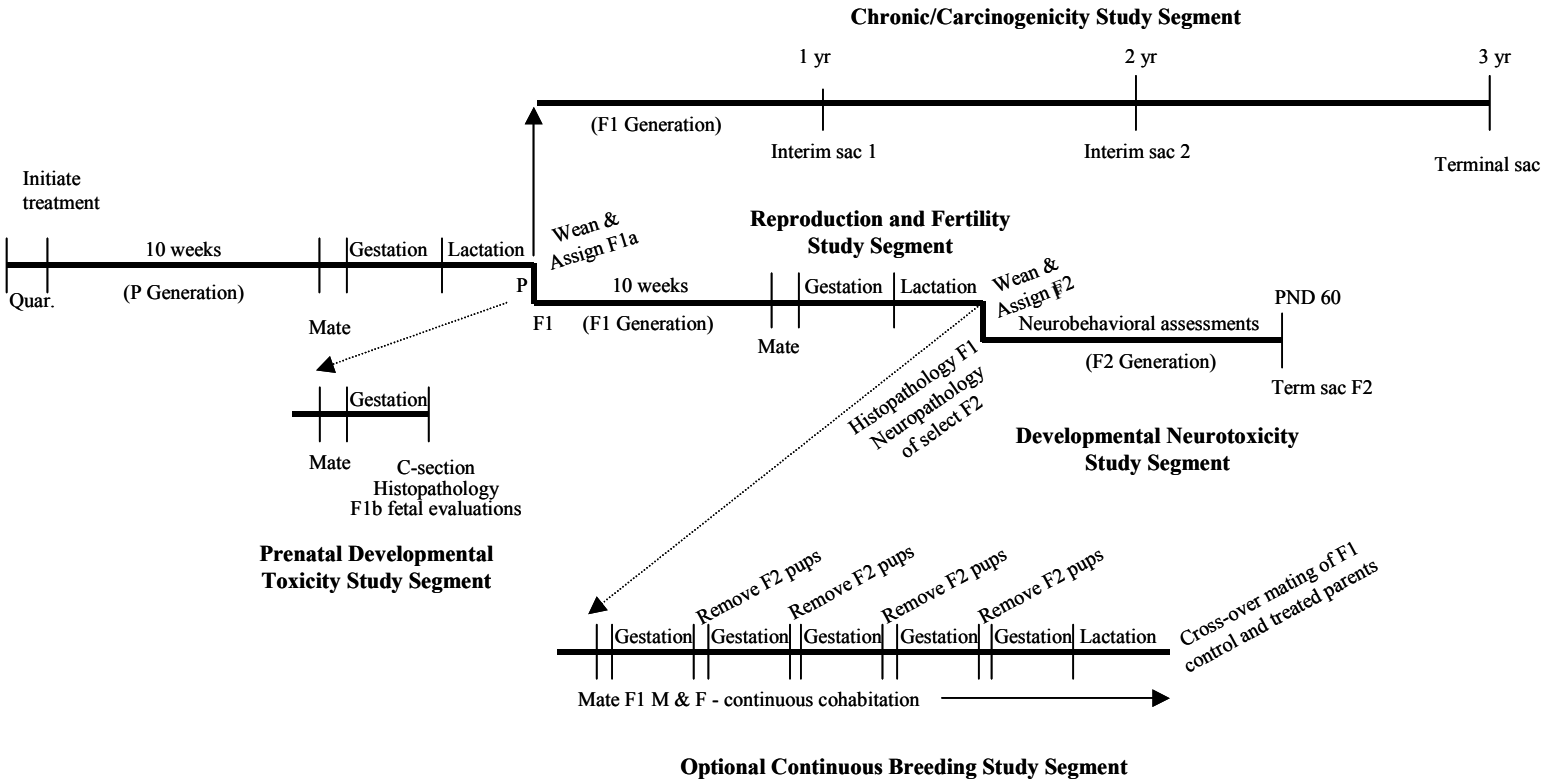
chemistry and hematology, ophthalmology, gross pathology, and histopathology). In addition, neurological and immunological evaluations are performed in the adult animals at multiple intervals into old age, which, along with the fact that the animals are exposed to the chemical during all life stages, contributes to the superiority of this study design.

Although the temporal linear nature of this study protocol makes it less complicated to conduct in the laboratory, this attribute also results in the inability to easily assess some other important endpoints, such as prenatal development, reproduction and endocrine function, and DNT. Additionally, by 3 years of age, when this study would be terminated, survival in laboratory rats may be compromised; therefore, it may be necessary to consider using feed restriction to maximize the number of animals available for in vivo and post mortem assessment of aged animals. In addition, housing from birth in specific-pathogen-free facilities may be necessary to maintain sufficient viable animals for such an extended period of time (see the background white paper on aging, Versar Inc., 2001a).

3.3.2.2. *The Unified Screening Study*

An alternative study design, the unified screening study, is illustrated in Figure 3-8. This study is composed of at least four segments: a two-generation reproduction and fertility study, an expanded chronic/carcinogenicity study, a developmental toxicity study, and a DNT study. Each of these is currently conducted as a separate study. An optional continuous-breeding study segment could be added to the design. When conducted in the rat, the unified screening study assesses all life stages of the animals and provides a means to evaluate prenatal developmental toxicity, DNT, reproduction, and endocrine function, all within animals that are derived from the same gene pool and are evaluated within two generations of the progenitor rodents that are initially placed on study.

The unified screening study begins as a typical two-generation reproduction and fertility study, with 10 weeks of treatment, mating, gestation, and lactation phases conducted according to OPPTS 870.3800. The F1 weanlings are selected for either the second generation of the reproduction and fertility study or the expanded chronic/carcinogenicity study. (As a point of clarification, at any point that animals are selected and/or assigned to a different study segment, it is assumed that the treatment group remains constant for each animal.) The parental (P) animals from the first generation are not immediately terminated; rather, they are transferred to a prenatal developmental study segment. After a short rest, they are mated. The P males can be terminated at any time point; the P females are continued through to caesarian section on approximately GD 20. The resulting F1b fetuses are processed and examined for external, soft tissue, and skeletal



Quar. = Quarantine P = Parental F1 and F2 = 1st and 2nd filial generations

Figure 3-8. Unified screening study. Study lines are not drawn to scale.

abnormalities, as is typical in an OPPTS 870.3700 study. At necropsy, however, the P-generation animals receive an extended postmortem examination, according to the procedures for the two-generation reproduction and fertility study, that includes sperm measures for the males and extensive histopathology of the reproductive and other organ systems for both sexes.

The expanded chronic/carcinogenicity study segment, using F1 animals, would continue as described above concurrently with all other segments of the unified screening study but continuing well past the time that the others have been terminated. The other F1 pups that are selected as second-generation parental animals in the reproduction and fertility study segment are treated for 10 weeks and then undergo the standard reproductive functional assessments, as specified in the OPPTS 870.3800 guideline. Because a number of F2 pups from this generation will continue on into the DNT study segment, some additional observations are required during the lactation segment of the second generation. Specifically, F2 pups are selected and assigned for neurobehavioral assessments on PND 4 (at the time of litter standardization). Preweaning observations include weekly age-appropriate clinical/functional behavioral observations conducted outside of the home cage and motor activity assessments on PNDs 13 and 17. Additional assessments of physical, reflex, and sensory development may also be conducted during this period.

At the time of weaning of the F2 pups on PND 21, those preselected for neurobehavioral assessment continue into the DNT segment and other weanlings are sacrificed for postmortem evaluations that address the considerations of both the reproduction protocol (including organ weight data) and the DNT protocol (requiring in situ perfusion fixation of tissues and neuropathology, including morphometric analysis). The DNT-segment F2 animals are evaluated as per OPPTS 870.6300, which includes multiple assessments of clinical and functional observations, motor activity, auditory startle habituation, and learning and memory. They are maintained until termination (with postmortem evaluations, including neuropathology following perfusion fixation) at approximately PND 60.

Also at the time of weaning of the F2 pups, a decision could be made to either sacrifice the F1 parental animals immediately (with the usual sperm measures and postmortem evaluations) or to maintain them through a continuous-breeding reproduction study segment, sequentially mating the F1 adults for the production of five litters (the pups from these litters are terminated in early lactation). This continuous-breeding study segment, which would extend the reproduction study for about 100 additional days, uses a standardized assessment protocol that has been well characterized in the peer-reviewed literature (Lamb, 1985; Lamb et al., 1985; Morrissey et al., 1989) but does not have a corresponding OPPTS guideline.

As previously stated, in this unified study protocol, the animals are both exposed and assessed during all life stages, and the evaluation of both structural and functional endpoints for multiple organ systems are maximized in the overall design, for example, by the inclusion of immunotoxicity and neurotoxicity endpoints. There is one notable exception to this statement in that reproductive senescence is not standardly examined. Nevertheless, if the two-generation reproduction study segment identifies problems with fertility or cyclicity, this could be pursued more rigorously by the addition of testing during the second or third year of the expanded chronic/carcinogenicity study, for example, evaluating cyclicity in aged female rats and/or evaluating ovarian follicular counts and atrophy at sacrifice.

Another benefit of using the unified screening study design is that it results in the purchase and use of many fewer naive animals for study initiation and it increases the efficient utilization of animals, particularly of the F2 offspring from the reproduction study.

Although there are obvious benefits in using a unified screening study, there are also a number of concerns or potential problems involved with its conduct. Although it is assumed that treatment levels and route of administration will remain constant across all study segments, this approach to dose-setting and route selection may not always be optimal for every phase. Generally, a temporal nonlinear design of this nature is more difficult to manage in the laboratory. The strain of rat generally used in toxicity studies is the Sprague-Dawley, whereas the Fischer 344 rat is often used in the standard chronic/carcinogenicity study. Fischer 344 rats have not typically been used in reproductive and developmental toxicity studies. The use of either strain for the unified study could compromise the use of historical data for comparison, for example, for the chronic/carcinogenicity study if the Sprague-Dawley is used and for the reproductive and developmental toxicity studies if the F344 is used.

As study complexity increases, so does the opportunity for error. In some cases, a serious technical error in one study segment could compromise subsequent study segments and result in an extensive waste of animals and resources. As with the expanded chronic/carcinogenicity study discussed above (section 3.3.2.1.), survival during the chronic/carcinogenicity study segment in this design may need to be enhanced via feed restriction. Also, if the test substance interferes with reproduction or results in increased mortality, the number of offspring that are available for assignment to subsequent study segments (e.g., the selection of F2 animals for the DNT phase) may be critically reduced. An additional but similar problem could arise when selecting F1 animals for the expanded chronic/carcinogenicity study segment at the same time as for the second generation of the reproduction and fertility study segment, because a large number of offspring needs to be available all at one time. Additionally, the offspring that are assigned to the

chronic/carcinogenicity segment should be genetically diverse within each dose group and should originate from as many litters as possible (i.e., not be siblings).

A number of possible solutions that could be used alone or in combination to increase the number of F1 pups available for selection in other study phases include the following:

1. Reducing the number of animals needed for the expanded chronic/carcinogenicity segment by examining fewer animals at each serial sacrifice or by abandoning the final year of evaluation, as described above.

2. Reducing the number of animals assigned to the second generation of the reproduction and fertility study segment; however, this could compromise the number of F1 offspring that would be available for the DNT study segment.

3. Standardizing litters to 10 rather than 8 pups per sex and assuming that no litter has less than 10 pups and that no pups die during lactation. Because some small litters and neonatal pup deaths almost always occur, even in controls, it is wiser to design the study more conservatively in order to avoid discovering that there are not enough F1 pups to assign to the later segment(s).

4. Assigning additional females to the two-generation reproduction and fertility study segment in order to produce extra F1 pups for selection. Although even a modest increase in the number in each group would increase the probability of producing a sufficient number of F1 pups, a larger number of litters would generally be required in order to ensure genetic diversity among the weanlings that are assigned to the chronic/carcinogenicity study segment. This could be accomplished by placing additional P-generation females or breeding pairs on study, perhaps combined with 2:1 mating procedures, or by mating the males with the reproduction and fertility study-segment females first and then with an extra set of females. One adverse consequence of placing additional females on study so that their litters can be used for selection of genetically diverse offspring for the chronic/carcinogenicity study segment is that this method results in a larger number of excess F1 weanling pups that would not be used for evaluations in this protocol. However, these pups could be used for other evaluations, such as immunotoxicity, specialized neurotoxicity tests, or adult onset disease or diseases of aging.

Some of the above options appear to be more advantageous and preferable than others; however, no recommendation is proffered because the list is presented only to illustrate some of the many possibilities that could be used in a customized study design. It should be noted that simply combining the reproduction and fertility study and the DNT study when a two-generation reproduction and fertility study has not already been conducted greatly reduces the total number of animals that would be required to conduct the two studies individually. No additional animals are required over the reproduction and fertility study alone, and there is greater efficiency in the use of the F2 offspring when the DNT study is conducted in that group.

4. FRAMEWORK FOR SETTING ACUTE, SHORT-TERM, LONGER-TERM, AND CHRONIC REFERENCE VALUES

As noted in Chapter 2, the Technical Panel is recommending that EPA begin deriving acute, short-term, and longer-term reference values in addition to chronic reference values. The approach to reference values discussed here is intended for use in risk assessments for health effects known or assumed to be produced through a nonlinear and/or threshold mode of action. Although there has been a dichotomy between cancer and noncancer risk assessment in terms of the underlying assumption about the linearity or nonlinearity of the dose-response curve, there is a move toward harmonization among approaches for all health effects (Butterworth and Bogdanffy, 1999; Bogdanffy et al., 2001). This includes recognition of the possibility that some carcinogenic agents may work through nonlinear mechanisms (U.S. EPA, 1999a), whereas some agents that produce effects other than cancer may work through linear mechanisms (see discussion in U.S. EPA, 1998d). Thus, the decision to use a linear extrapolation approach or a reference value approach should take into consideration the underlying mode of action and presumed dose-response relationship.

The approach described here is the default approach to be used when the assumption is a nonlinear and/or threshold mode of action, except for cases where other methods have been developed (e.g., in support of the NAAQS). This approach can and should be improved upon or replaced when more specific data on toxicokinetics and mode of action are available to allow the development of a chemical-specific or a biologically based dose-response model for prediction of risks to humans and to susceptible individuals within the population. The acute, short-term, longer-term, and chronic reference values derived on the basis of the recommendations in this report should be included in IRIS after appropriate internal, external, and consensus review. These values would then be available for use by program offices, where appropriate.

In this chapter, we discuss the definitions of the exposure durations and the proposed changes in the definition of the corresponding reference values. In addition, several issues are discussed regarding the adequacy of studies and characterization of the extent of the database with regard to sufficiency of data for deriving reference values. The derivation of reference values also is discussed with regard to dosimetric adjustment and application of UFs. A number of recommendations are made with regard to this process. In particular, the Technical Panel recommends incorporating the concept of life stage and expanding the endpoints evaluated as well as consideration of duration and timing of exposure and latency to response in characterizing the extent of the database used for setting reference values. The Technical Panel

strongly encourages the use of a narrative description of the database, including strengths and limitations, rather than a single confidence statement for support of a reference value.

The adjustments required for derivation of the human equivalent dose (HED) for oral and dermal exposure and the human equivalent concentration (HEC) for inhalation exposure are described and discussed. This is followed by recommendations about the evaluation and comparison of data for the POD, based on an analysis of each potentially limiting endpoint carried through the reference value derivation process, followed by selection of the appropriate health-protective reference value.

Finally, the Technical Panel emphasizes that considerable use of scientific judgment is advisable and necessary in practically all phases of the process, especially in the application of UFs. This review and its recommendations build on the principles in the Agency's handbook on risk characterization (U.S. EPA, 2000b), which calls for transparency, clarity, consistency, and reasonableness in the risk assessment process.

4.1. DEFINITIONS OF EXPOSURE DURATIONS FOR USE IN SETTING REFERENCE VALUES

The Technical Panel proposes the following definitions of exposure duration as a first step in the development of consistent approaches for the Agency. These definitions are based on exposure durations for humans; analogous exposure durations for rodents are indicated for the longer-term and chronic durations. The definitions are not intended to be rigid specifications, but simply general descriptions of the relevant exposure time period. Their application is meant to be flexible, so that, for example, if a 4-month animal study is available, it may be used as the basis for both a longer-term and a chronic reference value.

The definitions were developed on the basis of the review of values currently set by various program offices (see Chapter 2), and they have been standardized to be compatible with those definitions currently used by various program offices within the Agency. The definitions for various durations, as follow, were discussed at an EPA Risk Assessment Forum Colloquium (CDM Group Inc, 2000).

Acute: Exposure by the oral, dermal, or inhalation route for 24 hours or less.

Short-term: Repeated exposure[4] by the oral, dermal, or inhalation route for more than 24 hours, up to 30 days.

Longer-term: Repeated exposure by the oral, dermal, or inhalation route for more than 30 days, up to approximately 10% of the life span in humans[5] (more than 30 days up to approximately 90 days in typically used laboratory animal species[6]).

Chronic: Repeated exposure by the oral, dermal, or inhalation route for more than approximately 10% of the life span in humans (more than approximately 90 days to 2 years in typically used laboratory animal species).

The Technical Panel believes there is an advantage in having a central source of consensus reference values of various exposure durations available to risk assessors throughout EPA. EPA Program Offices could use the values for risk assessments in which the known or assumed exposure duration approximated the exposure duration in the appropriate reference value definition. The Panel recognizes that Program Offices may make further adjustments to the reference values depending on circumstances that are unique in their assessments.

The Technical Panel recommends that the principles of sound science be used when the expanded array of reference values are developed. The Panel cautions that the exposure-response relationships for all durations of exposure and issues of latency need to be carefully considered to ensure that there are no obvious conflicts in the series of recommended reference values for any specific chemical. This analysis can become complex in a case where the toxicological endpoint may differ for the different durations of exposure.

The Technical Panel is aware that there will be data limitations for an individual chemical that may preclude development of all four reference values. For example, currently, a chronic RfD or RfC would not ordinarily be considered for inclusion in the IRIS database unless

---

[4]A repeated exposure may be either continuous, periodic, or intermittent. A continuous exposure is a daily exposure for the total duration of interest. A periodic exposure is one occurring at regular intervals, e.g., inhalation exposure 6 hrs/day, 5 days/wk or oral exposure 5 days/wk. An intermittent exposure is one in which there is no effect of one exposure on the effect of the next; this definition implies sufficient time for the chemical and its metabolites to clear the biological system before the subsequent exposure, that is, noncumulative pharmacokinetics. A periodic exposure may or may not be intermittent.

[5]The lifespan value used depends on the situation under consideration. For example, an average of 70 years has been the typical default used for chronic exposures, but the average life span based on U.S. census data is 75.5 years (U.S. EPA, 1997a).

[6]Typically used laboratory animal species refers to rats, mice, and rabbits, for example.

a subchronic or chronic study were available. Similarly, where data of the type needed for deriving acute, short-term, or longer-term reference values are not available, theses values would not ordinarily be considered for inclusion in the IRIS database. In situations where an acute, short-term, or longer-term reference value is needed but appropriate data fitting the definition for duration are not available, then the Program Office may wish to consider several options. One option would be to not develop a reference value for that particular duration of exposure. Another option would be to use the reference value for the next longer duration of exposure as a conservative estimate of a reference value that would be protective for a short-term exposure duration. For example, the Office of Water (see Chapter 2) will use a longer-term health advisory for a child as a conservative estimate for a 10-day exposure in the absence of data to derive a 10-day health advisory. Other program-specific options might also be considered.

The Technical Panel is aware that time and resources need to be considered when implementing its recommendations. The IRIS program has begun to implement a pilot program to test whether development of the expanded array of reference values is practical and can be accomplished without unduly delaying the completion of an IRIS file. As a part of the pilot, the IRIS program will need to identify the methods to be used in deriving these additional values.

4.2. PROPOSED CHANGES IN THE REFERENCE VALUE DEFINITIONS

In the process of considering definitions for different duration reference values, the Technical Panel discussed several issues that have been raised about the current definitions of the chronic RfD and RfC (Box 4-1). The following items describe the issues and the recommended changes.

1. The parenthetical statement in the current RfD and RfC definitions—"with uncertainty spanning perhaps an order of magnitude"—has been variously used by risk assessors and risk managers to mean that the estimate is at the upper end, the lower end, or the middle of the range of an

Box 4-1. Current definitions for the chronic oral RfD and inhalation RfC

RfD: an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or BMD, with UFs generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments.

RfC: an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or BMD, with UFs generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments.

order of magnitude. This statement has been removed from the proposed revision of the definition for reference value (Box 4-2), and it is recommended that issues of uncertainty/variability be discussed qualitatively as part of the weight of evidence and characterization of the database. Attempts to adapt such a qualitative derivation process to formal quantitative procedures for prediction of accuracy presents major difficulties. A particularly obvious difficulty is that the same definition and phrase were applied to different reference values that may have varied markedly in their underlying data and, thus, their potential for accuracy. For example, the same "order-of-magnitude" range applied equally to a robust reference value with known exposures plus observable and quantifiable dose-response data derived from a segment of the human population and to a marginal reference value based only on animal data with minimal supporting information.

Box 4-2. Proposed revisions in the reference value definitions

Reference Value: an estimate of an exposure, designated by duration[a] and route, to the human population (including susceptible subgroups[b]) that is likely to be without an appreciable risk of adverse health effects over a lifetime. It is derived from a BMDL, a NOAEL, a LOAEL, or another suitable point of departure, with uncertainty/variability factors[c] applied to reflect limitations of the data used.

---

[a] The generalized durations are similar to those given in Section 4.1. for acute (≤24 hours), short-term (up to 30 days), longer-term (up to 10% of average lifespan), and chronic (up to a lifetime), all considered to be continuous exposures throughout the duration specified.

[b] Susceptible subgroups may refer to life stages, e.g., children or the elderly, or to other segments of the population, e.g., asthmatics or the immune-compromised, but they are likely to be somewhat chemical specific and they may not be consistently defined in all cases. See below (Section 4.3.2.3) for further discussion.

[c] See discussion later in this chapter (Section 4.4.5) on application of uncertainty/variability factors.

The Technical Panel notes a lack of support from the external reviewers of this document for any such prediction and recommends that the database characterizations for reference values be approached in a comprehensive way, as discussed in Section 4.3 to ensure that they are authoritative and as complete as possible in order to yield qualitative information about the range that could be predicted around the individual estimates rather than attempting quantitative evaluations of accuracy and ranges.

2. The term "deleterious" is considered ambiguous by some, so it has been replaced with the term "adverse," because the latter is more commonly understood in the context of data evaluation and selection of endpoints for setting reference values.

3. In the spirit of harmonization of risk assessment approaches for human health effects, it has been recommended that health effects no longer be categorized as "cancer" or "noncancer" for the purposes of hazard characterization and dose-response analysis (U.S. EPA, 1997b, 1998d; Bogdanffy et al., 2001). As indicated earlier, the approach to reference values discussed here is intended for risk assessments for any type of health effect known or assumed to be produced through a nonlinear and/or threshold mode of action (which may include U-shaped or other nonmonotonic dose-response curves as well as thresholds). In light of this recommendation, the term "noncancer" has been removed from the definition, denoting the move toward defining approaches for low-dose estimation or extrapolation based on mode of action. It is recommended that this issue be considered further in the deliberations by the Risk Assessment Forum's Technical Panel on a framework for harmonization of approaches for human health risk assessment.

To fulfill the need for consistency in the designation of various duration reference values, the Panel recommends that the terminology for reference values be standardized. Rather than continuing to use RfD and RfC only to denote chronic oral and inhalation reference values, respectively, standardized terminology should be developed that denotes both duration and route of exposure. Although Technical Panel members did not come to agreement on the best way to do this (and we welcome alternative suggestions), the terminology shown below is offered as an example of the way in which consistent labels could be developed and used. Either new standard terminology, (e.g., reference value) could be used, or RfD and RfC could continue to be used, but they would always need to be accompanied by the qualifying duration of exposure and, in the case of the RfD, by the route of exposure. Thus, the following alternatives for terminology are offered:

Acute (Oral, Dermal) Reference Value or Dose, Acute (Inhalation) Reference Value or Concentration: RfV_{AO}, RfV_{AD}, RfV_{AI}; RfD_{AO}, RfD_{AD}, RfC_{AI} or RfC_{A}

Short-term (Oral, Dermal) Reference Value or Dose; Short-term (Inhalation) Reference Value or Concentration: RfV_{SO}, RfV_{SD}, RfV_{SI}; RfD_{SO}, RfD_{SD}, RfC_{SI} or RfC_{S}

Longer-term (Oral, Dermal) Reference Value or Dose; Longer-term (Inhalation) Reference Value or Concentration: RfV_{LO}, RfV_{LD}, RfV_{LI}; RfD_{LO}, RfD_{LD}, RfC_{LI} or RfC_{L}

Chronic (Oral, Dermal) Reference Value or Dose; Chronic (Inhalation) Reference Value or Concentration: RfV_{CO}, RfV_{CD}, RfV_{CI}; RfD_{CO}, RfD_{CD}, RfC_{CI} or RfC_{C}

The Panel recommends that endpoint- or life stage-specific reference values such as the RfD_{DT} (reference dose for developmental toxicity), which were originally proposed in *Guidelines for Developmental Toxicity Risk Assessment* (U.S. EPA, 1991), not be derived. Rather, a sample reference value should be calculated for each relevant and appropriate endpoint and these should then be considered in the derivation of various duration reference values. Reference values should be derived to be protective of all types of effects for a given duration of exposure and are intended to protect the population as a whole, including potentially susceptible subgroups. Thus, the RfD_{DT} concept of a critical window of exposure for some health effects is addressed in the adoption of the less-than-chronic reference values. This recommendation does not preclude, however, using specific common endpoints in the assessment of cumulative risk for mixtures or chemicals that have a common mode of action or for risk management purposes.

4.3. CHARACTERIZATION OF THE EXTENT OF THE HEALTH-RELATED DATABASE FOR SETTING REFERENCE VALUES

A necessary first step in hazard characterization is the critical evaluation of all pertinent and relevant human and animal data that are available in the open literature as well as data submitted to the Agency in response to various regulatory standards, data call-ins, or other requirements and agreements.

4.3.1. Review of Studies

Data will be available from a wide variety of sources, including studies conducted according to EPA guidelines, studies conducted by industry using Organization for Economic Cooperation and Development or other protocols, experimental studies conducted by academic researchers, epidemiology studies, case reports or series, and controlled clinical studies in

volunteers.[7] These studies will be of widely differing quality; EPA must evaluate each study to determine whether it is of acceptable quality.

4.3.1.1. *Adequacy of Studies*

The following list of questions could be helpful in the process of evaluating data from animal and human studies.

All types of studies:

- What was the purpose of the study and is there a clearly delineated hypothesis?

- Is there sufficient description of the protocol, statistical analyses, and results to make an evaluation?

- Were the appropriate endpoints assessed in the study?[8] Were the techniques used for the assessment scientifically sound?

- Were appropriate statistical techniques applied for each endpoint? Was the power of the study adequate to detect effects?

- Did the study establish dose-response relationships? Was a BMD lower confidence level (BMDL), LOAEL or NOAEL established?

- Is the shape of the dose-response curve consistent with the known toxicokinetics of the test compound?

---

[7]Currently, OPP is reviewing its policy concerning use of human data from studies in which there is intentional pesticide exposure, and it has asked the National Academy of Sciences for input on the acceptability of such studies and ethical criteria for their use under the Protection of Human Subjects Rule (the "Common Rule") (EPA, 2001c).

[8]A chemical may cause a variety of toxic effects depending on the amount, duration, timing, and pattern of exposure (i.e., continuous, periodic, or intermittent). These effects may range from severe—such as death—to more subtle biochemical, physiological, or pathological changes in one or more organ systems. In addition, the effects will vary depending on their latency following exposure and when the observations are made. Primary attention is given in risk assessment to those effects in the lower exposure range and/or the effects most biologically appropriate for a human health risk assessment.

- Do effects fit with what is known about mode of action?

- Is the dose-response curve for precursor events consistent with the dose-response curve for clinical effects?

- Are the results of the study biologically plausible?

- What uncertainties exist? Do the results of the study indicate the need for follow-up studies to reduce uncertainties?

- Are the study conclusions supported by the data?

Human studies:

- What were the data sources for exposure, health status, and risk factors (e.g., questionnaires, biological measurements, exposure/work history record reviews, or exposure/disease registries) and what were their strengths and limitations?

- What methods were used to control, measure, or reduce various forms of error (e.g., misclassification or interviewer bias, confounding factors and potential effect modifiers) and their potential impact on the findings? What is the validity (accuracy) and reliability (reproducibility) of the methods used to determine exposure and outcome? What were the response rates?

- What major demographic and other personal factors were examined (e.g., age, sex, ethnic group, socioeconomic status, smoking status, and occupational exposure)? What other climate or life stage factors were important for the endpoints and exposures assessed?

- Were the findings examined for biologic plausibility, internal and external consistency of the findings, and the influence of limitations of the design, data sources, and analytic methods?

Animal studies:

- Was the study sufficiently documented (e.g., conducted in accordance with good laboratory practices)?

- Were appropriate analytical techniques used to measure the stability, homogeneity, and actual level of the test substance in the study (in the water, feed, air, etc.)?

- Was an appropriate animal species used?[9] Was an appropriate number of animals used? Were sex and age considered?

- Were the dose levels appropriate? What was the basis for choosing the dose levels?

- Was an appropriate method used to assign the animals to dose groups?

- Was an appropriate route and matrix of exposure employed?[10]

- Was the duration of exposure adequate for the particular study design?

- Were possible alterations in metabolism considered at the higher exposure levels?

---

[9]The laboratory animals used most often are the rat, mouse, rabbit, guinea pig, hamster, dog, or monkey. When reviewing these studies, the risk assessor makes judgments about the ability of the study to predict the potential for toxicity in humans and tries to select data from the species that is most relevant to humans using the most defensible biological rationale. When available, comparative toxicokinetics can be used to support this decision. Absent a clearly most-relevant species, the most sensitive mammalian species is used, that is, the species that shows toxicity at the lowest exposure level.

[10]The most appropriate route of exposure is the route for which an evaluation is to be made. The toxicity of the chemical may differ with route of exposure because of differences in mechanism of action or toxicokinetics (absorption, distribution, metabolism, and excretion). Development of data to establish dosimetry for the purpose of route-to-route extrapolation is encouraged; however, route-to-route extrapolation is inappropriate when based exclusively upon default assumptions regarding exposure and toxicokinetics. Even within the same route of exposure, responses may differ due to alterations in toxicokinetics, for example, dietary or water exposure versus oral gavage.

Professional judgment is required to decide, on the basis of a thorough review of all available data and studies, whether any observed effect is adverse and how the results fit with what is known about the underlying mode of action. These judgments require the input of experts trained in toxicology, statistics, and epidemiology and, often, of specialists in the structure and function of the target organ systems. Both the biological and the statistical significance of the effects are considered when making these judgments. Biological significance is the determination that the observed effect (a biochemical change, a functional impairment, or a pathological lesion) is likely to impair the performance or reduce the ability of an individual to function or to respond to additional challenge from the agent. Biological significance is also attributed to effects that are consistent with steps in a known mode of action. Statistical significance quantifies the likelihood that the observed effect is not due to chance alone. Precedence is given to biological significance, and a statistically significant change that lacks biological significance is not considered an adverse response.

For many discrete or quantal endpoints (e.g., birth defects, tumors, or some discrete pathological changes), this judgment is more straightforward because criteria have been established for deciding what type and incidence of effects are to be considered to be adverse, and an increase above the background rate can be judged using statistical tools. In the case of continuous measures (e.g., body weight, enzyme changes, physiological measures), this tends to be more difficult, because the amount of change to be considered adverse has not been defined by toxicologists or health scientists. Consequently, the endpoint is often decided in the context of the endpoint itself, the study, and the relationship of changes in that endpoint to other effects of the agent.

Decisions about the amount of change to consider adverse must always be made using professional judgment and must be viewed in light of all the data available on the endpoint of concern. All toxicological data on a chemical must be reviewed before deciding whether an effect is biologically significant and adverse. Using a default cutoff value to define adversity for continuous measures may result in an inappropriate interpretation of data and less than optimum evaluation of a chemical's effects.

4.3.2. Issues to be Considered in Characterizing the Database for Risk Assessment

4.3.2.1. *The Weight-of-Evidence Approach*

A weight-of-evidence approach such as that provided in EPA's RfC Methodology (U.S. EPA, 1994) or in EPA's proposed guidelines for carcinogen risk assessment (U.S. EPA, 1999a) should be used in assessing the database for an agent. This approach requires a critical

evaluation of the entire body of available data for consistency and biological plausibility. Potentially relevant studies should be judged for quality and studies of high quality given much more weight than those of lower quality. When both epidemiological and experimental data are available, similarity of effects between humans and animals is given more weight. If the mechanism or mode of action is well characterized, this information is used in the interpretation of observed effects in either human or animal studies. Weight of evidence is not to be interpreted as simply tallying the number of positive and negative studies, nor does it imply an averaging of the doses or exposures identified in individual studies that may be suitable as PODs for risk assessment. The study or studies used for the POD are identified by an informed and expert evaluation of all the available evidence.

4.3.2.2. *Use of Human and Animal Data in Risk Assessment*

Adequate human data are the most relevant for assessing risks to humans. When sufficient human data are available to describe the exposure-response relationship for an adverse outcome(s) that is judged to be the most sensitive effect(s), reference values should be based on human data. Much more data on a wide range of endpoints typically are required to establish confidence that there are no effects of exposure. If sufficient human data are not available to provide the basis for reference values, data from animal studies must be employed. It is advantageous if some human data are available to compare with effects observed in animals, even if the human data are not adequate for quantitative analysis. Availability of data on effects in humans at least allows qualitative comparison with effects observed in animals for determining whether toxicity occurs in the same organ systems and whether the nature of the effects is similar or different. If no human data are available, reliance must be exclusively on animal data. In that case, attention should be paid to whether data are available in more than one species and, if so, whether the same or similar effects occur in different species and possible sources of any observed differences.

One of the major default assumptions in EPA's risk assessment guidelines is that animal data are relevant for humans (e.g., U.S. EPA, 1991, 1996, 1998c). Such defaults are intended to be used in the absence of experimental data that can provide direct information on the relevance of animal data.

Several types of information should be considered when determining the relevance or nonrelevance of effects observed in animal models for humans. This information is used in a variety of ways, from determining the role of metabolism in toxicity (Is the parent chemical or a metabolite responsible for toxicity?), to assessing whether homologous activity would be

expected across species (Do humans share the sensitivity of the animal model, or is the response due to some species-specific idiosyncratic reaction?), to determining whether or not a threshold is likely to exist for the response (Are repair mechanisms capable of maintaining a homeostatic process?). All of this information must be weighed in light of the known heterogeneity of the human population versus the relatively inbred status of laboratory animals used in toxicity testing studies and housed under carefully controlled environmental conditions.

Table 4-1 presents several factors to consider when evaluating the weight of evidence about the likelihood of the occurrence of effects in humans that is based on animal data (in conjunction with human data, if available). The table is not necessarily intended to delineate all factors that may need to be considered, but rather to provide a framework for evaluation and interpretation. It is important to evaluate the database in a holistic manner, determining strengths and weaknesses that are relevant to the overall assessment. Each chemical and database presents a unique set of issues that must be evaluated critically and thoughtfully.

The dose-response nature of the data is an important characteristic of the database or individual study. When data are **dose related**, that is, when the incidence and/or intensity of response changes in an orderly manner as a function of dose, the effect should be considered to be of greater importance than when there is no apparent association between exposure and toxicity. Note, however, that the dose-response relationship need not be monotonic. U-shaped (or inverted U-shaped) dose-response functions are not uncommon in toxicology. For example, a chemical may induce an enzyme at low doses and inhibit it at high doses. Similarly, many solvent-like chemicals (including alcohol) produce increased motor activity at lower doses and depressed activity at high doses.

Similarly, comparative **toxicokinetic/metabolism** data that suggest qualitative and quantitative comparability to that in humans would support the relevancy of animal data. Evidence suggesting a difference in toxicokinetics/metabolism would require additional exploration regarding whether the difference(s) results in a major qualitative or quantitative difference in internal dose in humans.

The **similarity of effects** between species is also an important aspect in characterizing the database. Similar effects in more than one species indicate that the effect provides increased weight of evidence for the risk assessment process, even if such data are not available in humans. In contrast, response data that show inconsistency of effects among studies and/or species that cannot be explained by differences in toxicokinetics/metabolism or timing and/or magnitude of exposure, may suggest that less emphasis be placed on the effect. "Similarity" does not necessarily require identical effects between species. For example, changes in motor activity in

Table 4-1. Factors for evaluation of the weight of evidence regarding the likelihood of effects in humans

Factor	Increased weight	Decreased weight
Dose-response relationship	Orderly change in effect as a function of exposure (need not be monotonic)	No identified relationship between exposure and magnitude of effect
Toxicokinetics/ metabolism	Qualitative and quantitative comparability between humans and animals	Qualitative and quantitative differences between humans and animals
Similarity of effects	Similar effects in more than one animal species or in animals and humans	Inconsistency of effects among studies and/or species that cannot be explained by differences in timing and/or magnitude of exposure or toxicokinetics/metabolism
Mode of action	Demonstration of homologous mode of action in animal model and humans	Evidence suggesting that the mode of action is species specific and irrelevant to humans
Temporal relationship	Consistent temporal relationship between exposure and effect	Lack of temporality between exposure and effect

animals evaluated in the neurotoxicity screening test and cognitive effects in humans would generally be considered similar, because both are indicative of changes in nervous system function.

Mode of action information is also important in understanding whether a particular effect may be important for humans. For example, a transient reduction in anogenital distance in the postnatal animal following perinatal exposure to an anti-androgen has increased weight if the chemical is also known to act as an anti-androgen in humans. Likewise, the interpretation of increased skeletal variants observed following exposure to many chemicals would be enhanced by data indicating that the mechanistic pathways for these agents and the overall biological significance defined were also a possibility in humans. Mode of action data are also important in determining whether various chemicals work by common modes or mechanisms of action, which would then be considered in a cumulative risk assessment.

Another criterion that is important in evaluating data is the **temporal relationship** between exposure and effect. The exposure should precede the effect at an interval that is consistent with what is known about the toxicokinetics and mode of action of the agent. It may be the case, however, that higher doses produce a shorter latency to effect than do lower doses.

4.3.2.3. *Characterization of Effects in Potentially Susceptible Subpopulations*

A dose-response analysis for potentially susceptible subpopulations should be done as part of the overall dose-response analysis for health effects in general. "Susceptible" in this context means a differential (greater) response at the same internal dose in a particular segment of the population due to intrinsic (possibly unknown) factors. "Susceptible subpopulations" is used here to refer both to life stages and to other factors that may predispose individuals to greater response to an exposure. Life stages may include the developing individual before and after birth up to maturity (e.g., embryo, fetus, young child, adolescent), adults, or aging individuals. Other susceptible subpopulations may include people with specific genetic polymorphisms that render them more vulnerable to a specific agent or people with specific diseases or pre-existing conditions (e.g., asthmatics). The term may also refer to gender differences, lifestyle choices, or nutritional state.

It is important to recognize that little basis currently exists for a priori identification of susceptible subpopulations for many chemicals. Without other data to raise suspicions, only the evaluation of effects in various segments of the population such as those mentioned above can identify susceptible subpopulations for a particular chemical and a particular set of exposure conditions. In some situations, differential exposure rather than differential susceptibility per se may be the critical issue (e.g., hand-to-mouth activity in toddlers). Economic differences may also result in differential exposure and susceptibility.

A great deal of attention has been given in recent years to the issue of children as a susceptible subpopulation. Several approaches have been proposed for characterizing the database concerning the potential pre- and postnatal toxicity of a particular chemical and providing some guidance as to the weight of evidence or degree of concern for children's health. However, each approach has been developed for a slightly different purpose and, as such, is generally complementary to, but not the same as, the other approaches.

EPA's developmental toxicity (U.S. EPA, 1991) and reproductive toxicity (U.S. EPA, 1996) risk assessment guidelines describe an approach that characterizes the database as sufficient or insufficient to judge whether a chemical does or does not pose a hazard within the context of dose, route, duration, and timing of exposure. The International Programme on Chemical Safety (IPCS) (IPCS, 1995) proposed an approach based on the quality of information gathered in developmental and reproductive toxicity studies and the types of data that were not available from these studies. EPA's draft 10X toxicology report (U.S. EPA, 1999b) further extended the recommendations for characterizing risks to children's health within the context of the FQPA by discussing issues that would increase or decrease the level of concern.

The present report endorses and extends the recommendations of the 10X Toxicology Working Group's report by incorporating the issues dealing with level of concern into a framework for evaluating the evidence regarding the identification and characterization of susceptible subpopulations (see below). A workshop was held recently to discuss aspects of a framework for children's health risk assessment and to emphasize a broader perspective on the issues that should be considered in hazard characterization, dose-response assessment, exposure assessment, and risk characterization for children as a susceptible subpopulation (ILSI RSI, 2001).

In contrast with the attention paid to children and asthmatics as potentially susceptible subpopulations in recent years, little attention has been focused on risk assessment for other potentially susceptible subgroups. As outlined in Chapter 3, there currently are no requirements in EPA animal study protocols for exposure during old age or for outcome evaluations near the end of the life span following earlier life stage exposures. Similarly, healthy animals that are more genetically homogeneous than humans are used in standard toxicity testing protocols, and information on pre-existing conditions or genetic polymorphisms is largely unavailable from animal studies.

Human studies also usually employ healthy nonelderly individuals, although some studies in more susceptible populations have been conducted, such as studies of the effects of air pollutants in asthmatics. Individuals who have identified risk factors that are not the focus of a study are usually excluded from the study sample. It is important to consider such characteristics of the database if human data are used as the basis for the risk assessment.

As can be seen in Table 4-2, several issues must be considered in assessing the potential for some subpopulations, including different life stages, to have greater susceptibility than others to a chemical. These include the **timing (life stage)-response relationship,** indicating greater susceptibility to exposure at some life stages than at others; whether effects are of a **different type** in identifiable subgroups of the population; and the **dose-response relationship**, that is, whether effects are observed at different levels of exposure in different subpopulations.

Another important consideration is whether effects are observed at the same dose but with a shorter **latency** in different subpopulations. Additionally, differences among groups in terms of the **seriousness** and **reversibility of effects** must be considered. For example, an agent may produce relatively mild and reversible neurological effects in adults but produce permanent behavioral impairment following in utero exposure. It is also important to keep in mind that effects that may initially appear to be reversible may re-appear later or be predictive of later adverse outcomes. This is probably best exemplified by certain outcomes following a

Table 4-2. Factors for evaluating evidence regarding identification and characterization of susceptible subpopulations[a]

Factor	Increased weight	Decreased weight
Timing (life stage) - response relationship	Effects occur at greater magnitude at one or more life stage(s)	No difference in effects at different life stage(s)
Type of effect	Different types of effects in specific subpopulations	Same effect(s) across all potential subpopulations
Dose-response relationship	Effect occurs at lower exposures in one or more subpopulation(s)	No evidence for differential dose-response across different subpopulations
Latency of effect	Latency to observed effect different in specific subpopulations	No difference between subpopulations in latency to effect
Seriousness/ reversibility of effects	Effects different in seriousness or degree of reversibility in specific subpopulations and/or differences in later consequence of an initially reversible effect	No differences between subpopulations in seriousness and/or reversibility of effects, or in later consequences of an initially reversible effect

[a] Subpopulations may be defined by gender, individuals at different life stages (fetus, child, adult, elderly), differences in genetic polymorphisms, and/or pre-existing diseases or conditions that may result in differential sensitivity to adverse effects from exposure to a specific toxic agent.

developmental exposure; for example, an initial depression in birth weight or weight gain or subtle developmental retardation may be indicators of more serious abnormalities later in life.

4.3.3. Characterization of the Extent of the Database

The derivation of an RfD or an RfC is a multifaceted process that involves the coordination of data gathering and evaluation, analysis and judgment in varying proportions, and integration of all the information available. A vital part of the chronic RfD and RfC derivation process that relies heavily on judgment, for example, is the current approach to characterizing the database. For example, the minimum dataset for low-confidence and high-confidence RfDs and RfCs has been specifically defined as follows (U.S. EPA, 1994, 2002c): *minimum dataset for a low confidence chronic RfD or RfC* is a single subchronic study. The *minimum dataset for*

a high confidence chronic RfD or RfC is a chronic study in two species, a single two-generation reproductive toxicity study, and a developmental toxicity study in two species by the appropriate route of exposure.

The Technical Panel is recommending a somewhat different approach. Instead of specifying particular studies, this approach emphasizes the types of data needed (in terms of both human and animal data) for deriving reference values and recommends the use of a narrative description of the extent of the database rather than a single confidence statement. The Technical Panel believes that this approach encourages the use of a wider range of information in deriving reference values that take into consideration the issues of duration and route of exposure, the timing of exposures, the types and extent of endpoint assessment (i.e., structural and function), the susceptible subpopulations evaluated, and the potential for latent effects and/or reversibility of effects. In addition, this approach encourages the identification of data that would be needed or useful for improving the risk assessment for a particular chemical or group of chemicals.

To characterize the database, the Technical Panel has developed a description of a "minimal" database and a "robust" database as a way of describing the range of data that can be used for deriving a reference value (Box 4-3). A great deal of scientific judgment is necessary when evaluating the extent of the database for a particular chemical. Defining the extent of the database requires an overall evaluation and judgment as to where in the minimal–robust continuum the available database should be characterized. The Technical Panel purposely did not define additional categories between minimal and robust (moderate), and the Panel has serious concerns about developing such categories because of the tendency to try to characterize a database with single word descriptors. Instead, we strongly support a narrative description of the extent of the database, with emphasis on the strengths and limitations of the data. It should also be noted that a database that is less than minimal should not be used to derive a reference value.

Rather than presenting separate "minimal" and "robust" database descriptions for each type of reference value that might be derived, the descriptions in Box 4-3 are intended to apply generally across the various reference value types (e.g., acute, short-term, longer-term , or chronic durations for oral, dermal, or inhalation routes of exposure). Additionally, it is expected that the different types of reference values for a particular chemical will be developed within the same assessment. In this manner, the entire database for a chemical may be relied upon in the development of each of the different values (e.g., important and relevant insights may be gleaned

from toxicity studies for exposure durations other than those directly corresponding to the type of reference value being developed).

A minimal database as defined above can be used to set reference values, but the limitations of such a database should be clearly recognized and discussed in the narrative description. For example, a minimal database may provide data on only one duration or route of exposure or it may be specific to only one endpoint or organ system. Thus, the uncertainties related to such a database will be great and should be reflected in the size of the UFs applied for reference value derivation (see further discussion below).

Box 4-3. Description of minimal and robust databases

Minimal Database: no human data available, route-specific toxicity data are limited to dose-response data applicable to the duration in question with assessment of endpoints other than mortality. A study showing only effect levels for mortality or other extremely severe toxicity would not be sufficient to set a reference value.

Robust Database: includes extensive human and/or animal toxicology data that cover route-specific information on many health endpoints, durations of exposure, timing of exposure, life stages and susceptible subpopulations. In the absence of complete human data, mechanistic and other data show the relevance of the animal data for predicting human response. Specifically, the dose-response data for the reference value in question includes endpoint-specific data (e.g., developmental toxicity, neurotoxicity) coupled with toxicokinetic information as needed for route-to-route extrapolation. The toxicity studies include the evaluation of a variety of endpoints (e.g., hematological, clinical, histology of target organs) and endpoints specific to any known hazard characterization. The database for a reference value of less-than-chronic duration has also addressed the issue of reversibility of effects and latency to response, taking into consideration the possibility that less-than-chronic exposure may lead to effects at some period of time after exposure. Biological and chemical characteristics of the exposure and outcomes, as well as known limits on reserve capacities and repair of damage, form the basis for determining the appropriate length of follow-up.

On the other hand, a robust database would address issues of potential toxicity in humans and animals and include data on several durations and routes of exposure as well as a thorough assessment of a variety of health endpoints. It would also include sufficient data on toxicokinetics and mode action to provide extensive information for extrapolation of effects to humans, including potentially susceptible subpopulations. A complete database on a single health endpoint that does not contain information on other endpoints of possible relevancy would not necessarily constitute a robust database, nor would a database that provides complete information on one route and/or duration of exposure be considered robust.

It is clear that a robust database represents a "gold standard" that will rarely, if ever, be available. However, a lack of robustness does not mean that the database is deficient to the extent that a reference value could not be derived or that large UFs would need to be applied. Sound scientific judgement will be required to determine which UFs are appropriate in each case.

A critical assessment of the extent and quality of the database will inform the selection of the endpoints to be used to derive the reference values and the appropriate UFs. A reference value based on a single study would likely have a high degree of uncertainty. As more information from additional toxicology studies, toxicokinetic studies, structure-activity relationships, and human data becomes available, EPA can have greater assurance that the appropriate species, route of exposure, and target organ system(s) are known for each duration reference value needed for a human health risk assessment. As this additional information becomes available, the use of UFs will likely decrease. The ultimate objective is to account for all human health endpoints resulting from exposures over all life stages from before conception to the elderly adult.

The optimum assessment considers subtle effects that impact an individual's quality of life as well as so-called "frank" effects (death and major disease). The evaluation should encompass immediate health outcomes as well delayed responses to an exposure (i.e., latent responses), although most current testing guidelines do not explicitly evaluate latency to response.

4.3.3.1. *Extent of the Database*

The following series of questions regarding the extent of the database can help guide the assessment process:

- Have adequate studies been conducted to establish the target organs/endpoints?
- Have the effects been characterized for both sexes and all life stages?
- Are data pertaining to potentially susceptible subpopulations available?
- Are the responses consistent across species? Are the results of the studies biologically plausible?
- Is the route and matrix of exposure relevant to the specific reference value being derived?
- Is the duration of exposure appropriate for the specific reference value being derived?

- Is the animal species and strain appropriate for extrapolation to humans?

- To what degree may the biological endpoints be extrapolated (qualitatively and quantitatively) to humans?

- Are toxicokinetic data available? Are they available for both sexes, for relevant life stages, for other susceptible subpopulations?

- Is the shape of the dose-response curve consistent with the known toxicokinetics of the test compound?

- Are the metabolism and toxicokinetics in the animal species similar to those of humans?

- Has the dose-response curve been replicated by or is it consistent with data from other laboratories and other test species?

- Have the data for all relevant endpoints been adequately modeled by the BMD or other appropriate quantitative analysis to determine the most sensitive endpoint(s)?

- How well is the toxicity characterized? Do the results of all the studies indicate the possibility of effects on particular systems that have not yet been explored sufficiently or do they indicate that additional studies may reveal effects not yet characterized?

4.4. DERIVATION OF REFERENCE VALUES

After the database has been thoroughly evaluated for quality and extent, as outlined above, several decisions must be made and procedures applied before the final derivation of a reference value. This section summarizes the current procedures and points out assumptions made and areas for improvement and clarification. A variety of factors related to the derivation of reference values is discussed, including the selection of relevant endpoints for the POD for various duration reference values (Section 4.4.1). Adjustment of the study dose/exposure for duration is described in Section 4.4.2, and derivation of a HED or HEC is discussed in Section 4.4.3.

Other issues are discussed briefly in Section 4.4.4, such as varying levels of response at the BMDL, BMCL (lower confidence limit on the BMC), or NOAEL due to varying study designs and test sensitivity and considerations of adversity and severity (i.e., nature of the response) for choosing the benchmark response (BMR) level. The nature and application of uncertainty/variability factors and MFs are discussed and critiqued in Section 4.4.5, and future directions are briefly discussed in Section 4.4.6. Section 4.4.7 summarizes key points from two case studies that are presented in detail in Appendix B.

4.4.1. Sample Reference Values and Selection of Endpoints to Use as the POD for Reference Values

Currently, the "critical effect" is used as the basis for the POD, and various UFs are applied to the dose at the critical effect to derive the RfD or the RfC. The critical effect is defined as "the first adverse effect, or its known precursor, that occurs to the most sensitive species as the dose rate of an agent increases" (U.S. EPA, 2002c). The underlying assumption is that if the RfD or the RfC is derived to prevent the critical effect from occurring, then no other effects of concern will occur; in addition, this approach assumes that the relationship of various health effects for a particular chemical is maintained across species.

The Technical Panel is concerned that presenting only a single critical effect and the critical study from which it was derived in the IRIS summary table that appears at the beginning of each RfD or RfC file may not provide enough information to the reader who is unfamiliar with risk assessment and thus could be misleading. Presenting a single endpoint as a POD for a systemic effect, for example, cannot capture the nature of the dose-response curve for that particular endpoint, nor does it convey the possibility that other more serious endpoints may have a dose-response character markedly different from the less serious endpoint. For example, an agent may have a clear progression of responses with increasing dose that is seen as one type of effect at the lowest exposure level (e.g., proteinuria in the case of cadmium), but at a higher level it produces additional effects (proteinuria PLUS GFR decrements) and at the highest level even more types of effects (proteinuria PLUS GFR decrements PLUS osteomalacia). Each of these effects could have a markedly different dose-response character.

Focusing on a single critical effect also does not reflect the situation in which other types of effects may be found at similar levels of exposure or the variety of health outcomes that may result when an exposure significantly exceeds the RfD or the RfC. Most importantly, in light of the Technical Panel's recommendations for deriving an expanded number of reference values for different durations and routes of exposure, the limitations of focusing only on the critical effect

become apparent because the most sensitive endpoint may be different for different durations or routes of exposure.

Layered upon this complex consideration of dose-response is the further complication that all of the exposure levels producing these effects are or should be adjusted to a human equivalent exposure at the time of their comparison. These adjustments may profoundly affect what is considered the most sensitive organ or system. Effects that occur at the same external inhaled concentration but in different organs in the same exposed animals (e.g., effects in the liver and the nasal cavity) may have quite different HECs, based on the current RfC methodology (U.S. EPA, 1994), because the underlying basis for the adjustment used for systemic effects is markedly different from that used for portal-of-entry effects between animals and humans. This adjustment procedure is discussed further below but is noted here because of its interrelationship with identifying what is to be considered a critical effect.

These aspects all support the case that a more comprehensive approach to setting reference values requires a more extensive and systematic analysis of endpoints than has typically been conducted in the past. In the approach proposed here, the selection of the POD would be similar to the current critical effect approach (e.g., U.S. EPA, 1994) and would include the use of sound scientific judgment in evaluating the strength and validity of studies and the extent of the database, as described in Section 4.3. In this approach, however, the selection of the POD would be based on consideration of all relevant and appropriate endpoints carried through the derivation of sample reference values, with selection of the limiting value(s) protective of all endpoints as the final step (the same approach would be used for deriving a POD for low-dose modeling, as discussed in the proposed cancer risk assessment guidelines [U.S. EPA, 1999a]).

For example, the dose-response curves would be modeled for several adverse endpoints and the corresponding BMDs and BMCs and their lower 95% confidence limits (BMDLs/BMCLs) calculated (U.S. EPA, 2000c) or NOAELs determined if dose-response modeling is not possible. Next, duration adjustment to the continuous exposure scenario would be performed for each endpoint, with further adjustment to the corresponding HECs using the RfC methodology (U.S. EPA, 1994) or adjusted BMDLs or NOAELs for oral or dermal exposures (see Section 4.4.3 for further discussion). These adjusted values would represent the POD for each relevant endpoint. Then, uncertainty/variability factors that take into account a variety of issues, including chemical-specific data, such as known toxicokinetic differences between the laboratory animal species tested and humans, and mode of action information would be applied to the adjusted values for each relevant endpoint. The sample reference values would

then be compared across endpoints and organ systems to determine which are the most relevant for use in deriving the final reference value for each exposure duration that will be protective of the human population (including susceptible subgroups).

The Technical Panel recommends the use of a more visual and graphic exposure-response array to depict the PODs for all relevant endpoints for various routes and durations of exposure, somewhat like those shown in the ATSDR toxicology profiles but with appropriate changes for the purpose of deriving reference values. The exposure-response array of the PODs would facilitate the evaluation and comparison of relevant endpoints and values. (See examples of the proposed approach discussed in Section 4.4.7 and in two case studies in Appendix B.)

4.4.2. Dose Adjustment for Duration of Exposure

Available studies from which reference values are derived seldom if ever match the intent of the reference value regarding species or duration. For example, chronic RfD and RfC values are intended by definition to be for "a continuous exposure to the... human population." Doses or exposures from studies in which animals are exposed for less than a lifetime or in which worker populations are exposed only during working hours require adjustment to continuous exposure in order to be concordant with the intended duration of the reference value (see Rozman and Doull, 2000; Rozman et al., 2001, for further discussion). This section describes various procedures that are currently used by the Agency to adjust a LOAEL, a NOAEL, or a BMDL with regard to duration. The basis for these adjustments is discussed, as is the applicability of these procedures to various routes of exposure.

The Agency has invested considerable time and effort into exploring these aspects for the inhalation route. A major point that will become apparent in this discussion is that methodologies for duration adjustment via the inhalation route are currently in place as part of the existing methodology for the chronic RfC and as proposed for ARE derivations, whereas no comparable documents yet exist for the oral or dermal routes of exposure.

4.4.2.1. *Duration Adjustment Procedures for Inhalation Exposures to Continuous-Exposure Scenarios*

Adjustment of duration to a continuous exposure scenario is regularly applied as a default procedure to studies with repeated exposures but not to single-exposure inhalation toxicity studies in animals and humans (U.S. EPA, 1994). Operationally, this is accomplished by

applying a C^n x t product[11] for both the number of hours in a daily exposure period and the number of days per week that the exposures are performed. In an inhalation study in which animals are exposed to 100 mg/m^3 for 6 hours, 5 days per week, the adjustment to a continuous exposure concentration would consider both hours per day and days per week:

$$100 \text{ mg/m}^3 \times 6/24 \text{ hrs x } 5/7 \text{ days/wk} = 17.9 \text{ mg/m}^3,$$

with 17.9 mg/m^3 being the concentration adjusted for continuous exposure. Study designs that include exposures 7 days/wk, for example, prenatal developmental toxicity studies and DNT studies, do not require the 5/7 days/wk adjustment.

Exposures from human occupational studies are most often reported as 8-hr time-weighted averages (TWAs) and are therefore also discontinuous. Adjustment of these exposures to derive a HEC is explained below in Section 4.4.3.

These adjustment procedures imply that the C × t product and not C is associated with the endpoints observed; this may be restated as implying that the area under the curve (AUC), C × t, rather than the peak concentration, C, is the dosimeter associated with toxicity. Although neither of these dosimeters may be demonstrable experimentally to be the appropriate measure of dose, the Agency uses adjustment to a continuous inhalation exposure based on the C × t relationship as a matter of policy.

When applied to a discontinuous inhalation exposure regimen from an experimental study, adjustment to a continuous exposure will always result in a lower value of C and maintain a measure of total exposure, that is, C × t. Thus, application of this procedure provides an automatic margin of protectiveness for chemicals for which C alone may be appropriate, and it reflects the maximum dose for agents for which total or cumulative dose is the appropriate measure. When considered in this way, this policy can be regarded as being protective of public health. However, assessors are encouraged to look for data on specific chemicals that support the use of C x t or that offer alternative models for adjustment of exposure duration.

4.4.2.2. *Duration Adjustment for Inhalation Developmental Toxicity Studies—A Current Exception*

A notable exception to duration adjustment of inhalation exposures is for inhalation developmental toxicity studies in which this practice historically has not been done. The current

[11]Where $C^n = C^1$, as described in Section 4.4.2.3.

guidelines for developmental toxicity risk assessment (U.S. EPA, 1991) recommend against duration adjustment (i.e., from a discontinuous to a continuous exposure) as a default procedure unless toxicokinetic data are available to indicate an accumulation with continuous exposure. This is contrary to the default approach used for other types of studies in which duration adjustment is done without a requirement for toxicokinetic information. In fact, for other types of studies, toxicokinetic information is often used as the basis for moving away from the default adjustment.

Furthermore, although the effects of some agents that cause developmental toxicity have been shown to be more a function of peak concentration (Nau, 1991), the effects of other agents have been shown to be related to either AUC or C, depending on the timing of exposure and the developmental timing of the organ system affected (Terry et al., 1994). In addition, recent studies have shown that the developmental effects of certain agents that have a short half-life, such as all-trans-retinoic acid (Tzimas et al., 1997) and ethylene oxide (Weller et al.,1999), or a very discrete exposure period, for example, hyperthermia (Kimmel et al., 2002), are a function of AUC.

On the basis of this information and the rationale used for duration adjustment for other health effects (i.e., that exposure adjustment based on $C \times t$ tends to be more health protective), the Technical Panel recommends that duration adjustment procedures to continuous exposures based on $C \times t$ be used as a default procedure for inhalation developmental toxicity studies as it is for other health effects from inhalation exposure. The Technical Panel also urges continued development of data, modeling, and improved procedures for dose-duration adjustments related to developmental toxicity.

4.4.2.3. *Duration Adjustment for Acute Reference Values—Discontinuous Scenarios of 24 Hours or Less*

As discussed above, the magnitude of response to a toxic chemical exposure usually depends on both the concentration and the duration of the exposure, such that the combination of these components, $C \times t$, determines the response and, by logical extension, the internal dose of a chemical at the target tissue. In deriving acute, short-term, or longer-term reference values, there may be a need to specifically adjust or present these values under alternative $C \times t$ combinations. For example, an acute reference value may be required for both a 1-hour duration and an 8-hour duration, but the available data are from a 4-hour exposure. The current guidance on this issue is contained in the draft methodology for development of AREs (U.S. EPA, 1998a). This section presents the adjustment procedures recommended in the draft ARE methodology.

Because of the recognized limitations of the C x t model, a modification has been developed such that $C^n \times T = k$, with n being empirically derived. The consequences of varying the values of the "n" exponent are shown in Figure 4-1. This figure, which was derived from the

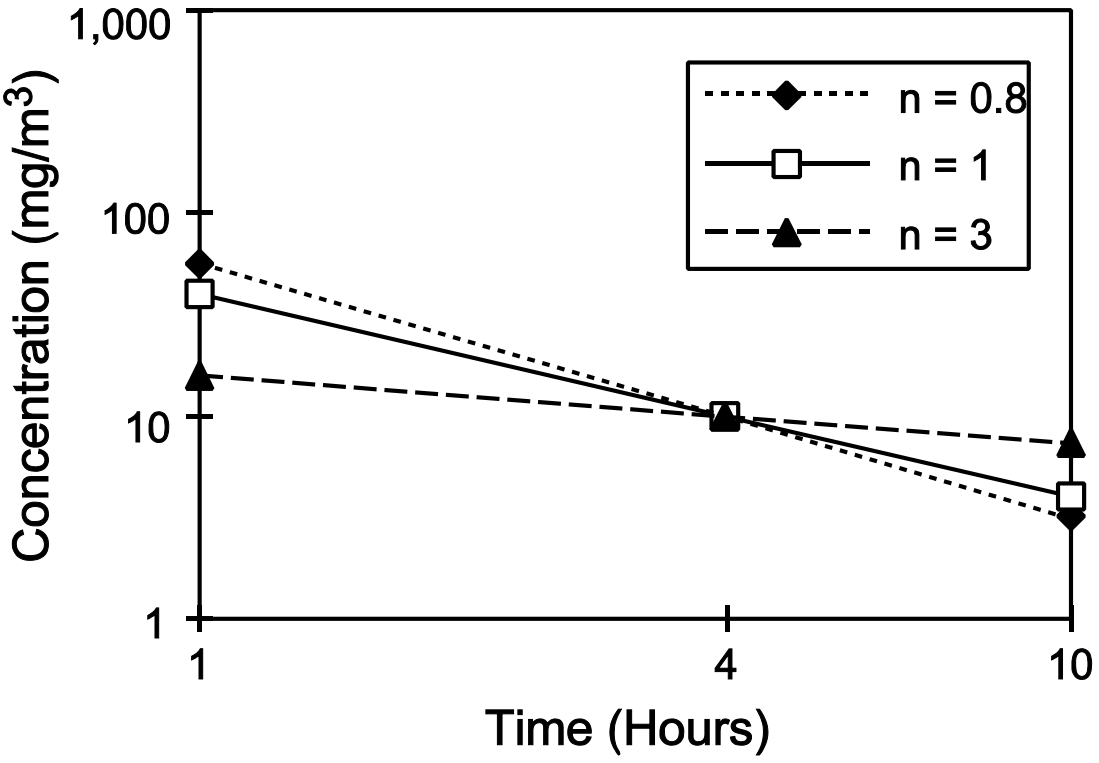


Figure 4-1. Concentration-by-duration plot showing the effect of the exponent in the C^n x T = k on extrapolation across time.

Source: Adapted from ten Berge et al., 1986.

current version of the Agency's ARE methodology, is based on the data of ten Berge et al. (1986). These investigators were able to empirically derive values of "n" that ranged from 0.8 to 3.5 for a number of chemicals on the basis of acute lethality. A value of 1 for the exponent "n" would indicate that the relationship described by Haber's law holds and that the response is related to total dose.

Note that for any degree of downward slope with increasing duration (lines marked with n = 1 or n = 0.8), an extrapolation from a longer to a shorter duration (i.e., from right to left) would result in a higher value for C. Extrapolating from a shorter to a longer duration (i.e., from left to right), however, would have a different consequence in that with any degree of downward slope, C would always be lower for the longer duration. Several possible approaches for extrapolation in this situation could be envisioned. One approach would be to assume a value of 1 for "n," such that $C^n \times T = k$ and lower values of C would always result; this approach is likely to be the actual case, because the value of "n" for most chemicals so far examined has shown an appreciable downward slope (e.g., $0.8 < n < 3.5$ [ten Berge et al., 1986]).

The optimal approach for extrapolating from one dose-duration response situation to another is the use of a physiologically based pharmacokinetic model (PBPK) model. The principle of using PBPK models as the basis for describing the correlations between level and duration of exposure, internal dose, and biological effect has been stated clearly by Andersen et al. (1987). Integration of information using PBPK models requires a chemical database that is rich in toxicity data; therefore, this approach is not applicable to most chemicals for which toxicokinetic data are scarce or nonexistent.

In the absence of such a database to support the development of a PBPK model, the approach recommended by the draft ARE methodology is the use of chemical-specific data on duration dependence from other adequate but longer-duration data, if they exist (e.g., in extrapolating to 28 days using 7-day data, the 90-day repeated-dose data should also be considered). This is considered a conservative approach, because the duration adjustment approach (i.e., averaging to continuous exposure), when applied to multiple exposure studies always results in decreased values for C (i.e., extrapolation would be from shorter to longer durations on the curves in Figure 4-1).

In the absence of chemical-specific data to inform duration adjustment, the response has most often been related to the simple $C \times t$ product. This is also the default in the draft ARE methodology for adjustment to longer durations. For adjustment to shorter durations, the ARE methodology conservatively recommends that there be no change in concentration.

Further investigation would increase confidence in the basic assumptions made for the latter two methods of duration adjustment, including the applicability of the C x t relationship over spans of exposure from months to years and assessing the "conservativeness" of these approaches in relation to public health. Further investigation of C x t relationships relative to life stage is also recognized as a research need.

4.4.3. Derivation of a HEC or a HED

Animal data often form the basis for dose-response assessment. By definition, the IRIS risk values are for humans, thereby making animal-to-human extrapolation requisite. The specific point of this extrapolation is to estimate from animal exposure information the human exposure scenario that would result in the same response. The simplest manner in which this may be done is application of an animal-to-human UF (discussed further below), typically with a value of 10; in application this means that humans are assumed to be more sensitive to effects than are animals by a factor of 10.

Much of the RfC methodology (U.S. EPA, 1994) focused on improving the science underlying the animal-to-human UF, segregating it into toxicokinetic and toxicodynamic components and providing generalized procedures to derive dosimetric adjustment factors (DAF). Application of DAFs to the animal airborne exposure values yields estimates of the concentration that would result in the same concentration to humans, that is, the HEC. Application of a DAF in the calculation of a HEC is considered to address the toxicokinetic aspects of the animal-to-human UF (i.e., to estimate from animal exposure information the human exposure scenario that would result in the same dose to a given target tissue).

Current Agency practice is to accommodate uncertainty about the remaining toxicodynamic component through application of a partial animal-to-human UF ($10^{0.5}$, which is typically rounded to 3). The theoretical basis for deriving DAFs used in calculating HECs, along with recommendations for improvement of this process, is discussed in this section.

Exposures from human occupational studies are most often reported as 8-hr TWAs for exposures during work days (5 days/wk). As with discontinuous exposures of animal studies (e.g., 6 hrs/day, 5 days/wk), exposures from occupational studies are also adjusted to derive continuous HECs relevant to the human population (U.S. EPA, 1994). As described below for animal data, the optimal approach is to use a biologically motivated mathematical, or PBPK, model. An occupational exposure can be extrapolated in the same fashion as intermittent exposure regimens from experimental laboratory animals, using particle deposition or PBPK models with human exertion (work) ventilation rates and exposure durations appropriate to the occupational setting.

In the event that a PBPK model or required physicochemical and physiological parameters are not available, the default approach for human exposure scenarios is to adjust by the default occupational ventilation rate and for the intermittent work week schedule. The ventilation rate adjustment is based on the assumed amount of air used by a worker during the work period, that is, half of the daily ventilatory capacity of an adult male human is assigned (10

m^3 of 20 m^3 total) to the 8-hour occupational exposure (i.e., instead of 1/3 or 8/24 hrs) (ICRP, 1994). By basing this adjustment on a functioning physiological parameter, that is, a fractional ventilatory capacity based on the assumption that activity levels are higher in this setting than in others, such as at rest or asleep, this adjustment may be considered to have a toxicokinetic basis. The 8-hour TWA concentrations are multiplied by this factor, 10/20 m^3, and the product is considered to be an average continuous airborne concentration.

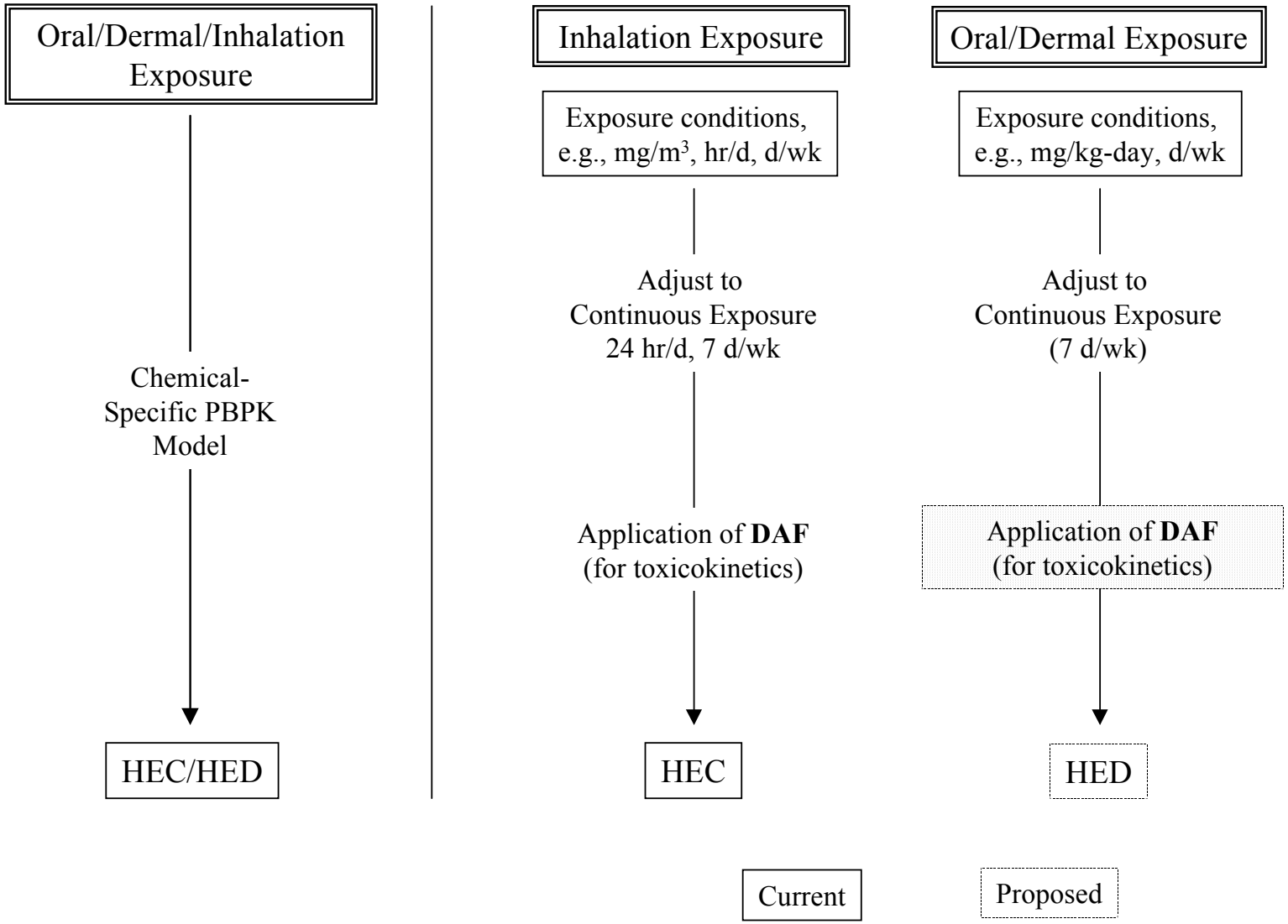
In parallel with the animal studies, an adjustment for days per week (usually 5/7 days/wk) is also made, if applicable. This adjusted airborne concentration is considered to be a HEC. This default calculation, as with those described below for extrapolation from animal data, was developed for the general human population. It may be appropriate to further evaluate this approach or to develop an alternate default approach to ensure adequate consideration of intrahuman variation.

Currently, no procedures parallel to the inhalation RfC methodology exist for deriving either oral or dermal human equivalents from animal data. Default factors (usually of 10) are routinely applied to address the issue of animal-to-human extrapolation. Thus, no parallel to the HEC, that is, a HED, is derived nor are other adjustments applied to the animal oral or dermal dose.

This section recommends that dose adjustments similar to those by which HECs are estimated be explored in deriving HEDs for oral and dermal exposures. This would be accomplished in a manner parallel to the HEC derivation, by instituting and applying a DAF to animal oral or dermal exposures. Specific recommendations are also presented and discussed concerning the basis for deriving DAFs for HED calculation. These recommendations, along with current procedures for estimating human equivalent values, are illustrated in Figure 4-2. This figure also demonstrates how calculation of the HEC through application of a DAF is considered to address the toxicokinetic but not the toxicodynamic component of the animal-to-human extrapolation. Procedures outlined in this figure for deriving a HEC may be applied to any animal inhalation exposure, regardless of whether it is a BMDL, a NOAEL, a LOAEL, or another effect level.

4.4.3.1. *PBPK Models and Derivation of HEDs and HECs: Estimating Internal Dose*

The preferred option for calculating a HED or a HEC is to use a chemical-specific PBPK model parameterized for the species and regions (e.g., respiratory tract) involved in the toxicity, as shown on the left-hand side in Figure 4-2. When sufficiently parameterized, a PBPK model is capable of calculating internal doses to a target organ from any exposure scenario in an animal



DAF = Dosimetric Adjustment Factor

Figure 4-2. Current and proposed generalized procedures for deriving HECs or HEDs from animal exposures.

and then estimating what human exposure would result in this same internal dose, that is, the HED or the HEC. A formal DAF is not calculated in this process; rather, the model itself serves as a DAF in estimating HECs or HEDs. However, constructing a PBPK model is an information-intensive process that requires much chemical-specific data, including route-specific data. Such sophisticated data and models are available usually for only a subset of chemicals that have extensive databases.

It should be noted that even these sophisticated models are often parameterized on the basis of adult members of the species. Many of the parameters critical to PBPK model solutions are sensitive to life stages, such as lung function/development in humans (Pinkerton and Joad, 2000), for which no or few data are available. Thus, these models are available but often cannot specifically address species differences at life stages other than mature adults (and then usually males). The Technical Panel encourages research and data gathering to support the construction of PBPK models, it endorses attempts to produce PBPK models that are sensitive to life stages, and it supports fully attempts to produce template models for suites of related chemicals, as has recently been done by Barton et al. (2000).

4.4.3.2. *Default Procedures and Derivation of HECs from the RfC Methodology: Derivation and Application of DAFs*

The next lower level of complexity in deriving HECs is less data intensive than the PBPK approach. As shown in Figure 4-2, this procedure involves the use of species-specific physiologic and anatomic factors relevant to the form of pollutant (e.g., particle or gas) and categorized with regard to elicitation of response either locally (i.e., within the respiratory tract) or remotely. These factors are all employed in determining the appropriate DAF. For HECs, DAFs are applied to the "duration-adjusted" concentration to which the animals were exposed (e.g., to a weekly average). The generalized DAF procedures may also employ chemical-specific parameters, such as mass transport coefficients, when available. In lieu of such data, however, default procedures that yield generalized adjustments are recommended. Although these generalized procedures were developed from the existing scientific understanding of the relevant processes, they have not been comprehensively evaluated (e.g., using data from humans and animals). They are explained fully in the RfC methodology (U.S. EPA, 1994).

For example, the manner in which a HEC is calculated for a reactive gas that elicits an effect in the extrathoracic region of the respiratory tract (i.e., the nasal tract) of a rat is by creating a surface area/ventilation ratio for both humans and rats and applying it to the external exposure concentration for rats. The current default values used for both the human and the rat

extrathoracic surface area are single estimates from the literature and are apparently estimated from adult specimens. The ventilation measure for humans is set at a default value of 20 m^3, and the ventilation measure for rats is based on an algorithm of body weight (from U.S. EPA, 1988).

A major assumption made in this particular adjustment is that the distribution of a gas in the region of interest is uniform, although it is known to be highly nonuniform (Kimbell et al., 1993, 1997). Data are not available to address this simplified assumption directly. Use of the method, for example on effects in the extrathoracic region, results in a DAF of about 0.2, such that the resultant HECs are about 20% of the animal-duration-adjusted concentration. Although information is not yet available to address this assumption, indications are that resolution with actual data may produce DAFs that are much closer to unity, that is, that are near the animal-adjusted concentration.

In comparison to the procedure for gases that elicit respiratory effects, calculation of a HEC for a category 3 gas (i.e., a gas that is relatively water-insoluble and unreactive in the respiratory tract and for which the site of toxicity is generally remote to the site of absorption in the pulmonary region) is usually accomplished by creating a ratio of the blood:gas partition coefficient for the laboratory animal species to the human value. The ratio is used as the DAF and applied to the experimental exposure concentration. In lieu of data on the values for blood:gas partition coefficients for the chemical or when the data indicate the ratio to be >1, the default assumption is that the ratio of animal coefficient to human coefficient is 1, and therefore the DAF would be 1. However, available data on partition coefficients for a number of compounds indicate that the animal/human ratio is usually >1 (Gargas et al., 1989; Jepson et al., 1994) such that the DAF would also be >1. In the context of substituting data-derived values for UFs, the Technical Panel recommends further investigation into using data-derived values in constructing the animal/human ratios—even when much greater than 1—in place of the default.

The default dosimetric adjustment procedure for particulate substances is an empirical model that estimates regional deposition only, although it is recognized that with the development of the relevant data, clearance and the retained dose may be used as a DAF (U.S. EPA, 1994). The DAF for particles is more specifically termed the regional deposited dose ratio and is derived from a normalizing factor (surface area being the recommended factor for all three regions of the respiratory tract), the ratio of animal-to-human minute volumes (where the human default value is the traditional adult value of 13.8 L vs. the adult value for the relevant animal), and the ratio of animal-to-human regional fractional deposition. Physiological parameters used in estimating the regional deposition include body weight, minute volume, and surface area for the three areas of the respiratory tract. Defaults for the human values are based on adult data

(e.g., 70 kg body weight, 13.8 L minute ventilatory volume, etc); the animal values are also traditionally based on adult data. To evaluate protectiveness of these default calculations for different life stages, it may be appropriate to perform ratio calculations using data for other life stages.

As a general recommendation, the Technical Panel encourages further consideration of the existing animal-to-human extrapolation procedures described in current methodologies (e.g., the chronic RfC methodology [U.S. EPA, 1994]) and the development of procedures for inhalation adjustment to incorporate the most current scientific thought and data to address, as needed, issues of variability due to life stage and other intrinsic factors. This consideration would include examining the extent to which calculating a HEC (or any recommended HED) addresses cross-species toxicokinetics as well as identification and parallel investigation into issues of toxicodynamics.

4.4.3.3. *HECs and Children—A Special Case?*

Children are often characterized as constituting a potentially susceptible subgroup because they could be at greater risk than adults for inhaled toxic agents (including both gases and particulates) for reasons relating to either toxicokinetics or toxicodynamics. It is clear for any of a variety of reasons related to toxicokinetics that an adult and a child breathing the same concentration of an agent such as a reactive gas may receive different doses to the body or to the lungs. A generalized theoretical approach to judging whether children would receive greater doses than would adults when both breathe the same concentration of a reactive gas, for example, would be to compare the amount of gas breathed in (which would be directly proportional to the ventilatory volume) with the overall surface area in the respiratory tract on which the gas may impinge. The current Agency default assumption used in deriving HECs for particles and reactive gases that elicit respiratory effects is that the surface area of the total respiratory tract of an adult male, estimated at 54.3 m^2, is exposed to a total daily air intake of 20 m^3, a volume for an adult male derived from a combination 24-hour activity pattern in ICRP (1994) of sitting awake for 8 hours, exercising lightly for 8 hours, and sleeping for 8 hours.

It has been well established that the human respiratory system passes through several distinct stages of maturation and growth that involve branching morphogenesis and cellular differentiation during the first several years of life and into adolescence (Pinkerton and Joad, 2000). The proportion of surface area to ventilation volume may be markedly different during these developmental stages. The significance of these disproportions with regard to toxicant exposure overall or to the sites of active cellular differentiation have yet to be elucidated.

The Technical Panel recommends that issues involving dose to the young from inhalation exposures be pursued both theoretically and experimentally in order to establish the basis on which children should be considered as a susceptible subpopulation for inhalation exposures. It should also be reiterated that this is an estimate of the toxicokinetic aspect of dose only, and toxicodynamic differences between the lungs of young children and adults are not addressed.

4.4.3.4. *Deriving a HED for Oral and Dermal Exposure—Use of $BW^{3/4}$ as a Cross-Species DAF*

As indicated above, the Agency currently does not provide a procedure for calculating a HED for oral or dermal exposure scenarios that would parallel calculation of the inhalation HEC. Instead, assumptions are made regarding the comparability of ingested or applied dose, based on a mg/kg body-weight basis, and there is no adjustment for portal-of-entry alterations to internal dose or on portal-of-entry versus systemic effects. The Technical Panel recognizes the work of an interagency workgroup to develop and propose dosimetric adjustment procedures for both dermal and oral routes of exposure in order to address those aspects of cross-species dosimetric adjustment that are missing in Figure 4-2. Some of these proposals have already appeared in abstract form (Jarabek, 2000; Hanna and Jarabek, 2000; Hubal et al., 2000; Rigas et al., 2000).

Figure 4-2 demonstrates that dosimetric adjustment procedures for estimating human equivalents from animal values are not consistent for different exposure routes. Other procedures, both from within and external to the Agency, could be explored for the purposes of deriving a DAF and employing it to estimate a HED. For example, in the absence of more sophisticated physiologically based models, the Agency has endorsed scaling of doses for carcinogens between species according to body mass raised to the 3/4 power ($BW^{3/4}$) (U.S. EPA, 1992). This procedure presumes that equal doses in these units (i.e., in $mg/kg^{3/4}/day$) when administered daily over a lifetime, will result in equal lifetime cancer risks across mammalian species. This same relationship (i.e., $BW^{3/4}$) has been affirmed to apply across entire phyla, including plants (Gillooly et al., 2001), for general metabolic rates.

The basis for the less-than-full-power relationship for general metabolic processes (i.e., $< BW^{1}$) is thought to be related to species differences in exchange surfaces and distribution networks that constrain concentration and flux of metabolic reactants (West et al., 1997; Enquist et al., 1998). Thus, when this procedure is applied to animal data, the resulting scaled human dose may be viewed as a valid cross-species relationship not only of cancer potency but also for general metabolic processes and, by extension, for other phenomena involving the fundamental determinants of concentration and flux, the same ones that drive basic toxicokinetics.

This brief analysis of the $BW^{3/4}$ cross-species relationship and toxicokinetic processes and the Agency's endorsement of this procedure for carcinogenic agents makes this process a possible candidate for estimating cross-species toxicokinetic relationships in the absence of adequate toxicokinetic information. That is, $BW^{3/4}$ factors could be applied as DAFs for deriving a HED. This procedure would parallel the one used for deriving the HEC. As with the HEC, however, this process applies only to toxicokinetic aspects of cross-species extrapolation and does not address toxicodynamic differences that may exist between species. As with the HEC, consideration of toxicodynamics is proposed to be through application of a portion of the animal-to-human extrapolation ($10^{0.5}$, which is typically rounded to 3). Table 4-3 shows the general magnitude of the DAFs that would be applied to various species to obtain the HED along with the default UF of 3 to cover toxicodynamic differences.

Table 4-3. DAFs based on $BW^{3/4}$ for various species

Species	Weight (kg)	DAF[a]
Mouse	0.03	7
Rat	0.25	4
Guinea pig	0.5	3
Rabbit	2.5	2
Human	70	1

[a] Derived on the basis of $BW^{3/4}$ relationship. All variables in $BW^{3/4}$ relationship containing time will scale $BW^{-1/4}$, such that animal $BW^{-1/4}$/human $BW^{-1/4}$ = DAF.

The Technical Panel encourages consideration of cross-species extrapolation procedures for oral and dermal reference values, including evaluation of the most current scientific thought and data to address, as needed, issues of variability due to life stage and other intrinsic factors. This consideration would include examination of the extent to which calculation of a HED addresses cross-species toxicokinetics and identification and parallel investigation into issues of toxicodynamics.

4.4.4. Other Issues

The Technical Panel considered several other issues related to the application of a factor (data-derived or default) to the BMDL, the BMCL, the NOAEL, or the LOAEL selected as the POD from data considered adequate for risk assessment. In particular, there was controversy about the application of such a factor on the basis of the level of response at the BMD, the BMC, the NOAEL, or the LOAEL. For example, the use of a quantitative dose-response modeling approach results in the calculation of a BMD or a BMC, which is based on a particular level of response, that is, the BMR. The BMR is usually selected to be at the low end of the observable range of the data, which is dependent on the power of the study to detect changes from control values. The limit of sensitivity for most long-term bioassays is in the range of 10%, as determined from both the typical number of animals used in bioassays (~50/group) and a low spontaneous background rate (e.g., 0.1%) for a given effect (Haseman, 1984; Haseman et al., 1989).

For other types of studies, however, the limit of sensitivity may be lower or higher than 10%. For example, in an analysis of a large number of standard prenatal developmental toxicity studies with an average sample size of 15–20 litters, the limit of sensitivity averaged 5% for the proportion of pups affected per litter, whereas when the quantal endpoint (i.e., the number of litters affected) was analyzed in dams from the same studies, the limit of sensitivity averaged 30% (Allen et al., 1994). For data from some human studies (e.g., high-quality, large epidemiology studies), the limit of sensitivity may be in the range of 1 to 5%.

In the BMD guidance document (U.S. EPA, 2000c), the BMDL or BMCL is recommended for the POD in order to ensure that a majority of the population is below the selected BMR. However, a concern has been raised that a BMD or BMC based on a response rate of $\geq$10% may not be appropriate to use in deriving an exposure to the human population (including sensitive or susceptible subgroups) *that is likely to be without appreciable risk of deleterious or adverse effects* (from current and proposed reference value definitions [Boxes 4-1 and 4-2]) without application of a factor to extrapolate to a lower dose/exposure level considered to reflect a more appropriate level of risk (e.g., <10%).

Similarly, the NOAEL is not necessarily a no-effect level, and it depends on the study design, including sample size, background rate, and response variability, which can be used to determine the limit of detection for a particular study. Thus, a NOAEL may be equivalent to no response or it may actually represent a substantial response rate. Previously, there has been no attempt to apply a factor to the NOAEL on the basis of power calculations, sample size, or response variability for deriving a POD, although professional judgment is recommended in deciding whether the study is acceptable for use in deriving a POD.

Adjustment for the steepness of the dose-response curve has been noted as another critical aspect of the dose-response character that is not currently considered in the choice of a response level using either a BMD/BMC or a NOAEL approach.

The Technical Panel was unable to fully evaluate these issues or to reach agreement about any recommendation for change to the current methodology, and it recommends that they be considered further by the Agency. The Technical Panel also recommends that factors such as the response rates at the BMD or the NOAEL, the power of the study, and the slope of the dose-response curve be included in the description of the database, where possible, as part of risk characterization.

Box 4-4. Variability and Uncertainty

Variability refers to true heterogeneity or diversity. For example, among a population that drinks water from the same source and with the same contaminant concentration, the risks from consuming the water may vary. This may be due to differences in exposure (i.e., different people drinking different amounts of water and having different body weights, different exposure frequencies, and different exposure durations) as well as differences in response (e.g., genetic differences in resistance to a chemical dose). Those inherent differences are referred to as variability. Differences among individuals in a population are referred to as inter-individual variability, while differences for one individual over time is referred to as intra-individual variability.

Uncertainty occurs because of a lack of knowledge. It is not the same as variability. For example, a risk assessor may be very certain that different people drink different amounts of water but may be uncertain about how much variability there is in water intakes within the population. Uncertainty can often be reduced by collecting more and better data, while variability is an inherent property of the population being evaluated. Variability can be better characterized with more data, but it cannot be reduced or eliminated. Efforts to clearly distinguish between variability and uncertainty are important for both risk assessment and risk characterization.

Source: U.S. EPA, 1997b.

4.4.5. Application of Uncertainty/Variability Factors

Reference values are derived in a way that attempts to account for both the uncertainty and the variability in the data available (see Box 4-4). The existing definition of UF in the IRIS glossary mixes the above concepts. The present definition for UF is as follows.

> **Uncertainty Factor**: One of several, generally 10-fold, factors used in operationally deriving the RfD and RfC from experimental data. UFs are intended to account for (1) the variation in sensitivity among the members of the human population (i.e., interhuman or intraspecies variability); (2) the uncertainty in extrapolating animal data to humans (i.e., interspecies variability); (3) the uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime exposure (i.e., extrapolating from subchronic to chronic exposure); (4) the uncertainty in extrapolating from a LOAEL rather than from a NOAEL; and (5) the uncertainty associated with extrapolation from animal data when the database is incomplete.

Following the logic above, the LOAEL-to-NOAEL extrapolation, the subchronic-to-chronic extrapolation, and the database deficiency factors are UFs. The variation in susceptibility among members of the human population is a variability factor. When a default factor is used for intrahuman variability, however, this factor also contains some degree of uncertainty, because the range of uncertainty is not really known, although it is presumed to be no more than 10-fold. Rather than adding a new definition of variability factor, we propose to modify the wording of the UF definition as follows.

> **Uncertainty/Variability Factor:** One of several, generally 10-fold, default factors used in operationally deriving the RfD and the RfC from experimental data. The factors are intended to account for (1) the variation in sensitivity among the members of the human population (i.e., inter-individual variability); (2) the uncertainty in extrapolating animal data to humans (i.e., interspecies uncertainty); (3) the uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime exposure (i.e., extrapolating from subchronic to chronic exposure); (4) the uncertainty in extrapolating from a LOAEL rather than from a NOAEL; and (5) the uncertainty associated with extrapolation when the database is incomplete.

In setting pesticide tolerances, the FQPA directs EPA to use an additional 10-fold margin of safety to protect infants and children, taking into account the potential for pre- and postnatal toxicity and the completeness of the toxicology and exposure databases. The statute authorizes EPA to replace this additional 10X factor with a factor of a different value (higher or lower, including 1) only if, on the basis of reliable data, the resulting level of exposure would be safe for infants and children. The Agency use of this FQPA safety factor has been discussed in several documents (U.S. EPA, 1999b, c, 2002b).

The Agency has concluded that in many cases, concerns regarding pre- and postnatal toxicity can be addressed by calculating an RfD or an MOE using pre- or postnatal developmental endpoints and applying traditional UFs to account for deficiencies in the toxicity data (U.S. EPA, 2002b). These traditional UFs include extrapolation from the LOAEL when a NOAEL is not available, extrapolation from a subchronic study to a chronic-exposure scenario when no chronic study data are available, and application of a database UF when there are gaps in the data considered essential for setting a reference value, including lack of data on children.

In addition to considering these FQPA-relevant areas of uncertainty, which are addressed in the development of an RfD/RfC, OPP assessments of pesticide risk to children also consider applying part or all of the FQPA factor in certain situations to account for areas of residual

uncertainty that the traditional UFs do not address or for which they are believed to be insufficient. These areas of residual uncertainty include exposure uncertainties and high concern for an observed susceptibility. This risk management approach is consistent with procedures used in the past for managing potential risks, although the FQPA has brought a significant new focus on improving the process of risk assessment relative to children's health risks from environmental exposures.

In considering the robustness of the RfC/RfD methodology and its adequacy for assessing hazards to infants and children, the Technical Panel also recognized the overlap of areas covered by the FQPA factor and those addressed by the traditional UFs. For example, the database UF may be invoked where data are unavailable or are insufficient to explicitly consider the potential sensitivity of the developing organism. The Technical Panel agrees with the 10X Task Force draft Toxicology Working Group report (U.S. EPA, 1999b) that the current interspecies, intraspecies, LOAEL-to-NOAEL, subchronic-to-chronic, and database-deficiency UFs, if appropriately applied using the approaches recommended in this review, will be adequate in most cases to cover concerns and uncertainties regarding the potential for pre- and postnatal toxicity and the completeness of the toxicology database. In other words, an additional UF is not needed in the RfC/RfD methodology because the currently available factors are considered sufficient to account for uncertainties in the database from which the reference values are derived (and it does not exclude the possibility that these UFs may be decreased *or* increased from the default value of 10).

Guidance is needed on the use of developmental toxicity data in all reference values, including the appropriate application of UFs, because of the assumption that a single exposure during development may produce an effect (U.S. EPA, 1991) and the concomitant recognition that multiple exposures may result in effects at lower doses in many cases or cause tolerance in other cases. These issues are chemical specific, and scientific judgement about when and how to apply UFs must include consideration of toxicokinetics/metabolism as well as the mode of action for each agent.

4.4.5.1. *Recommendations for Application of UFs*

The exact value of the UFs chosen should depend on the quality of the studies available, the extent of the database, and scientific judgment. It is imperative that the IRIS documentation contain a justification for the individual UFs selected for a particular agent. The default factors typically used cover a single order of magnitude (i.e., 10^1). By convention, in the Agency, a value of 3 is used in place of one-half power (i.e., $10^{0.5}$) when appropriate. The Technical Panel recommends that these half-power values be factored as whole numbers when they occur singly

but as powers or logs when they occur in tandem. A composite UF of 3 and 10 would be expressed as 30 (3×10^1), whereas a composite UF of 3 and 3 would be expressed as 10 ($10^{0.5} \times 10^{0.5} = 10^1$). It should be noted, in addition, that rigid application of log or ½ log units for UFs could lead to an illogical set of reference values; therefore, the Technical Panel emphasizes that application of scientific judgment is critical to the overall process.

It is imperative that the IRIS documentation contain a justification for the individual factors selected for each chemical or assessment and for each duration reference value. Although default factors of 10 are recommended, with 3 used in place of half-power values (i.e., $10^{0.5}$) when occurring singly, the exact value of the UF chosen should depend on the quality of the studies available, the extent of the database, and scientific judgment. Sound scientific judgment should be used in the application of UFs to derive reference values that are applied to the value chosen for the POD derived from the available database (BMDL, NOAEL, or LOAEL).

The Technical Panel recognizes that there is overlap in the individual UFs and believes that the application of five UFs of 10 for the chronic reference value (yielding a total UF of 100,000) is inappropriate. In fact, in cases where maximum uncertainty exists in all five areas, it is unlikely that the database is sufficient to derive a reference value. Uncertainty in four areas may also indicate that the database is insufficient to derive a reference value. In the case of the RfC, the maximum UF would be 3000, whereas the maximum would be 10,000 for the RfD. This is because the derivation of RfCs and RfDs have evolved somewhat differently. The RfC methodology (U.S. EPA, 1994) recommends dividing the interspecies UF in half, one-half ($10^{0.5}$) each for toxicokinetic and toxicodynamic considerations, and it includes a DAF to account for toxicokinetic differences in calculating the HEC, thus reducing the interspecies UF to 3 for toxicodynamic issues. RfDs, however, do not incorporate a DAF for deriving a HED, and the interspecies UF of 10 is typically applied.

The Technical Panel recommends limiting the total UF applied for any particular chemical to no more than 3000 and avoiding the derivation of a reference value that involves application of the full 10-fold UF in four or more areas of extrapolation. This maximum of 3000 applies only to the UFs discussed in the following sections and does not include the various adjustment factors that have been discussed previously (Sections 4.4.2. and 4.4.3.). Similar concerns would need to be considered for the less-than-lifetime reference values, taking into account those UFs that are appropriate for each duration reference value.

4.4.5.2. *Interspecies UF*

The interspecies UF is applied to account for the extrapolation of laboratory animal data to humans, and it generally is presumed to include both toxicokinetic and toxicodynamic aspects. The toxicokinetic aspects of this factor were addressed in the section on deriving HEDs and HECs (Section 4.4.3). This UF is intended also to account for differences in species sensitivity (i.e., toxicodynamics) between the laboratory animal species used for testing and humans. Seldom are there data available to inform toxicodynamic differences. One-half the default 10-fold interspecies UF (i.e., $10^{0.5}$) is assumed to account for such differences, but more specific data should be used when available (see discussion of chemical-specific adjustment factors, Section 4.4.6.1 below), and the flexibility for applying a factor greater than 10 should be recognized. Unless data support the conclusion that the test species is more or equally as susceptible to the pollutant as are humans, and in the absence of any other specific toxicokinetic or toxicodynamic data, a default factor of 3 (in conjunction with HEC derivation) or 10 is applied.

4.4.5.3. *Intraspecies UF*

The intraspecies UF is applied to account for variations in susceptibility within the human population (interhuman variability) and the possibility (given a lack of relevant data) that the database available is not representative of the dose/exposure-response relationship in the subgroups of the human population that are most sensitive to the health hazards of the chemical being assessed. As the reference concentration/dose is defined to be applicable to "susceptible subgroups," this UF was established to account for uncertainty in that regard. In general, the Technical Panel reaffirms the importance of this UF, recommending that reduction of the intraspecies UF from a default of 10 be considered only if data are sufficiently representative of the exposure/dose-response data for the most susceptible subpopulation(s).

Various authors who have evaluated the intraspecies UF using data from animal or human studies (as summarized by Dourson et al. [1996]) have concluded that the 10-fold default factor appears to be protective when starting from a median response—by inference a NOAEL assumed to be from an average group of humans. Renwick and Lazarus (1998) considered data on toxicokinetics and toxicodynamics to support the idea that the 10-fold intraspecies factor can be divided into two factors to account for kinetics and dynamics. When they evaluated the composite 10-fold factor to account for variability in both kinetics and dynamics, they concluded that a 10-fold factor would cover the vast majority (>99%) of the population. These evaluations, however, did not specifically consider children as part of the range of human variability when evaluating the adequacy of the intraspecies UF.

In papers that have evaluated this factor for the general population as well as for specific subpopulations, including children (Renwick and Lazarus,1998; Renwick, 1998) and the elderly (Abdel-Mageed et al., 2001), the 10-fold intraspecies factor appears to be sufficient in most cases, and chemical-specific factors often indicate a requirement for less than a l0-fold factor. Renwick (1998) indicated that the 10-fold factor is more likely to be sufficient if developmental toxicity data are available on the specific agent. Calabrese (2001) reviewed the data available on a number of chemical classes and concluded that the young are often more susceptible than adults but that there is a not-infrequent occurrence of greater susceptibility in adults. The sometimes greater sensitivity among the elderly than among mature adults appears to be related primarily to reduced renal clearance (Abdel-Mageed et al., 2001; Skowronski and Abdel-Rahman, 2001).

The Technical Panel urges continued research and evaluation of the similarities and differences between the general population and susceptible subpopulations—particularly children and the elderly—in their responses to specific agents. From such evaluations, the protectiveness of the 10-fold default factor can continue to be assessed.

The cases on IRIS in which the intraspecies UF has been reduced from the default of 10-fold have been documented by Dourson et al. (1996). These include 2/46 RfCs and 13/346 RfDs (overall frequency 3.6%). In those cases where developmental effects were the most sensitive endpoint (0 RfCs, 6 RfDs), reduction of the intraspecies UF from 10 to 3 was based on data derived either from human data showing which age groups or time periods were most susceptible (e.g., methyl mercury exposure to the developing fetus) or from an animal study with support from strong human or other data (e.g., Aroclor 1016 in utero exposure in monkeys, strontium-induced rachitic bones in young rats). In three cases the intraspecies UF was reduced to 1, based on very specific data about the particular vulnerability of infants and children within certain age ranges to an agent (e.g., nitrate, nitrite, fluorine/soluble fluoride). However, even within these populations it is possible that some variability exists, based on genetics, lifestyle, or other factors.

In cases where the susceptible subpopulation is quite specifically defined (e.g., through knowledge of the chemical's mode of action) so that the resultant RfC is truly applicable to the susceptible subpopulation (although not necessarily to hypersensitive individuals), reduction of the intraspecies UF is warranted. Thus, the Technical Panel supports and expands the recommendation of the Toxicology Working Group of the 10X Task Force (U.S. EPA, 1999b) that reduction of the intraspecies UF from a default of 10 be considered only if data are sufficient to support the conclusion that the data set on which the POD is based is representative of the exposure/dose-response data for the susceptible subpopulation(s). Given this, whether and how

much the intraspecies UF may be reduced must be linked to how completely the susceptible subpopulation has been identified and their sensitivity described (vs. assumed). At the other extreme, a 10-fold factor may sometimes be too small because of factors that can influence large differences in susceptibility, such as genetic polymorphisms. The Technical Panel urges the development of data to support the selection of the appropriate size of this factor, but recognizes that often there are insufficient data to support a factor other than the default.

4.4.5.4. *LOAEL-to-NOAEL UF*

A UF (default 10) is typically applied to the LOAEL when a NOAEL is not available. The size of the LOAEL-to-NOAEL UF may be altered, depending on the magnitude and nature of the response at the LOAEL. It is important to consider the slope of the dose-response curve in the range of the LOAEL in making the determination to reduce the size of the LOAEL-to-NOAEL UF. Several papers have described the magnitude of the difference between the dose at the LOAEL and at the NOAEL. For example, Lewis et al. (1990) and Faustman et al. (1994) showed that the ratio of the LOAEL-to-NOAEL in many cases was approximately threefold, but in a few cases the difference was as much as 10-fold.

In general, the ratio of the doses at the LOAEL and the NOAEL is likely to vary considerably among studies and may not be informative. This is because the lowest dose in a study is often selected to ensure that no statistically significant response above control is observed and the next higher dose is selected to ensure that some significant response is observed, rather than selecting doses that will give a maximum NOAEL and a minimum LOAEL. Data should be carefully evaluated, taking into consideration the level of response at the LOAEL and the NOAEL and the slope of the dose-response curve before reducing the size of the UF applied to the LOAEL.

4.4.5.5. *Database UF*

The database UF is intended to account for the potential for deriving an underprotective RfD/RfC as a result of an incomplete characterization of the chemical's toxicity. In addition to identifying toxicity information that is lacking, review of existing data may also suggest that a lower reference value might result if additional data were available. Consequently, in deciding to apply this factor to account for deficiencies in the available data set and in identifying its magnitude, the assessor should consider both the data lacking and the data available for particular organ systems as well as life stages.

In many respects, the additional 10-fold factor for infants recommended by the National Research Council (NRC, 1993) and by Schilter et al. (1996) and called for in the 1996 FQPA is

similar to the database UF. If the RfD/RfC is based on animal data, a factor of 3 is often applied if either a prenatal toxicity study or a two-generation reproduction study is missing, or a factor of 10 may be applied if both are missing (Dourson et al., 1996). Dourson et al. (1992) examined the use of the database UF by analyzing ratios of NOAELs for chronic dog, rat, and mouse studies and reproductive and developmental toxicity studies in rats. They concluded that reproductive and developmental toxicity studies provide useful information for establishing the lowest NOAEL, and if one or more bioassays are missing, a factor should be used to address this scientific uncertainty in deriving a chronic RfD.

If data from the available toxicology studies raise suspicions of developmental toxicity and signal the need for developmental data on specific organ systems (e.g., detailed nervous system, immune system, carcinogenesis, or endocrine system), then the database factor should take into account whether or not these data are available and used in the assessment and their potential to affect the POD for the particular duration RfD or RfC under development.

If the RfD/RfC is based on human data, a similar assessment regarding the completeness of the database is necessary. Information on life stages and organ systems may come from either animal or human studies. If data on specific life stages or organ systems are unavailable or limited data suggest that availability of more extensive data might decrease the POD, this should be taken into account in assigning a database UF. For example, depending on the database and what is known about the chemical, the lack of a two-generation animal reproductive toxicity study might be considered a deficiency even if the reference value is based on human data. In any case, the size of the database factor to be applied will depend on other information in the database and on how much impact the missing data may have on determining the toxicity of a chemical and, consequently, the POD.

4.4.5.6. *Subchronic-to-Chronic-Duration UF*

As indicated earlier, a duration adjustment currently in use is the application of a UF when only a subchronic duration study is available to develop a chronic reference value such as the RfC or the RfD (U.S. EPA, 1994). A default value of 10 for this UF is applied to the NOAEL/LOAEL or BMDL/BMCL from the subchronic study on the assumption that effects from a given compound in a subchronic study occur at a 10-fold higher concentration than in a corresponding (but absent) chronic study. This factor would be applied subsequent to the adjustment of the exposures from intermittent to continuous, as above.

The specific use of a UF applied to a subchronic study in the derivation of a chronic reference value is reasonable. Some work has been published on this aspect of extrapolation (Lewis et al., 1990; Pieters et al., 1998). Guidance for replacement of the default factor of 10 by

CSAFs may be forthcoming. It would be appropriate to incorporate such data into applicable assessments. In the current practice, this factor is applied when a chronic reference value is derived from a database in which the critical study is of subchronic duration. No chronic reference value is derived if neither a subchronic nor chronic study is available. The application of a UF to less-than-subchronic studies is not part of the current practice, but further exploration of this issue may be appropriate. For short-term and longer-term reference values, the application of a UF analogous to the subchronic-to-chronic duration UF also needs to be explored, as there may be situations in which data are available and applicable but they are from studies in which the dosing period is considerably shorter than that for the reference value being derived.

4.4.5.7. *Modifying Factor (MF)*

A clear definition of intended usage for an MF is lacking. The only comments located about the MF are in the RfC methodology (U.S. EPA, 1994), and they indicate that the MF is intended to account for scientific uncertainties in the study or database that are not explicitly treated by other UFs. It is further stated that use of the factor depends principally on professional judgment and assessment. Some example applications are also given, such as accounting for small sample size or for poor exposure characterization in the principal study. The definition in the IRIS glossary gives similar examples.

The description of the database UF shows substantial similarity to that of the MF. Text on the database UF indicates that this factor attempts to recognize that without a comprehensive array of endpoints there is uncertainty as to whether all possible toxicologic endpoints at the various life stages are adequately addressed. Without this information, uncertainty remains as to whether the critical effect chosen for RfD or RfC derivation is either the most sensitive or the most appropriate. There are only seven cases in IRIS for which an MF has been applied: RfDs for chromium III, chromium VI, nitrite, 1,1-biphenyl, and manganese and RfCs for methyl ethyl ketone and acetonitrile. The rationale for these varies considerably but in all cases appeared to be for reasons that could be considered under other UFs.

Recent developments in the IRIS process include the obligation for risk characterization within the assessments. A central aspect of risk characterization includes discussing confidence and uncertainties in the quality of data used and the "clarity, transparency, consistency and reasonableness" of the assessment (U.S. EPA, 2000b). Within the risk characterization, the assessor has a pathway provided to discuss and analyze all aspects of uncertainty about the database, including the adequacy or limitations of the database, directly in the assessment.

The Panel considers the purpose of the MF to be sufficiently subsumed in the general database UF. The Panel also notes that the risk characterization section of assessments may be used to provide a full and complete characterization of all uncertainty, including any residual uncertainty that may not be addressed by the other UFs. In view of these factors, the Panel recommends that use of the MF be discontinued.

4.4.6. Future Directions

4.4.6.1. *Chemical-Specific Adjustment Factors (CSAFs)*

There is growing support for the use of CSAFs in place of DAFs (see Section 4.4.3.), and this will provide an incentive to fill existing data gaps (Murray and Andersen, 2001; Meek, 2001; Meek et al., 2001; Bogdanffy et al., 2001). Additional chemical-specific data permit the replacement of components of interspecies or inter-individual variation with data-derived values in the context of the traditional default framework as developed by Renwick (1993) and revised by IPCS (1994). The following is a brief discussion of available methodologies that promote the use of CSAFs in risk assessment.

Renwick (1993) described the use of toxicokinetic and toxicodynamic data as a means of replacing the traditional 10-fold safety factors for human sensitivity and experimental animal-to-human extrapolation in developing acceptable daily intakes. His data-derived approach assigns default values for both toxicokinetic and toxicodynamic differences within each traditional 10-fold safety factor. Specifically, Renwick proposed dividing both the interspecies and the inter-individual UFs into a factor of 2.5 for toxicodynamics and a factor of 4.0 for toxicokinetics. IPCS (1994) has adopted the data-derived approach initially proposed by Renwick (1993), with a slight modification in the UF for inter-individual variation (3.16 for toxicodynamics and 3.16 for toxicokinetics). IPCS has used this approach in several of its recent risk assessments (e.g., IPCS, 1998), and EPA is proposing a similar approach for boron (U.S. EPA, 2001b).

IPCS has developed a draft guidance document (IPCS, 2001) to assist risk assessors in the use of experimental data in deriving CSAFs for interspecies differences and human variability in dose/concentration response assessment. CSAFs have been adopted because they better describe the nature of the refinement to the usual default approach.

For several years, EPA used a more qualitative approach to modify the usual 10-fold default values (Dourson et al., 1996). Recently, it has used a data-derived approach as one of the methods to derive a UF for boron (U.S. EPA, 2001b).

EPA has not yet established guidance for the use of chemical-specific data for deriving UFs, but the division of UFs into toxicodynamic and toxicokinetic components is in the RfC methodology (U.S. EPA, 1994). EPA's assessments of data assume a division of both

interspecies and intraspecies UFs into toxicokinetic and toxicodynamic components that have assigned default values of 3.16 ($10^{0.5}$) each. The Agency will develop its own guidance for the use of CSAFs in risk assessment, based on some of the available methodologies (e.g., IPCS).

The Technical Panel would like to caution the user that for many substances there are relatively few data available to serve as an adequate basis to replace defaults for interspecies differences and human variability with more informative CSAFs. Currently, relevant data for consideration are often restricted to the component of uncertainty related to interspecies differences in toxicokinetics. Although there are fewer relevant data with which to address the other four components namely—interspecies (animal-to-human) differences in toxicodynamics, intraspecies (human) variability in toxicokinetics, intraspecies (human) variability in toxicodynamics, and adequacy of the database—it is anticipated that availability of such information will be needed to apply CSAFs. Specifically, the data-derived CSAF approach for any single substance is necessarily determined principally by the availability of relevant data. The extent of data available is, in turn, often a function of the economic importance of the substance, and this is frequently related to the extent of potential human exposure.

4.4.6.2. *Probabilistic Approaches*

Another approach to quantifying uncertainty in RfD or RfC derivation when data are not sufficient to develop a chemical-specific or biologically based dose-response model is probabilistic analysis. When the available data are sufficient to meaningfully characterize the distributions of interest, a probabilistic approach would provide results as a distribution rather than as a single measure for the dose/concentration-response. For example, distributions could be used for inputs into a toxicokinetic model to derive a distribution of internal dose metrics. Also, the approaches described in the draft IPCS guidance document (IPCS, 2001) are amenable to probabilistic analysis.

Probabilistic analysis for human health assessments generally has been confined to the exposure variables. In deriving human health toxicity reference values, inter-individual variability in toxicokinetics and toxicodynamics is usually represented with a UF because data are insufficient to support a more quantitative representation of these sources of inter-individual variability. Several studies have been published addressing the use of probabilistic data for health assessments (Baird et al., 1996; Maull et al., 1997; Slob and Pieters, 1998; Swartout et al., 1998; Brand et al., 1999; Gaylor and Kodell, 2000; Evans et al., 2001). The Technical Panel recommends that the Agency further evaluate approaches such as probabilistic analysis for characterizing variability and uncertainty in toxicity reference values.

4.4.7. Summary of Key Points from the Case Studies

Two case studies were developed to illustrate many of the recommendations in this report. The studies are for two hypothetical chemicals: Inhalate, a synthetic halogenated aliphatic alkene, and Luteinate, a new pesticide that acts via the neuroendocrine system.

The available database on Inhalate was considered adequate for deriving inhalation reference values for all four durations of exposure (acute, short-term, longer-term, and chronic). Very little is known about the mode of action for Inhalate except for the tumorigenic effects in liver, which are thought to be produced as a result of prolonged cytotoxicity caused by oxidative metabolism. Thus, a nonlinear mode of action is assumed for Inhalate carcinogenesis, and a chronic reference value is derived that takes into account these effects along with others seen after chronic exposure. Acute, short-term, longer-term, and chronic reference values were derived for Inhalate. This case study illustrates the use of a variety of types of data from toxicity testing studies in deriving a set of inhalation reference values, including carcinogenic effects assumed to have a nonlinear dose-response.

Luteinate belongs to a class of chemicals known to work through a neuroendocrine mode of action. In order to ascertain its potency and confirm a similar mode of action, a number of short-term studies were conducted, followed by testing in more traditional toxicology studies to establish its long-term effects and dose response relationships. The data were considered adequate to derive oral reference values for all four durations of exposure. This case study provides an example of the usefulness of mode-of-action information in establishing the short- and long-term effects of Luteinate on relevant target organ systems at different life stages. Such information enables the development of a targeted robust data set for use in establishing reference values for various durations of exposure.

A detailed summary of the case studies is provided in Appendix B. Several key points are described here that demonstrate the use of the proposed framework outlined in this chapter. First, the data are reviewed and characterized on the basis of the hazard and dose-response information, including consideration of the weight-of-evidence factors discussed in Section 4.3.2, above. A narrative statement is used to describe the extent of the database for each chemical as well as the gaps in information that would make the database more robust. Dosimetric adjustments were made to derive HECs in the case of Inhalate. For Luteinate, adjustments for oral exposure were made on a BW^1 basis and do not incorporate the $BW^{3/4}$ scaling factor or other DAF, as further work is needed on the harmonization of approaches for deriving of oral and dermal HEDs.

The data are presented both in tabular form and in graphical form as an exposure response array to provide a visualization of the data applicable to each duration of exposure.

Then, the reference values are derived by considering all of the relevant data for each duration reference value, weighing the evidence in the database, developing sample values on the basis of various endpoints considered for each duration, and selecting a final reference value for each duration on the basis of an evaluation of each of the relevant endpoints rather than on a single critical study and critical effect.

The approaches illustrated by the case studies showing derivation of multiple-duration reference values are not without precedent. Several offices within EPA, as well as ATSDR and the AEGL committee, derive multiple duration values for various purposes (see review in Chapter 2). The derivation of sample reference values in selecting the final reference value also is not a new idea. For example, EPA's assessment for methylmercury included the derivation of sample RfDs from prospective longitudinal studies of the effects of in utero exposure to methylmercury (Table 4-4) in deriving a chronic RfD (U.S. EPA, 2001d).

Sample RfDs were derived from a number of neuropsychological endpoints from two studies in which an association was observed (New Zealand and the Faroe Islands) as well as an integrative analysis of those studies plus a study in the Seychelles Islands in which no association between in utero methylmercury exposure and deficits in neuropsychological function were reported. The sample RfDs converged on 0.1 μg/kg/day, providing strong support for the appropriateness of this value. This RfD is not a developmental RfD per se, and its use is not restricted to pregnancy or developmental periods. The RfD, derived from an overall evaluation of the database, is applicable to lifetime daily exposure for all populations, including sensitive subgroups.

In a recently released health assessment document on 1,3-butadiene (U.S. EPA, 2002d), sample RfCs were derived for determining the chronic RfC.

Table 4-4. BMDLs, ingested dose, and RfDs for various endpoints from the Faroe Islands, New Zealand, and the NRC integrative analysis

Test[b]	BMDL[a] (ppb mercury cord blood)	Ingested dose[b] (µg/kg/day)	RfD[c] (µg/kg/day)
BNT Faroes			
Whole cohort	58	1.081	0.1
PCB adjusted[d]	71	1.323	0.1
Lowest PCB tertile	40	0.745	0.1
CPT Faroes			
Whole cohort	46	0.857	0.1
PCB adjusted	49	0.913	0.1
Lowest PCB tertile	28	0.522	0.05
CVLT Faroes			
Whole cohort	103	1.920	0.2
PCB adjusted	78	1.454	0.1
Lowest PCB tertile	52	0.969	0.1
Finger Tap Faroes			
Whole cohort	79	1.472	0.1
PCB adjusted	66	1.230	0.1
Lowest PCB tertile	24	0.447	0.05
Geometric mean			
Whole cohort	68	1.268	0.1
PCB adjusted	65	1.212	0.1
Lowest PCB tertile	34	0.634	0.1
Median values			
Faroes	48	0.895	0.1
New Zealand	24	0.447	0.05
Smoothed values			
BNT Faroes	48	0.895	0.1
CPT Faroes	48	0.895	0.1
CVLT Faroes	60	1.118	0.1
Finger Tap Faroes	52	0.969	0.1
MCCPP New Zealand	28	0.522	0.05
MCMT New Zealand	32	0.596	0.1
Integrative			
All endpoints	32	0.596	0.1

[a] $BMDL_{05}$s from NRC (2000), Tables 7-4, 7-5, 7-6. Hair mercury was converted to blood mercury using a 250:1 ratio and an assumption of equivalent maternal and cord levels.
[b] Calculated using a one-compartment model.
[c] Calculated using an UF of 10.
[d] There was significant co-exposure to PCBs in the Faroe Islands study, with PCB cord tissue concentrations available for about half of the whole cohort. Analyses were performed adjusted for PCBs (half the cohort) as well as unadjusted for PCBs in those individuals in the lowest PCB tertile (i.e., one-sixth of the whole cohort).

BNT = Boston Naming Test; CPT = Continuous Performance Test; CVLT = California Verbal Learning Test; MCCPP = McCarthy Perceived Performance; MCMT = McCarthy Motor Test.

5. RECOMMENDATIONS

A number of recommendations have been made in other parts of this report. This chapter summarizes those recommendations, based on the Technical Panel's review of the RfD and RfC process. The Technical Panel assumes that it will be possible to implement some of the recommendations in the near future, given adequate resources and personnel, whereas others will require additional effort. In particular, testing strategies are needed that consider toxicokinetic and mode of action information early in the process, as well as when to implement new testing guidelines in the process of developing a data package on a particular chemical. OPPTS, together with scientists in other parts of the Agency, will consider the recommendations to develop additional or alternative testing guidelines as part of the Harmonized Health Effects Test Guidelines (870 Series).

As part of its deliberations, the Technical Panel considered the recommendations of the Toxicology Working Group of the 10X Task Force (U.S. EPA, 1999b, and Appendix A). The Technical Panel endorses those recommendations and extends and expands them to deal with a broader view of life stages, timing and duration of exposure, and evaluation of endpoints, both structural and functional. The recommendations are presented here in the order of the chapters in which they appear. Further discussion of the specific recommendations can be found in the earlier chapters.

Chapter 2

The Technical Panel concurred with the recommendation of the 10X Task Force that reference values should be derived, where possible, for acute, short-term, and longer-term as well as chronic exposures for oral, dermal, and inhalation routes and that they be included in the IRIS database for use by EPA programs, where applicable. The definitions for duration should be standardized but left flexible so they can be adjusted depending on the exposure situation of concern.

Chapter 3

The Technical Panel reviewed and evaluated current testing guidelines and testing approaches as a follow-up to its recommendation in Chapter 2 concerning the derivation of less-than-lifetime reference values. This review was undertaken to determine what information is currently gathered with regard to life stage assessment, endpoint assessment, route and duration of exposure, and latency to response. The intent of this review is not to suggest that additional

testing be conducted for each and every chemical in order to fill in the information gaps identified for those organ systems evaluated. Nor is it suggested that the alternative testing protocols discussed in this chapter be conducted for every chemical or become part of current toxicology testing requirements or that these alternative protocols are the only options available. Rather, it is the goal of this document to provide a basis for the development of innovative alternative testing approaches and the use of such data in risk assessment. The recommendations include:

- Develop a strategy for alternative approaches to toxicity testing, with guidance on how and when to use existing and newly recommended guidelines.

- Develop guidelines or guideline study protocols that will provide more systematic information on toxicokinetics and toxicodynamics (i.e., mechanism or mode of action), including at different life stages.

- Develop protocols for acute and short-term studies that provide more comprehensive data for setting reference values.

- Modify existing guideline study protocols to provide more comprehensive coverage of life stages for both exposure and outcomes.

- Collect more information from less-than-lifetime exposures to evaluate latency to effect and reversibility of effect.

- Develop guidelines or guideline study protocols to assess immunotoxicity, carcinogenicity, and cardiovascular toxicity at different life stages.

- Explore the feasibility of setting dermal reference values for direct toxicity at the portal of entry, including sensitization.

Chapter 4

The Technical Panel discussed a number of modifications to the existing framework for reference value derivation, both for the current chronic reference values (RfD and RfC) and for the acute, short-term, and longer-term reference values. In addition, two case studies that

illustrate many of these concepts are summarized in Chapter 4 and discussed in detail in Appendix B. The recommendations for improvement and expansion of the existing approaches are aimed at taking a broader approach to the characterization of the entire database and what impact that will have on the dose-response assessment and risk characterization of a chemical. Included are recommendations for setting several less-than-lifetime reference values, broader characterization of the database instead of using a checklist of a minimum set of studies for setting a reference value, using an exposure-response array and carrying appropriate and relevant endpoints through the derivation of sample reference values before deciding which endpoint(s) to use for the POD, and deriving reference values in a way that is protective of all relevant endpoints rather than setting reference values on particular endpoints (e.g., the RfD_{DT}) but using a process that facilitates the evaluation of risk to particular subgroups for specific program office needs, including cumulative risk assessment.

The specific recommendations follow:

1. Include the acute, short-term, longer-term, and chronic reference values derived on the basis of the recommendations in this report in IRIS after appropriate internal, external, and consensus review.

2. Use consistent definitions for the duration of exposure in deriving acute, short-term, longer-term, and chronic reference values.

3. Use the revised definition for reference values shown in Chapter 4. This definition is aimed at clarifying that the approach to reference values discussed here is intended for risk assessments for any type of health effect known or assumed to be produced through a nonlinear and/or threshold mode of action (which may include U-shaped or other nonmonotonic dose-response curves as well as thresholds). Thus, the term "noncancer" has been removed from the definition in the spirit of overall harmonization of risk assessment approaches for human health effects because it has been recommended that health effects no longer be categorized as "cancer" or "noncancer" for the purposes of hazard characterization and dose-response analysis. This change denotes the move toward defining approaches for low-dose estimation or extrapolation based on mode of action.

The term "deleterious" has been replaced with the term "adverse," because the latter is more commonly used and understood in data evaluation and selection of endpoints for setting reference values. The parenthetical statement in the current RfD and RfC definitions, "with uncertainty spanning perhaps an order of magnitude," has been removed from the proposed revision of the definition for reference value, and it is recommended that issues of uncertainty/variability be discussed qualitatively as part of the weight of evidence and characterization of the database.

4. For consistency in the designation of various duration reference values, the Panel recommends that the terminology for reference values be standardized; this standardized terminology should reflect both duration and route of exposure. Consistent terminology recommendations for reference values are proposed in this report, but additional suggestions are welcome.

5. The Technical Panel recommends that endpoint-specific reference values per se not be developed, including the RfD_{DT}, which was originally proposed in *Guidelines for Developmental Toxicity Risk Assessment* (U.S. EPA, 1991). Rather, a sample reference value should be calculated for each relevant and appropriate endpoint and these should then be considered in the derivation of various duration reference values. The reference values should be derived to be protective of all types of effects for a given duration of exposure.

6. An expanded approach to the evaluation of studies and characterization of the extent of the database as a whole is recommended; in particular, several factors are discussed that should be considered in a weight-of-evidence approach for characterizing hazard for the population as a whole as well as for potentially sensitive subpopulations. Those considerations for assessing level of concern raised by the Toxicology Working Group of the 10X Task Force (U.S. EPA, 1999b) have been incorporated into this approach.

7. A narrative approach rather than a confidence ranking of high, medium, or low should be used in describing the extent of the database. The extremes for the extent of the database (i.e., minimal or robust) are defined in Chapter 4. The narrative approach is intended to emphasize the types of data available (both human and animal) as well as research needed to fill the data gaps that could improve the derivation of reference

values, and it should encourage the use of a wider range of information in deriving reference values, taking into consideration the life stages evaluated; the issues of timing, duration, and route of exposure; the types and extent of endpoint assessment (i.e., structure and function); and the potential for latent effects and/or reversibility of responses.

8. Duration adjustment procedures to continuous exposures for inhalation developmental toxicity studies should be done in the same way as for other health endpoints.

9. Additional consideration of the HEC and HED derivation methodology is needed to confirm or assess the relevance for all population subgroups (particularly children).

10. An exposure-response array should be used as a visual display of all relevant endpoints and durations of exposure, as shown in the case studies. This type of array can be used to evaluate the range of exposure-response data for different durations of exposure in order to determine the range of numerical values available for each route and duration reference value.

11. The POD should be selected on the basis of an evaluation of all appropriate and relevant endpoints carried through to sample reference value derivation, with selection of the limiting value(s) as the final step, rather than on a single "critical study" and "critical effect."

12. It is imperative that the IRIS documentation contain a justification for the individual factors selected for each chemical or assessment, because rigid application of UFs could lead to an illogical set of reference values. Although default factors of 10 are recommended, with 3 used in place of half-power values (i.e., $10^{0.5}$) when occurring singly, the exact value of the UF chosen should depend on the quality of the studies available, the extent of the database, and scientific judgment. Sound scientific judgment should be used in the application of UFs to derive reference values that are applied to the value chosen for the POD derived from the available database (BMDL, NOAEL, or LOAEL).

13. The Technical Panel recommends that if there is uncertainty in more than four areas of extrapolation, it is unlikely that the database is sufficient to derive a reference value. Even when there is uncertainty in four areas, the database should be carefully evaluated to determine whether the derivation of a reference value is appropriate. In addition, the Technical Panel recommends limiting the total UF applied to a chronic reference value for any particular chemical to 3000. This maximum of 3000 applies only to the UFs and does not include the various adjustment factors discussed in Chapter 4.

14. The intraspecies UF is applied to account for variations in susceptibility within the human population (interhuman variability) and the possibility (given a lack of relevant data) that the database available is not representative of the dose/exposure-response relationship in the subgroups of the human population that are most sensitive to the health hazards of the chemical being assessed. As the reference concentration/dose is defined to be applicable to "susceptible subgroups," this UF was established to account for uncertainty in that regard. In general, the Technical Panel reaffirms the importance of this UF, recommending that reduction of the intraspecies UF from a default of 10 be considered only if data are sufficiently representative of the exposure/dose-response data for the most susceptible subpopulation(s).

At the other extreme, a 10-fold factor may sometimes be too small because of factors that can influence large differences in susceptibility, such as genetic polymorphisms. The Technical Panel urges the development of data to support the selection of the appropriate size of this factor, but it recognizes that often there are insufficient data to support a factor other than the default.

15. The Technical Panel urges continued research and evaluation of the similarities and differences between the general population and sensitive subpopulations in their responses to particular agents, particularly children and the elderly. From such evaluations, the protectiveness of the 10-fold default factor can continue to be assessed.

16. Given that several UFs can be used to deal with data deficiencies as part of the current reference value process, and given that these are assumed to overlap to some extent, the Technical Panel agrees with the 10X Task Force Toxicology Working Group (U.S. EPA, 1999b) that the current interspecies, intraspecies, and database deficiency UFs, if appropriately applied using the approaches recommended in this review, will be

adequate in most cases to cover concerns and uncertainties about children's health risks. Any residual concerns about toxicity and/or exposure can be dealt with in risk characterization/risk management (e.g., by retention of all or part of the FQPA safety factor for pesticides).

17. The Panel considers the purpose of the MF to be sufficiently subsumed in the general database UF. Therefore, the Panel recommends that use of the MF be discontinued.

18. EPA has not yet established guidance for the use of specific data to replace UFs (i.e., CSAFs), but the division of the interspecies UF into toxicodynamic and toxicokinetic components is in the RfC methodology (U.S. EPA, 1994) and may apply to the intraspecies UF as well. The Agency is encouraged to develop its own guidance, based on some of the available methodologies (e.g., IPCS).

The following issues were discussed by the Technical Panel but were considered more appropriate for discussion and recommendation by other panels/committees:

1. There have been inconsistencies in the use of BMD modeling approaches to deriving RfDs and RfCs currently in IRIS. The Technical Panel was unable to fully evaluate these issues or to reach agreement about any recommendation for change to current methodology and recommends that they be considered further by the Agency. The Technical Panel also recommends that factors such as the response rates at the BMD or NOAEL, the power of the study, and slope of the dose-response curve be included in the description of the database, where possible, as part of risk characterization.

2. The Technical Panel recommends harmonization of the approaches for HEC and HED derivation for all types of health effects. Development of the appropriate adjustment procedure is referred to the Harmonization Framework Technical Panel.

3. The Technical Panel recommends that the Agency further evaluate approaches such as probabilistic analysis for characterizing variability and uncertainty in toxicity reference values.

4. The Technical Panel recommends further evaluation of appropriate adjustment of doses for duration of exposure. The method derived from ten Berge et al. (1986) is raised as a possibility for acute exposures on the basis of its recommendation in the ARE methodology. Duration adjustment for short-term and longer-term reference values analogous to the subchronic-to-chronic duration UF for chronic reference values is raised in the case study and should be explored further.

APPENDIX A: ISSUES RAISED BY THE 10X TASK FORCE

A number of issues were raised by the 10X Task Force[1] in its discussions of the requirements for protecting children's health and application of an additional 10X safety factor, as mandated by the 1996 FQPA. The Task Force felt that these issues, which include the following, should be discussed on a broader Agency-wide basis as well as with the outside community for both pesticides and other agents.

1. *Appropriate application of the database modifying factor for additional required developmental and adult toxicity studies.* It appears from the data available that the default intraspecies 10-fold uncertainty factor may be adequate in the majority of cases for protecting children's health. However, when data specific to children's health are missing or inadequate for a particular agent, application of the database modifying factor in addition to the intraspecies variability factor may be sufficient to account for the possibility that children may be significantly more sensitive than adults. This issue needs further examination.

2. *How to account for the level of concern in the RfD/RfC process.* Criteria for assessing the level of concern for children's health were developed by the Toxicology Working Group of the 10X Task Force and include factors such as (a) human data on pre- and postnatal toxicity; (b) pre- and postnatal toxicity in animal studies, including effects of a different or similar type as those in adults; (c) dose-response nature of the experimental animal data, including the dose-related incidence of response, relative potency of response, slope of the dose-response curve when the margin of exposure is small, and how well the NOAEL or BMD is defined; and (d) relevance of the experimental animal data to humans, including toxicokinetics, similarity of the biological response, and knowledge of the mechanism of action. For each of these areas, criteria are given for estimating a level of concern for children's health as high, moderate, or low. The level of concern may be taken into account in the uncertainty and modifying factors applied to the RfD, although there is currently no formal process for doing so.

---

[1]See 10X Task Force documents: *Toxicology Data Requirements for Assessing Risks of Pesticide Exposure to Children's Health* (U.S. EPA, 1999b) and *Exposure Requirements for Assessing Risks from Pesticide Exposure to Children's Health* (U.S. EPA, 1999c).

3. As indicated in the toxicology document appended to the Task Force report, *the current default recommended for using developmental toxicity data for different duration reference values is to apply most endpoints for all durations*. This is because it is assumed that most endpoints of developmental toxicity can be caused by a single exposure. If, however, developmental effects are more sensitive than those seen after longer-term exposures, then even the chronic RfD/RfC should be based on such effects to reduce the risk of potential greater sensitivity in children. Because the standard studies currently conducted for developmental toxicity involve repeated exposures, data are not often available on which endpoints may be induced by acute, subacute, subchronic, or chronic dosing regimens and, therefore, on which should be used in setting various duration reference values. Further consideration of the appropriate application of developmental toxicity endpoints to various duration reference values is recommended. As part of this recommendation, an in-depth review of the HED document on Hazard Identification—Toxicology Endpoint Selection System, should be undertaken.

4. *Appropriate setting of intermediate RfDs/RfCs for pesticides and other agents.* The focus of the RfD and the RfC has been on chronic exposure reference values. Acute RfDs are also set for pesticides, and intermediate reference values are set for residential exposures as well as for drinking water. Data on developmental toxicity will often be a greater factor in calculating the acute and intermediate reference values, and exposures to children are more often of this type as well. Consideration should be given to setting intermediate reference values for environmental agents. In addition, the question of whether or when to set RfDs/RfCs specific for children should be considered.

5. *Appropriate adjustment of the NOAEL or the BMD from inhalation exposure studies for extrapolation of developmental toxicity data using less-than-continuous exposure to a continuous-exposure scenario.* Currently, NOAELs/BMDs from inhalation exposure studies other than those for developmental toxicity using, for example, a 6-hr/day exposure regimen, are adjusted to a continuous (24 hr/day) exposure for calculating RfDs/RfCs. The developmental toxicity risk assessment guidelines (U.S. EPA, 1991) recommended against making this adjustment because it was assumed that there was a threshold above which exposure would have to occur before an effect would result. This recommendation needs to be reconsidered, along with the adjustment of NOAELS/BMDs in general.

Several improvements in testing approaches were also proposed for consideration in the 10X Task Force report as a way to improve the assessment of potential risks to children. The Technical Panel was asked to consider the need for such tests, when they should be required, and interpretation of the data for risk assessment purposes. The improvements to be considered include

- pharmacokinetics that include data from different developmental stages, perhaps done in a tiered approach as suggested in Kimmel and Francis (1990);

- direct dosing of neonates, especially when early exposure is of concern, because this is the time when differences in metabolic capability are greatest;

- perinatal carcinogenesis studies and appropriate triggers for when they should be required;

- developmental immunotoxicity testing and appropriate triggers; and

- advanced DNT testing, in particular, cognitive testing that is more similar to that used in humans.

An additional issue was how to make exposure assessments compatible with the dose-response assessment. For example, how should the appropriate durations of exposure be determined for toxic endpoints of concern? Should standard exposure durations be used?

APPENDIX B: CASE STUDIES—EVALUATING AND SELECTING HEALTH ENDPOINTS FOR DERIVING REFERENCE VALUES

Two case studies were developed by the Technical Panel to illustrate many of the points discussed in this report. The first case study is of a hypothetical volatile chemical for which limited data are available and little information is known about the mode of action except that there is support for a nonlinear mode of action for cancer.

The second case study is of a hypothetical endocrine disruptor for which the mode of action is known or assumed from other chemicals in the same class. This case study is used to illustrate in part how the information on mode of action can inform a more focused collection of data as well as the interpretation of the data and its use in risk assessment.

In both case studies, NOAELs and LOAELs rather than BMDLs or BMCLs are used to derive reference values, in large part because the data are fictitious and were not developed to the point that they could be readily modeled. However, the Technical Panel strongly encourages the use of dose-response modeling and calculation of BMDLs or BMCLs for selection of the POD to be used as the basis for deriving reference values.

CASE STUDY 1: INHALATE

Inhalate, a synthetic halogenated aliphatic alkene, is a nonflammable volatile liquid at room temperature. The chemical enters the air through its industrial and commercial use, primarily as a solvent. It is also found in surface and ground water and soil upon disposal. The most important route of human exposure is inhalation of the chemical in the ambient and indoor air, although there is a lower possibility of ingestion through contaminated drinking water. Because of its high volatility, dermal exposure to the chemical is expected to be minimal.

This case study illustrates the use of single or multiple health endpoints for deriving reference values for different durations of exposure following inhalation exposure. It also illustrates the harmonized approach for all effects (including cancer) that are known or assumed to be produced through a nonlinear or threshold mode of action. For the purpose of illustration, results of key studies are summarized in Table B-1, including dose-response data for different health endpoints relevant to different durations of exposure via inhalation exposure. Although oral data for this chemical were available, a brief description is included here only to show the consistency with which effects were seen after either inhalation or oral exposure.

SUMMARY OF HEALTH EFFECTS INFORMATION

Absorption, Distribution, Metabolism, and Elimination

There is very little information on the absorption and distribution of Inhalate in humans and laboratory animals following oral, inhalation, or dermal exposure. However, similar effects are seen by oral and inhalation exposures, suggesting that Inhalate or its metabolites reach their target sites after absorption from either exposure route. Available in vitro metabolic studies indicate that Inhalate is extensively metabolized in target tissues including the liver and kidney of rats and mice. Limited in vitro studies with human tissues show a similar pattern of metabolism. As discussed below, much of Inhalate-induced toxicity appears to be due to its metabolites. These metabolites have been detected in the urine of rats and mice following inhalation and oral exposure to the parent chemical.

Postulated Mode of Action

No information is available on mode of action except for the carcinogenicity of Inhalate. The carcinogenic effects of Inhalate in rodent liver are attributed to oxidative metabolism-mediated cytotoxicity in the target organ. The oxidative metabolism produces highly tissue-reactive metabolites that lead to tissue injury and cell death. The persistent cell proliferation

Table B-1. Summary results of major inhalation exposure studies on Inhalate

Species	Sex	Exposure duration and frequency	Concentrations (mg/m^3)	LOAEL/ NOAEL (mg/m^3)	HEC [a] (mg/m^3)	Responses
Human	M/F	2 hrs	4, 40, 400	40/4	40/4, 2-hr	Headache, dizziness, incoordination, drowsiness, anesthesia at 2000 mg/m^3
		Accidental exposure	NA	NA	-	Narcosis, proteinuria, hematuria
		6 hrs/day for 7 days (clinical exposure)[b]	10, 20, 100, 150	20/10	4/2	Headache, dizziness, incoordination, drowsiness
		Occupational (>15 yrs)	TWA of 56	20/NA	20/NA	Dizziness, forgetfulness; changes in serum liver enzymes; increased urinary levels of lysozymes, beta-glucuronidase
	F	Occupational	NA	NA	-	Menstrual disorders; spontaneous abortion; cardiac anomalies in children of workers
Rat/ mouse	M/F	4 hrs	0, 13, 24, 50	13/NA	13/NA, 4-hr	Dose-related hyperactivity, ataxia, hypoactivity, narcosis
		6 hrs/day, 5 days/wk for 2 wks	0, 50, 100, 200	50/NA	9/NA	Dose-related hyperactivity, ataxia, hypoactivity, narcosis
Rat	M/F	6 hrs/day, 5 days/wk for 13 wks	0, 30	30/NA[c]	–	Changes in fatty acid composition of the brains

Table B-1. Summary results of major inhalation exposure studies on Inhalate (continued)

Species	Sex	Exposure duration and frequency	Concentrations (mg/m^3)	LOAEL/ NOAEL (mg/m^3)	HEC [a] (mg/m^3)	Responses
Rat (cont)	M/F	6 hrs/day, 5days/wk for 13 wks	0, 22, 39, 88	39/22	7/4	Dose-related hyperactivity, ataxia; liver hypertrophy, vacuolization of hepatocytes, necrosis; cytomegaly, toxic nephrosis of tubular epithelial cells
Mouse	M/F	6 hrs/day, 5 days/wk for 13 wks	0, 28, 50, 100	50/28	9/5	Dose-related hyperactivity, ataxia; liver hypertrophy, vacuolization of hepatocytes, necrosis; cytomegaly, toxic nephrosis of tubular epithelial cells
Rat	M/F	6 hrs/day, 5 days/wk for 104 wks	0, 11, 22, 44	22/11	4/2	Clinical signs of neurotoxicity; dose-related liver hypertrophy, vacuolization of hepatocytes, necrosis, hepatocellular carcinoma; cytomegaly, toxic nephrosis of tubular epithelial cells
Mouse	M/F	6 hrs/day, 5 days/wk for 78 wks[e]	0, 11, 28, 56	28/11	5/2	
Rat	F	GDs 70–13, 6 hrs/day	0, 10, 90	90/10	22.5/2.5	Decreased motor activity in pups; ataxia in dams at high dose
	M/F	Two-generation reproductive study, 6 hrs/day, 5 days/wk	0, 28, 100	100/28	18/5	Reduced litter size, reduced survival of offspring, sedation at high dose

[a] These values are approximate HECs derived in accordance with the RfC methodology (U.S. EPA, 1994) and Chapter 4 (this report) for category 3 Gases using pharmacokinetic data for Inhalate.

[b] Data included here for comparison; data from intentional human exposures are not currently used by EPA in risk assessment.

[c] Special study for neurotoxicity, no HEC calculated.

[d] Considered to be a lifetime exposure

HEC = human equivalent conentration

NA = not available

TWA = time-weighted average

GD = gestation day

presumably would lead to higher probabilities of cell mutation and subsequent cancer. Liver tumors were produced only at dose levels that resulted in repeated or sustained cytotoxicity and regenerative cell proliferation. This postulated mode of action is further supported by the observation in specialized studies that neither the cytotoxicity nor cell proliferation occurred in the CYP2E1 null mouse or in the wild type treated with a P450 inhibitor at the same exposure. The weight of the evidence indicates that a mutagenic mode of action via DNA reactivity is not a significant component of Inhalate-induced liver tumors in rats and mice. There are no data indicating that the mode of action observed in rodents is not also likely to apply to humans.

Nervous System

Inhalate has been found to elicit dose-dependent clinical signs of CNS effects in adult humans following acute inhalation exposure and accidental ingestion. CNS symptoms have been reported in several studies of occupational exposure of workers to Inhalate. Dose-dependent clinical signs of CNS effects have been observed in adult rats and mice exposed to Inhalate by inhalation following different duration of exposures. There are limited data indicating that prenatal exposure to Inhalate adversely affects the developing nervous system in rats and mice (see Growth and Development below). There are no data on the ability of Inhalate to affect the nervous system at other life stages (e.g., during the fetal period, infancy, childhood, or old age). The mechanism of action for the CNS effects has not been clearly established, but it is believed to be related to effects of the parent compound on lipid and fatty acid composition of the membranes.

Inhalation Exposure

Several reports available in the open literature indicate dose-dependent clinical signs of CNS symptoms in adults exposed acutely and subacutely via inhalation to Inhalate. Males and females exposed acutely to high concentrations (40–400 mg/m^3 for 2 hours) showed dose-dependent effects, including headache, dizziness, incoordination, drowsiness, and anesthesia. No effect was reported following acute exposure to 4 mg/m^3. Similar effects were observed in adult human volunteers at lower concentrations (10, 20, 100, 150 mg/m^3) for 6 hours per day for up to 7 days, with a NOAEL of 10 mg/m^3. An acute accidental exposure of a small group of workers to an unknown (presumably high) concentration resulted in narcosis.

Long-term and chronic neurotoxic effects have been reported in several studies of occupational exposure of workers to Inhalate in different industries. Exposure data were not provided in these reports; however, it can be presumed that these workers were exposed to a

daily TWA exposure of 56 mg/m^3. Subjective neurological symptoms, including dizziness and forgetfulness, were consistently reported across studies. No other information on possible neurological effects was collected in these studies.

Concentration-dependent clinical signs of neurological effects, including hyperactivity, ataxia, hypoactivity, and finally loss of consciousness, have also been reported in adult rats and mice following acute (13, 24, 50 mg/m^3 for 4 hours) and short-term inhalation exposure (50,100, 200 mg/m^3 6 hrs/day for 2 weeks) to Inhalate at high concentration. Similar effects were observed in rats exposed at 39 and 88 mg/m^3 and in mice at 50 and 100 mg/m^3 for 13 weeks. The subchronic NOAELs for rats and mice were 22 and 28 mg/m^3, respectively.

Chronic exposure to Inhalate at lower concentrations resulted in less serious clinical signs of CNS effects in rats (22 or 44 mg/m^3) and mice (28 or 56 mg/m^3). The chronic NOAELs for rats and mice in these studies were both 11 mg/m^3. It should be noted that neurological endpoints examined in these animal studies are limited to clinical signs and histopathology. In a special study, changes in fatty acid composition of the brain were observed in rats exposed at 30 mg/m^3 (the only tested concentration) for 90 days.

Pregnant rats were exposed by inhalation to Inhalate at 0, 10, or 90 mg/m^3 for 6 hrs/day on days 7–13 of gestation. Decreased motor activity was observed in 21- or 60-day-old pups from dams exposed to 90 mg/m^3. A NOAEL of 10 mg/m^3 for developmental neurotoxicity was identified in this study.

Oral Exposure

Acute neurological effects in adult humans after ingestion of Inhalate are similar to those seen after inhalation. Accidental exposure to approximately 6–8 mL (or about 100 mg/kg/day) resulted in narcotic effects. Neural tube defects and eye anomalies were reported in studies of offspring of residents exposed to drinking water contaminated with Inhalate and other solvents. Exposure levels were not determined in this study.

Single oral gavage administration of Inhalate to adult rats (1000 mg/kg) caused ataxia. Ataxia was also observed in pregnant rats treated by gavage at 900 mg/kg on gestation days (GDs) 6–19. No CNS effects were reported in a chronic oral gavage study in rats and mice at 50, 100, or 300 mg/kg/day. However, neurological endpoints examined in these studies were limited to clinical signs and histopathology.

In a study that investigated the effect of Inhalate on the developing nervous system, male mouse pups were treated by gavage at 50 or 300 mg/kg/day for 7 days (postnatal days [PNDs]

10–17). Hyperactivity was reported in animals during adulthood at the high dose. No studies with exposure throughout development (prenatal and postnatal) were available.

Liver

Two reports provided suggestive evidence of liver effects in workers exposed to Inhalate in different industries. Inhalate has been shown to induce dose-dependent liver toxicity in adult rats and mice following subchronic and chronic exposure by inhalation and oral gavage. Liver tumors were also observed in chronic studies of rats and mice. Available data support the conclusion that liver tumors were produced only at dose levels that resulted in repeated or sustained cytotoxicity and regenerative cell proliferation. Inhalate carcinogenic effects in rodent liver are attributed to oxidative metabolism-mediated cytotoxicity in the target organ. The oxidative metabolism produces highly tissue-reactive metabolites that lead to tissue injury and cell death. The persistent cell proliferation presumably would lead to higher probabilities of cell mutation and subsequent cancer.

This postulated mode of action is further supported by the observation in specialized studies that neither the cytotoxicity nor cell proliferation occurred in the CYP2E1 null mouse or in the wild type treated with a P450 inhibitor at the same exposure. The weight of the evidence indicates that a mutagenic mode of action via DNA reactivity is not a significant component of Inhalate-induced liver tumors in rats and mice. No data exist indicating that the mode of action observed in rodents is not also likely to apply to humans. There are no data that provide any insights into possible differential sensitivity across life stages.

Inhalation Exposure

One study reported changes in serum levels of liver enzymes in workers exposed to the chemical at a daily TWA exposure concentration of about 56 mg/m^3 over an 8-hour work shift. These workers, however, did not exhibit any clinical symptoms of liver dysfunction.

Dose-related liver effects (liver hypertrophy, vacuolization of hepatocytes, necrosis) have been observed in mice following subchronic exposure (13 weeks) to Inhalate at 50 or 100 mg/m^3, with a NOAEL of 28 mg/m^3. Dose-related liver toxicity and hepatocellular carcinomas were also found in mice following chronic exposure at 28 and 56 mg/m^3. The NOAEL for liver toxicity in mice in this chronic study was 11 mg/m^3.

Rats showed similar liver responses but at higher exposure concentrations following subchronic exposure (39 or 88 mg/m^3), with a NOAEL of 22 mg/m^3. Liver toxicity and hepatocellular carcinomas were also observed at 22 or 44 mg/m^3 in a chronic study in rats. The

NOAEL for liver effects in rats was 11 mg/m^3. It should be noted that liver effects examined in these subchronic and chronic studies were limited to clinical chemistry, morphology, and histopathology.

Oral Exposure

Liver effects were observed in mice and rats treated subchronically (100, 300, 500 mg/kg/day) or chronically (50, 100, 300 mg/kg/day) with Inhalate via oral gavage and were similar to those seen after inhalation exposure. Mice showed more severe effects than did rats. Dose-related hepatocellular carcinomas were also found in treated mice in a chronic study.

Kidney

Available human and animal studies indicate that Inhalate also has the potential to cause renal toxicity in adults. The mechanisms for the development of kidney effects in humans and animals are not known. No data were available to evaluate the effects of Inhalate at life stages other than in adults (i.e., during development or old age).

Inhalation Exposure

Symptoms of renal dysfunction (proteinuria, hematuria) have been associated with accidental human exposure to anesthetic concentrations of Inhalate. Subtle or no renal effects were reported in workers exposed chronically. Increased urinary levels of lysozyme and beta-glucuronidase suggestive of mild renal tubular damage have been observed in workers exposed for an average of 15 years to a daily TWA concentration of 56 mg/m^3.

Dose-related renal toxicity (cytomegaly, toxic nephrosis of tubular epithelial cells in the inner renal cortex) was induced in rats (39, 88 mg/m^3) and mice (50, 100 mg/m^3) exposed to Inhalate for 13 weeks. Subchronic NOAELs for renal effects in rats and mice were 22 mg/m^3 and 28 mg/m^3, respectively. Similar renal effects were observed in a chronic study in rats (22, 44 mg/m^3) and mice (28, 56 mg/m^3). Chronic NOAELs for renal effects in rats and mice were both 11 mg/m^3.

Oral Exposure

Dose-related toxic nephropathy characterized by degenerative changes in the proximal convoluted tubules and necrosis of the tubular epithelium were found in rats and mice treated with Inhalate via oral gavage for 90 days at 100, 300, or 500 mg/kg/day and for 2 years at 50,

100, or 300 mg/kg/day. Subchronic and chronic NOAELs for renal effects in rats and mice were at 100 and 50 mg/kg/day, respectively.

Growth and Development

Available studies in humans and animals indicate that Inhalate has the potential to cause developmental effects by inhalation and oral ingestion. Limitations of human studies could not resolve whether the observed developmental effects were causally related to the chemical or were a result of chance or bias. However, the epidemiologic findings are supported by animal studies with exposure to Inhalate by inhalation and oral gavage that show that the developing nervous system is the most sensitive target in rats and mice.

Inhalation Exposure

Epidemiologic studies of women occupationally exposed to Inhalate and other related solvents have reported elevated risk of spontaneous abortion and cardiac anomalies in their offspring. Due to limitations of these studies, an exposure-response could not be established. No other health endpoints were investigated in these studies.

Pregnant rats were exposed by inhalation to Inhalate at 0, 10, or 90 mg/m^3 for 6 hrs/day on days 7–13 of gestation. Decreased motor activity was observed in 21- and 60-day-old pups from dams exposed to 90 mg/m^3. A NOAEL of 10 mg/m^3 for developmental effects was identified in this study. No studies included exposure throughout gestation and lactation to determine effects at other developmental stages. Data from a two-generation reproduction study indicated a reduction in litter size and survival of offspring in rats exposure to Inhalate at 100 mg/m^3, a concentration that also resulted in sedation and renal effects. No effects were reported at 28 mg/m^3.

Oral Exposure

Neural tube defects and eye anomalies have been reported in studies of residents exposed to drinking water contaminated with Inhalate and other solvents. Exposure levels were not determined in this study.

An increased incidence of micro/anophthalmia were observed in the offspring of rats treated with Inhalate by gavage at 900 mg/kg/day on GDs 6–19. In a study that investigated the effect of Inhalate on the developing nervous system, male mouse pups were treated by gavage with Inhalate at 50 or 300 mg/kg/day for 7 days (age 10–17 days). Hyperactivity was reported in animals during adulthood at the high dose. No effects were found at the low dose. No studies

included exposure throughout gestation and lactation to determine effects at other developmental stages.

Reproductive System

Available studies in humans and animals suggest that Inhalate may have the potential to cause reproductive effects. The underlying mechanism of action for potential reproductive effects is not known.

Inhalation Exposure

There is suggestive evidence of spontaneous abortion and menstrual disorders among women occupationally exposed to Inhalate. However, no definitive conclusions can be made because of the limitations associated with these studies.

Reduced litter size and reduced survival of offspring were reported in rats exposed to Inhalate at 100 mg/m^3 in a two-generation reproduction inhalation study; this concentration also resulted in sedation and renal effects. No effects were identified at 28 mg/m^3. The protocol used, however, was not the most recent one, in which reproductive development (e.g., timing of puberty or anogenital distance) and adult reproductive function (semen quality, estrous cyclicity) are evaluated, nor were organ weights measured. No effects on the reproductive system were noted in any other studies.

Oral Exposure

No information is available on the potential reproductive effects of Inhalate in animals via oral exposure.

SELECTION OF HEALTH ENDPOINTS AND DERIVATION OF REFERENCE VALUES

Narrative Description of the Extent of the Database

No information is available on possible modes of action except for liver carcinogenic effects in rats and mice. In this case, the mode of action is attributed to oxidative metabolism-mediated cytotoxicity and persistent cell proliferation in the liver. Persistent regenerative cell proliferation presumably would lead to higher probabilities of cell mutation and subsequent cancer. No data exist to indicate that the mode of action observed in rodents is not also likely to apply to humans. Pharmacokinetic data indicate that Inhalate is extensively metabolized and that much of its toxicity is due to the metabolites.

The database for inhalation exposure is limited but adequate for deriving reference values. Some human data on acute, short-term, and longer-term exposures are available, although the range of endpoints evaluated and the dose-response information for different durations of exposure are limited. The animal data include acute, short-term, longer-term, and chronic studies with exposures beginning in young adult animals. The acute and short-term data are limited to clinical signs of morbidity and mortality, whereas the longer-term and chronic studies include some histopathology as well.

A developmental neurotoxicity (DNT) study was conducted in rats with prenatal exposure limited to GDs 7–13 (as opposed to more extensive exposure throughout a major part of CNS development, e.g., GD 6 to PND 11 or 20 in the standard DNT study testing protocol). No other studies of prenatal or postnatal developmental toxicity were done except for evaluations of survival and growth in a two-generation reproduction study in rats. However, the protocol used was one in which reproductive development (e.g., timing of puberty or anogenital distance) and adult reproductive function (semen quality, estrous cyclicity) were not evaluated, nor were organ weights measured. No studies were conducted that considered issues related to the toxicity of the agent in old age, either from earlier exposures or from exposures in aged animals.

The database for oral exposure is much more limited than the database for inhalation exposure, with acute accidental ingestion data in humans at a single, high-dose level resulting in narcosis and chronic drinking water exposure (no dose information) associated with an increase in birth defects. The animal data are likewise very limited, with a single-dose acute toxicity study in rats in which clinical signs of morbidity and mortality were evaluated and subchronic (90-day) and chronic toxicity data in rats and mice indicating effects similar to those seen with inhalation exposure. Prenatal developmental toxicity data were available in rats following exposure on GDs 6–19, and an evaluation of adult neurotoxicity was conducted in mice following developmental exposure on PNDs 10–17. No other developmental toxicity data, and no information on reproductive toxicity or adult neurotoxicity were available. No studies were conducted that considered issues related to the toxicity of the agent in old age, either from earlier exposures or from exposures in aged animals.

Exposure-Response Array

In addition to displaying the data in tabular form (Table B-1), an exposure-response array can be a useful way of visually displaying the data (see Figures B-1 through B-4) to show what data are available for each duration of exposure. The data shown in the graphs are the human equivalent concentrations (HECs) based on the dosimetric adjustments discussed in Chapter 4 of

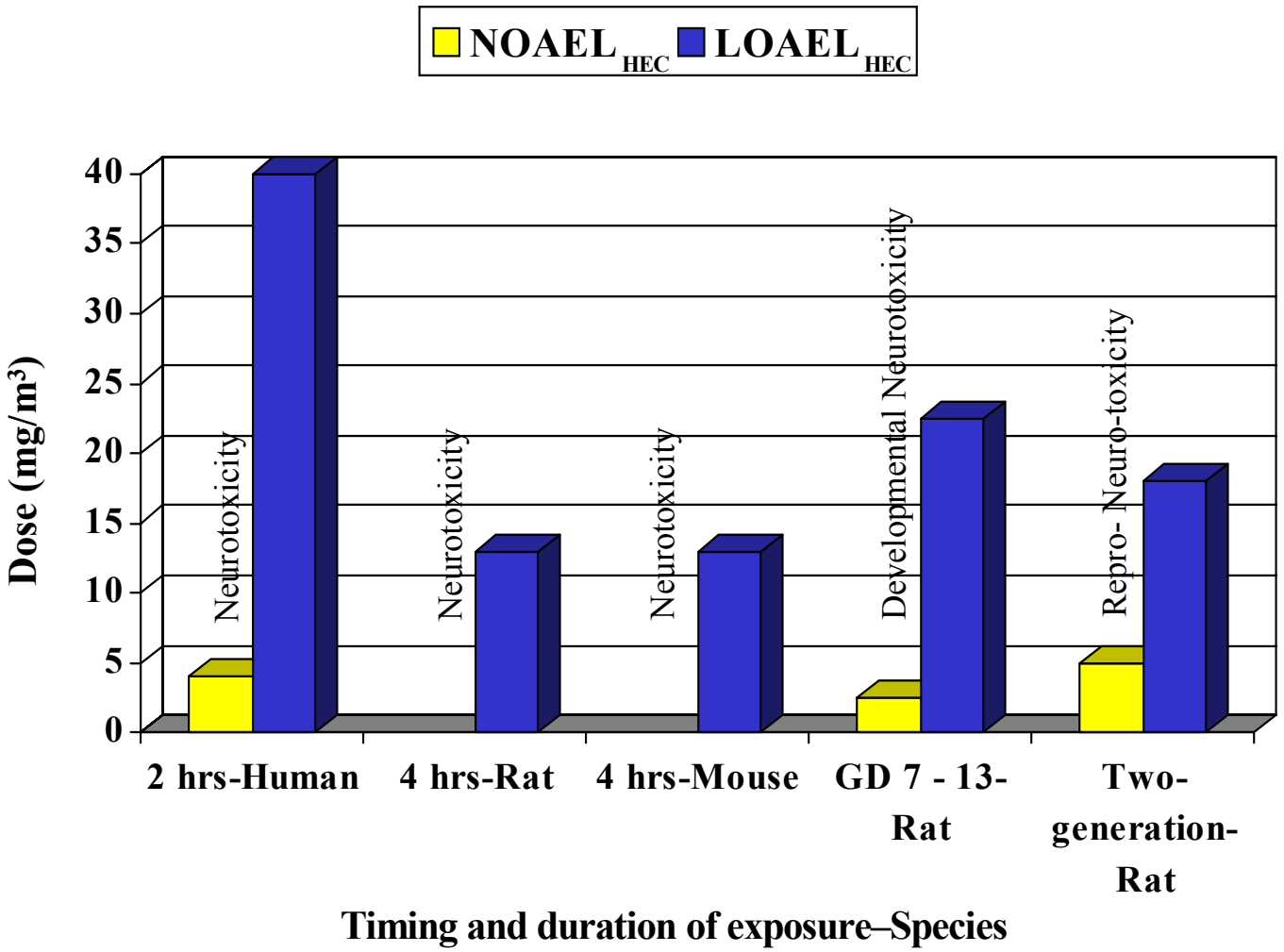


Figure B-1. Exposure-response array of data considered for the Inhalate acute reference value.

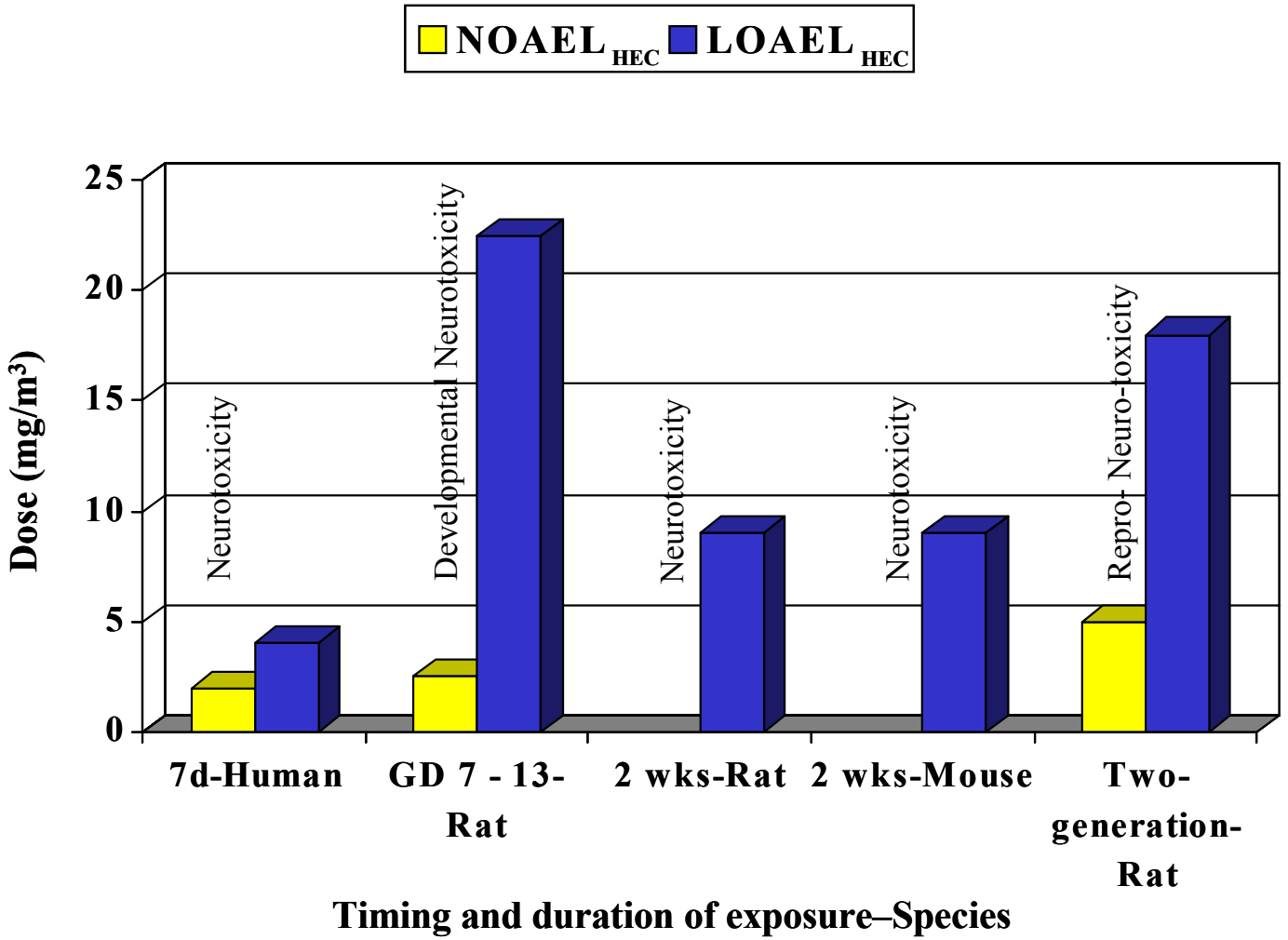


Figure B-2. Exposure-response array of data considered for the Inhalate short-term reference value.

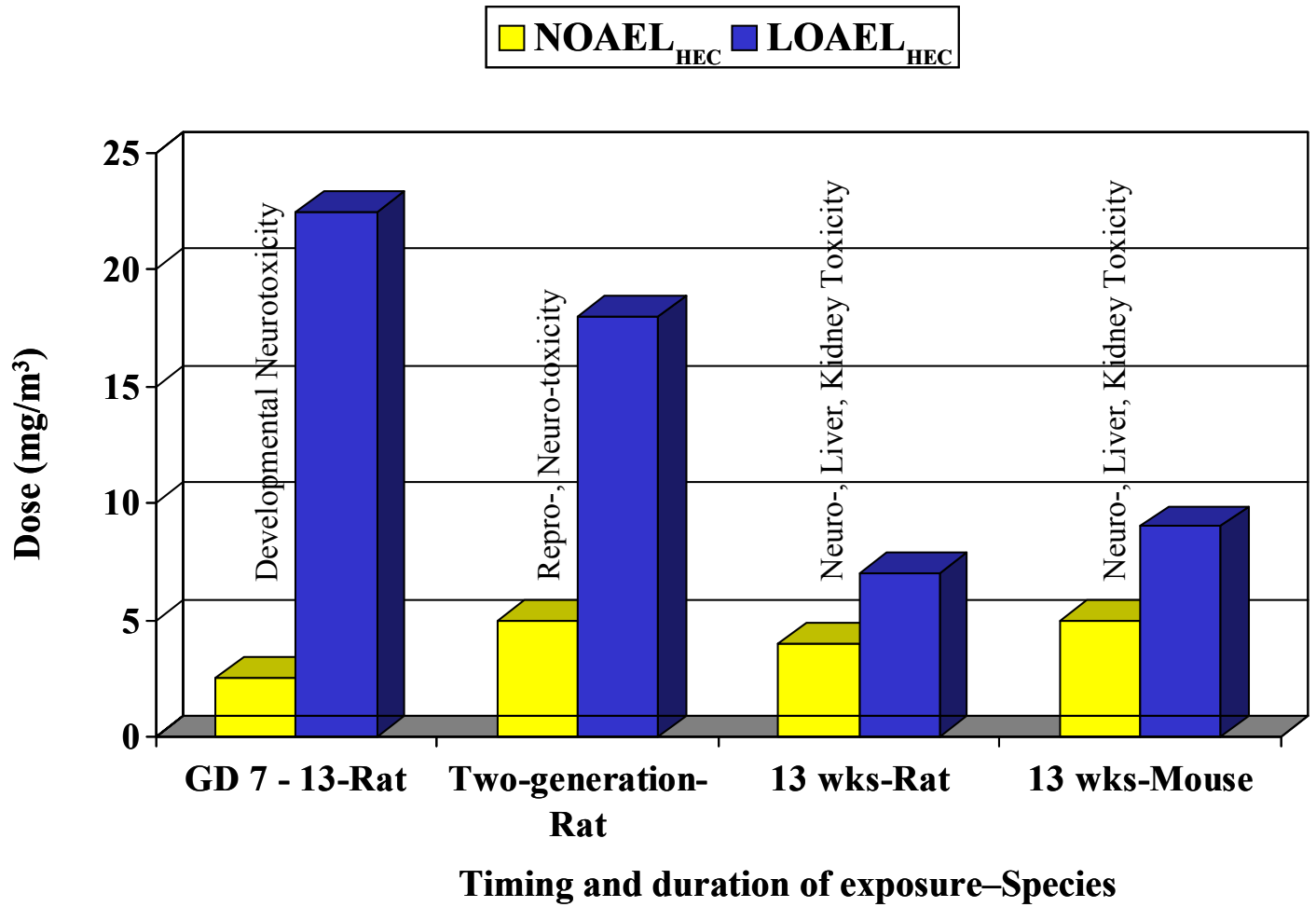


Figure B-3. Exposure-response array of data considered for the Inhalate longer-term reference value.

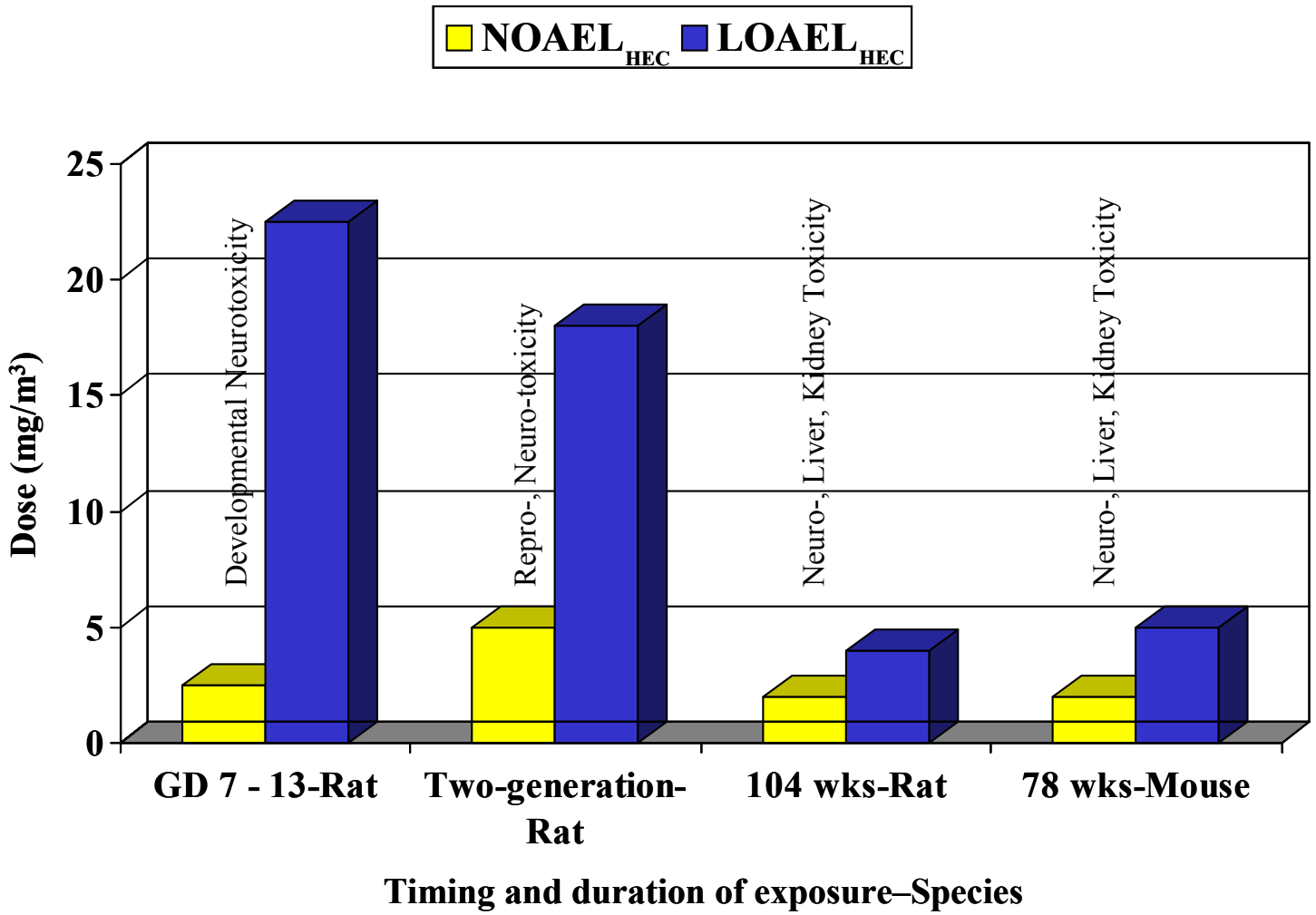


Figure B-4. Exposure-response array of data considered for the Inhalate chronic reference value.

this report, including dosimetric adjustment of the developmental toxicity data in the same manner as for other types of toxicity data, as recommended in U.S. EPA (1994) and Chapter 4.

No toxicokinetic model is available for Inhalate, so adjustments to the applied concentrations for dose adjustments and calculation of the HEC in this case study were done using the default procedures discussed in U.S. EPA (1994) and Chapter 4 of this document. As the effects observed are systemic, with no indications of portal-of-entry effects, the specific default procedures are based on the vapors of Inhalate being a category 3 gas. The principal parameter for interspecies extrapolation of category 3 gases, the blood:gas (air) partition coefficient ($H_{b/g}$), is unknown for both humans and animals and assumed to be 1. Therefore, the dosimetric adjustment factor (DAF) (see Chapter 4, Section 4.4.3) applied to the duration adjusted concentrations is 1, such that the duration adjusted values then become the HECs.

Duration adjustment was accomplished by factoring the exposure concentration (in mg/m^3) by 6/24 (for hours of exposure) and, where applicable, by 5/7 for the number of days per week exposed. As noted in Chapter 4, Section 4.4.2.1, duration adjustments were not made to acute (i.e., less than 24-hr) single exposures.

Uncertainty Factors (UFs)

An interspecies UF (A in Table B-2) of $10^{1/2}$ was applied in all cases to animal studies because the dosimetric adjustment procedures applied to Inhalate are considered to address the toxicokinetics portion of this UF, leaving the rest of the UF to cover interspecies toxicodynamics (U.S. EPA, 1994). An intraspecies UF (H in Table B-2) of 10 was applied in all cases because the data did not allow the estimation of within-human variability and the most sensitive life stage and/or susceptible subpopulation was not clearly identified in the database.

A LOAEL-to-NOAEL UF (L in Table B-2) was applied to the human data for neurotoxicity for the longer-term and chronic reference values because only a LOAEL was identified in the data for those two durations of exposure. A subchronic-to-chronic (duration) UF (S in Table B-2) was not applied, as there were data available for the appropriate duration of exposure in each case. For the longer-term and chronic reference values, a duration UF was not applied to the data from the developmental toxicity study, as this is not the usual practice when considering these data for longer-term and chronic exposures. However, as noted in the report (Chapter 4, Section 4.4.5.6), the application of a UF analogous to the subchronic-to-chronic-duration UF should be explored, as there may be situations in which data are available and applicable but they are from studies in which the dosing period is considerably shorter than that for the reference value being derived.

Table B-2. Derivation of reference values for Inhalate—inhalation exposure

Reference value duration	Exposure duration	HEC (mg/m^3)	Species	Type of effect[a]	Uncertainty factors[b]						Reference value (ppm)[c]	
					Total	A	H	L	S	D	Sample	Final
Acute	2 hrs	4	Human	NT	30	1	10	1	1	3	0.13	0.03
	GDs 7–13	2.5	Rat	DNT	100	3	10	1	1		0.03	
	4 hrs	13L[d]	Rat/ Mouse	NT	1000	3	10	10	1		0.01	
	13–16 wks	5	Rat	RT	100	3	10	1	1		0.05	
Short-term	7 days	2	Human	NT	30	1	10	1	1	3	0.07	0.03
	14 days	9L[d]	Rat/ Mouse	NT	1000	3	10	10	1		0.01	
	GDs 7–13	2.5	Rat	DNT	100	3	10	1	1		0.03	
	13–16 wks	5	Rat	RT	100	3	10	1	1		0.05	
Longer-term	>15 yrs	20L[d]	Human	NT	300	1	10	10	1	3	0.07	0.03
	13 wks	4	Rat	NT	100	3	10	1	1		0.04	
	13 wks	5	Mouse	NT	100	3	10	1	1		0.05	
	GDs 7–13	2.5	Rat	DNT	100	3	10	1	1[e]		0.03	
	13–16 wks	5	Rat	RT	100	3	10	1	1		0.05	
	13 wks	4	Rat	LT	100	3	10	1	1		0.04	
	13 wks	5	Mouse	LT	100	3	10	1	1		0.05	
	13 wks	4	Rat	KT	100	3	10	1	1		0.04	
	13 wks	5	Mouse	KT	100	3	10	1	1		0.05	
Chronic	>15 yrs	20L[d]	Human	NT	300	1	10	10	1	3	0.07	0.02
	104 wks	2	Rat	NT	100	3	10	1	1		0.02	
	78 wks	2	Mouse	NT	100	3	10	1	1		0.02	
	GDs 7–13	2.5	Rat	DNT	100	3	10	1	1[e]		0.03	
	13–16 wks	5	Rat	RT	100	3	10	1	1		0.05	
	104 wks	2	Rat	LT	100	3	10	1	1		0.02	
	78 wks	2	Mouse	LT	100	3	10	1	1		0.02	
	104 wks	2	Rat	KT	100	3	10	1	1		0.02	
	78 wks	2	Mouse	KT	100	3	10	1	1		0.02	

[a] NT = neurotoxicity; DNT = developmental neurotoxicity; RT = reproductive toxicity; LT = liver toxicity; KT = kidney toxicity

[b] A = animal-to-human (interspecies); H = inter-individual (intraspecies); L = LOAEL-to-NOAEL; S = subchronic-to-chronic duration; D = database deficiency

[c] Sample = reference value based on that particular endpoint, species, duration; Final = reference value for the entire database for a particular duration of exposure.

[d] L indicates that this value is the HEC based on the LOAEL.

[e] A duration UF was not applied to the data from the developmental neurotoxicity study for either the longer-term or chronic reference value; however, the adjustment should be considered when extrapolating from shorter to longer durations of exposure.

GD = gestation day

A database UF of $10^{1/2}$ was applied in all cases because there were data indicating that one of the main target organs for Inhalate was the nervous system, that there was neurotoxicity in adults (humans and animals), and that there were some data on developmental neurotoxicity, but with exposure limited to only a portion of the developmental period. In addition, there was no study in which pregnancy outcomes were evaluated (i.e., fetal survival, growth, and structural development) except for a possible association with spontaneous abortion and cardiac anomalies in occupationally exposed workers. These data gaps were not considered likely to reduce the NOAEL by more than a factor of 3 because there were some data on developmental exposures from a two-generation reproductive toxicity study. There were no reports of effects on fertility or reproduction per se, except for a possible association with menstrual disturbances in occupationally exposed workers.

Acute Inhalation Exposure

Results of available studies indicate that acute inhalation exposure to Inhalate can result in neurotoxic effects in human adults with a $LOAEL_{HEC}$ of 40 mg/m^3 and a $NOAEL_{HEC}$ of 4 mg/m^3 in a 2-hour exposure. Animal studies also show that Inhalate has the potential to cause neurotoxicity in adults with a $LOAEL_{HEC}$ of 13 mg/m^3 in a 4-hr exposure, and developmental neurotoxicity and other reproductive effects with $LOAEL_{HEC}$s of 22.5 mg/m^3 and 18 mg/m^3 and $NOAEL_{HEC}$s of 2.5 mg/m^3 and 5 mg/m^3, respectively.

Default UFs of $10^{1/2}$ (animal-to-human extrapolation), 10 (inter-individual differences), and $10^{1/2}$ (database deficiencies: no adequate prenatal developmental toxicity studies in two species, no adequate developmental neurotoxicity study) were applied to all the $NOAEL_{HEC}$s to derive sample reference values. In addition, a UF of 10 (LOAEL to NOAEL) was applied to the 4-hr rat and mouse $LOAEL_{HEC}$s. Human and animal studies indicate that the nervous system is vulnerable to Inhalate exposure. Although the sample reference values for the 4-hr adult rat and mouse exposures were lower than that based on the developmental neurotoxicity study, the values were within a similar range and the sample reference value for developmental neurotoxicity had less overall uncertainty. Therefore, the resultant reference value chosen for acute inhalation exposure is 0.03 mg/m^3 (Table B-2).

Short-term Inhalation Exposure

The reference value for short-term inhalation exposure is based on the human data $NOAEL_{HEC}$s of 2 mg/m^3 as well as the animal developmental neurotoxicity NOAEL of 2.5 mg/m^3 and

reproductive toxicity NOAEL of 5 mg/m^3 ($LOAEL_{HEC}$s of 4 mg/m^3, 22.5 mg/m^3, and 18 mg/m^3, respectively). In addition, the 14 day exposures of rats and mice resulted in a $LOAEL_{HEC}$ of 9 mg/m^3. For the human $NOAEL_{HEC}$ of 2 mg/m^3, applying a 10-fold default UF for intraspecies uncertainty and variability and a $10^{1/2}$-fold UF for database deficiencies results in a sample reference value for short-term inhalation exposure of 0.07 mg/m^3. Sample reference values based on animal data include default factors of $10^{1/2}$ (interspecies), 10 (intraspecies) and $10^{1/2}$ (database deficiencies) applied to the $NOAEL_{HEC}$s for developmental neurotoxicity and reproductive toxicity (2.5 and 5 mg/m^3, respectively) and result in sample reference values of 0.03 and 0.05 mg/m^3. An additional factor of 10 (LOAEL to NOAEL) applied to the adult neurotoxicity data in rats and mice results in a sample reference value of 0.01 mg/m^3. Given the close range of values and the lower overall uncertainty in the sample reference value for developmental neurotoxicity than that for adult neurotoxicity, the final reference value of 0.03 mg/m^3 is chosen (Table B-2).

Longer-term Inhalation Exposure

Subchronic and chronic inhalation exposure to Inhalate can result in multiple health effects. Available studies demonstrate neurotoxicity in adult humans. However, dose-response information is not available, and the presumed LOAEL (20 mg/m^3) for neurotoxicity in humans is somewhat higher than the HECs for other health endpoints (developmental, reproductive, liver, and renal effects) observed in animal studies, where the $LOAEL_{HEC}$s range from 7 mg/m^3 to 22.5 mg/m^3 and the $NOAEL_{HEC}$s range from 2.5 mg/m^3 to 5 mg/m^3. Dose-response data for these health endpoints in animal studies can be used as the basis for deriving a longer-term inhalation reference value for Inhalate.

UFs of $10^{1/2}$ (interspecies), 10 (intraspecies), and $10^{1/2}$ (database deficiencies) were applied to $NOAEL_{HEC}$s for the various endpoints in deriving sample reference values. If an additional factor of 3 were applied to the rat developmental toxicity data to account for the marked difference in exposure duration in the study itself (7 days of exposure: GDs 7–13) versus the duration covered by this reference value (up to 10% of the life span), a longer-term sample reference value of 0.01 mg/m^3 would result. Without this additional factor, the sample reference value from the developmental toxicity study was still the lowest value (0.03 mg/m^3), although all values from the animal studies were in a similar range (0.03–0.05).

The final reference value chosen was 0.03 mg/m^3 to be protective of the developing individual as well as adults (Table B-2). Whether an additional factor should be applied to the developmental toxicity data or to other data of much shorter duration should be explored further.

Chronic Inhalation Exposure

For the chronic inhalation reference value, the $LOAEL_{HEC}$s range from 4 mg/m^3 to18 mg/m^3 and the $NOAEL_{HEC}$s range from 2 mg/m^3 to 5 mg/m^3. UFs of $10^{1/2}$ (interspecies), 10 (intraspecies), and $10^{1/2}$ (database deficiencies) applied to the chronic exposure $NOAEL_{HEC}$s for neurotoxicity and liver, kidney, and reproductive toxicity data result in sample reference values of 0.02–0.05 mg/m^3 (Table B-2). Applying these UFs to the $NOAEL_{HEC}$ of 2.5 mg/m^3 for developmental toxicity yields a sample reference value of 0.03 mg/m^3, which falls within the range of chronic study-based values. If, contrary to current practice, an additional 3- or 10-fold UF for subchronic to chronic duration were applied to the developmental $NOAEL_{HEC}$, the sample reference value would be 0.01 or 0.003. As mentioned in Section 4.4.5.6, this issue may need further exploration. In this example, several endpoints result in a sample chronic inhalation reference value of 0.02 mg/m^3, the value chosen for the chronic inhalation reference value.

With regard to the liver effects of Inhalate, the $NOAEL_{HEC}$s and reference values for liver histopathology were the same as for the tumorigenic effects; thus, the reference value based on liver toxicity should be protective of the carcinogenic effects of Inhalate.

Overall Evaluation of Reference Values

The reference values for Inhalate were similar across all durations of exposure. This is because the same data were used as the basis for the acute, short-term, and longer-term reference value, that is, the effects on developmental neurotoxicity. Although there were human data appropriate for consideration for all four durations of exposure, the endpoints examined in these studies or reports were limited and were not indicative of effects on the developing nervous system. To be protective of developmental life stages, it was considered appropriate to base the reference values on the developmental neurotoxic effects, for which the sample reference value was slightly lower than for other endpoints. For the chronic reference value, a number of sample reference values, including that for the carcinogenic effects of Inhalate, were clustered in the same range, at 0.02 mg/m^3, slightly lower than that for developmental neurotoxicity. In this case, it was considered appropriate to use this lower reference value to be protective of all potential effects for lifetime exposures.

CASE STUDY 2: LUTEINATE

INTRODUCTION

Luteinate is a new pesticide that was developed for use as an herbicide. Environmental fate studies have shown that it will persist in soils and will therefore likely to move into ground and surface water. The general population may be exposed to Luteinate through consumption of food and drinking water.

Luteinate belongs to a class of pesticides for which the neuroendocrine mode of action is known. It was designed to be less potent than other members of the class. In order to ascertain its potency, a number of short-term studies were first conducted; they confirmed a similar mode of action and the fact that Luteinate was less potent. Following this, Luteinate was tested in more traditional toxicology studies to establish its long-term effects and dose-response relationships.

This case study provides an example of the usefulness of mode of action information in establishing the short- and long-term effects of Luteinate on relevant target organ systems at different life stages. Such information enables the development of a targeted robust data set for use in establishing reference values for various durations of exposure.

Postulated Mode of Action

Other members of this class of pesticides have been shown to act on the hypothalamic-pituitary-ovarian axis. These pesticides affect the hypothalamus, leading to a decreased secretion of hypothalamic norepinephrine (NE). Decreased NE levels result in decreased release of gonadotropin releasing hormone (GnRH) from the hypothalamus. GnRH is the hormone responsible for inducing the pituitary gland to release luteinizing hormone (LH). Thus, the decrease in GnRH leads to a suppression of the pituitary LH release. These compounds also decrease the neurotransmitter dopamine, which in turn leads to a decrease in pituitary prolactin.

Decreased LH and prolactin levels have the potential to impact several organ systems at different life stages. In humans, there are robust pulses of the LH surge in the fetus and prior to birth. The LH pulsatility continues with diminishing amplitude during the early months of postnatal life, and then LH secretion becomes barely detectable during much of the first decade of postnatal life. Around the age of 10, there is the reemergence or re-awaking of LH pulses while sleeping.

The natural progression from prepubertal to postpubertal status is dependent on the normal function of the hypothalamic-pituitary-gonadal axis. Likewise, many of the same hypothalamic mechanisms that control pituitary function and the pituitary hormones themselves

(especially LH and prolactin) play a key role in pubertal development. In the adult, ovulation is dependent on sufficient LH levels. Therefore, this class of pesticides impacts critical reproductive processes, including puberty, ovarian cyclicity, pregnancy, and lactation (milk quality/production). Given the role of NE and dopamine in the development of the CNS, this class of pesticides may also affect the developing CNS; however, relevant toxicology studies have not yet been conducted. In addition, suppression of prolactin during the early postnatal period in rodents may lead to prostate inflammation in male offspring.

This class of pesticides has also been shown to lead to mammary tumors in Sprague-Dawley (SD) rats but not in other strains of rats or mice. Mammary tumors result from the prolonged decrease in serum LH, which leads to a cessation of ovulation and eventually causes the ovarian follicles to continue to secret estradiol. In concert with prolactin, estrogen acts on the mammary gland and leads to the formation of mammary tumors. However, the induction of the tumors has been shown to be due to unique features of the reproductive aging process in Sprague-Dawley rats; because humans do not age in the same manner, this mode of action is unlikely to be operative in humans.

SUMMARY OF HEALTH EFFECTS INFORMATION

Studies with other pesticides in this chemical class have found negligible differences in response (with the exception of the induction of mammary tumors in SD rats) among various strains of rats and between rats and mice. Therefore, all studies on Luteinate were conducted on SD rats.

Luteinate exposure results in a spectrum of effects that are related to decreases in serum LH and decreases in prolactin. The toxicology studies are summarized below in the context of the specific effects associated with either decreases in LH or prolactin. In addition, the effect levels for specific toxicological endpoints are summarized in Table B-3. It is important to note that as the spectrum of effects are mechanistically related to decreases in serum LH or prolactin, there is great similarity in the doses that affect LH and prolactin levels and those that cause the related effects.

Absorption, Distribution, Metabolism, and Elimination

Studies in adult rats have shown that Luteinate is rapidly absorbed following oral exposure. It is not metabolized and is eliminated in the urine. No other information is available. The toxicokinetic profiles of other pesticides in this chemical class have not been extensively examined either.

Table B-3. Summary of endpoints and effect levels for rat studies of Luteinate

Response	Exposure period	Dose levels (mg/kg/day)	NOAEL/LOAEL (mg/kg/day)
Decreased LH	Pregnant females: GDs 1–8	 50, 100, 200	 100/200
	PNDs 22–41	25, 50, 100	50/100
	Adults: 3 days 28 days 6 months	 50, 100, 200 40, 80, 160 4, 8, 16	 100/200 40/80 4/8
	Two-generation reproduction study: 10 weeks - F0 13 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 5/10 5/10
Disruption of estrous cyclicity	PND 22-41	25, 50, 100	50/100
	Adults: 28 days 6 months	 40, 80, 160 4, 8, 16	 40/80 4/8
	Two-generation reproduction study: 10 weeks - F0 13 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 5/10 5/10
Altered pregnancy maintenance	Pregnant females: GDs 1–8 GDs 6–10 GDs 1–20	 50, 100, 200 100, 200, 400 50, 100, 200	 100/200 200/400 100/200
	Two-generation reproduction study: 11–13 weeks - F0 14–16 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 5/10 5/10
Delayed parturition	Pregnant females: GDs 6–10	 100, 200, 400	 100/200
	Two-generation reproduction study: 14 weeks - F0 17 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 5/10 5/10
Delayed vaginal opening	PNDs 22–41	25, 50, 100	50/100
	Two-generation reproduction study: 4–5 weeks - F1[a]	 5, 10, 25, 50	 25/50
Delayed preputial separation	PND 23–53	25, 50, 100	50/100
	Two-generation reproduction study: 6–7 weeks - F1[a]	 5, 10, 25, 50	 25/50

Table B-3. Summary of endpoints and effect levels for rat studies of Luteinate (continued)

Response	Exposure period	Dose levels (mg/kg/day)	NOAEL/LOAEL (mg/kg/day)
Attenuation of prolactin release	Lactating females: PNDs 1–4	 25, 50, 100	 25/50
	Adults: 3 days 28 days 6 months	 50, 100, 200 40, 80, 160 4, 8, 16	 100/200 40/80 4/8
	Two-generation reproduction study: 10 weeks - F0 13 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 5/10 5/10
Increased prostatitis in offspring	PNDs 1–4	25, 50, 100	25/50
	Adults: 28 days 6 months	 40, 80, 160 4, 8, 16	 >200/NA[b] >160/NA[b]
	Two-generation reproduction study: 16 weeks - F0 16 weeks - F1[a]	 5, 10, 25, 50 5, 10, 25, 50	 >50/NA[b] 5/10
Reduced weight of seminal vesicles and ventral prostate	PNDs 23–53	25, 50, 100	50/100
	Adults: 28 days 6 months	 40, 80, 160 4, 8, 16	 >200/NA[b] >160/NA[b]
	Two-generation reproduction study: 16 weeks - F1[a]	 5, 10, 25, 50	 5/10
Fetus: Delayed ossification Reduced fetal weight	 GDs 1–20 GDs 1–20	 50, 100, 200 50, 100, 200	 100/200 >200/NA[b]

[a] F1 exposures are indicated for the duration of postnatal exposure, but it is assumed that 3 weeks of prenatal exposure also occurred.

[b] When no effect on a particular endpoint was noted in a study, the NOAEL is indicated as > the highest dose, and the LOAEL as NA (not applicable).

Decreased LH Surge and Related Effects

Exposure to Luteinate resulted in a significant decrease in serum LH that was dependent on the dose and duration of exposure. In a 3-day gavage study of SD male and female rats exposed to levels of 0, 50, 100, and 200 mg/kg/day, there were significant decreases in serum LH at 200 mg/kg/day. Serum LH was significantly reduced at doses of 80 mg/kg/day and above in a 28-day study and at doses of 8 mg/kg/day and above in a 6-month study of SD rats. Levels in the latter study were similar to those observed in the F0 rats in a two-generation reproductive toxicity study in which significant decreases in serum LH were observed at doses of 10 mg/kg/day and higher following 10 weeks of exposure.

Serum LH was also measured in pregnant dams exposed to doses of 0, 50, 100, and 200 mg/kg/day on GDs 1–8 and was significantly reduced at 200 mg/kg/day. Exposure of weanling SD rats on PNDs 22–41 to doses of 0, 25, 50, and 100 mg/kg/day of Luteinate resulted in a significant decrease in serum LH at 100 mg/kg/day. In a two-generation reproductive toxicity study, serum LH was measured in the F1 generation prior to mating (13 weeks postnatal exposure) and was significantly reduced at doses of 10 mg/kg/day and higher.

Disruption of Estrous Cyclicity

In the adult female, ovulation is dependent on sufficient levels of LH. Because exposure to Luteinate suppresses LH, it would be anticipated to also disrupt the normal estrous cycle. Estrous cyclicity was abnormal in SD rats exposed to 80 mg/kg/day for 28 days and 8 mg/kg/day for 6 months. In the two-generation reproductive toxicity study, estrous cyclicity was abnormal in the F0 females exposed to 10 mg/kg/day for 10 weeks and in the F1 females exposed to 10 mg/kg/day for 13 weeks postnatally (as well as prenatally). In addition, exposure of weanlings to 100 mg/kg/day on PNDs 22–41 resulted in abnormal estrous cycles.

Pregnancy

Altered LH levels also impact the ability of the female to maintain pregnancy as well as the timing of parturition. Thus, in a variety of exposure scenarios, an increase in pre- and postimplantation loss was observed. Implantation was affected following exposure to 200 mg/kg/day on GDs 1–8 or 1–20 and after exposure to 400 mg/kg/day on GDs 6–10. Parturition was delayed following exposure to 200 mg/kg/day and above on GDs 6–10. In the two-generation reproductive toxicity study, there was an increase in pre- and postimplantation loss and a delay in parturition in the F0 and F1 females at doses of 10 mg/kg/day and higher.

Sexual Maturation

At the time of puberty, the CNS and pituitary respond to increased concentrations of estradiol in a positive feedback fashion culminating in the first LH surge. Thus, exposure to Luteinate would likely impact sexual maturation. Vaginal opening was delayed in female SD rats following exposure to 100 mg/kg/day on PNDs 22–41 and preputial separation was delayed in males exposed to 100 mg/kg/day on PNDs 23–53. In the two-generation reproductive toxicity study, vaginal opening was delayed in the F1 females exposed to 50 mg/kg/day and preputial separation was delayed in the F1 males exposed to 50 mg/kg/day.

Decreased Prolactin and Related Effects

Prolactin levels were significantly decreased in adult SD rats following exposure to 200 mg/kg/day of Luteinate for 3 days, to 80 mg/kg/day for 28 days, and to 8 mg/kg/day for 6 months. In a study in which lactating dams were exposed during PNDs 1–4, prolactin levels were reduced at doses of 50 mg/kg/day. In the two-generation reproductive toxicity study, prolactin levels were reduced in the F0 and F1 animals at doses of 10 mg/kg/day.

Prostatitis

As a consequence of the reduced prolactin levels, an increased incidence of prostatitis was observed in males following maternal exposure to 50 mg/kg/day during lactation days 1–4, and in the F1 males exposed to 10 mg/kg/day and higher. Prostatitis was not observed in males exposed during adulthood only (i.e., F0 males, males in 28-day and 6-month studies).

Organ Weights

In males exposed to doses of 100 mg/kg/day Luteinate during PNDs 23–53, there was a decrease in absolute and relative weights of the seminal vesicles and ventral prostate. This was also observed in the F1 males exposed to doses of 10 mg/kg/day or greater at the time of terminal sacrifice following mating. Histopathological examination revealed no lesions in either study. No organ weight changes or histopathological lesions were noted in any of the other studies.

Effects Unrelated to LH or Prolactin

A prenatal developmental toxicity study was conducted in which pregnant SD rats were exposed to doses of 0, 50, 100, or 200 mg/kg/day Luteinate on GDs 1–20. In addition to the increase in implantation loss noted above, there was an increase in delayed ossification at 200

mg/kg/day but no effect on fetal body weight. It is unlikely that this effect is related to LH levels. However, the mode of action is unknown.

SELECTION OF HEALTH ENDPOINTS AND DERIVATION OF REFERENCE VALUES

Narrative Description of the Extent of the Database

The database for oral exposure is quite robust and is adequate for deriving reference values. Because the mode of action was known for other pesticides in this chemical class, it was possible to study known targets during relevant life stages. Luteinate was shown to interfere with the pituitary-hypothalamic axis, resulting in a decrease in serum LH and prolactin. The decrease was shown to be dependent on dose and duration of exposure. Normal LH levels are known to be required for ovulation, maintenance of pregnancy, timing of parturition, and sexual maturation. Similarly, decreases in prolactin can lead to prostatitis and effects on the male reproductive organs. Each of these events were examined in several short-term studies as well as in traditional prenatal developmental toxicity studies and a two-generation reproductive toxicity study.

An acute toxicity study of Luteinate was not conducted. However, information was available on serum LH and prolactin levels from other short-term exposures that was informative for acute exposures. The database as a whole demonstrates that there is a clear dose-duration relationship for Luteinate on serum LH levels. Serum LH levels were reduced following a 3-day exposure in adult rats and in pregnant rats on GDs 1–8 at doses of 200 mg/kg/day; longer exposures to weanlings on PNDs 22–41 or to adults for 28 days required a dose of 100 or 80 mg/kg/day, respectively. The effective dose of 10 mg/kg/day was still lower following 10 to 13 weeks exposure. Therefore, it is unlikely that serum LH levels and related effects would occur at doses less than 200 mg/kg/day following an acute exposure, but higher doses may actually be necessary.

The situation is less clear for the decrease in prolactin and related effects. Prolactin levels were reduced following a 3-day exposure to adult rats at 200 mg/kg/day and following exposure to lactating dams at 50 mg/kg/day. A 28-day exposure to adult rats resulted in decreased prolactin levels at 80 mg/kg/day, whereas 10 mg/kg/day was an effective dose at 10 to 13 weeks of exposure. Thus, although there is a clear dose-duration relationship between the 4- and 10-week periods, the relationship is less clear for durations of less than 28 days. However, it is unlikely that serum prolactin levels and related effects would have NOAELs less than 25 mg/kg/day following an acute exposure.

Knowledge about the effect of Luteinate on serum LH and prolactin levels from these other studies was considered in conjunction with the prenatal developmental toxicity study for deriving the acute reference value. It is assumed that effects resulting from developmental exposures (both prenatal and postnatal) may be the result of a single exposure. Therefore, the fetal and offspring effects resulting from exposures during gestation and postnatally to the time of sexual maturation were also considered for the acute reference value. Data from the two-generation study were not considered because there were always shorter-term studies showing effects at higher doses. Taken as a whole, the lack of an acute toxicity study was not considered to be a major data gap for Luteinate.

A chronic study of Luteinate was not conducted. However, serum LH was decreased at 10 mg/kg/day in the F0 and F1 animals in the two-generation reproductive toxicity study and at 8 mg/kg/day in the 6-month study. Because there is essentially no change in the effective dose levels following 13 weeks of exposure and 6 months of exposure, it is unlikely that a longer exposure period would substantially lower the effective dose. Thus, the lack of a chronic study was not considered a major database deficiency. Although there is no knowledge of the effects of continued LH and prolactin suppression on reproductive aging, this is probably a qualitative, rather than a quantitative data gap. It is known from chronic bioassays of other pesticides in this chemical class that lifetime exposure results in mammary tumors in SD rats, but this mode of action is unlikely to be operative in humans due to differences in the aging process. Thus, this was not considered to be a data gap for Luteinate.

A developmental neurotoxicity study was not conducted. Given that Luteinate interferes with two neuroreceptors, norepinephrine and dopamine, the potential for developmental neurotoxicity exists. Thus, the lack of knowledge regarding the developing nervous system is considered to be a data gap.

Exposure-Response Array

In addition to the display in tabular form (Table B-3), the data considered useful for deriving each reference value were displayed in an exposure-response array. The arrays for acute, short-term, longer-term, and chronic exposures are shown in Figures B-5 through B-8, respectively. Although data on delayed ossification were considered in deriving the longer-term and chronic reference values, the data were not included in Figures B-7 and B-8 because the doses were substantially higher than those for all other endpoints, and inclusion would have altered the y-axis substantially. Figure B-9 is a composite of the NOAELs for each endpoint

with each exposure duration. In addition, the NOAELs for each relevant endpoint, the uncertainty factors (UFs), and the sample and final reference values are shown in Table B-4.

Uncertainty Factors

An interspecies UF (A in Table B-4) of 10 was applied in all cases because the data were insufficient to characterize toxicokinetic and toxicodynamic differences between rodents and humans. An intraspecies UF (H in Table B-4) of 10 was applied in all cases because the data did not allow the estimation of human variability. A UF for LOAEL to NOAEL (L in Table B-4) was not applied because the NOAEL was known for each exposure duration. A subchronic-to-chronic UF (S in Table B-4) was not applied because there were data available for each exposure duration. A database UF of $10^{1/2}$ was applied in all cases because there are concerns for the potential developmental neurotoxicity of Luteinate and there were no available data.

A database UF of $10^{1/2}$ (rather than a UF of 10) was applied on basis of the observation that for other pesticides in this chemical class, the doses that result in a decrease in norephinephrine and dopamine are very similar to the doses that lead to decreases in serum LH. Therefore, it is likely that any developmental neurotoxicity effects would be observed at similar doses. Because the effects of Luteinate on serum LH levels have been well characterized, additional developmental neurotoxicity information is unlikely to dramatically affect the reference values. However, a UF of $10^{1/2}$ was retained to reflect underlying uncertainties of this data gap.

Acute Exposure

For the derivation of the acute reference value, a standard prenatal developmental toxicity study provided information on the effects of Luteinate on fetal ossification in the absence of an effect on fetal weight. There was no information on serum LH or prolactin and related effects following an acute exposure. As described above in the discussion of the extent of the database, it is unlikely that the NOAEL for serum LH levels and related effects would be less than 100 mg/kg/day following an acute exposure, and it could be higher. Similarly, it is unlikely that the NOAEL for serum prolactin levels and related effects would be less than 25 mg/kg/day following an acute exposure. The effects on sexual maturation occurred at a dose of 100 mg/kg/day (NOAEL of 50 mg/kg/day) and effects on ossification occurred at a dose of 200 mg/kg/day (NOAEL of 100 mg/kg/day).

Thus, derivation of the acute reference value based on these endpoints would also be protective of effects on serum LH and related effects, but may not cover effects on prolactin and

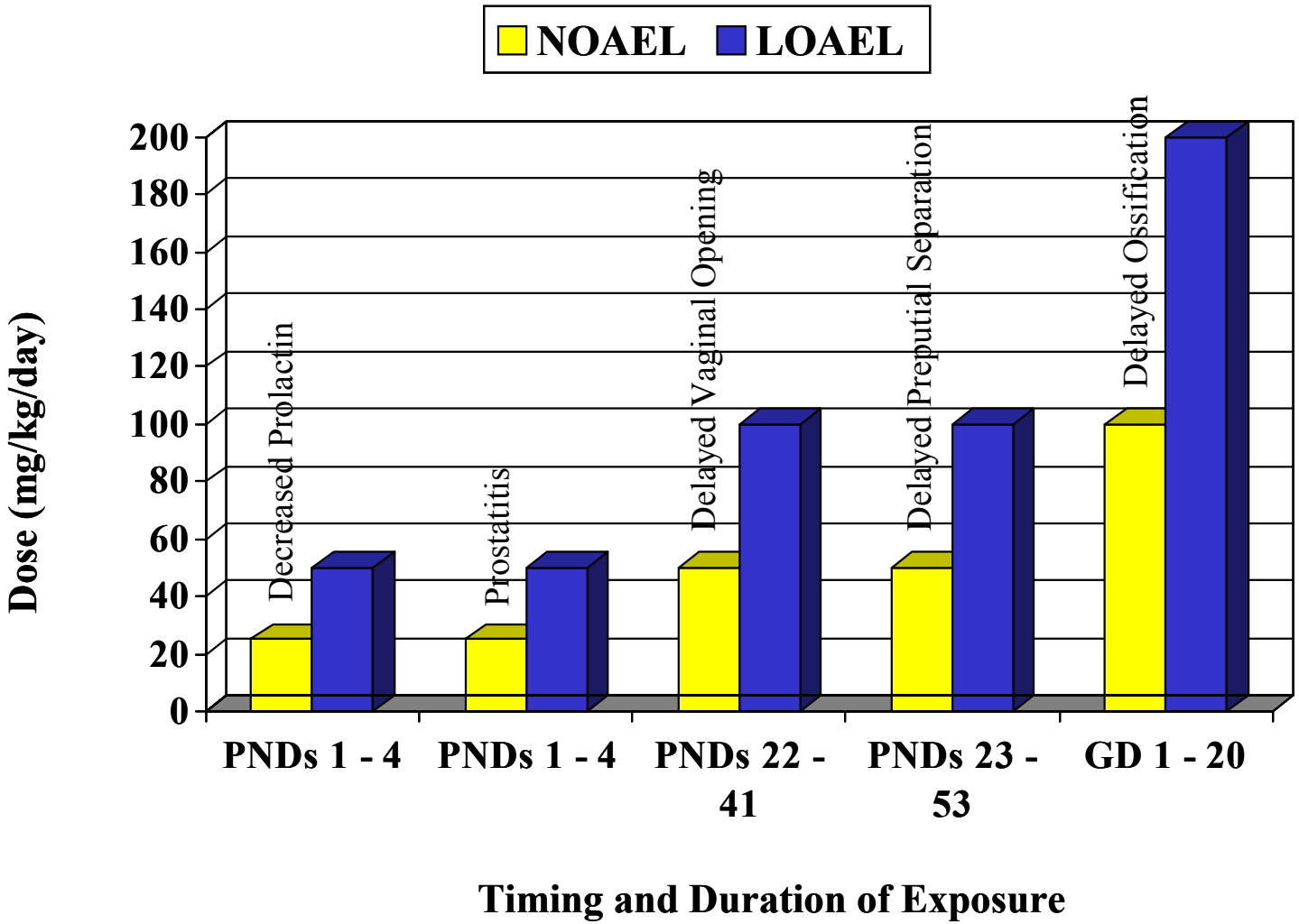


Figure B-5. Exposure-response array of data considered for the Luteinate acute reference value.

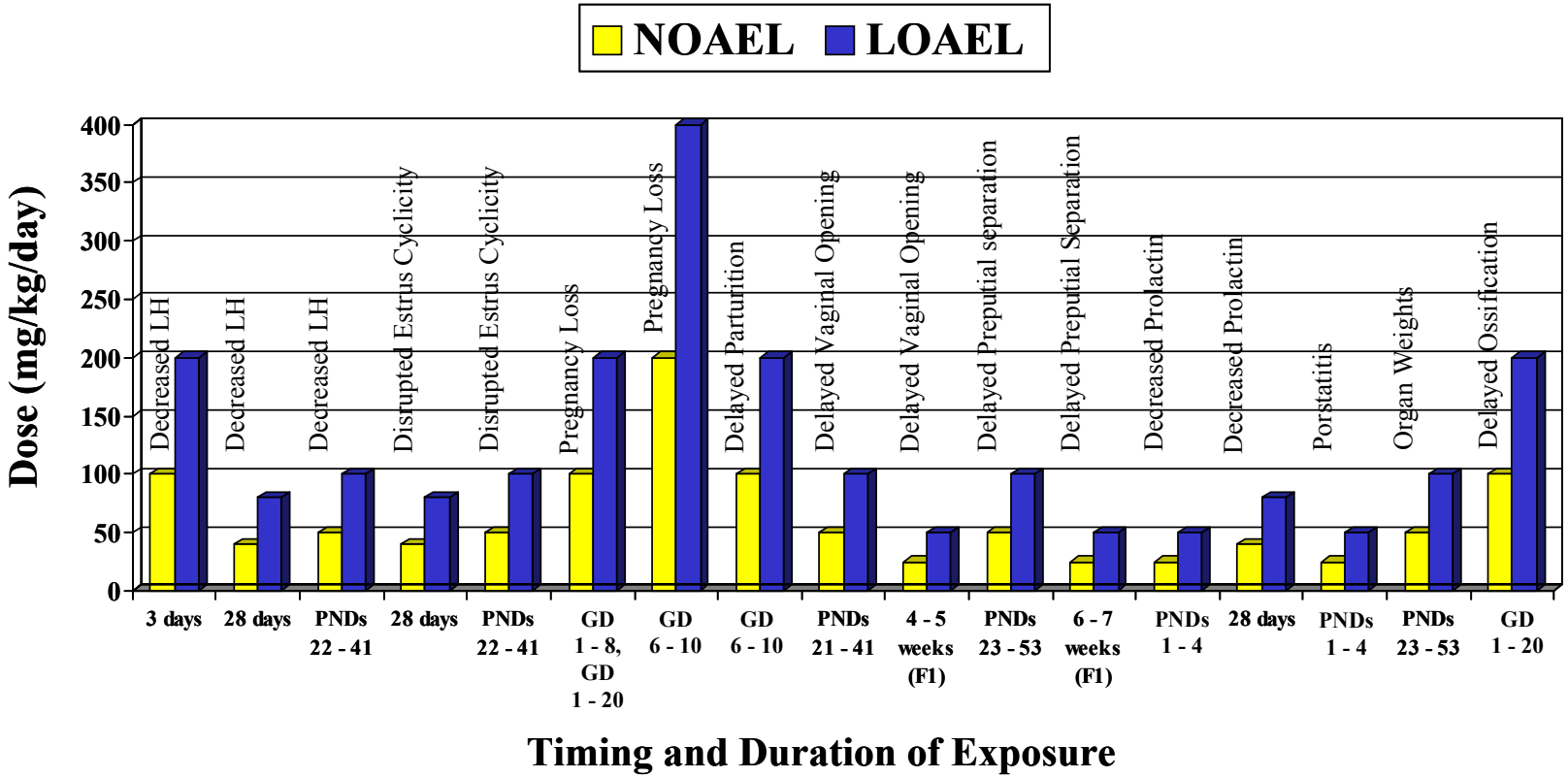


Figure B-6. Exposure-response array of data considered for the Luteinate short-term reference value.

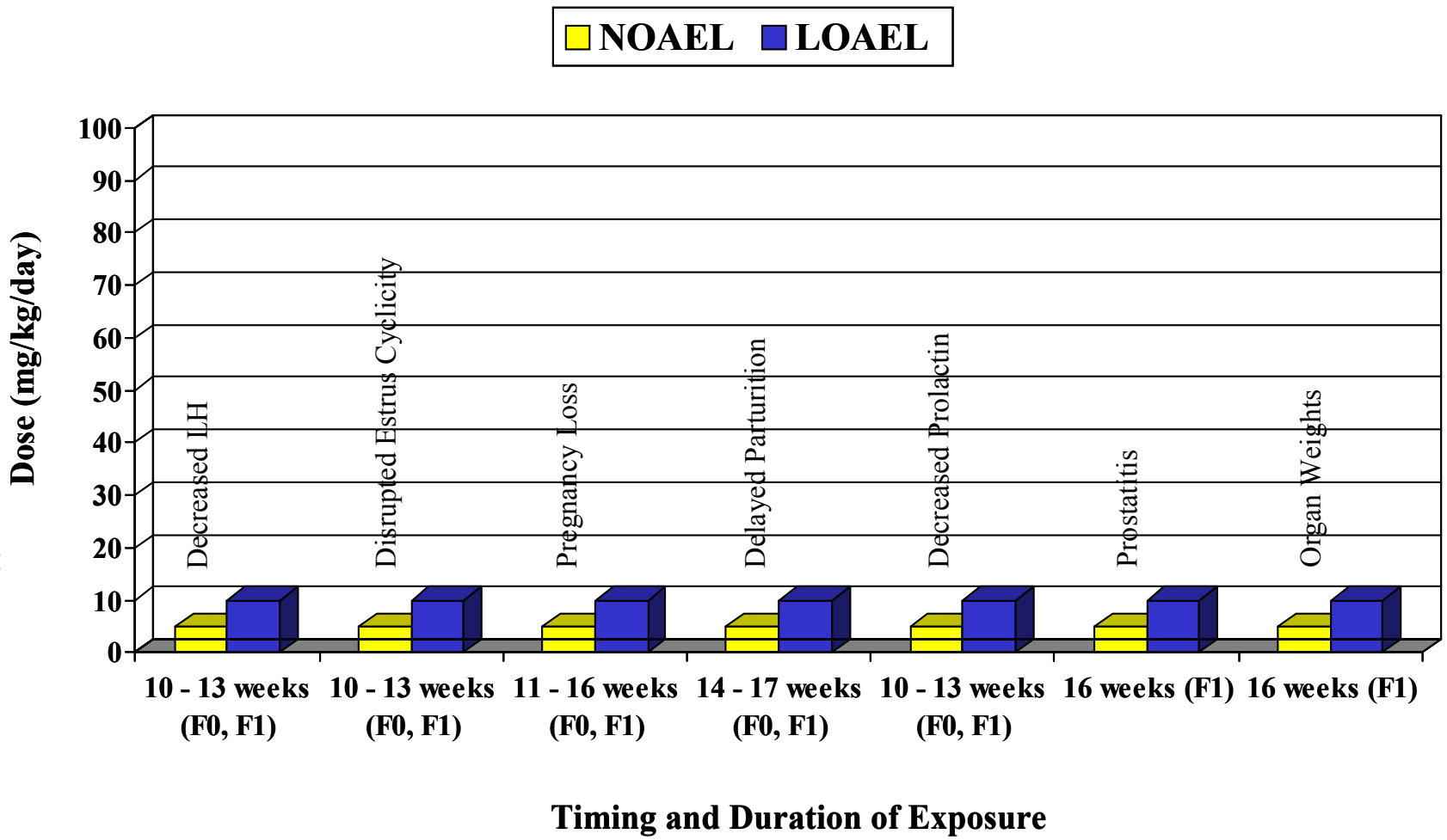


Figure B-7. Exposure-response array of data considered for the Luteinate longer-term reference value.

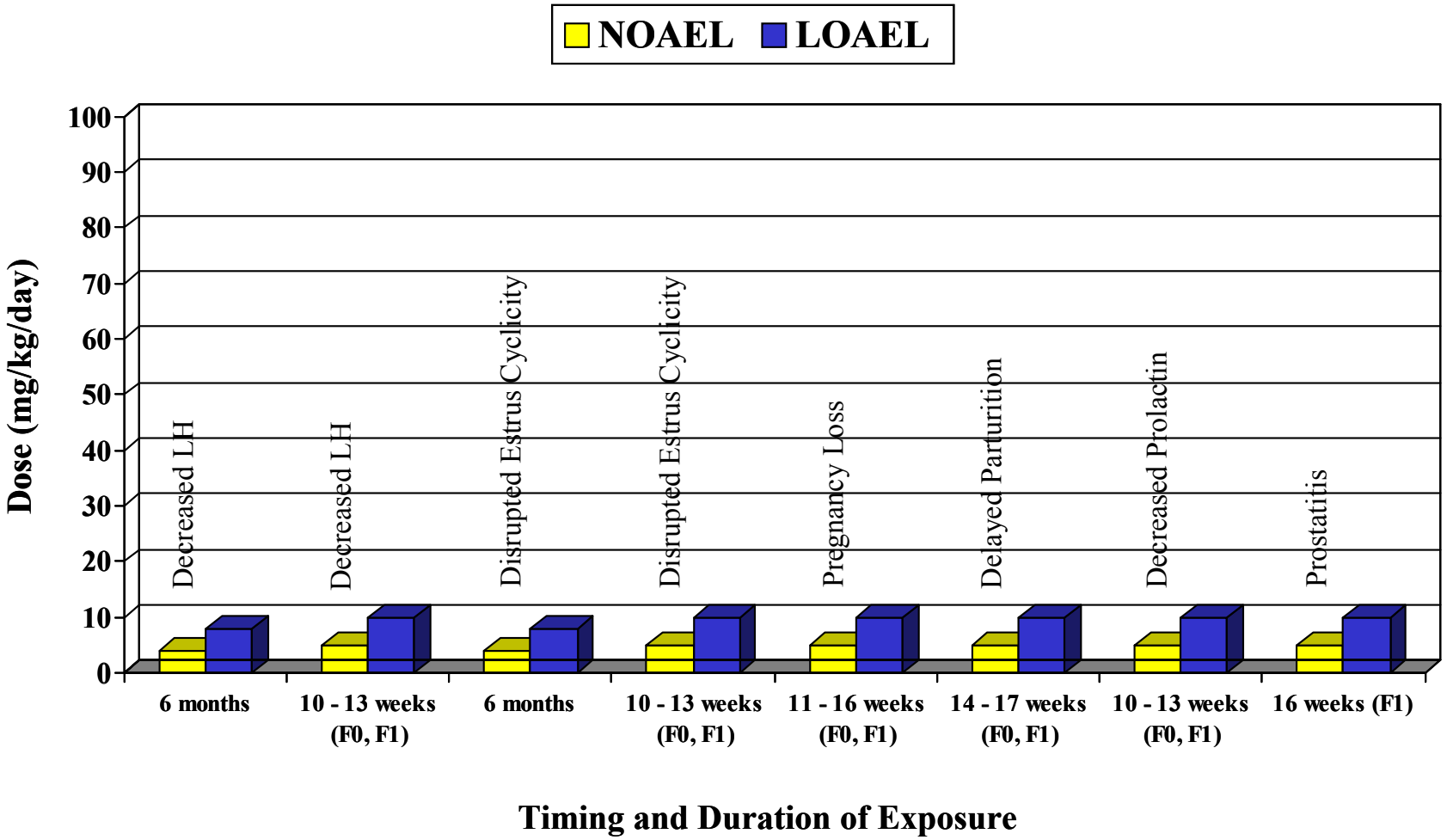


Figure B-8. Exposure-response array of data considered for the Luteinate chronic reference value.

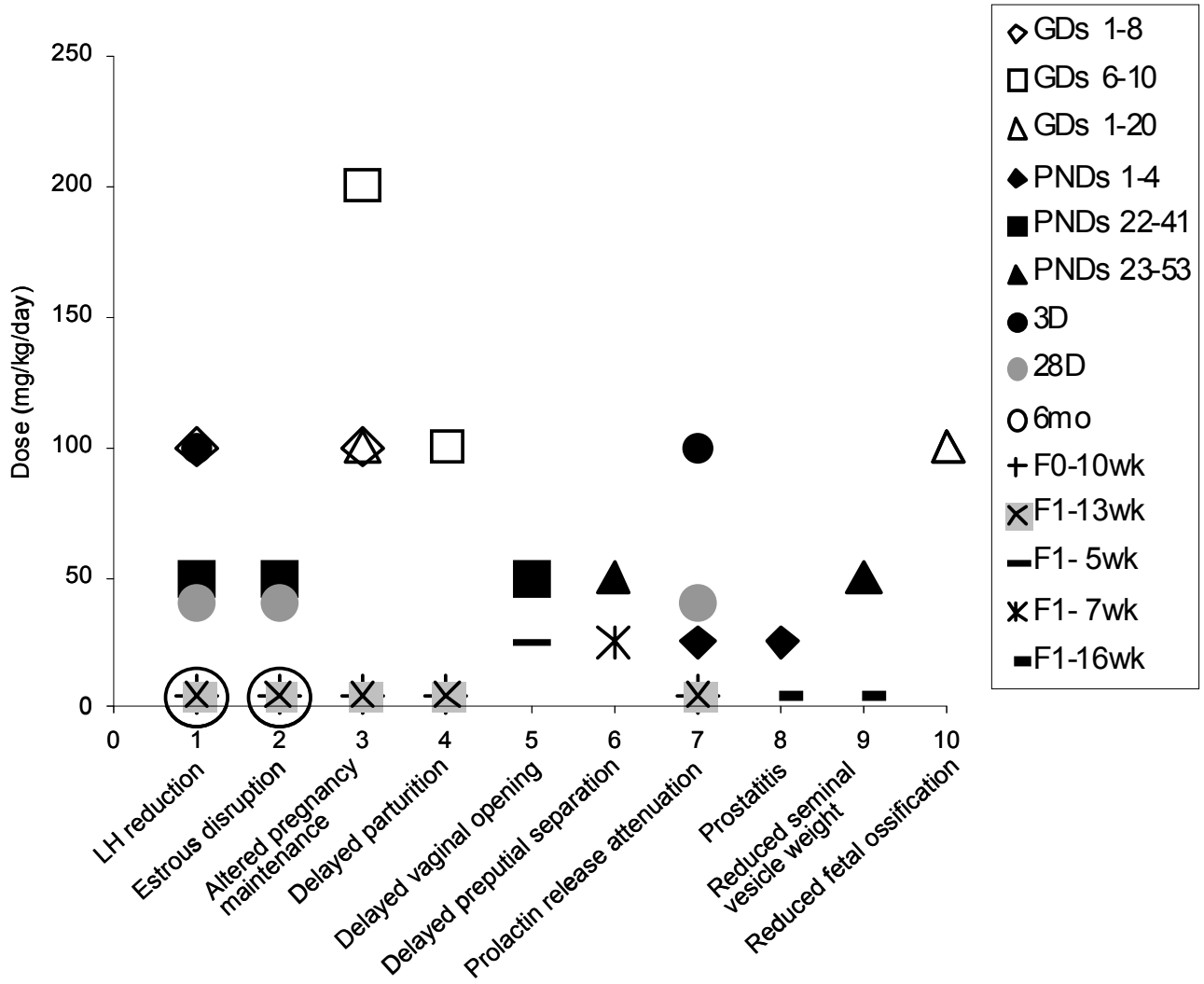


Figure B-9. NOAELs for Luteinate with timing/duration of exposure.

Table B-4. Summary of reference values for Luteinate

Reference value duration	Exposure duration	NOAEL (mg/kg/day)	Type of effect	Uncertainty factors[a]						Reference value (mg/kg/day)	
				A	H	L	S	D	Total	Sample	Final
Acute	PNDs 1–4	25	Prolactin	10	10	1	1	3	300	0.08	0.08
	PND 1–4	25	Prostatitis	10	10	1	1		300	0.08	
	PNDs 22–41	50	Sexual Maturation	10	10	1	1		300	0.17	
	PNDs 23–53	50	Sexual Maturation	10	10	1	1		300	0.17	
	GDs 1–20	100	Fetal ossification	10	10	1	1		300	0.33	
Short-term	3-day adult	100	LH	10	10	1	1	3	300	0.33	0.08
	28-day adult	40	LH	10	10	1	1		300	0.13	
	PNDs 22–41	50	LH	10	10	1	1		300	0.17	
	GDs 1–8	100	LH	10	10	1	1		300	0.33	
	28-day adult	40	Estrus	10	10	1	1		300	0.13	
	PNDs 22–41	50	Estrus	10	10	1	1		300	0.17	
	GDs 1–8	100	Pregnancy Maintenance	10	10	1	1		300	0.33	
	GDs 1–20	100	Pregnancy Maintenance	10	10	1	1		300	0.33	
	GDs 6–10	200	Pregnancy Maintenance	10	10	1	1		300	0.67	
	GDs 6–10	100	Parturition	10	10	1	1		300	0.33	
	PNDs 22–41	50	Sexual Maturation	10	10	1	1		300	0.17	
	PNDs 23–53	50	Sexual Maturation	10	10	1	1		300	0.17	

Table B-4. Summary of reference values for Luteinate (continued)

Reference value duration	Exposure duration	NOAEL (mg/kg/day)	Type of effect	Uncertainty factors[a]						Reference value (mg/kg/day)	
				A	H	L	S	D	Total	Sample	Final
Short-term	4–7 weeks	25	Sexual Maturation	10	10	1	1	3	300	0.08	0.08
	PNDs 1–4	25	Prolactin	10	10	1	1		300	0.08	
	3-day adult	100	Prolactin	10	10	1	1		300	0.33	
	28-day adult	40	Prolactin	10	10	1	1		300	0.13	
	PNDs 1–4	25	Prostatitis	10	10	1	1		300	0.08	
	PNDs 23–53	25	Organ Wt	10	10	1	1		300	0.08	
	GDs 1–20	100	Fetal ossification	10	10	1	1		300	0.33	
Longer-term	10–13 weeks	5	LH	10	10	1	1	3	300	0.02	0.02
	10–13 weeks	5	Estrus	10	10	1	1		300	0.02	
	11–16 weeks	5	Pregnancy Maintenance	10	10	1	1		300	0.02	
	14–17 weeks	5	Parturition	10	10	1	1		300	0.02	
	4–7 weeks	25	Sexual Maturation	10	10	1	1		300	0.08	
	10–13 weeks	5	Prolactin	10	10	1	1		300	0.02	
	16 weeks	5	Prostatitis	10	10	1	1		300	0.02	
	16 weeks	5	Organ Wt	10	10	1	1		300	0.02	
	GDs 1–20	100	Fetal ossification	10	10	1	1		300	0.33	

Table B-4. Summary of reference values for Luteinate (continued)

Reference value duration	Exposure duration	NOAEL (mg/kg/day)	Type of effect	Uncertainty factors[a]						Reference value (mg/kg/day)	
				A	H	L	S	D	Total	Sample	Final
Chronic	6 months	4	LH	10	10	1	1	3	300	0.01	0.01
	6 months	4	Estrus	10	10	1	1		300	0.01	
	11–16 weeks	5	Pregnancy Maintenance	10	10	1	1		300	0.02	
	14–17 weeks	5	Parturition	10	10	1	1		300	0.02	
	10–13 weeks	5	Prolactin	10	10	1	1		300	0.02	
	16 weeks	5	Prostatitis	10	10	1	1		300	0.02	
	16 weeks	5	Organ Wt	10	10	1	1		300	0.02	
	GDs 1–20	100	Fetal ossification	10	10	1	1		300	0.33	

[a] A = animal-to-human (interspecies); H = inter-individual (intraspecies); L = LOAEL-to-NOAEL; S = subchronic-to-chronic duration; D = database deficiency

related effects. For this reason, the NOAEL of 25 mg/kg/day for prolactin and related effects following a 4-day developmental (neonatal) exposure was used as the basis for the acute reference value, with the assumption that a single exposure during a critical period of development would be sufficient to produce these effects.

Although a single-exposure LOAEL, particularly for the reduction in prolactin and increase in prostatitis (which showed a dependence on exposure duration in adults), might be slightly higher than the 4-day LOAEL, lack of data on this endpoint from a single-day exposure leads us to rely on the 4-day value during what appears to be a particularly sensitive time in the early postnatal period. Effects on prolactin (and therefore dopamine) also indicate a concern for developmental neurotoxicity (Figure B-5). Because there is a strong relationship between the data gaps of the acute toxicity study and the developmental neurotoxicity study, a database UF of $10^{1/2}$ was applied. In addition, UFs of 10 for interspecies and intraspecies uncertainty/variability were applied. The resulting reference value was 0.08 mg/kg/day.

Short-term Exposure

A variety of endpoints were examined for calculating the short-term reference value. Information was available on serum LH from the 3-day and 28-day adult studies as well as the study in which pregnant dams were exposed on GDs 1–8. Information on estrous cyclicity was available from the 28-day study in adult rats and in the study in which female weanlings were exposed on PNDs 22–41. Information on pregnancy maintenance was available from studies in which pregnant dams were exposed on GDs 1–8, 6–10, and 1–20, and information on parturition was available for exposures on GDs 6–10.

Information on sexual maturation in females was available from the study in which female weanlings were exposed on PNDs 22–41, and information on sexual maturation in males was available from the study in which weanling male rats were exposed on PNDs 23–53. In addition, information on sexual maturation from the two-generation reproductive toxicity study was also considered relevant. The latter study also provided information on the effects of Luteinate on the seminal vesicles and ventral prostate. Information on prolactin was available from the 28-day study in adults and the study in which lactating dams were exposed on PNDs 1–4; information on prostatitis was also available from the latter study (Figure B-6).

Interspecies and intraspecies UFs of 10 were applied to each of these endpoints (Table B-4). For the entire database available for short-term exposure, a database UF of $10^{1/2}$ was applied in the derivation of the final reference value for the lack of a developmental neurotoxicity study. The resulting reference values ranged from 0.08 to 0.67 mg/kg/day, and the final reference value was 0.08 mg/kg/day.

Longer-term Exposure

For derivation of the longer-term reference value, the databases available for the acute and shorter-term exposures were considered. In addition, the two-generation reproductive toxicity study provided information following longer exposures for serum LH, estrous cyclicity, pregnancy maintenance, parturition, prolactin, and prostatitis (Fig. B-7). Interspecies and intraspecies UFs of 10 were applied to each of these endpoints (Table B-4). The resulting sample reference values were 0.33 mg/kg/day, based on delayed ossification and 0.08 mg/kg/day, based on sexual maturation; all the other endpoints yielded a value of 0.02 mg/kg/day. The latter was chosen as the final value for the longer-term reference value.

Chronic Exposure

Information from all of the exposure scenarios described above was considered in deriving the chronic reference value. In addition, information on serum LH and estrous cyclicity was available from the 6-month study in adult rats. For endpoints where there was a clear exposure-dose-effect relationship, only information from studies of the longest exposure period was included in Table B-4 and Figure B-8. As noted above, decreases in serum LH were observed at 8 mg/kg/day in the 6-month study and at 10 mg/kg/day in the two-generation reproductive toxicity study; the NOAELs were 4 and 5 mg/kg/day, respectively. Estrous cyclicity was affected at these same dose levels in the two studies. The NOAEL for all other endpoints in the two-generation reproductive toxicity study was also 5 mg/kg/day.

Given the similarity in effect levels in the two studies, it is unlikely that longer exposures would alter the effect level. For this reason, although there is no information on the effect of Luteinate on reproductive aging, this is considered to be a qualitative gap in hazard identification, but was not considered to be a database deficiency for the purposes of deriving a chronic reference value. UFs of 10 for interspecies and intraspecies variability and uncertainty were applied to each of these endpoints (Table B-4). The reference values ranged from 0.01 to 0.02 mg/kg/day, and 0.01 mg/kg/day was chosen as the final value.

GLOSSARY

NOTE*: The following terms are used in this document. To the extent possible, definitions were taken from other EPA sources, e.g., IRIS, the Children's Health Research Strategy, the RfC Methodology. In some cases, the definitions have been revised from the originals in IRIS for the sake of clarity or to be consistent with usage in this document. Those terms and definitions that are changed and/or newly proposed in this document to be added to IRIS are shown in italics and the definition(s) they are proposed to replace are indicated in brackets. A number of other terms are included in the IRIS glossary that are not listed here, simply because they were not used in this document.*

Acute Exposure: One dose or multiple doses of short duration spanning less than or equal to 24 hours. [Current IRIS definition.]

Acute Exposure*: Exposure by the oral, dermal, or inhalation route for 24 hours or less.* [Proposed definition to replace the current Acute Exposure definition on IRIS.]

Adverse Effect: A biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism or reduces an organism's ability to respond to an additional environmental challenge.

Benchmark Dose (BMD) or Concentration (BMC): A statistical lower confidence limit on the dose that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background. [current IRIS definition]

Benchmark Dose (BMD) or Concentration (BMC): *A dose or concentration that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background.* [Proposed definition to replace the current definition on IRIS.]

BMDL or BMCL: *A statistical lower confidence limit on the dose or concentration at the BMD or BMC, respectively.* [A new definition to be added to IRIS.]

Benchmark Response (BMR): An adverse effect used to define a benchmark dose from which an RfD (or RfC) can be developed. The change in response rate over background of the BMR is usually in the range of 5 to 10%, which is the limit of responses typically observed in well-conducted animal experiments.

Bioassay: An assay for determining the potency (or concentration) of a substance that causes a biological change in experimental animals.

Bioavailability: The degree to which a substance becomes available to the target tissue after administration or exposure.

Biologically Based Dose Response (BBDR) model: A predictive tool used to estimate potential human health risks by describing and quantifying the key steps in the cellular, tissue, and organismal responses as a result of chemical exposure. [Current IRIS definition.]

Biologically Based Dose Response (BBDR) Model: *A predictive model that describes biological processes at the cellular and molecular level linking the target organ dose to the adverse effect.* [Proposed definition to replace the current definition on IRIS.]

Blood-to-air Partition Coefficient: A ratio of a chemical's concentration between blood and air when at equilibrium.

Chronic Exposure: Multiple exposures occurring over an extended period of time or a significant fraction of the animal's or the individual's lifetime. [Current IRIS definition.]

Chronic Exposure*: Repeated exposure by the oral, dermal, or inhalation route for more than approximately 10% of the life span in humans (more than approximately 90 days to 2 years in typically used laboratory animal species).* [Proposed definition to replace the current definition for Chronic Exposure on IRIS.]

Chronic Study: A toxicity study designed to measure the (toxic) effects of chronic exposure to a chemical.

Critical Effect: The first adverse effect, or its known precursor, that occurs to the most sensitive species as the dose rate of an agent increases.

Critical Study: The study that contributes most significantly to the qualitative and quantitative assessment of risk; also called Principal Study.

Developmental Toxicity: Adverse effects on the developing organism that may result from exposure prior to conception (either parent), during prenatal development, or postnatally until the time of sexual maturation. The major manifestations of developmental toxicity include death of the developing organism, structural abnormality, altered growth, and functional deficiency.

Dose*: The amount of a substance available for interactions with metabolic processes or biologically significant receptors after crossing the outer boundary of an organism. The* ***potential dose*** *is the amount ingested, inhaled, or applied to the skin. The* ***applied dose*** *is the amount presented to an absorption barrier and available for absorption (although not necessarily having yet crossed the outer boundary of the organism). The* ***absorbed dose*** *is the amount crossing a specific absorption barrier (e.g., the exchange boundaries of the skin, lung, and digestive tract) through uptake processes.* ***Internal dose*** *is a more general term denoting the amount absorbed without respect to specific absorption barriers or exchange boundaries. The amount of the chemical available for interaction by any particular organ or cell is termed*

*the **delivered** or **biologically effective dose** for that organ or cell.* [New definition proposed to be added to IRIS.]

Dose-Response Assessment: A determination of the relationship between the magnitude of an administered, applied, or internal dose and a specific biological response. Response can be expressed as measured or observed incidence, percent response in groups of subjects (or populations), or as the probability of occurrence within a population. [Current IRIS definition.]

Dose-Response Assessment: *A determination of the relationship between the magnitude of an administered, applied, or internal dose and a specific biological response. Response can be expressed as measured or observed incidence or change in level of response, percent response in groups of subjects (or populations), or the probability of occurrence or change in level of response within a population.* [Proposed definition to replace the current definition on IRIS.]

Dose-Response Relationship: The relationship between a quantified exposure (dose) and the proportion of subjects demonstrating specific, biological changes (response). [Current IRIS definition.]

Dose-Response Relationship: *The relationship between a quantified exposure (dose) and the proportion of subjects demonstrating specific biological changes in incidence or in degree of change (response).* [Proposed definition to replace the current definition on IRIS.]

Endpoint: An observable or measurable biological event or chemical concentration (e.g., metabolite concentration in a target tissue) used as an index of an effect of a chemical exposure.

Epidemiology: The study of disease patterns in human populations. [Current IRIS definition.]

Epidemiology: *The study of the distribution and determinants of health-related states or events in specified populations and the application of this study to the control of health problems.* [Proposed definition to replace the current definition on IRIS.]

Exposure: Contact made between a chemical, physical, or biological agent and the outer boundary of an organism. Exposure is quantified as the amount of an agent available at the exchange boundaries of the organism (e.g., skin, lungs, gut).

Exposure Assessment: An identification and evaluation of the human population exposed to a toxic agent, describing its composition and size, as well as the type, magnitude, frequency, route, and duration of exposure.

Exposure Pathway: The physical course an environmental agent takes from the source to the individual exposed.

Extrapolation, Low Dose: An estimate of the response at a point below the range of the experimental data, generally through the use of a mathematical model.

Hazard: A potential source of harm.

Hazard Assessment: The process of determining whether exposure to an agent can cause an increase in the incidence of a particular adverse health effect (e.g., cancer, birth defect) and whether the adverse health effect is likely to occur in humans.

Hazard Characterization*: A description of the potential adverse health effects attributable to a specific environmental agent, the mechanisms by which agents exert their toxic effects, and the associated dose, route, duration, and timing of exposure.* [New definition proposed to be added to IRIS]

Human Equivalent Concentration (HEC): The human concentration (for inhalation exposure) of an agent that is believed to induce the same magnitude of toxic effect as the experimental animal species concentration. This adjustment may incorporate toxicokinetic information on the particular agent, if available, or use a default procedure.

Human Equivalent Dose (HED): The human dose (for other than the inhalation routes of exposure) of an agent that is believed to induce the same magnitude of toxic effect as the experimental animal species dose. This adjustment may incorporate toxicokinetic information on the particular agent, if available, or use a default procedure, such as assuming that daily oral doses experienced for a lifetime are proportional to body weight raised to the 0.75 power.

Incidence: The number of new cases of a disease that develop within a specified population over a specified period of time.

Incidence Rate: The ratio of new cases within a population to the total population at risk given a specified period of time.

Latency Period: The time between exposure to an agent and manifestation or detection of a health effect of interest.

Linear Dose Response: A pattern of frequency or severity of biological response that varies proportionately with the amount of dose of an agent. [Current IRIS definition.]

Linear Dose Response: *A pattern of frequency or severity of biological response that varies directly with the amount of dose of an agent. This linear relationship holds only at low doses in the range of extrapolation.* [Proposed definition to replace the current definition on IRIS.]

Longer-term Exposure: *Repeated exposure by the oral, dermal, or inhalation route for more than 30 days, up to approximately 10% of the life span in humans (more than 30 days up to*

approximately 90 days in typically used laboratory animal species). [Proposed new definition to be used relative to the Longer-term Reference Value. Similar to the current definition for Subchronic Exposure. Because subchronic exposure studies will continue to be used in risk assessment, the latter term should be retained as well but replaced with the definition for Longer-term Exposure.]

Lowest-Observed-Adverse-Effect Level (LOAEL): The lowest exposure level at which there are statistically or biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control group. Also referred to as lowest-effect level (LEL). [Current IRIS and RfC Methodology definition.]

Lowest-Observed-Adverse-Effect Level (LOAEL): *The lowest exposure level at which there are biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control group.* [Proposed to replace the current definition in IRIS and the RfC methodology, U.S. EPA, 1994]

Margin of Exposure (MOE): The LED10 or other point of departure divided by the actual or projected environmental exposure of interest.

Mechanism of Action: The complete sequence of biological events that must occur to produce the toxic effect.

Mode of Action (MOA): A less-detailed description of the mechanism of action in which some but not all of the sequence of biological events leading to a toxic effect is known.

Modifying Factor (MF): A factor used in derivation of a reference dose or reference concentration. The magnitude of the MF reflects the scientific uncertainties of the study and database not explicitly treated with standard uncertainty factors (e.g., the completeness of the overall database). A MF is greater than zero and less than or equal to 10, and the default value for the MF is1. [Current definition in IRIS; this report recommends that its use be discontinued.]

No-Observed-Adverse-Effect Level (NOAEL): The highest exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effect between the exposed population and its appropriate control; some effects may be produced at this level, but they are not considered adverse, nor precursors to adverse effects. [Current IRIS and RfC Methodology definition.]

No-Observed-Adverse-Effect Level (NOAEL): *The highest exposure level at which there are no biologically significant increases in the frequency or severity of adverse effect between the exposed population and its appropriate control; some effects may be produced at this level, but they are not considered adverse or precursors to adverse effects.* [Proposed to replace the current definition in IRIS and the RfC methodology, U.S. EPA, 1994.]

Nonlinear Dose Response: A pattern of frequency or severity of biological response that does not vary proportionately with the amount of dose of an agent. When mode of action information indicates that responses may not follow a linear pattern below the dose range of the observed data, non-linear methods for determining risk at low dose may be justified. [Current IRIS definition.]

Nonlinear Dose Response: *A pattern of frequency or severity of biological response that does not vary directly with the amount of dose of an agent. When mode of action information indicates that responses may fall more rapidly than dose below the range of the observed data, nonlinear methods for determining risk at low dose may be justified.* [Proposed definition to replace the current definition on IRIS.]

Physiologically Based Pharmacokinetic (PBPK) Model: Physiologically based compartmental model used to characterize pharmacokinetic behavior of a chemical. Available data on blood flow rates, and metabolic and other processes which the chemical undergoes within each compartment are used to construct a mass-balance framework for the PBPK model. [Current IRIS definition.]

Physiologically Based Pharmacokinetic (PBPK) Model: *A model that estimates the dose to a target tissue or organ by taking into account the rate of absorption into the body, distribution among target organs and tissues, metabolism, and excretion*. [Proposed definition to replace the current definition on IRIS.]

Point of Departure: The dose-response point that marks the beginning of a low-dose extrapolation. This point is most often the upper bound on an observed incidence or on an estimated incidence from a dose-response model. [Current IRIS definition.]

Point of Departure: *The dose-response point that marks the beginning of a low-dose extrapolation. This point can be the lower bound on dose for an estimated incidence or a change in response level from a dose-response model (BMD), or a NOAEL or LOAEL for an observed incidence, or change in level of response.* [Proposed definition to replace the current definition on IRIS.]

Ppb: A unit of measure expressed as parts per billion. Equivalent to 1×10^{-9}.

Ppm: A unit of measure expressed as parts per million. Equivalent to 1×10^{-6}.

Prevalence: The proportion of disease cases that exist within a population at a specific point in time relative to the number of individuals within that population at the same point in time.

Reference Concentration (RfC): An estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a

lifetime. It can be derived from a NOAEL, LOAEL, or benchmark concentration, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments. [Current IRIS definition.]

Reference Dose (RfD): An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or benchmark dose, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments. [Current IRIS definition.]

Reference Value (RfV): *An estimate of an exposure for [a given duration] to the human population (including susceptible subgroups) that is likely to be without an appreciable risk of adverse effects over a lifetime. It is derived from a BMDL, a NOAEL, a LOAEL, or another suitable point of departure, with uncertainty/variability factors applied to reflect limitations of the data used.* [Durations include acute, short-term, longer-term, and chronic and are defined individually in this glossary. This definition is proposed to replace those for the Reference Dose (RfD) and Reference Concentration (RfC). A subscript would be used with the RfV to denote route and duration, e.g., RfV_{AO} for the Acute Oral Reference Value.]

Regional Deposited Dose (RDD): The deposited dose of particles calculated for a respiratory tract region of interest (r) as related to an observed toxicity. For respiratory effects of particles, the deposited dose is adjusted for ventilatory volumes and the surface area of the respiratory region affected (mg/min-sq. cm). For extrarespiratory effects of particles, the deposited dose in the total respiratory system is adjusted for ventilatory volumes and body weight (mg/min-kg).

Regional Deposited Dose Ratio (RDDR): The ratio of the regional deposited dose calculated for a given exposure in the animal species of interest to the regional deposited dose of the same exposure in a human. This ratio is used to adjust the exposure-effect level for interspecies dosimetric differences to derive a human equivalent concentration for particles.

Regional Gas Dose: The gas dose calculated for the region of interest as related to the observed effect for respiratory effects. The deposited dose is adjusted for ventilatory volumes and the surface area of the respiratory region affected (mg/min-sq.cm).

Regional Gas Dose Ratio (RGDR): The ratio of the regional gas dose calculated for a given exposure in the animal species of interest to the regional gas dose of the same exposure in humans. This ratio is used to adjust the exposure effect level for interspecies dosimetric differences to derive a human equivalent concentration for gases with respiratory effects.

Risk (in the context of human health): The probability of injury, disease, or death from exposure to a chemical agent or a mixture of chemicals. In quantitative terms, risk is expressed in values ranging from zero (representing the certainty that harm will not occur) to one

(representing the certainty that harm will occur). The following are examples of how risk is expressed within IRIS: E-4 or 10-4 = a risk of 1/10,000; E-5 or 10-5 = 1/100,000; E-6 or 10-6 = 1/1,000,000. Similarly, 1.3 E-3 or 1.3 x 10-3 = a risk of 1.3/1,000=1/770; 8 E-3 or 8 x 10-3 = a risk of 1/125 and 1.2 E-5 or 1.2 x 10-5 = a risk of 1/83,000. [Current IRIS definition.]

Risk: *The probability of adverse effects resulting from exposure to an environmental agent or mixture of agents.* [Proposed definition to replace the current definition on IRIS.]

Risk Characterization*: The integration of information on hazard, exposure, and dose-response to provide an estimate of the likelihood that any of the identified adverse effects will occur in exposed people.* [New definition proposed to be added to IRIS.]

Risk Assessment (in the context of human health): The determination of potential adverse health effects from exposure to chemicals, including both quantitative and qualitative expressions of risk. The process of risk assessment involves four major steps: hazard identification, dose-response assessment, exposure assessment, and risk characterization. [Current IRIS definition.]

Risk Assessment: *The evaluation of scientific information on the hazardous properties of environmental agents (hazard characterization), the dose-response relationship (dose-response assessment), and the extent of human exposure to those agents (exposure assessment). The product of the risk assessment is a statement regarding the probability that populations or individuals so exposed will be harmed and to what degree (risk characterization).* [Proposed definition to replace the current definition on IRIS.]

Short-term Exposure: Multiple or continuous exposure to an agent for a short period of time, usually one week. [Current IRIS definition.]

Short-term Exposure: *Repeated exposure by the oral, dermal, or inhalation route for more than 24 hours, up to 30 days.* [Proposed definition to replace the current definition for Short-term Exposure on IRIS.]

Statistical Significance: The probability that a result [sic] likely to be due to chance alone. By convention, a difference between two groups is usually considered statistically significant if chance could explain it only 5% of the time or less. Study design considerations may influence the a priori choice of a different statistical significance level. [Current IRIS definition.]

Statistical Significance: *The probability that a result is not likely to be due to chance alone. By convention, a difference between two groups is usually considered statistically significant if chance could explain it only 5% of the time or less. Study design considerations may influence the a priori choice of a different level of statistical significance.* [Proposed definition to replace the current definition on IRIS.]

Subchronic Exposure: Exposure to a substance spanning approximately 10% of the lifetime of an organism. [See note for Longer-term Exposure.]

Subchronic Study: A toxicity study designed to measure effects from subchronic exposure to a chemical.

Supporting Studies: Studies that contain information useful for providing insight and support for conclusions.

Susceptible Subgroups*: May refer to life stages, for example, children or the elderly, or to other segments of the population, for example, asthmatics or the immune-compromised, but are likely to be somewhat chemical-specific and may not be consistently defined in all cases.* [New definition proposed to be added to IRIS.]

Susceptibility*: Increased likelihood of an adverse effect, often discussed in terms of relationship to a factor that can be used to describe a human subpopulation (e.g., life stage, demographic feature, or genetic characteristic).* [New definition proposed to be added to IRIS.]

Systemic Effects or Systemic Toxicity: Toxic effects as a result of absorption and distribution of a toxicant to a site distant from its entry point, at which point effects are produced. Not all chemicals that produce systemic effects cause the same degree of toxicity in all organs. [Current IRIS definition.]

Systemic Effects or Systemic Toxicity: *Toxic effects as a result of absorption and distribution of a toxicant to a site distant from its entry point.* [Proposed definition to replace the current definition on IRIS.]

Target Organ: The biological organ(s) most adversely affected by exposure to a chemical substance. [Current IRIS definition.]

Target Organ: *The biological organ(s) most adversely affected by exposure to a chemical or physical agent.* [Proposed definition to replace the current definition on IRIS.]

Threshold: The dose or exposure below which no deleterious effect is expected to occur.

Toxicity: The degree to which a chemical substance elicits a deleterious or adverse effect upon the biological system of an organism exposed to the substance over a designated time period. [Current IRIS definition.]

Toxicity: *Deleterious or adverse biological effects elicited by a chemical, physical, or biological agent.* [Proposed definition to replace the current definition on IRIS.]

Toxicodynamics*: The determination and quantification of the sequence of events at the cellular and molecular levels leading to a toxic response to an environmental agent (also called pharmacodynamics).* [New definition proposed to be added to IRIS.]

Toxicokinetics*: The determination and quantification of the time course of absorption, distribution, biotransformation, and excretion of chemicals (also called pharmacokinetics).* [New definition proposed to be added to IRIS.]

Toxicology: The study of harmful interactions between chemicals and biological systems. [current IRIS definition]

Toxicology: *The study of harmful interactions between chemical, physical, or biological agents and biological systems.* [Proposed definition to replace the current definition on IRIS.]

Toxic Substance: A chemical substance or agent which may cause an adverse effect or effects to biological systems. [Current IRIS definition.]

Toxic Substance: *A chemical, physical, or biological agent that may cause an adverse effect or effects to biological systems*. [Proposed definition to replace the current definition on IRIS.]

Uncertainty*: Uncertainty occurs because of a lack of knowledge. It is not the same as variability. For example, a risk assessor may be very certain that different people drink different amounts of water but may be uncertain about how much variability there is in water intakes within the population. Uncertainty can often be reduced by collecting more and better data, whereas variability is an inherent property of the population being evaluated. Variability can be better characterized with more data but it cannot be reduced or eliminated. Efforts to clearly distinguish between variability and uncertainty are important for both risk assessment and risk characterization.* [New definition proposed to be added to IRIS.]

Uncertainty Factor (UF): One of several, generally 10-fold, factors used in operationally deriving the RfD and RfC from experimental data. UFs are intended to account for (1) variation in sensitivity among the members of the human population, i.e., interhuman or intraspecies variability; (2) the uncertainty in extrapolating animal data to humans, i.e., interspecies variability; (3) the uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime exposure, i.e., extrapolating from subchronic to chronic exposure; (4) the uncertainty in extrapolating from a LOAEL rather than from a NOAEL; and (5) the uncertainty associated with extrapolation from animal data when the data base is incomplete. [current IRIS definition]

Uncertainty/Variability Factors (UFs): *One of several, generally 10-fold, default factors used in operationally deriving the RfD and RfC from experimental data. The factors are intended to account for (1) variation in sensitivity among the members of the human population (i.e., inter-individual variability); (2) uncertainty in extrapolating animal data to humans (i.e.,*

interspecies uncertainty); (3) uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime exposure (i.e., extrapolating from subchronic to chronic exposure); (4) uncertainty in extrapolating from a LOAEL rather than from a NOAEL; and (5) uncertainty associated with extrapolation when the database is incomplete. [Proposed definition to replace the current one for Uncertainty Factor on IRIS.]

Variability*: Variability refers to true heterogeneity or diversity. For example, among a population that drinks water from the same source and with the same contaminant concentration, the risks from consuming the water may vary. This may be due to differences in exposure (i.e., different people drinking different amounts of water and having different body weights, different exposure frequencies, and different exposure durations) as well as differences in response (e.g., genetic differences in resistance to a chemical dose). Those inherent differences are referred to as variability. Differences among individuals in a population are referred to as inter-individual variability; differences for one individual over time is referred to as intra-individual variability.* [New definition proposed to be added to IRIS.]

REFERENCES

Abdel-Mageed, MA; Suh, DH; Abdel-Rahman, MS. (2001) Intra-species extrapolation in risk assessment of different classes of antimicrobials. Human Ecol Risk Assess 7:15–36.

Allen, BC; Kavlock, RJ; Kimmel, CA; et al. (1994) Dose-response assessment for developmental toxicity: II. Comparison of generic benchmark dose estimates with NOAELs. Fundam Appl Toxicol 23:487–495.

Andersen, ME; Gargas, ML; Clewell, HJ, III; et al. (1987) Quantitative evaluation of the metabolic interactions between trichloroethylene and 1,1-dichloroethylene in vivo using gas uptake methods. Toxicol Appl Pharmacol 89:149–57.

ATSDR (Agency for Toxic Substances and Disease Registry). (1996) Minimal risk levels for priority substances and guidance for derivation. Federal Register 61:25873–25881.

Baird, SJS; Cohen, JT; Graham, JD; Shlyakhter, AI; Evans, JS. (1996) Noncancer risk assessment: a probabilistic alternative to current practice. Human Ecol Risk Assess 2:79-102.

Barnes, DG; Dourson, M. (1988) Reference dose (RfD): description and use in health risk assessments. Regul Toxicol Pharmacol 8:471-486.

Barton, HA; Deisinger, PJ; English, JC; et al. (2000) Family approach for estimating reference concentrations/doses for series of related organic chemicals. Toxicol Sci 54(1):251–61.

Bogdanffy, MS; Daston, G; Faustman, EM; et al. (2001) Harmonization of cancer and non-cancer risk assessment: proceedings of a consensus-building workshop. Toxicol Sci 61:18–31.

Brand, KP; Rhomberg, L; Evans, JS. (1999) Estimating noncancer uncertainty factors: are ratios NOAELs informative? Risk Anal 19:295–308.

Butterworth, BE; Bogdanffy, MS. (1999) A comprehensive approach for integration of toxicity and cancer risk assessments. Regul Toxicol Phamacol 29:23-36.

Calabrese, EJ. (2001) Assessing the default assumption that children are always at risk. Human Ecol Risk Assess 7:37–59.

CDM Group, Inc. (2000) Colloquium on setting acute and intermediate reference values, summary report, September 22. Contract number 68-D6-0060. Chevy Chase, MD.

Creason, J; Neas, L; Walsh, D; et al. (2001) Particulate matter and heart rate variability among elderly retirees: the Baltimore 1998 PM study. J Expo Anal Environ Epidemiol 11(2):116–122.

Dourson, ML; Knauf, LA; Swartout, JC. (1992) On reference dose (RfD) and its underlying toxicity data base. Toxicol Ind Health 8:171–189.

Dourson, ML; Felter, SP; Robinson, D. (1996) Evolution of science-based uncertainty factors in noncancer risk assessment. Regul Toxicol Pharmacol 24:108–120.

Enquist, BJ; Brown, JH; West, GB. (1998) Allometric scaling of plant energetics and population density. Nature 95(6698):163–165.

Evans, JS; Rhomberg, LR; Williams PL; Wilson, AM; Baird, SJS. (2001) Reproductive and developmental risks from ethylene oxide: a probabilistic characterization of possible regulatory thresholds. Risk Anal 21:697-717.

Faustman, EM; Allen, BC; Kavlock, RJ; et al. (1994) Dose-response assessment for developmental toxicity: I. Characterization of data base and determination of NOAELs. Fundam Appl Toxicol 23:478–486.

Gad, SC; Chengelis, CP. (1998) Acute toxicology testing. 2nd ed. San Diego: Academic Press.

Gargas, ML; Burgess, RJ; Voisard, DE; et al. (1989) Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. Toxicol Appl Pharmacol 98:87–99.

Gaylor, DW; Kodell, RL. (2000) Percentiles of the product of uncertainty factors for establishing probabilistic reference doses. Risk Anal 20:245–250.

Gillooly, JF; Brown, JH; West, GB; et al. (2001) Effects of size and temperature on metabolic rate. Science 293:2248–2251.

Hanna, LM; Jarabek AM. (2000) Proposed inhalation route model suites: mode-of-action-motivated dosimetry project. Presented at: Society of Risk Analysis annual meeting; Dec. 3–6, 2000; Arlington, VA.

Haseman, JK. (1984) Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environ Health Perspect 58:385–92.

Haseman, JK; Huff, JE; Rao, GN; et al. (1989) Sources of variability in rodent carcinogenicity studies. Fundam Appl Toxicol 12:793–804.

Hoyer, PB; Sipes, IG. (1996) Assessment of follicle destruction in chemical-induced ovarian toxicity. Annu Rev Pharmacol Toxicol 36:307–331.

Hubal, EA; Bunge, A; Hanna, LM; et al. (2000) Proposed suite of models for estimating dose resulting from exposures by the dermal route. Presented at: Society of Risk Analysis annual meeting; Dec. 3–6, 2000; Arlington, VA.

ICRP (International Commission on Radiological Protection). (1994) Human respiratory tract model for radiological protection: a report of a task group of the International Commission on Radiological Protection. ICRP Publ. no. 66, Annals of the ICRP 24(1–3). Tarrytown, NY: Elsevier Science.

ILSI RSI (International Life Sciences Risk Sciences Institute). (2001) Workshop to develop a framework for assessing risks to children from exposure to environmental agents; July 31–Aug. 2, 2000; Stowe, VT.

IPCS (International Programme on Chemical Safety). (1994) Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. (Environ. Health Criteria no. 170). Geneva: World Health Organization.

IPCS. (1995) IPCS/OECD workshop on the harmonization of risk assessment of reproductive and development toxicity. BIBRA International, Carshalton, Surrey, UK, October, 1994, IPCS/95.25.

IPCS. (1998) Boron. (Environ. Health Criteria no. 204). Geneva: World Health Organization.

IPCS. (2001) Guidance document for the use of data in development of chemical-specific adjustment factors (CSAF) for inter-species differences and human variability in dose/concentration response assessment. WHO/ IPCS/01.4, draft final. Available from: <http://www.ipcsharmonize.org/CSAFsummary.htm>.

Jarabek, AM. (2000) Mode of action: framework for dosimetry model development. Presented at: Society of Risk Analysis annual meeting; Dec. 3–6, 2000, Arlington, VA.

Jepson, GW; Hoover, DK; Black, RK; et al. (1994) A partition coefficient determination method for nonvolatile chemicals in biological tissues. Fundam Appl Toxicol 22:519–24.

Jick, H; Porter, J; Morrison, AS. (1977) Relation between smoking and age of natural menopause. Lancet 1:1354–55.

Kimbell, JS; Gross, EA; Joyner, DR; et al. (1993) Application of computational fluid dynamics to regional dosimetry of inhaled chemicals in the upper respiratory tract of the rat. Toxicol Appl Pharmacol 121:253–263.

Kimbell, JS; Godo, MN; Gross, EA; et al. (1997) Computer simulation of inspiratory airflow in all regions of the F344 rat nasal passages. Toxicol Appl Pharmacol 145:388–398.

Kimmel, CA; Francis, EZ. (1990) Proceedings of the workshop on the acceptability and interpretation of dermal developmental toxicity studies. Fundam Appl Toxicol 14:386–398.

Kimmel, GL; Williams, PL; Claggett, TW; et al. (2002) Response-surface analysis of exposure-duration relationships: the effects of hyperthermia on embryonic development of the rat in vitro. Toxicol Sci. 69:391-399.

Lamb, JC. (1985) Reproductive toxicity testing: evaluating and developing new testing systems. J Am Coll Toxicol 4:163–171.

Lamb, JC, IV; Maronpot, RR; Gulati, DK; et al. (1985) Reproductive and developmental toxicity of ethylene glycol in the mouse. Toxicol Appl Pharmacol 81:169–173.

Leopold, PW; Rivens, ML; Leach, MO. (1997) Noninvasive diagnostic techniques in vascular disease. J Vasc Technol 11:183–186.

Lewis, SC; Lynch, JR; Nikiforov, AI. (1990) A new approach to deriving community exposure guidelines from "no-observed-adverse-effect-levels." Regul Toxicol Pharmacol 11:314–330.

Maull, EA; Cogliano, VJ; Scott, CS; et al. (1997) Trichloroethylene health risk assessment: a new and improved process. Drug Chem Toxicol 20:427–442.

Meek, ME. (2001) Categorical default uncertainty factors: interspecies variation and adequacy of database. Human Ecol Risk Assess 7:157–163.

Meek, ME; Renwick, A; Ohanian, EO; et al. (2001) Guidelines for application of compound specific adjustment factors (CSAF) in dose/concentration response assessment. Comments on Toxicol: 7:575-590.

Morrissey, RE; Lamb, JC, IV; Morris, RW; et al. (1989) Results and evaluations of 48 continuous breeding reproduction studies conducted in mice. Fundam Appl Toxicol 13:747–777.

Murray, FJ; Andersen, ME. (2001) Data derived uncertainty factors; boric acid (BA) as a case study. Human Ecol Risk Assess 7:125–138.

NAS (National Academy of Sciences). (1977) Drinking water and health. Safe Drinking Water Committee. Washington, DC: National Academy Press.

Nau, H. (1991) Pharmacokinetic considerations in the design and interpretation of developmental toxicity studies. Fundam Appl Toxicol 16(2):219–221.

NRC (National Research Council). (2000) Acute exposure guideline levels for selected airborne chemicals, Vol. 1. Washington, DC: National Academy Press.

Orme, J; Ohanian, EV. (1991) Health advisories for pesticides. In: Richardson, ML, ed. Chemistry, agriculture and the environment. Cambridge, UK: Royal Society of Chemistry; pp. 429–443.

Pieters, MN; Kramer, HJ; Slob, W. (1998) Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. Regul Toxicol Pharmacol 27:108–111.

Pinkerton, KE; Joad, J. (2000) The mammalian respiratory system and critical windows of exposure for children's health. Environ Health Perspect 108(Suppl 3):457–462.

Renwick, AG. (1993) Data-derived safety factors for the evaluation of food additives and environmental contaminants. Food Addit Contam 10:275–305.

Renwick, AG. (1998) Toxicokinetics in infants and children in relation to the ADI and TDI. Food Addit Contam 15(Suppl):17–35.

Renwick, AG; Lazarus, NR. (1998) Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. Regul Toxicol Pharmacol 27:3–20.

Rigas, ML; Roth, WL; Jarabek, AM. (2000) Proposed models for estimating relevant dose resulting from exposures by the gastrointestinal route. Presented at: Society of Risk Analysis annual meeting; Dec. 3–6, 2000; Arlington, VA.

Rozman, KK; Doull, J. (2000) Dose and time as variables of toxicity. Toxicology 144:169-178.

Rozman, KK; Doull, J; Hayes, WJ, Jr. (2001) Dose, time, and other factors influencing toxicity. In: Handbook of Pesticide Toxicology, vol. 1, Principles. Ed. R. Kreiger, Academic Press, pp. 1-93.

Schilter, B; Renwick, AG; Huggett, AC. (1996) Limits for pesticide residues in infant food: a safety-based approach. Regul Toxicol Pharmacol 24:126–140.

Skowronski, G; Abdel-Rahman, MS. (2001) Relevance of the 10X uncertainty factor to the risk assessment of drugs used by children and geriatrics. Human Ecol Risk Assess 7:139–152.

Slob, W; Pieters, MN. (1998) A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk Anal 18:787–798.

Smith, TL; Koman, LA; Mosberg, AT. (1994) Cardiovascular physiology and methods for toxicology. In: Hayes, AW, ed. Principles and methods of toxicology. New York: Raven Press; pp. 917–935.

Swartout, JC; Price, PS; Dourson, ML; et al. (1998) A probabilistic framework for the reference dose (probabilistic RfD). Risk Anal 18:271–282.

ten Berge, WF; Zwart, A; Appelman, LM. (1986) Concentration-time mortality response relationship of irritant and systemically acting vapours and gases. J Hazard Mater 13:301–309.

Terry, KK; Elswick, BA; Stedman, DB; et al. (1994) Developmental phase alters dosimetry-teratogenicity relationship for 2-methoxyethanol in CD-1 mice. Teratology 49:218–227.

Tzimas, G; Thiel, R; Chahoud, I; et al. (1997) The area under the concentration-time curve of all-trans-retinoic acid is the most suitable pharmacokinetic correlate to the embryotoxicity of this retinoid in the rat. Toxicol Appl Pharmacol 143:436–444.

U.S. EPA (Environmental Protection Agency). (1986) Guidelines for the health risk assessment of chemical mixtures. Federal Register 51(185):34014–34025 and <http://www.epa.gov/ncea/raf/>.

U.S. EPA. (1988) Recommendations for and documentation of biological values for use in risk assessment. EPA/600/6-87/008. Available from: National Technical Information Service, Springfield, VA.

U.S. EPA. (1991) Guidelines for developmental toxicity risk assessment. Federal Register 56(234):63798–63826 and <http://www.epa.gov/ncea/raf/rafguid.htm>.

U.S. EPA. (1992) Draft deport: a cross-species scaling factor for carcinogen risk assessment based on equivalence of $mg/kg^{3/4}$ /day. Federal Register 57(109):24152.

U.S. EPA. (1994) Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry. EPA/600/8-90/066F. Available from: National Technical Information Service, Springfield, VA.

U.S. EPA. (1996) Guidelines for reproductive toxicity risk assessment. Federal Register 61(212):56274–56322 and <http://www.epa.gov/ncea/raf/rafguid.htm>.

U.S. EPA. (1997a) Exposure factors handbook, Vols. I, II, and III. EPA/600/P-95/002Fa. Prepared by the National Center for Environmental Assessment, Washington, DC. Available from: <http://www.epa.gov/ncea/>.

U.S. EPA. (1997b) Summary of the U.S. EPA Colloquium on a Framework for Human Health Risk Assessment, Vol. 1. EPA/600/R-99/001. Available from: <http://www.epa.gov/nceawww1/colloquium.htm>.

U.S. EPA. (1997c) Guiding principles for Monte Carlo analysis. EPA/630/R-97/001. Risk Assessment Forum, Washington, DC. Available from: National Technical Information Service, Springfield, VA; PB97-188106 and <http://www.epa.gov/ncea/raf/>.

U.S. EPA. (1998a) Methods for exposure-response analysis for acute inhalation exposure to chemicals: development of the acute reference exposure. EPA/600/R-98/051.

U.S. EPA. (1998b) Hazard identification: toxicology endpoint selection process. Office of Pesticide Programs.

U.S. EPA. (1998c) Guidelines for neurotoxicity risk assessment. Federal Register 63(93):26926–26954 and <http://www.epa.gov/ncea/raf/rafguid.htm>.

U.S. EPA. (1998d) Summary of the U.S. EPA Colloquium on a Framework for Human Health Risk Assessment, Vol. 2. EPA/600/R-98/155. Available from: <http://www.epa.gov/nceawww1/colloquium.htm>.

U.S. EPA. (1999a) Guidelines for carcinogen risk assessment [review draft]. NCEA-F-0644. Risk Assessment Forum, Washington, DC. Available from: <http://www.epa.gov/ncea/raf/cancer.htm>.

U.S. EPA. (1999b) Toxicology data requirements for assessing risks of pesticide exposure to children's health: report of the Toxicology Working Group of the 10X Task Force [April 28, 1999 draft]. Available from: <http://www.epa.gov/scipoly/sap/1999/index.htm#may>.

U.S. EPA. (1999c) Exposure data requirements for assessing risks of pesticide exposure to children's health: report of the Exposure Working Group of the 10X Task Force [March 8, 1999 draft]. Available from: <http://www.epa.gov/scipoly/sap/1999/index.htm#may>.

U.S. EPA. (2000a) Supplementary guidance for conducting health risk assessment of chemical mixtures. EPA/630/R-00/002. Available from: <http://cfpub.epa.gov/ncea/cfm/>.

U.S. EPA. (2000b) Science policy handbook on risk characterization. EPA/100/B-00-002. Available from: <http://www.epa.gov/ORD/spc/2riskchr.htm>.

U.S. EPA. (2000c) Benchmark dose technical guidance document [external review draft]. EPA/630/R-00/001. Available from: <http://www.epa.gov/ncea/bnchmrk/bmds_peer.htm>.

U.S. EPA. (2001a) Scientific applicability for use of an alternative test guideline for dermal sensitivity. FIFRA Scientific Advisory Panel review. Available from: <http://www.epa.gov/scipoly/sap/2001/index.htm>.

U.S. EPA. (2001b) Toxicological review of boron and compounds (CAS No. 7440-42-8) [external review draft]. National Center for Environmental Assessment, Washington, DC. Available from: <http://www.epa.gov/ncea/boron.htm>.

U.S. EPA. (2001c) Press release: Agency requests National Academy of Sciences input on consideration of certain human toxicity studies; announces interim policy. December 14. (Contact David Deegan, 202-564-7839; deegan.david@epa.gov).

U.S. EPA (2001d) Water Quality Criterion for the Protection of Human Health: Methylmercury. EPA-823-R-01-001, January. Office of Science and Technology, Office of Water, Washington, DC.

U.S. EPA. (2002a) Framework for cumulative risk assessment [external review draft]. Risk Assessment Forum, Washington, DC. Available from: <http://cfpub.epa.gov/ncea/cfm/nceahome.cfm>.

U.S. EPA. (2002b) Determination of the appropriate FQPA safety factor(s) for use in the tolerance-setting process. Office of Pesticide Programs. Available from: <http://www.epa.gov/oppfead1/trac/science/#10_fold>.

U.S. EPA. (2002c) Integrated risk information system (IRIS). National Center for Environmental Assessment, Washington, DC; <http://www.epa.gov/iris/>.

U.S. EPA. (2002d) Health assessment of 1,3-butadiene. EPA/600/P-98/001F. Available from: <http://cfpub.epa.gov/ncea/cfm/nceahome.cfm>.

Versar, Inc. (2001a) Report on exploration of aging and toxic response issues. Risk Assessment Forum, Washington, DC. EPA Contract 68-C-99-238. Available from: http://cfpub.epa.gov/ncea/raf.

Versar, Inc. (2001b) Report on exploration of perinatal pharmacokinetic issues. Risk Assessment Forum, Washington, DC. EPA Contract 68-C-99-238. Available from: http://cfpub.epa.gov/ncea/raf.

Weller, E; Long, N; Smith, A; et al. (1999) Dose-rate effects of ethylene oxide exposure on developmental toxicity. Toxicol Sci 50:259–270.

West, GB; Brown, JH; Enquist, BJ. (1997) A general model for the origin of allometric scaling laws in biology. Science 276(5309):122–126.