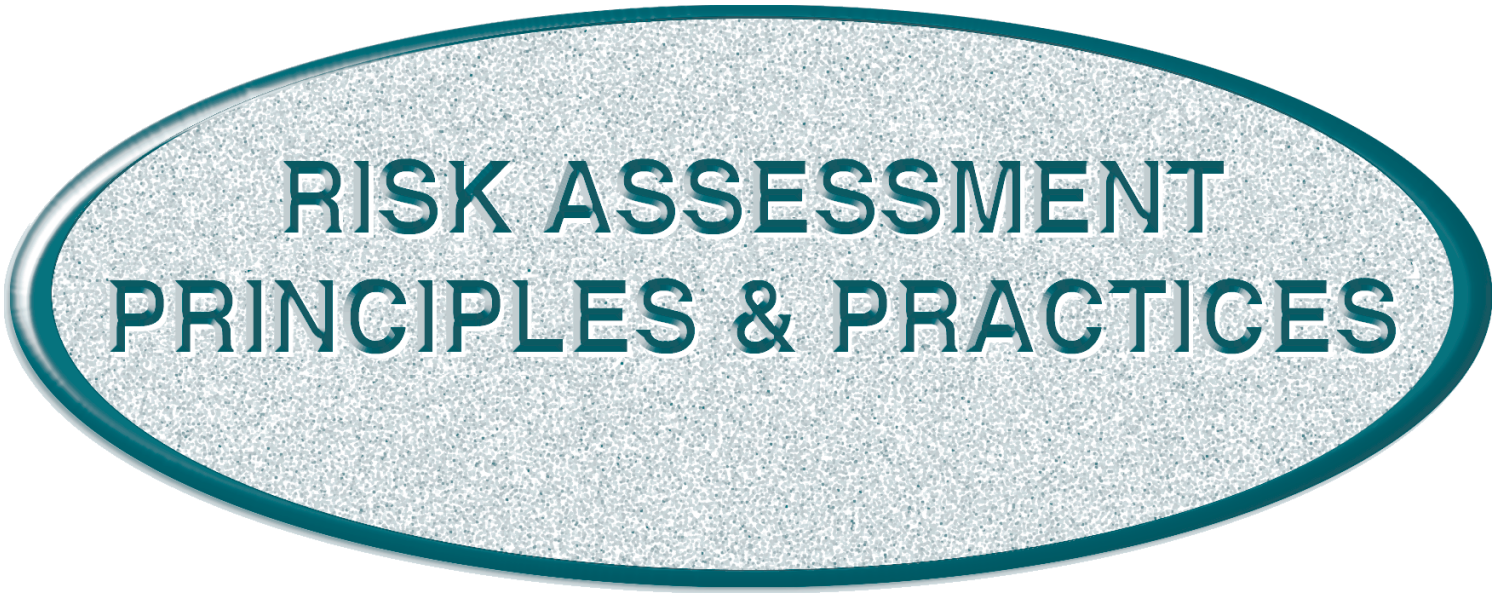
United States Environmental Protection Agency

# Office of the Science Advisor

# STAFF PAPER



EPA/100/B-04/001
March 2004

# U.S. Environmental Protection Agency

# AN EXAMINATION OF EPA RISK ASSESSMENT PRINCIPLES AND PRACTICES

**Staff Paper Prepared for the U.S. Environmental Protection Agency by members of the Risk Assessment Task Force**

**Office of the Science Advisor**
**U.S. Environmental Protection Agency**
**Washington, DC 20460**

# DISCLAIMERS

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

This document presents an analysis of EPA's general risk assessment practices, based on typical historic and current practice. The document does not establish new Agency policy or guidance or amend any existing Agency policy or guidance. Nor does the document attempt to present binding prospective requirements, necessarily applicable to future agency actions. The use of the words "should," "can," "would," and "may" in this document means that something is suggested or recommended, but not required.

A particular risk assessment practice described in this document may not apply to an individual situation based upon the circumstances. Interested parties are free to raise questions and objections about the substance of the practices discussed in this document and the propriety of the application of those practices to a particular situation. Any individual or site-specific risk management decision will be based on the applicable statute and regulations, and on facts specific to the circumstances at issue. Variance from the approaches outlined in this document does not necessarily have any significance. EPA and other decision makers retain the discretion to adopt approaches on a case-by-case basis that differ from those described in this document where appropriate.

Risk assessments discussed in this staff paper reflect a "snapshot" in time and may not be reflective of any further assessment activity past the time of a particular description. For example, the Integrated Risk Information System (IRIS) descriptions, particularly of past assessments, may not be reflective of the current IRIS data base, as assessments are continuously updated. Further, it is important to note that current IRIS health assessments are conducted using the 1999 draft cancer guidelines (as of this examination), and not the 2003 draft final cancer guidelines.

# TABLE OF CONTENTS

ACKNOWLEDGMENTS . . . . . Page ix

AN EXAMINATION OF EPA RISK ASSESSMENT PRINCIPLES AND PRACTICES . . . . . Page 1

1. INTRODUCTION TO EPA RISK ASSESSMENT . . . . . Page 2
- 1.1 Overview . . . . . Page 2
  - 1.1.1 What Is Risk Assessment? . . . . . Page 2
  - 1.1.2 Why Does EPA Conduct Risk Assessments? . . . . . Page 2
  - 1.1.3 How Does EPA Use Risk Assessments in Decision Making? . . . . Page 3
  - 1.1.4 What Is Some Historical Perspective Relevant to EPA Risk Assessment Practices? . . . . . Page 4
- 1.2 EPA Process for Evaluation of Risk Assessment Principles and Practices . Page 6
  - 1.2.1 Why Are We Conducting an Evaluation of Our Risk Assessment Principles and Practices Now? . . . . . Page 6
  - 1.2.2 What Is the Purpose and Intent of This Staff Paper? . . . . . Page 6
  - 1.2.3 Through What Process Is EPA Conducting This Evaluation? . . . . . Page 7
  - 1.2.4 How Does the 2003 OMB Draft Report to Congress Relate to Our Evaluation? . . . . . Page 7
  - 1.2.5 What Is the General Nature of the Comments on EPA Risk Assessment Practices Submitted to OMB? . . . . . Page 8
- 1.3 Organization of This Document . . . . . Page 8
  - 1.3.1 How Are the General Comments Addressed? . . . . . Page 8
  - 1.3.2 How Are the Specific Comments and Examples Addressed? . . . . . Page 9
- 1.4 Other Components Impacting Risk Assessment, But Not Addressed in This Document . . . . . Page 9
  - 1.4.1 What Impact Do the OMB and EPA Information Quality Guidelines Have on EPA Risk Assessment Practices? . . . . . Page 9
  - 1.4.2 Why Is Peer Review So Important? . . . . . Page 10

2. EPA RISK ASSESSMENT AND PUBLIC AND ENVIRONMENTAL HEALTH PROTECTION . . . . . Page 11
- 2.1 Public and Environmental Health Protection (“Public Health Protection”) . . . . . Page 11
  - 2.1.1 What Is EPA’s General Approach for Developing Risk Assessments? . . . . . Page 11
  - 2.1.2 What Is the “Conservatism” Issue in Terms of Public Health Protection? . . . . . Page 11

2.1.3 Is Science Policy Utilized Within the Risk Assessment Process? . . . . . Page 12
2.1.4 What Impact Does Statutory Language Have on EPA Risk Assessment Practices for Public Health Protection? . . . . . Page 13
2.1.5 Does EPA Take a Reasonable Approach to Public Health Protection? . . . . . Page 16
2.1.6 What Happens When Default Assumptions Are Combined? . . . . Page 17
2.2 General Risk Assessment Approaches Used by EPA for Public and Environmental Health Protection ("Public Health Protection") . . . . . Page 20
2.2.1 How Comprehensive Are EPA's Risk Assessments? . . . . . Page 20
2.2.2 Whom Is EPA Trying To Protect? . . . . . Page 22
2.2.3 Are Risk Assessment and Risk Management Separate? . . . . . Page 22
2.2.4 How Do "Planning and Scoping" Help Environmental Risk Assessment? . . . . . Page 23
2.2.5 How Does EPA Use a Screening Risk Assessment? . . . . . Page 24
2.2.6 What Happens if EPA Identifies a Potential Risk That Needs To Be Addressed After a Screening Risk Assessment? . . . . . Page 25
2.2.7 How Are High-End Exposures Reflected in EPA Evaluations? . . Page 26

3. UNCERTAINTY AND VARIABILITY . . . . . Page 30
3.1 Overview . . . . . Page 30
3.2 Uncertainty and Variability . . . . . Page 30
3.2.1 What Is Uncertainty? . . . . . Page 30
3.2.2 What Is Parameter Uncertainty? . . . . . Page 31
3.2.3 What Is Model Uncertainty? . . . . . Page 32
3.2.4 How Does Variability Differ From Uncertainty? . . . . . Page 32
3.3 Characterizing Uncertainty and Variability . . . . . Page 33
3.3.1 Why Is Characterizing Uncertainty and Variability Important? . . Page 33
3.3.2 When Should EPA Conduct Uncertainty Analysis? . . . . . Page 34
3.3.3 What Is an Appropriate Level of Uncertainty Analysis? . . . . . Page 35
3.3.4 With What Precision Should EPA Results Be Reported? . . . . . Page 36
3.4 Issues in Characterizing Uncertainty and Variability . . . . . Page 36
3.4.1 What Is the Issue of Deterministic Versus Probabilistic Approaches? . . . . . Page 37
3.4.2 What Is the Importance of Quantitative Characterization of Uncertainty in Dose-Response? . . . . . Page 38
3.4.3 How Does EPA Use Probabilistic Analysis? . . . . . Page 40
3.4.4 How Does Research Help Reduce Uncertainty? . . . . . Page 41
3.5 Inherent Variability in Biological Response . . . . . Page 42
3.5.1 Why Consider Populations and Life-Stages? . . . . . Page 42

3.5.2 Are Certain Populations and Life-Stages Always at Greater Risk? . . . . . Page 43
3.5.3 How Are Sensitivities to Toxic Effects Considered? . . . . . Page 43
3.5.4 Are Unique or Disproportionate Exposures Considered? . . . . . Page 45
3.5.5 Is Human Variability Considered in Occupational Cohorts? . . . . . Page 47
3.5.6 What Needs To Be Done To Address the Risk Assessment of Populations and Life-Stages? . . . . . Page 48
3.6 What Is EPA Doing To Improve Assessment of Uncertainty and Variability? . . . . . Page 49
3.6.1 How Can EPA Move Forward Toward Quantitative Characterization of Uncertainty in Dose-Response? . . . . . Page 49
3.6.2 Will EPA Move Toward Integrated Uncertainty Analysis? . . . . . Page 49

4. CONSIDERING INFORMATION GAPS IN HEALTH ASSESSMENTS: USE OF DEFAULT AND EXTRAPOLATION ASSUMPTIONS . . . . . Page 51
4.1 Default Assumptions . . . . . Page 51
4.1.1 How Does EPA Use Default Assumptions? . . . . . Page 51
4.1.2 Are Default Assumptions Science Policy? . . . . . Page 52
4.1.3 Does Any Change Seen in Animals Indicate There Will Be a Problem for Humans? . . . . . Page 53
4.1.4 Are Benign (Histologically Non-Malignant) Tumors Presumed To Have Potential To Progress to Malignant Tumors, and Are They Counted as if They Were Malignant? . . . . . Page 54
4.1.5 Should We Expect Target Organ Concordance Between Effects Observed in Animals and Those Expected in Humans? . . Page 55
4.2 Extrapolations . . . . . Page 56
4.2.1 Why Does EPA Use Experiments With Animals To Predict Effects in Humans? . . . . . Page 56
4.2.2 How Does EPA Adjust for the Differences Between Animals and Humans? . . . . . Page 57
4.2.3 Does EPA Use the Most Sensitive Animal Results To Predict Effects in Humans? . . . . . Page 57
4.2.4 Is an Additional Body Scaling Factor for Children Considered? . . Page 59
4.2.5 How Are Effects Observed at High Exposures Used To Predict Responses to Much Lower Exposures? . . . . . Page 59
4.2.6 Why Is the Maximum Tolerated Dose Used? . . . . . Page 61
4.2.7 Does the Use of a Maximum Tolerated Dose Affect Cancer Potency Estimates? . . . . . Page 61
4.2.8 Are the Cancer Risks Estimated by EPA the Expected Incidence of Cancer for a Given Exposure? . . . . . Page 62

4.2.9 How Does EPA Evaluate Incidence Versus Mortality for Cancer Rates? . . . . . Page 63
4.2.10 If Data Are Only Available for One Route of Exposure, Does EPA Evaluate Other Routes of Exposure? . . . . . Page 64
4.3 Use of Uncertainty Factors . . . . . Page 67
4.3.1 How Are NOAELs or LOAELs Used for Extrapolations? . . . . . Page 67
4.3.2 What Uncertainty Factors Does EPA Use To Reduce the Experimental NOAEL in Health Assessments? . . . . . Page 68
4.3.3 Does EPA Consider the Effects of Combining Several UFs? . . . . Page 69
4.3.4 Does EPA Use UFs Derived for Specific Chemicals? . . . . . Page 71
4.4 Weight of Evidence . . . . . Page 71
4.4.1 How Does EPA Consider Weight of Evidence? . . . . . Page 71
4.4.2 Does EPA Give More Weight to Positive Animal Studies Than to Negative Animal Studies? . . . . . Page 72
4.4.3 Does EPA Sometimes Weight Animal Data More Than Human Data? . . . . . Page 72
4.5 Toxicity Equivalency Factor Approach . . . . . Page 83
4.5.1 What Are TEFs? . . . . . Page 83
4.5.2 How Did the TEF Approach Evolve? . . . . . Page 85
4.5.3 Is the Additivity Assumption for TEF Justified? . . . . . Page 87
4.5.4 What Is EPA's Experience With TEFs? . . . . . Page 87
4.6 Model Use . . . . . Page 88
4.6.1 Why Does EPA Use Environmental Models? . . . . . Page 88
4.6.2 How Does EPA Approach Environmental Models? . . . . . Page 89
4.6.3 What Is EPA Doing To Improve the Use of Environmental Models? . . . . . Page 89
4.6.4 How Does EPA Approach Fate and Transport Modeling and Account for Uncertainty? . . . . . Page 92
4.6.5 Does EPA Consider or Review Existing Data on Environmental Concentrations? . . . . . Page 94
4.6.6 Does EPA Consider and Use Case-Specific Data in Models? (Or: How Does EPA Decide To Modify Default Assumptions in Later Tiers Using Models?) . . . . . Page 95
4.6.7 How Will EPA Improve Pesticide Fate and Transport Modeling in the Future? . . . . . Page 95
4.6.8 Does EPA Use Worst-Case Assumptions in Modeling That Are Unreasonable and Do Not Account for Degradation, Partitioning, and "Sinks" in the Environment? . . . . . Page 96

5. SITE- AND CHEMICAL-SPECIFIC ASSESSMENTS . . . . . Page 99
5.1 Overview . . . . . Page 99
5.1.1 What Are the Superfund Program's Purpose and Regulatory Requirements? . . . . . Page 101
5.1.2 Whom Does the Superfund Program Seek To Protect? The Reasonable Maximum Exposure Scenario . . . . . Page 101
5.1.3 Is the RME Overly Conservative? . . . . . Page 102
5.2 Superfund Risk Assessment Guidance . . . . . Page 102
5.2.1 What Are Some Types of Risk Assessment Under Superfund? . Page 102
5.2.2 How Does Superfund Use Site-Specific Concentration and Maximum Concentration Data in Risk Assessment? . . . . . Page 103
5.2.3 Has the Superfund Guidance Been Updated and Externally Peer Reviewed Since It Was First Released in 1989? . . . . . Page 104
5.3 Default Assumptions and Site-Specific Information . . . . . Page 105
5.3.1 What Is the Role of Default Assumptions and Site-Specific Information in Superfund Risk Assessments? . . . . . Page 105
5.3.2 What Is the Drinking Water Consumption Rate? . . . . . Page 106
5.3.3 What Is the Inhalation Rate? . . . . . Page 106
5.3.4 What Is the Exposure Duration for Residences? . . . . . Page 106
5.3.5 What Is the Exposure Duration for Workers? . . . . . Page 107
5.3.6 What Is the Ingestion Rate for Construction Workers? . . . . . Page 108
5.3.7 What Are the Incidental Soil Ingestion Rates for Children? . . . . Page 108
5.4 Site-Specific Information . . . . . Page 109
5.4.1 When Does EPA Use Site-Specific Information? . . . . . Page 109
5.4.2 How Is Bioavailability Addressed? . . . . . Page 110
5.4.3 How Are Fish Ingestion Rates Evaluated? . . . . . Page 111
5.5 Other Factors in Superfund Assessments . . . . . Page 115
5.5.1 What Is the Role of Stakeholders in Superfund Assessments? . . Page 115
5.5.2 How Are Superfund Assessments Factored in Superfund Risk Management Decisions? . . . . . Page 116
5.6 Superfund Site-Specific Conclusions . . . . . Page 119
5.7 Chemical Mixtures Risk Assessment Methods and Practice . . . . . Page 120
5.7.1 What Are the Issues Regarding Chemical Mixture Site Assessments? . . . . . Page 120
5.7.2 What Are Dose Addition and Response Addition? . . . . . Page 121
5.7.3 How Are Dose Addition and Response Addition Applied in Practice? . . . . . Page 123
5.7.4 Are Interaction Data Used in Mixture Risk Assessment? . . . . . Page 124
5.7.5 Are Data on Whole Mixtures Used in Risk Assessment? . . . . . Page 124
5.7.6 What Other EPA Applications Exist for Mixtures Risk Assessment? . . . . . Page 125

6. ECOLOGICAL ASSESSMENT . . . . . Page 127
6.1 Overview . . . . . Page 127
6.1.1 What Is the EPA Ecological Risk Assessment Approach? . . . . . Page 128
6.2 Organism-Level Versus Population-Level Ecological Risk Assessments . . . . . Page 129
6.3 Conservatism and Ecological Risk Assessments . . . . . Page 131
6.3.1 How Is Conservatism Addressed in Ecological Risk Assessments? . . . . . Page 131
6.3.2 How Are Toxicity Reference Values Developed in Ecological Risk Assessments? . . . . . Page 131
6.3.3 How Do Screening-Level and Definitive (Baseline) Assessments Differ? . . . . . Page 132
6.3.4 How Is Site-Specific Information Used To Derive Exposure Concentrations in Ecological Risk Assessments? . . . . . Page 132
6.4 Water Quality Criteria . . . . . Page 133
6.4.1 What Is the Agency Standard Methodology for Deriving Aquatic Life Water Quality Criteria? . . . . . Page 133
6.4.2 What Is the GLI Tier II Methodology? . . . . . Page 134
6.4.3 Was the GLI Tier II Methodology Made Available for Public Comment or Vetted Through Other Public Processes? . . . . . Page 134
6.5 Uncertainty and Pesticide Ecological Assessment . . . . . Page 135
6.5.1 How Does EPA Address Uncertainty Analysis in Pesticide Ecological Assessments? . . . . . Page 135
6.5.2 How Does OPP Address Rodenticide Use Patterns? . . . . . Page 136
6.6 Summary and Conclusions . . . . . Page 139

7. SUMMARY AND RECOMMENDATIONS . . . . . Page 141
7.1 Summary . . . . . Page 141
7.2 Several Current EPA Activities That Will Enhance EPA Risk Assessment Principles and Practices . . . . . Page 142
7.3 Recommendations for Further Improvements to Risk Assessment Principles and Practices . . . . . Page 143

LIST OF USEFUL ABBREVIATIONS AND ACRONYMS . . . . . Page 148

GENERAL REFERENCES . . . . . Page 154

REFERENCES OF EPA RISK ASSESSMENT GUIDELINES . . . . . Page 180

ADDITIONAL USEFUL WEB SITES . . . . . Page 182

# ACKNOWLEDGMENTS

Many people worked on and contributed to the examination of risk assessment principles and practices that ultimately resulted in this document. Many EPA employees from the Agency's Offices and Regions provided input and we would like to specifically acknowledge the efforts and contributions made by the following individuals:

Daniel Axelrad (OPEI)
Timothy Benner (ORD)
Michael Beringer (Region 7)
George Bollweg (Region 5)
David Bussard (ORD)
Michael Callahan (Region 6)
Catherine Campbell (OPEI)
Laurel Celeste (OGC)
Weihsueh Chiu (ORD)
Patricia Cirone (Region 10)
Jenny Craig (OAR)
Larry Cupitt (ORD)
Julie Damon (ASPH/EPA Fellow)
Kerry Dearfield (OSA) - Chair, Task Force
Vicki Dellarco (OPP)
Lynn Flowers (ORD)
Brenda Foos (OCHP)
Stiven Foster (ORD)
Gary Foureman (ORD)
Jack Fowle (ORD)
Herman Gibb (ORD)
Andrew Gillespie (detail from Dept of Agriculture)
David Guinnup (OAR)
Karen Hammerstrom (ORD)
Belinda Hawkins (OEI)
Tala Henry (OW)
Oscar Hernandez (OPPT)
Richard Hertzberg (ORD)
Robert Hetes (ORD) - Chapter 2 lead
James Hetrick (OPP)
Ann Johnson (OPEI)
Myra Karstadt (OPPT)
Aparna Koppikar (ORD)
Michael Kravitz (ORD)
Stephen Kroner (OSWER)
Patricia Lafornara (OEI)
Elizabeth Leovey (OPP)
John Lipscomb (ORD)
Mary Manibusan (ORD)
Carl Mazza (OAR)
Robert McGaughy (ORD)
Michael Metzger (OPP)
Jayne Michaud (OSWER) - Chapter 5 lead
Amy Mills (ORD)
Kenneth Mitchell (OAR)
Jacqueline Moya (ORD)
Debbie Newberry (OSWER)
Susan Norton (ORD)
Marian Olsen (Region 2)
Jennifer Orme-Zavaleta (ORD)
Nicole Paquette (OEI)
Pasky Pascual (ORD)
Resha Putzrath (ORD) - Chapter 4 lead
Mary Reiley (OW)
Susan Reith (ORD)
Glenn Rice (ORD)
Charles Ris (ORD)
Bruce Rodan (ORD)
John Schaum (ORD)
Rita Schoeny (OW)
Brad Schultz (Region 5) - Chapter 3 lead
Cheryl Scott (ORD)
Suhair Shallal (SAB)
Michael Slimak (ORD) - Chapter 6 lead
Daniel Stralka (Region 9)

Glenn Suter (ORD)
Jeffrey Swartout (ORD)
Christina Swartz (OPP)
Linda Teuschler (ORD)
Nelson Thurman (OPP)
Douglas Urban (OPP)
Winona Victery (Region 9)
Donn Viviani (OPEI)
Katie Warwick (ORD)

U.S. Environmental Protection Agency

# AN EXAMINATION OF EPA RISK ASSESSMENT PRINCIPLES AND PRACTICES

# 1. INTRODUCTION TO EPA RISK ASSESSMENT

## 1.1 Overview

### 1.1.1 What Is Risk Assessment?

The most common basic definition of risk assessment used within the U.S. Environmental Protection Agency (EPA) is paraphrased from the 1983 report *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983), by the National Academy of Sciences' (NAS's) National Research Council (NRC):

> **Risk assessment is a process in which information is analyzed to determine if an environmental hazard might cause harm to exposed persons and ecosystems.**

This process is highly interdisciplinary in that it draws from such diverse fields as biology, toxicology, ecology, engineering, geology, statistics, and the social sciences to create a rational framework for evaluating environmental hazards. While this definition has been somewhat enhanced and elaborated upon through subsequent NAS writings, it still basically describes risk assessment as it is performed within EPA. EPA uses risk assessment as a tool to integrate exposure and health effects or ecological effects information into a characterization of the potential for health hazards in humans or other hazards to our environment.

### 1.1.2 Why Does EPA Conduct Risk Assessments?

The mission of the EPA is to protect human health and to safeguard the natural environment — air, water, and land — upon which life depends. EPA fulfills this mission by, among other things, developing and enforcing regulations that implement environmental laws enacted by Congress. The implementation of environmental laws may include grants and other financial assistance to state and tribal governments carrying out environmental programs approved, authorized, or delegated by EPA.

Determining environmental standards, policies, guidelines, regulations, and actions requires making decisions. Environmental decision making is often a controversial process involving the interplay among many forces: science, social and economic factors, political considerations, technological feasibility, and statutory requirements. There are often conflicting interests regarding these various forces than can have a bearing on environmental decisions. Setting an environmental standard that is too lax may threaten public health, while a standard that is unnecessarily stringent may impose a significant marginal economic cost for small marginal gain. Environmental decisions are often time-sensitive, for example when public health is known or suspected to be at risk. The decisions must frequently be made with incomplete or imperfect information and many times under the additional pressure of heightened public

scrutiny and concern. And, once made, the decisions are often challenged in court and subject to high levels of public and scientific scrutiny.

*EPA conducts risk assessment to provide the best possible scientific characterization of risks based on a rigorous analysis of available information and knowledge* — that is, a description of the nature and magnitude of the risk, an interpretation of the adversity of the risk, a summary of the confidence or reliability of the information available to describe the risk, areas where information is uncertain or lacking completely, and documentation of all of the evidence supporting the characterization of the risk. EPA then incorporates this risk characterization with all of the other relevant information — social, economic, political, and regulatory — in making decisions (policies, regulations) about how to manage the risk. Risk assessment, therefore, informs decision makers about the science implications of the risk in question. Risk assessments that meet their objectives can help guide risk managers to decisions that mitigate environmental risks at the lowest possible cost and which will stand up if challenged in the courts.

### 1.1.3 How Does EPA Use Risk Assessments in Decision Making?

The primary purpose of a risk assessment is to inform the risk manager's decision making process. The primary purpose of a risk assessment is not to make or recommend any particular decisions; rather, it gives the risk manager information to consider along with other pertinent information. EPA uses risk assessment as a key source of scientific information for making good, sound decisions about managing risks to human health and the environment. Examples of such decisions include deciding permissible release levels of toxic chemicals, granting permits for hazardous waste treatment operations, and selecting methods for remediating Superfund sites.

The use of credible science in risk assessment helps make and support risk management decisions, but it is not the only factor that the risk manager considers. It is generally recognized — by the science community, by the regulatory community, and by the courts — that it is important to consider other factors along with the science when making decisions about risk management. In some regulations, the consideration of other factors is mandated (e.g., costs). Some of these other factors include:

a. Economic factors — the costs and benefits of risks and risk mitigation alternatives.

b. Laws and legal decisions — the framework that prohibits or requires some actions.

c. Social factors — attributes of individuals or populations that may affect their susceptibility to risks from a particular stressor.

d. Technological factors — the feasibility, impact, and range of risk management options

e. Political factors — interactions among and between different branches and levels of government and the citizens they represent.

f. Public factors — the attitudes and values of individuals and societies with respect to environmental quality, environmental risk, and risk management.

### 1.1.4 What Is Some Historical Perspective Relevant to EPA Risk Assessment Practices?

EPA was involved with risk assessment practices since EPA's early days, although risk assessment per se was not a formally recognized process then. EPA completed its first risk assessment document in December 1975: *Quantitative Risk Assessment for Community Exposure to Vinyl Chloride* (Kuzmack and McGaughy, 1975). The next significant document appeared in 1976: *Interim Procedures and Guidelines for Health Risk and Economic Impact Assessments of Suspected Carcinogens* (Train, 1976). The preamble of this document, signed by the Administrator, signaled the Agency's intent that "rigorous assessments of health risk and economic impact will be undertaken as part of the regulatory process." A general framework described a process to be followed in analyzing cancer risks of pesticides, and the document recommended that the health data be analyzed independently of the economic impact analysis. Later, in 1980, EPA announced the availability of water quality criteria documents for 64 contaminants (USEPA, 1980). This was the first application of quantitative procedures developed by EPA to a large number of carcinogens, and the first EPA document describing quantitative procedures used in risk assessment.

Then in 1983, the NAS published *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983; commonly referred to as the "Red Book"). EPA has integrated the principles of risk assessment from this groundbreaking report into its practices to this day. The following year, EPA published *Risk Assessment and Management: Framework for Decision Making* (USEPA, 1984), which emphasizes making the risk assessment process transparent, describing the assessment's strengths and weaknesses more fully, and providing plausible alternatives within the assessment.

Shortly after the publication of the Red Book, EPA began issuing a series of guidelines for conducting risk assessments (e.g., in 1986 for cancer, mutagenicity, chemical mixtures, developmental toxicology, and in 1992 for estimating exposures). Although EPA efforts focused initially on human health risk assessment, the basic model was adapted to ecological risk assessment in the 1990s to deal with a broad array of environmental risk assessments in which human health impacts are not directly at issue. EPA continues to make a substantial investment

in advancing the science and application of risk assessment through continual updates of these guidelines and the development of newer guidelines as needed. Refer to the section in the References set aside for the listing of these EPA guidelines.

Over time, the NAS expanded on its risk assessment principles in a series of subsequent reports, including *Pesticides in the Diets of Infants and Children* (NRC, 1993), *Science and Judgment in Risk Assessment* (NRC, 1994; also known as the "Blue Book"), and *Understanding Risk: Informing Decisions in a Democratic Society* (NRC, 1996). For example, the NAS places equal emphasis on fully characterizing the scope, uncertainties, limitations, and strengths of the assessment and on the social dimensions of interacting with decision makers and other users of the assessment in an iterative, analytic-deliberative process. The purpose of this process is to ensure that the assessments meet the intended objectives and are understandable. EPA risk assessment practices have evolved over time along with this progression of thought, and in many cases helped drive the evolution of thinking on risk assessment.

In 1995, EPA updated and issued the current Agency-wide Risk Characterization Policy (USEPA, 1995a). The Policy calls for all risk assessments performed at EPA to include a risk characterization to ensure that the risk assessment process is transparent; it also emphasizes that risk assessments be clear, reasonable, and consistent with other risk assessments of similar scope prepared by programs across the Agency. Effective risk characterization is achieved through **t**ransparency in the risk assessment process and **c**larity, **c**onsistency, and **r**easonableness of the risk assessment product — TCCR. EPA's Risk Characterization Handbook (USEPA, 2000a) was developed to implement the Risk Characterization Policy.

The Congressional/Presidential Commission on Risk Assessment and Risk Management (CRARM) was created by the Clean Air Act Amendments of 1990 and formed in 1994. Its mandate was to make a full investigation of the policy implications and appropriate uses of risk assessment and risk management in regulatory programs, under various federal laws, designed to prevent cancer and other chronic health effects that may result from exposure to hazardous substances. More specifically, its mandate was to provide guidance on how to deal with residual emissions from Section 112 hazardous air pollutants (HAPs) after technology-based controls have been placed on stationary sources of air pollutants. In 1997, the Commission published its report in two volumes (CRARM, 1997a; CRARM, 1997b). These discussed the importance of better understanding and quantification of risks, as well as the importance of evaluating strategies to reduce human and ecological risks.

EPA's risk assessment principles and practices build on our own risk assessment guidances and policies — such as the Risk Characterization Policy; *Guidance for Cumulative Assessment, Part 1: Planning and Scoping* (USEPA, 1997a); the Risk Assessment Guidance for Superfund, or RAGS (USEPA, 1989a, and subsequent updates); EPA's Information Quality Guidelines (USEPA, 2002a); and *A Summary of General Assessment Factors for Evaluating the*

*Quality of Scientific and Technical Information* (USEPA, 2003a) — as well as the NAS, the CRARM, and others' concepts. It is understood that risk assessment provides important information about the nature, magnitude, and likelihood of possible environmental risks to inform decisions — principles that evolved out of these many efforts.

## 1.2 EPA Process for Evaluation of Risk Assessment Principles and Practices

### 1.2.1 Why Are We Conducting an Evaluation of Our Risk Assessment Principles and Practices Now?

EPA constantly evaluates its risk assessment principles and practices, mostly via a gradual refinement of particular practices that may not be overtly visible to the public. There are times when EPA takes a concentrated, focused approach: for example, when revising a major risk assessment guideline such as for cancer assessment. EPA conducts a wider, general review of its risk assessment principles and practices occasionally to help strengthen core values and increase its ability to make better decisions.

In early 2002, the position of the EPA Science Advisor was established. The Science Advisor's overarching responsibility is to coordinate and oversee the scientific activities of the program and regional offices at EPA. Part of this responsibility is to ensure the best use of science at the Agency and in its decisions. At the Science Advisor's request, EPA staff began looking at the Agency's risk assessment practices and training with an eye to update them. When the Office of Management and Budget (OMB) solicited comments on risk assessment practices across the federal government (see section 1.2.4), we (EPA staff) took this as an opportunity to concentrate on a wider review to evaluate current risk assessment practices across programs and regions.

### 1.2.2 What Is the Purpose and Intent of This Staff Paper?

This staff paper was developed to give the EPA scientific and technical professional staff an opportunity to present what we (EPA staff) believe are the current EPA risk assessment principles and practices. The practices are presented in the context of the public comments submitted to OMB. The paper's purpose is first to open a dialogue among EPA risk assessors and risk managers about Agency risk assessment practices. Then, as we engage the public, we will continue the dialogue about how we can move forward together to clarify and, where appropriate, strengthen our risk assessment practices.

The staff paper is intended for a wide audience of people who are very familiar with risk assessment principles and practices -- risk assessors and risk managers within the EPA as well as those outside EPA with knowledge of risk assessment practices. The discussion contained here is not meant to be a primer on risk assessment or an introduction for those not very familiar with

these practices. Since the comments address complex and sometimes subtle nuances of risk assessment, the staff paper attempts to deal with those comments and concepts at that level of discussion.

### 1.2.3 Through What Process Is EPA Conducting This Evaluation?

Specifically for this evaluation, on June 17, 2003, an Agency-wide memorandum was issued from three EPA senior managers (Jessica Furey, Associate Administrator for OPEI; Paul Gilman, EPA Science Advisor and Assistant Administrator for ORD; and Stephen Johnson, Assistant Administrator for OPPTS) to start this evaluation. The memorandum called for the establishment of an Agency-wide workgroup, the Risk Assessment Task Force, to review risk assessment principles and practices at the Agency. The Risk Assessment Task Force focused on the practices, assumptions, defaults and principles identified in the comments sent to OMB (see section 1.2.3 below), as well as issues identified from within EPA. Generation of this report on the initial analyses of the Task Force (EPA staff) is the first step in a multi-step process. In the future, EPA expects to communicate with stakeholders about the results of this staff evaluation and give them opportunities for dialogue, in order to understand their concerns.

### 1.2.4 How Does the 2003 OMB Draft Report to Congress Relate to Our Evaluation?

On February 3, 2003, OMB published in the Federal Register (FR) a request for comments on its *Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations* (USOMB, 2003a). In this FR notice, OMB also sought the public's views on a number of important issues pertaining to the practice of risk assessment. OMB received many public comments and these were passed onto EPA. Our effort takes advantage of this information as we review risk assessment principles and practices at the Agency.

OMB issued its final report to Congress on regulatory policy in September 2003 (USOMB, 2003b) based on the comments it received on the February draft report. The final report presents findings on major federal rulemakings finalized over the previous 10 years, specific regulatory reforms, guidance on regulatory analysis, homeland security proposals, and Agency consultations with states and local governments. The report also deals with the concept of precaution in U.S. approaches to risk assessment and management. OMB concludes that precaution plays an important role in risk assessment and risk management, but precaution, coupled with objective scientific analysis, needs to be applied wisely on a case-by-case basis.

### 1.2.5 What Is the General Nature of the Comments on EPA Risk Assessment Practices Submitted to OMB?

The vast majority of the comments can be generally summarized from the American Chemistry Council (ACC) submission. ACC's basic recommendations are not radical or particularly new. They reiterate three points that can be found in the 1991 Executive Office of the President document *Regulatory Program of the United States Government* (EOP, 1991):

a) Risk assessments should not continue an unwarranted reliance on "conservative (worst-case) assumptions" that distort the outcomes of the risk assessment, "yielding estimates that may overstate likely risks by several orders of magnitude."

b) Risk assessments should "acknowledge the presence of considerable uncertainty" and present the extent to which conservative assumptions may overstate likely risks.

c) EPA risk assessments must not "intermingle important policy judgments within the scientific assessment of risk." Rather, the "choice of an appropriate margin of safety should remain the province of responsible risk-management officials, and should not be preempted through biased risk assessments."

On the other hand, various other comments submitted suggest that EPA risk assessments do not fully address all risks. Generally, these latter comments relate to issues of cumulative and aggregate risk. They state that EPA risk assessments concentrate on single chemical/stressor risks, failing to account for multiple chemical/stressor exposures and other factors as life-stage, lifestyle, and increased susceptibility of certain exposed populations. This practice would tend to underestimate risk in real-world scenarios.

## 1.3 Organization of This Document

### 1.3.1 How Are the General Comments Addressed?

This document examines EPA risk assessment principles and practices in light of the general comments outlined in the preceding section. The comments focus on issues of conservatism in risk assessment (e.g., overstated/understated risks), use of rigid default assumptions, poor transparency in the risk assessments we produce, and unacknowledged uncertainty. The following chapters are designed to address these general comments (issues). Chapter 2 deals with the issue of conservatism. Chapter 3 discusses the nature of uncertainty and variability and how EPA deals with these in risk assessment. Chapter 4, on defaults and extrapolations, discusses how assumptions can be used when chemical- and/or site-specific data

are not available or are inadequate to use in performing a risk assessment. The next chapter, chapter 5, approaches the use of site- and chemical-specific data. This chapter emphasizes the Superfund approaches — many of the specific comments we received were specific to Superfund.

Most of the comments EPA received focused on human health assessment. Many of the principles and practices, though, apply to ecological health assessments as well. Chapter 6 briefly discusses and evaluates our ecological health assessments.

The last chapter, chapter 7, provides a summary of the Risk Assessment Task Force conclusions and recommendations, with transparency in our risk assessment practices as a focus.

### 1.3.2 How Are the Specific Comments and Examples Addressed?

This document is organized to address the overarching issues of conservatism, uncertainty, and transparency. However, many comments focused on specific risk assessment principles and practices as well as specific chemical assessments as examples. We will not discuss every specific comment in a "one to one" discussion, although there are some instances in the following chapters for which specific examples are provided to illustrate a point. While we will mention specific chemicals and specific rulemakings in the context of the general issues discussion, we will not discuss individual chemical- or site-specific *decisions,* as these are generally of high importance to many, are complex, and (in many cases) are still in active interaction with interested stakeholders.

## 1.4 Other Components Impacting Risk Assessment, But Not Addressed in This Document

### 1.4.1 What Impact Do the OMB and EPA Information Quality Guidelines Have on EPA Risk Assessment Practices?

The OMB guidelines state that information needs to be "objective, realistic, and scientifically balanced" (USOMB, 2002). EPA embraced the OMB guidance when developing its Information Quality Guidelines (USEPA, 2002a). These guidelines point to our reliance on the extensive use of peer review, the practices found in the various risk assessment guidelines and the Risk Characterization Policy, and the Agency-wide use of EPA's Quality System. These EPA practices help ensure that information we use is objective, realistic, and scientifically balanced. As we stress, EPA uses scientific peer review to help ensure the quality of the risk assessments we generate and to keep the assessments as objective, and as consistent, as possible. EPA's Peer Review Handbook (USEPA, 2000b) stresses that peer review panels be balanced in terms of their expertise and biases so that a reasonable and scientifically balanced review results. Further, conflicts of interest need to be readily identified. The Risk Characterization Policy

encourages presentation of all plausible viewpoints that are realistic and scientifically supportable. In 1984, under EPA Order 5360.1, an Agency-wide Quality System was established. In 2000, Agency-wide policies were defined in EPA Order 5360.1 A2 and EPA Manual 5360 A1, which expanded the policy to accommodate current and evolving needs of the Agency. EPA's Quality System helps ensure that we stay objective in the use of information in any assessment we generate (USEPA, 2000c, 2000d). In association with the Information Quality Guidelines, we also developed general assessment factors related to quality issues regarding scientific and technical information (USEPA, 2003a). One should take these assessment factors into consideration as appropriate when evaluating the quality and relevance of information, regardless of source.

Many parties believe that EPA risk assessment practices do not follow this guidance. We believe EPA conducts risk assessments consistent with this guidance, as evidenced by some of the efforts detailed in the paragraph above. We continually strive to publicly present our risk assessment practices and guidance documents and subject them to peer review. Our Risk Characterization Policy directs us to consider all scientifically plausible and supportable viewpoints, but this information is only part of the full range of information risk managers use to make a decision.

### 1.4.2 Why Is Peer Review So Important?

The value of scientific peer review in ensuring the quality of our scientific and technical products is critical to EPA and is widely understood and accepted across the Agency. Conscientious use of peer review is essential to the credibility of the risk assessments EPA uses to support its decisions. Consistent Agency-wide application of peer review has been an EPA priority for many years and continues to be so. Since issuing the peer review policy in 1993, EPA has continually supported and strengthened the policy and practices of peer review, which in turn have supported and strengthened our risk assessment principles and practices. For example, peer review gave us an opportunity to seek expert review on many of the issues highlighted in this document, such as the reasonableness of assumptions and methodologies used in Agency risk assessments.

A peer review is a documented critical review of a specific scientific or technical work product, conducted by qualified individuals (or organizations) who are independent of those who performed the work (at a minimum, from a different office) but who are collectively equivalent in technical expertise (i.e., peers). EPA's Peer Review Handbook (2nd Edition; USEPA 2000b) is one of the most advanced treatments of peer review for intramural research and scientific/technical analysis of any federal agency. By utilizing the practices embodied in the Handbook, EPA ensures the credibility and quality of the risk assessments we generate to support Agency decisions.

# 2. EPA RISK ASSESSMENT AND PUBLIC AND ENVIRONMENTAL HEALTH PROTECTION

## 2.1 Public and Environmental Health Protection ("Public Health Protection")

EPA's risk assessments are conducted in support of its mission to protect public health and the environment. Given the uncertainty, variability, and data gaps encountered when conducting any risk assessment, a key objective for EPA's risk assessments is that they avoid both underestimation of risk and gross overestimation of risk.

### 2.1.1 What Is EPA's General Approach for Developing Risk Assessments?

Over the years, practices have been developed to assess risk based on available data and information, and EPA has been at the forefront of much of this development. Risk assessment is a complex process, requiring the integration of data and information across a broad range of activities and disciplines, including source characterization, fate and transport, modeling, exposure assessment, and dose-response assessment. EPA seeks to use the available information (data) in an objective, realistic, and scientifically balanced way. In each specific assessment, the Agency incorporates the relevant data and information to the extent possible (e.g., see chapter 5). Where relevant chemical- and/or site-specific data are not available, EPA uses specific default assumptions and extrapolations to fill in the data gaps and allow the risk assessment to proceed (see chapter 4) so that the Agency can ultimately make the decisions required under its mandates. This approach is consistent with the NRC's recommendation about EPA's use of defaults (NRC, 1994). In general, EPA's default assumptions are based on peer reviewed studies, empirical observations, extrapolation from related observations, or scientific theory.

### 2.1.2 What Is the "Conservatism" Issue in Terms of Public Health Protection?

Because of data gaps, as well as uncertainty and variability in the available data, risk cannot be known or calculated with absolute certainty. Further, as Hill (1965) noted, a lack of certainty or perfect evidence "does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time." Therefore, consistent with its mission, EPA risk assessments tend towards protecting public and environmental health by preferring an approach that does not underestimate risk in the face of uncertainty and variability. In other words, EPA seeks to adequately protect public and environmental health by *ensuring that risk is not likely to be underestimated*. However, because there are many views on what "adequate" protection is, some may consider the risk assessment that supports a particular protection level to be "too conservative" (i.e., it overestimates risk), while others may feel it is "not conservative enough" (i.e., it underestimates risk). This issue regarding the appropriate degree of "conservatism" in EPA's risk assessments has been a concern

from the inception of the formal risk assessment process and has been a major part of the discussion and comments surrounding risk assessment.

Even with an optimal cost-benefit solution, in a heterogeneous society, some members of the population will bear a disproportionate fraction of the costs while others will enjoy a disproportionate fraction of the benefits (Pacala et al., 2003). Thus, inevitably, different segments of our society will view EPA's approach to public health and environmental protection with different perspectives.

The NRC in its 1994 report (NRC, 1994) examined the conservatism issue, particularly as it relates to the default assumptions EPA uses. The Committee did not reach consensus on the degree to which "conservatism" should play a role in defining defaults. This issue was heavily debated to the point where divergent views were offered, one advocating the principle of "plausible conservatism" and the other the maximum use of scientific information in selecting defaults. In general, EPA uses defaults that guard against underestimating risk while also being scientifically plausible given existing uncertainty. EPA's use of various default assumptions is the basis for many of the differences of opinion about its risk assessment practices. This concern, identified by the NRC in 1994, continues today as indicated by many of the comments submitted to OMB.

The question of conservatism is heightened by the ambiguous definitions and uses of the term "conservatism" by the various concerned parties, including those that feel EPA overestimates risk and those that feel the Agency systematically understates risks. Some of the various concepts associated with term "conservatism" include prudence versus misestimation, conservatism as a response to uncertainty or variability, "level of conservatism," and "amount of conservatism" (NRC, 1994).

### 2.1.3 Is Science Policy Utilized Within the Risk Assessment Process?

Science policy positions and choices are by necessity utilized during the risk assessment process. Two major ways in which the risk assessment process uses science policy are described below in this section. Note that the utilization of science policy in the risk assessment process is not meant to "bury" or "hide" risk management decisions within the risk assessment. The use of any policy choice needs to be transparent in a risk assessment. In addition, although science policy is utilized in the risk assessment process, it is important to recognize that the policy positions themselves are developed outside the risk assessment. These policy positions are usually supported by scientific data and/or consensus and ensure that the risk assessment proceeds in a way that best serves the needs of the decision maker and the Agency.

First, there are some basic, fundamental policy positions that frame the risk assessment process to ensure that the risk assessments produced are appropriate for a particular decision.

For example, since EPA is a health and environmental protective agency, EPA's policy is that risk assessments should not knowingly underestimate or grossly overestimate risks. This policy position prompts risk assessments to take a more "protective" stance given the underlying uncertainty with the risk estimates generated. Another framing policy position is that EPA will examine and report on the upper end of a range of risks or exposures when we are not very certain about where the particular risk lies. For example, in a screening-level risk assessment for hazardous air pollutants (HAPs), the risk assessment starts with a 70-year exposure to assess an individual most exposed (see section 2.2.7 for more discussion on this specific example). Further, when several parameters are assessed, upper-end values and/or central tendency values are generally combined to generate a risk estimate that falls within the higher end of the population risk range. Currently, the use of the upper part of a range pertains more often to the exposure component of the risk assessment than the hazard/dose-response portion. Many comments to EPA suggest that the combining of upper ends leads to unreasonable estimates of risk. We generally believe otherwise (e.g., see section 2.1.6), and we feel that this practice should be explained clearly in the risk characterization to ensure that risk managers can make an appropriate decision consistent with the Agency's goal of public health and environmental protection. These policy positions not only shape the risk assessment process, but are also a factor in the decision making process outside the risk assessment.

Second, default assumptions utilized in any given risk assessment entail science policy positions or choices. These science policy choices are more specific than the framing science policies, but generally are consistent with the framing policies. For example, a change that is considered adverse (i.e., associated with toxicity) in an animal study is assumed to indicate a problem for humans unless data demonstrate otherwise. As discussed more fully in chapter 4 (particularly sections 4.1.1. and 4.1.2), default assumptions are generally supported by scientific data and/or scientific consensus. Their use in risk assessments is to allow the risk assessment to proceed when chemical- and/or site-specific data are missing or not useful.

Most importantly, any science policy position or choice used in the risk assessment process does not direct the risk assessment itself toward a specific risk management decision, e.g., the use of a specific risk estimate. Rather, the risk assessment informs the decision maker about the potential risks and uncertainties around the risk estimate(s). These characterized risks are then considered in light of the other factors before a decision is made (see section 1.1.3); science policy is one of the factors that help a risk manager make a decision.

### 2.1.4 What Impact Does Statutory Language Have on EPA Risk Assessment Practices for Public Health Protection?

The discretionary power afforded to agencies in making regulatory decisions varies greatly (OSTP, 1995). EPA faces regulatory, licensing, and other decisions covering a wide range of environmental issues and pollutants. These decisions are made within a number of EPA

program offices, each responding to a unique mixture of statutes, precedents, and stakeholders. Congress establishes legal requirements that generally describe the level of protectiveness that EPA regulations must achieve and, infrequently, Congress imposes specific risk assessment requirements. In addition, court precedents can affect how EPA considers assessments of risk. Given this apparent variety of goals, the constraints within which those goals must be reached, and the discretionary powers afforded, it is to be expected that decisions made in one EPA program office, for one particular environmental issue, may appear at least on the surface to be based on different risk considerations and/or public health protection goals from those used in other EPA program offices and for other environmental issues. For example, there may appear to be a variation from one EPA risk management decision to another in the way in that public health protection goals have been balanced against other considerations such as technological feasibility, precedent, and cost. However, overall, EPA generally interprets its statutes to be protective of public health and the environment.

Apparent inconsistencies in risk assessment practices across EPA can stem from differences in statutory language. For example, individual statutes identify varying risks to evaluate and to protect against (e.g., establish a margin of safety; protect sensitive resources; reduce overall risks) and mandate different levels of protection (e.g., protect public welfare; prevent unreasonable risk; reduce overall risks; function without adverse effects). Examples among major EPA program offices illustrate some of the different Congressional mandates regarding risk assessment and risk management practices:

a) In the case of threshold effects … an additional ten-fold margin of safety for the pesticide chemical residue shall be applied for infants and children … (OPPTS; FFDCA §408 (b)(2)(C))

b) The Administrator shall, in a document made available to the public in support of a regulation promulgated under this section, specify, to the extent practicable:

   1) Each population addressed by any estimate of public health effects;

   2) The expected risk or central estimate of risk for the specific populations;

   3) Each appropriate upper-bound or lower-bound estimate of risk … (OW; SDWA § 300g-1 (b)(3))

c) The Administrator shall … [add] pollutants which present, or may present, through inhalation or other routes of exposure, a threat of adverse human health effects...or adverse environmental effects through ambient concentrations, bioaccumulation, deposition, or otherwise but not including releases subject to

regulation under subsection (r) of this section as a result of emissions to air... (OAR; CAA §112(b)(2))

d) … Provide an ample margin of safety to protect public health or to prevent an adverse environmental effect (OAR; CAA §112(f)).

Similarly, individual statutory requirements regarding the appropriate level of protection can have a significant impact on the focus (the purpose and scope) of a risk assessment, which can lead to the appearance of inconsistency in risk assessment practices. Such requirements vary across Agency programs; for example:

a) … To assure chemical substances and mixtures do not present an unreasonable risk of injury to health or the environment (OPPTS; TSCA §2(b)(3)).

b) … Function without unreasonable and adverse effects on human health and the environment (OPPTS; FIFRA §3).

c) … Necessary to protect human health and the environment (OSWER; RCRA §3005 as amended).

d) … Provide the basis for the development of protective exposure levels (OSWER; NCP §300.430(d)).

e) … Adequate to protect public health and the environment from any reasonably anticipated adverse effects (OW; CWA §405(d)(2)(D)).

Even the statutory language used for different statutes administered within one major office, EPA's Office of Air and Radiation (OAR), shows differences:

a) … Protect public health with an adequate margin of safety (OAR; CAA §109).

b) … Provide an ample margin of safety to protect public health or to prevent an adverse environmental effect (OAR; CAA §112(f)).

c) … Protect the public welfare from any known or anticipated adverse effects (OAR; CAA §109).

d) … [Not] cause or contribute to an unreasonable risk to public health, welfare, or safety (OAR; CAA §202(a)(4)).

e) … Protect sensitive and critically sensitive aquatic and terrestrial resources (OAR; CAAA §404 (Appendix B)).

f) … Reduce overall risks to human health and the environment (OAR; Title VI of CAA).

g) … Actions to mitigate environmental and health risks (OAR; SARA Title IV).

For example, within OAR's National Ambient Air Quality Standards Program, the primary standards are health-based and do not consider costs or technological feasibility. The secondary standards consider the impacts of pollutants on human economic well-being such as visibility, agricultural productivity, and ecological impacts. Under the Clean Water Act, EPA's Office of Water (OW) publishes ambient water quality criteria based on protecting human health; these risk assessments do not consider the cost or technological feasibility of meeting these criteria. Under the Safe Drinking Water Act, OW conducts risk assessments to determine non-enforceable Maximum contaminant level goals (MCLGs). OW then sets enforceable Maximum contaminant levels (MCLs) as close as technically feasible to the MCLGs after taking costs into consideration.

In its report on chemical risk assessment, the General Accounting Office, or GAO, (USGAO, 2001) recognized how contextual differences in statutes may affect the focus of risk assessments. By taking an overall public health protective stance (as a response to uncertainty and variability), EPA's approach to risk assessment takes into account the variety of language found in the various statutes and allows some overall consistency in the Agency's risk assessment practices, while allowing specific EPA offices to follow their particular mandates.

### 2.1.5 Does EPA Take a Reasonable Approach to Public Health Protection?

Since uncertainty and variability are present in risk assessments, EPA usually incorporates a "high-end" hazard and/or exposure level in order to ensure an adequate margin of safety for most of the potentially exposed, susceptible population, or ecosystem. EPA's high-end levels are around 90% and above — a reasonable approach that is consistent with the NRC discussion (NRC, 1994). This policy choice is consistent with EPA's legislative mandates (e.g., adequate margin of safety, see section 2.1.4). Even with a high-end value, there will be exposed people or environments at greater risk and at lower risk. In addition to the high-end values, EPA programs typically estimate central tendency values for risk managers to evaluate. This provides a reasonable sense of the range of risk that usually lies on the actual distribution.

When EPA has usable data which show that protection at a higher level (e.g., 99%) or lower level (e.g., 90% or lower) is appropriate, then managers may use the information to make the risk management decision. For example, exposure data used for risk assessment are

frequently based on studies, such as controlled field trials for pesticides, that reflect high-end exposures. Usually, the relationship between these levels and the actual levels to which people are exposed is not known. This leads to uncertainty in estimates about the percentile of the population exposed in a given assessment. Therefore, the risk management decision regarding the percentile at which to regulate needs to consider how close the exposure data are likely to approximate actual exposures in the population being considered. Typically, percentiles chosen for pesticides regulation can range from possibly the $70^{th}$ or $80^{th}$ percentile in situations (e.g., occupational) where it is known that use of the available data will clearly overestimate exposure to the $99.9^{th}$ percentile when the data are likely to closely approximate the population exposures (e.g., specific dietary exposures).

Whatever hazard and exposure values are used to estimate risk, it is important to be transparent in characterizing the range of possible risks. For example, if we identify certain populations that will be at greater risk than the high end we propose, then it is important to highlight these populations so decision makers can make appropriate decisions regarding their possible risk. The risk assessment should not make that decision; it should characterize that risk and who or how much is at risk. Consequently, managers may decide on a greater level of action in a certain locale, if conditions warrant, than may be appropriate elsewhere. On the other hand, risk assessments should also characterize the range of risk, including the lower levels of risk. EPA's policy is to characterize the range of risk and to also highlight exceptionally susceptible populations (USEPA, 1995a, 2000a).

While the need for a full characterization of risk is stated in EPA's policies and guidance, the actual characterization in risk assessments may not be so explicit. Therefore, we may need to more consistently characterize the range of risk and highlight susceptible populations in our risk assessments. This greater transparency will help make the reasonableness of the estimated risk for the exposed population or ecosystem more understandable. Related to this, the NRC pointed out in 1994 that there is little empirical evidence to suggest that EPA's potency, exposure, or risk estimates are markedly higher than estimates embodying a reasonable degree of prudence (i.e., the conventional benchmarks of the $95^{th}$ or $99^{th}$ percentiles that statisticians use) (NRC, 1994).

### 2.1.6 What Happens When Default Assumptions Are Combined?

Some comments assert that EPA has so overemphasized conservatism that most risk estimates are alarmingly false, meaningless, and unscientific. It should be noted that the use of default assumptions does not render the process or results non-scientific. The basis of the argument that EPA's risk estimates are alarmingly false suggests that combining several values (e.g., use of values at a default level of 95%) results in excessive overestimates of risk. For example, it is implied that combining two $95^{th}$ percentile defaults results in an estimate above the $99^{th}$ percentile, that combining three $95^{th}$ percentile defaults results in an estimate above the $99.9^{th}$

percentile, and so forth. Suggestions have also been made that using the mean instead of the 95$^{th}$ percentile in these types of calculations would result in a less conservative risk estimate.

However, the final distribution in these types of calculations is influenced differently by the different inputs, and just multiplying the numbers as implied above will not lead to the answers above. In estimating exposure or risk using percentiles, how much "influence" using a particular 95$^{th}$ percentile value has in the ultimate exposure or risk calculation depends on more than just its position in the percentile rankings. It depends also on the variability of the data within the distribution for the input factors, the shape of the input distributions, and even the number of data points. If all the input variables show the same variability, shape, etc., then the above reasoning about compounding values is true. This is rarely the case in actual situations, however.

For illustrative purposes, consider a case where the calculated exposure is a product of several input factors. If most of these factors have little variability (e.g., the average lifetime of 70 years, or the volume of air breathed per day varies within plus or minus 25% for almost all the population subgroups) and the concentration of a chemical that the population is exposed to varies by three orders of magnitude, then the resulting percentile position of the calculated exposure within the resulting exposure distribution will be much more influenced by which value is selected from the input distribution of concentrations than it will be for which value is selected for the other factors. Selecting the mean value for the concentration input value and 95$^{th}$ percentile values for the others will result in a calculated exposure that is much closer to the mean of the resulting distribution than the 95$^{th}$ percentile (or higher), because the resulting distribution is heavily influenced by the concentration input. Conversely, selecting the 95$^{th}$ percentile from the concentration input distribution and the means of the others will result in a calculated exposure that is close to the 95$^{th}$ percentile of the resulting exposure distributions. Consequently, in the cases where all the input distributions are not the same in variability, where the final estimate falls on the combined distribution depends on *which* input variable is selected as 95$^{th}$ percentile.

In fact, the example above is similar to the RME (reasonable maximum exposure) calculated in Superfund assessments. Typically, the concentrations in environmental media are highly variable and the other parameters are less variable. Selecting the mean of the concentration distribution, while setting one or more of the other parameters at the 95$^{th}$ percentile value, usually results in an estimate that is reasonably conservative (with the acknowledgment that each case, due to its unique data set, will be different).

An additional factor within the RME calculation is that the 95% upper confidence limit (UCL) on the mean of the concentrations is used instead of the mean itself. Using the mean concentration for an area with contaminated soil is in effect setting a scenario where it is equally probable that the exposed individual will be at any given location in the area. When the site has

"attractive nuisances" with higher concentrations, the risk assessor factors this into the assessment. The coverage of the data set within the area affects the confidence in the mean calculated from the data. For example, confidence in calculating the true mean for a one-acre site with three data points will be much less than for a one-acre site with 300 data points. The 95% UCL allows the assessor to have some confidence that the mean is not being underestimated for a relatively sparse data set. On the other hand, for a robust data set, the 95% UCL will be quite close to the mean itself, so it does not introduce appreciable additional conservatism into the estimate of exposure. Further, the upper confidence limit on the mean is not the same as the 95$^{th}$ percentile of observed concentrations.

For many other exposure calculations that EPA performs, the Agency's 1992 Guidelines for Exposure Assessment (USEPA, 1992a) provide specific guidance on how to construct estimates using distributions of contributing factors, taking into consideration that the resulting exposure estimate will be more sensitive to some factors than others. These guidelines suggest that when exposure data or probabilistic simulations are not available, an exposure estimate that lies between the 90$^{th}$ percentile and the maximum exposure in the exposed population be constructed "by using maximum or near-maximum values for one or more of the most sensitive variables, leaving others at their mean values" (USEPA, 1992a).

Nor is just using the mean instead of the 95$^{th}$ percentile *necessarily* less "conservative." A percentile distribution, such as the collection of data associated with soil concentrations on a specific site, is an ordered series of data values. The values for the data may be such that the 95$^{th}$ percentile value may be higher or lower than the mean. While the mean of a data set is most often below the 95$^{th}$ percentile, consider a data set of 100 soil samples in which 97 samples are at the background level, 1 part per million (ppm), for a given pollutant, while the remaining samples are at 50 ppm, 100 ppm, and 1,700 ppm. The 95$^{th}$ percentile of this data set is 1 ppm, which is actually numerically indistinguishable from the minimum value; the mean, 1.947 ppm, lies between the 96$^{th}$ and 97$^{th}$ percentile. Consequently, use of the 95$^{th}$ percentile value in a calculation instead of the mean does not necessarily make the calculation more — or less — conservative. While the mean does indeed fall below the 95$^{th}$ percentile for the vast majority of cases EPA sees, the fact remains that mathematical manipulation of percentile data and selection of values from percentile distributions for purposes of estimating resulting exposures in a given scenario needs to be considered on an individual basis to determine the degree of "conservatism."

Risk estimates often involve adding risks from various chemicals. This can be done by adding doses for chemicals with the same mode of action (the approach used by EPA's Office of Pesticide Programs, or OPP), or it can be approximated by adding, for example, the resulting single-chemical cancer risks for various chemicals (the approach EPA uses in permitting and in its Superfund program). Adding individual chemical risks to estimate a combined, cumulative risk presents several complexities (USEPA, 2003b). For example, chemical interactions, which can lead to synergism or antagonism, are poorly understood. Also, there is a statistical

complexity because the risks themselves are expressed as upper bound estimates of risk and not most likely estimates. Cogliano (1997) showed that adding upper bound estimates may make the resulting risk sums more conservative, but that the actual resulting risk values are not misleading and are probably within a factor of **2 or 3** of the estimates that would result from calculating the 95% UCL of the sum of the most likely estimates. Cogliano states:

> (A)s the number of carcinogens increases, the sum of upper bounds becomes increasingly improbable as an estimate of overall risk. At the same time, however, the analysis shows that the sum of upper bounds is not a misleading estimate of overall risk. Obtaining similar results for different case studies suggests that these conclusions apply to more typical mixes of carcinogens.
>
> Central estimates of the overall risk can differ from the sum of upper bounds by a factor of 2 - 5, as the ratio between overall upper and lower bounds decreases …
>
> In conclusion, this analysis shows that sums of plausible upper bound risk estimates do provide useful information about the overall risk from several carcinogens. The overall risk depends on the independence, additivity, and number of risk estimates, as well as shapes of the underlying risk distributions.

In response to both uncertainty and variability, EPA develops risk estimates using default assumptions based on empirical evidence or based on scientifically sound extrapolations. Further, EPA risk assessments are in fact a combination of both high-end and central tendency estimates. Consequently, the resulting risk estimates are expected to be on the high end of the range of risks but within the range of plausible outcomes. The combination of default assumptions is therefore reasonable, especially for independent factors, and does not result in exaggerated estimates. On balance, while the resulting estimates are likely to be reasonable, without a detailed uncertainty analysis it is not possible to determine where on the range of plausible outcomes the estimates actually reside.

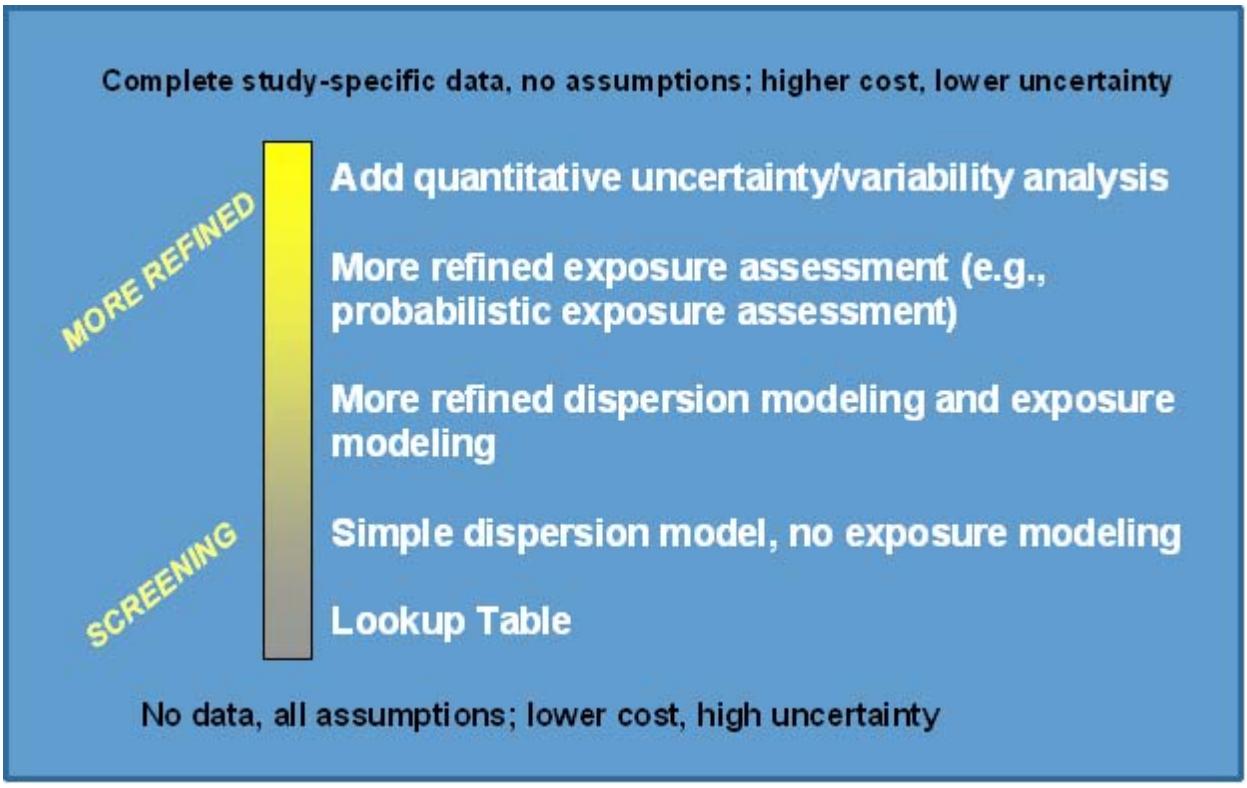
## 2.2 General Risk Assessment Approaches Used by EPA for Public and Environmental Health Protection ("Public Health Protection")

### 2.2.1 How Comprehensive Are EPA's Risk Assessments?

EPA cannot perform a time- and resource-intensive risk assessment for every situation and EPA decision. Consequently, for each risk assessment, EPA selects an approach that is consistent with the nature and scope of the decision being made. The appropriate approach depends on the needs of the decision maker and/or the role that risk information plays in the decision, balancing uncertainty and resources. Even using the best models and data, uncertainty is still inherent in the process. Given that uncertainty is inherent, there is a continuing tension between improving our understanding in order to make a decision and the reality of limited resources to perform the analysis and the desire for timely decision making. Figure 2-1 illustrates this risk assessment continuum and the balance of resources and uncertainty as the assessment becomes more complex.

**Figure 2-1. The Risk Assessment Continuum**

The following graphic illustrates that risk assessment can be performed with low levels of data and relatively little effort to develop conservative estimates of risk. Depending on the outcome and the needs of the risk manager, higher levels of analysis may be performed. Note that as one moves up the risk assessment continuum, the data needs and costs also rise. However, the quality of the result should also rise as well. (The following graphic is intended to be illustrative of the concept of tiered approaches. The actual modifications to the risk assessment that occur as the assessment is refined may vary from the sequence described here and are dependent on study-specific circumstances. In addition, a tiered approach is not prescriptive: one may begin with a screening-level assessment or begin with a higher level of analysis, as needed.)



### 2.2.2 Whom Is EPA Trying To Protect?

EPA typically cannot protect every individual but rather attempts to protect individuals who represent high-end exposures (typically around the 90$^{th}$ percentile and above) or those who have some underlying biological sensitivity; in doing so, EPA protects the rest of the population as well. In general, EPA tries to protect sensitive individuals based on normal distribution of sensitivities. EPA considers the most sensitive individuals where there are data, but does not necessarily attempt to protect "hypersensitive" individuals. The degree to which sensitive individuals are protected, or explicitly defined, may vary between programs based on factors such as the need to balance risk reductions and costs as directed and constrained by statutory authority. Programs may approach the problem semi-quantitatively (e.g., selecting individual parameter values at specified percentiles of a distribution) or qualitatively (e.g., making conservative assumptions to ensure protection for most individuals), though no overall degree of protection can be explicitly stated.

### 2.2.3 Are Risk Assessment and Risk Management Separate?

At EPA, risk assessment (evaluation of the science) and risk management (decision making, setting of policy) are not necessarily separate. We believe it is appropriate to involve decision makers from the beginning, as they typically initiate requests for risk assessments or analyses. Consequently, separating them entirely from the risk assessment process is neither logical nor desirable. Also, risk assessments typically are coordinated with the evaluations of economics, feasibility of remedies, and community concerns, for example, so that their results can be factored into decisions. EPA's Risk Characterization Handbook (USEPA, 2000a) describes in detail the roles of the risk assessor and risk manager in the risk assessment process. Further, the NRC report on understanding risk supports the concept that risk assessment is conducted for the purpose of supporting risk management, and risk management considerations shape what is addressed in a risk characterization (NRC, 1996).

Briefly, risk assessors are best qualified to understand the quality and nature of the data and to use that data to determine what the risk is, who/what is affected, the level of comfort with the conclusions, the uncertainty and variability inherent in the assessment, and the strengths and weaknesses of the assessment. In other words, they are the best qualified to make the scientific judgments necessary in risk assessments, including selecting models and data and assigning defaults where data gaps exist. Risk managers are responsible for valuing the risk and determine the amount of protection (as well as conservatism) to be applied in a decision. As a decision maker, the risk manager integrates the risk assessment with other considerations in order to make and justify regulatory decisions.

However, the comments point out that many of the default assumptions and policy choices inherent in the risk assessment may frequently not be apparent to the risk managers. It is

the role of the risk assessors to transparently characterize such details (e.g., default assumptions, data selected, policy choices) so as to make clear the range of plausible risk. The risk managers, in their role, should inquire about the use of defaults and the choices made, if not characterized clearly, in order to be fully informed before using the risk assessment in making a decision.

### 2.2.4 How Do "Planning and Scoping" Help Environmental Risk Assessment?

In environmental risk assessments, planning and scoping are performed prior to the main analytic work. This initial work defines the questions and issues to be addressed, the analysis needed to address these questions/issues, and the knowledge and information needs of the analysis.

Several general planning and scoping steps are relevant for many environmental risk assessments (see below), but specific circumstances are also important. EPA's recently completed *Framework for Cumulative Risk Assessment* (USEPA, 2003b) provides some general planning and scoping steps:

a) Defining the Purpose of the Assessment

b) Defining the Scope of Analysis and Products Needed

c) Agreeing on Participants, Roles and Responsibilities

d) Agreeing on the Depth of the Assessment and the Analytical Approach

e) Agreement on the Resources Available and Schedule

f) Problem Formulation

g) Developing the Conceptual Model

h) Constructing the Analysis Plan

Problem formulation is a systematic planning step, linked to the regulatory and policy context of the assessment, that identifies the major factors to be considered in a particular environmental assessment. The problem formulation process results in a conceptual model that identifies the sources, stressors, exposed populations, and the relationships among them.

The conceptual model and the associated narrative show the basic rationale for the decisions made concerning the course of action for the risk assessment. Specifically, the conceptual model and associated narrative provide: (1) the scientific rationale for selecting the

stressors, sources, receptors, exposed populations, exposure or environmental pathways, and endpoints/effects; (2) the scientific, technical, economic, or sociologic basis for the conceptual model; and (3) the scientific implications of additional data gathering.

The final stage in the planning and scoping process is the development of the analysis plan (see discussion in USEPA, 1998a). This plan describes how hypotheses about the relationships among the sources, stressors, exposure conditions, populations, and adverse effects/endpoints presented in the conceptual model and narrative will be considered during the risk analysis phase of the assessment. The plan includes a rationale for which relationships are to be addressed and which methods and models will be used, and discusses data gaps and uncertainties. The plan may also compare the level of confidence needed for the management decision with the confidence levels expected from alternative analyses in order to determine data needs and evaluate which analytical approach is best.

As stated in the EPA Science Policy Council's 2002 *Lessons Learned on Planning and Scoping for Environmental Risk Assessments* (USEPA, 2002b), formal planning and dialogue can improve the final risk assessment product by making it more specific to the needs of decision makers and other stakeholders. Many questions and issues may be candidates for consideration and analysis in a risk assessment. Planning and scoping help define the boundaries of the analytic work (i.e. "what's in and what's out" of the assessment), reducing ambiguities about what can and cannot be done with assessment results, and helping to reduce analytic work to manageable segments.

As important as how planning and scoping is done is *that it is done*. In general, planning and scoping are performed, to some degree, for all assessments. However, they may not be inclusive or explicit on all factors. For example, simplified analytical approaches have been developed and implemented over time. For each of these methods or analytical approaches, choices of what's in and what's out have already been made. These choices made may have been explicit when these approaches were developed, but may have become accepted over time and therefore no longer acknowledged. This may be especially true for the toxicity reference values. Some of the criticisms related to these components may stem from the lack of explicit characterizations of what they represent.

### 2.2.5 How Does EPA Use a Screening Risk Assessment?

Where data are sparse and uncertainty great, EPA carries out a screening risk assessment that tends to use default assumptions to avoid underestimating risk. These screening assessments typically provide high-end and bounding estimates. Pathways of trivial importance are then eliminated, and the remaining estimates are refined. This approach either demonstrates with minimal effort that no risk is large enough to consider reducing or, if that is not the case, it eliminates further work on refining estimates for pathways or chemicals that are clearly not

important. This is consistent with the NRC's recommendation that EPA should use bounding estimates for screening assessments to determine whether further analysis is necessary (NRC, 1994). If risks are not of concern even with these high-end or bounding screening estimates, then one can be fairly confident that the risks are not of concern. However, if these screening tests show that the risks are potentially of concern, then a more refined risk assessment may be warranted that uses more detailed data, models, etc., though at greater expense. An example is provided in section 2.2.7. These high-end screening assessments usually contain many default assumptions since data are generally not available. However, when usable data are available, they are considered instead of the defaults.

Many comments that focused on the screening assessments of risks may have misunderstood the purpose of these assessments. For example, the degree to which default assumptions are used in exposure assessments depends on the purpose of the assessment. By their very nature, screening assessments are broad in scope and based on relatively sparse data. In these cases, EPA attempts to clearly identify that the assessments are for screening purposes and to explain the meaning of the results and the utility of the assessment in the context of whatever decision is at hand.

### 2.2.6 What Happens if EPA Identifies a Potential Risk That Needs To Be Addressed After a Screening Risk Assessment?

When a screening assessment identifies the potential for a non-trivial risk, EPA decides if pursuing that risk is appropriate based on its current priorities and available resources. If the Agency decides to pursue the risk, more detailed, refined risk assessments are then performed, though the degree of refinement (i.e., where the risk assessment falls along the continuum shown in figure 2-1) depends on the type of decision, the available resources, and the needs of the decision maker. For those pathways or chemicals that were shown to be non-trivial by bounding estimates, we work to refine our estimates of exposure and dose. At this point we estimate exposures, doses, and responses that fall on the distribution of actual exposures pertinent to the population under study. In performing this continued analysis, we use a combination of data, ranges of data, distributions of data, and assumptions about each of the factors needed to estimate risk. Generally, we perform both central tendency and high-end estimates (and, increasingly, we develop fully probabilistic risk distributions). Each of these estimates is surrounded by uncertainty (perhaps unquantifiable); the degree of uncertainty depends on the quality and comprehensiveness of the available data.

For the more refined assessment, EPA would like to use appropriate and available data to generate a more data-based assessment. This poses great difficulty when the data are not available and/or adequate. If data are simply not available, then we usually employ basic default assumptions. More frequently, data are available but deemed inadequate by EPA. This creates a potentially confrontational situation if the generator of the data claims the data are adequate.

Peer review has been extremely helpful for determining the adequacy of the data, and EPA makes extensive use of this process. When peer review judges the data to be inadequate, we generally fall back on using defaults to help assess the risk. This situation probably is the basis for many of the comments that EPA does not use current data and/or chemical- or site-specific data. In fact, we do try to use the most relevant information as validated by peer review.

When exposure or dose estimates have sufficiently narrow uncertainty relative to the needs of the decision maker, we can develop the final risk assessment. Otherwise, the data or assumptions used usually have to be even further refined, if resources allow, in an attempt to further reduce uncertainty and bring the estimated exposure or dose closer to the actual values in the population. Refining the estimates usually requires that new data be considered. These data may come from other studies in the literature, information previously developed for a related purpose and adapted, or new survey, laboratory, or field data. The decision about which particular parts of the information base to refine should be based both on which data will most significantly reduce the uncertainty of the overall exposure or dose estimate of interest, and on which data are in fact obtainable either technologically or within resource constraints. After refinement of the estimate, we again determine whether the estimates provided will be sufficient to answer the questions posed to an acceptable degree, given the uncertainties that may be associated with those estimates. Refinements proceed iteratively until the assessment provides an adequate answer for the decision maker within the resources available.

### 2.2.7 How Are High-End Exposures Reflected in EPA Evaluations?

Although populations experience a range (or distribution) of exposures, a number of environmental statutes require that EPA consider those exposures at the high end of the distribution when making certain decisions. Utilizing high end exposures as one component of the risk decision making process helps to ensure equitable protection across an exposed population (i.e., protecting a high end person in a population helps ensure protection of most of the population).

One example of this is EPA's approach for evaluating a population's high end exposure to hazardous air pollutants (HAPs). This has sometimes been referred to as evaluation of "the porch potato" (i.e., the assumption that someone lives outdoors at the point of maximum concentration at or beyond the fence line of a facility for 24 hours a day for a lifetime). This section discusses the statutory and analytical frameworks that provide the basis for EPA's risk assessments for HAPs pursuant to certain provisions of the Clean Air Act (CAA).

*What Is the Statutory Basis for EPA's Selection of Exposure Scenarios When Assessing Hazardous Air Pollutant Risks*?

The Clean Air Act identifies the risk to the *individual most exposed* (IME) as the risk of interest when making certain decisions about the regulation of HAPs. Specifically, the 1990 CAA Amendments direct EPA to consider risk to the IME when determining whether a source category may be deleted from the list of sources of HAPs (Section 112(c)(9)(B)(i)), and when determining whether residual risk standards are necessary (Section 112(f)(2)(A)). Residual risk standards are the mechanism that Congress provided EPA for addressing public health and environmental risks that may remain after the regulated community has implemented technology-based standards for the control of HAPs.

In addition, Section 112(f)(2)(B) of the 1990 CAA Amendments incorporates by reference the use of a two-step risk assessment framework for setting residual risk standards under Section 112(f)(2)(A). (Note: EPA has not yet proposed any residual risk standards, but is working on a number of residual risk determinations and expects to issue the first proposal in 2004.) The two-step framework was articulated in EPA's 1989 Benzene NESHAP (54 Federal Register 38044) and consists of:

1) A first step, in which EPA ensures that risks are "acceptable." As explained in the Benzene NESHAP, in this step EPA generally limits the *maximum individual risk*, or "MIR," to no higher than approximately 1 in 10 thousand. The benzene NESHAP defines the MIR as the estimated risk that a person living near a plant would have if he or she were continuously exposed to the maximum pollutant concentrations for a lifetime (70 years).

2) A second step, in which EPA establishes an "ample margin of safety." In this step, EPA strives to protect the greatest number of persons possible to an estimated individual excess lifetime cancer risk level no higher than approximately 1 in 1 million.

In judging whether risks are acceptable and whether an ample margin of safety is provided, the benzene NESHAP states that EPA will consider not only the magnitude of individual risk, but "the distribution of risks in the exposed population, incidence, the science policy assumptions and uncertainties associated with the risk measures, and the weight of evidence that a pollutant is harmful to health." Therefore, decisions under Section 112(f)(2)(A) typically will include consideration of both population risk and individual risk, as well as other factors. (Note that the ample margin of safety analysis in the second step of the residual risk framework also includes consideration of additional factors relating to the appropriate level of control, "including costs and economic impacts of controls, technological feasibility, uncertainties, and any other relevant factors.")

*What Analytical Framework Does EPA Use to Estimate the "IME" and"MIR?"*

Risk assessment for air toxics involves using models to estimate HAP concentrations in air (and other media, as necessary), and then combining these HAP concentrations with other exposure assumptions and measures of pollutant toxicity to estimate risk (usually deterministically). As a result, OAR's risk estimates are ultimately a function of the values selected for numerous parameters, including: HAP-specific emission factors, meteorological parameters (e.g., wind speed, wind direction, precipitation, temperature), stack parameters (e.g., height, diameter, temperature, exit velocity), distance to receptor, and duration of exposure (e.g., residential occupancy period).

The exposure scenario that EPA uses to estimate the MIR is provided in the Benzene NESHAP. The Benzene NESHAP specifies that the MIR be based on exposure to the maximum pollutant concentrations for 70 years. The Benzene NESHAP indicates that it is appropriate to account for habitability when identifying the location of maximum pollutant concentrations. Therefore, EPA may characterize the maximum off-site annual average concentration in habitable areas (e.g., excluding such areas as lakes) to estimate the MIR.

OAR typically uses a tiered analytical approach when estimating risk to the IME. As discussed above, because there are many parameters that influence an individual's exposure, it is not possible to identify the actual individual in the population who is most exposed. OAR recognizes, however, that it is very unlikely that there exists an individual who should be characterized using the worst-case values of all exposure and toxicity parameters. Nevertheless, a tiered analysis often begins with a "worst-case" or bounding analysis that generally sets parameters at values that maximize the estimate of risk (e.g., exposure is assumed to continue for a lifetime). If risks estimated using such an analysis are not of concern, then there is no need to refine the analysis further, and EPA may proceed with the appropriate action (i.e., EPA may delete the source pursuant to 112(c)(9)(B)(i) or make a determination not to propose residual risk standards under 112(f)(2)(A)). If risks estimated in the initial analysis are of potential concern, analysts may make successive refinements in modeling methodologies and input data to derive successively less conservative, more realistic, estimates of the risk to the IME.

In refining the risk estimate for the IME, the exposure assumptions for which OAR has the best alternative information generally are modified first. For example, where facility-specific information is available, OAR will use the actual locations of residences (e.g., placing the receptor in the geographic center of a populated census block near a facility) to estimate the concentrations of HAPs in air. OAR also may conduct additional data gathering and analysis to further refine the characteristics of the actual emissions and emission points. The goal of successive tiers of the analysis is to start with a bounding or worst-case exposure estimate and then move away from the worst-case parameters until the combination of parameter values represents, in the judgment of the assessor, the exposure experienced by the individual in the

population who is most exposed. However, where the risk estimates are below the level of concern, there is no need to conduct additional analysis to refine the estimates further, even if the estimates might still overestimate the risk to the individual in the population who is most exposed.

As described above, the risk estimate is a function of the values of a number of parameters. In refining the bounding or worst-case risk analysis, OAR has not modified the assumption of 70-year, 24-hour per day, outdoor exposure. Although OAR recognizes that the majority of people do not reside outdoors and in one location for their entire lives, and that there is a general trend of increased population mobility, OAR believes that the data available for refining assumptions for exposure duration and frequency are less certain than the data available for refining other parameter values. For example:

a) Residential occupancy periods may be influenced by factors (e.g., economic, geographic) that may cause local population mobility patterns to differ from national estimates of population mobility.

b) If a source of concern occurs in the majority of communities in the country, then it is possible that an individual may be exposed to the source for a longer period of time than one might predict using national estimates of population mobility. That is, even though an individual moves, the individual's new residence may be located near a similar source of concern.

c) If a single source impacts a large geographic area, then it is possible that an individual may move or travel from one point of exposure to the source to another point of exposure to the same source.

d) Exposure to HAPs may not diminish when individuals are indoors. Empirical data for many pollutants show that long-term average indoor concentrations of outdoor air pollutants are roughly equivalent to long-term average outdoor concentrations of those pollutants (e.g., Sexton et al., 2004).

Moreover, OAR is cognizant that NRC acknowledged some of these issues in its 1994 report, which recommended that: "EPA should use the mean of current life expectancy as the assumption for the duration of individual residence time in a high-exposure area, or a distribution of residence times which accounts for the likelihood that changing residences might not result in significantly lower exposure. Similarly, EPA should use a conservative estimate for the number of hours a day an individual is exposed, or develop a distribution of the number of hours per day an individual spends in different exposure situations."

# 3. UNCERTAINTY AND VARIABILITY

## 3.1 Overview

Uncertainty and variability exist in all risk assessments. Even at its best, risk assessment does not estimate risk with absolute certainty. Thus, it is important that the risk assessment process handle uncertainties in a predictable way that is scientifically defensible, consistent with the Agency's statutory mission, and responsive to the needs of decision makers (NRC, 1994). Instead of explicitly quantifying how much confidence there is in a risk estimate, EPA attempts to increase the confidence that risk is not underestimated by using several options to deal with uncertainty and variability when data are missing. For example, in exposure assessment, the practice at EPA is to collect new data, narrow the scope of the assessment, use default assumptions, use models to estimate missing values, use surrogate data (e.g., data on a parameter that come from a different region of the country than the region being assessed), and/or use professional judgment. The use of individual assumptions can range from qualitative (e.g., assuming one is tied to the residence location and does not move through time or space) to more quantitative (e.g., using the 95$^{th}$ percentile of a sample distribution for an ingestion rate). This approach can also fit the practice of hazard assessment when data are missing. Confidence in ensuring that risk is not underestimated has often been qualitatively ensured through the use of default assumptions.

EPA has been increasingly concerned about characterizing uncertainty in its risk estimates. EPA's 1986 set of Risk Assessment Guidelines explicitly stated the importance of characterizing uncertainty. EPA's Exposure Assessment Guidelines developed this theme further for the exposure assessment part of risk assessment. EPA's Risk Characterization Policy provided even more direction for describing uncertainty in risk estimates. For probabilistic analysis specifically, EPA made significant efforts in recent years to use probabilistic techniques to characterize uncertainty; these include the March 1997 *Guiding Principles for Monte Carlo Analysis* (USEPA, 1997b), the May 1997 Policy Statement (USEPA, 1997c), and the December 2001 Superfund document *Risk Assessment Guidance for Superfund: Volume III — Part A, Process for Conducting Probabilistic Risk Assessment* (USEPA, 2001a)

## 3.2 Uncertainty and Variability

### 3.2.1 What Is Uncertainty?

Uncertainty can be defined as a lack of precise knowledge as to what the truth is, whether qualitative or quantitative.

Numerous schemes for classifying uncertainty have been proposed. The preferred approach of the NRC (1994) focused on two broad categories: parameter uncertainty and model uncertainty. These are defined below in sections 3.2.2 and 3.2.3.

### 3.2.2 What Is Parameter Uncertainty?

Risk assessments depict reality interpreted through mathematical representations that describe major processes and relationships. Process or mechanistic models use equations to describe the processes that an environmental agent undergoes in the environment in traveling from the source to the target organism. Mechanistic models have also been developed to represent the toxicokinetic and toxicodynamic processes that take place inside the organism, leading to the toxic endpoint. The specific parameters of the equations found in these models are factors that influence the release, transport, and transformation of the environmental agent, the exposure of the target organism to the agent, transport and metabolism of the agent in the body, and interactions on the cellular and molecule levels. Empirical models are also used to define relationships between two values, such as the dose and the response. Uncertainty in parameter estimates stem from a variety of sources, including:

a) Measurement errors:

   1) Random errors in analytical devices (e.g., imprecision of continuous monitors that measure stack emissions).

   2) Systemic bias (e.g., estimating inhalation from indoor ambient air without considering the effect of volatilization of contaminants from hot water during showers).

b) Use of surrogate data for a parameter instead of direct analysis of it (e.g., use of standard emission factors for industrialized processes).

c) Misclassification (e.g., incorrect assignment of exposures of subjects in historical epidemiologic studies due to faulty or ambiguous information).

d) Random sampling error (e.g., estimation of risk to laboratory animals or exposed workers in a small sample).

e) Non-representativeness (e.g., developing emission factors for dry cleaners based on a sample of “dirty” plants).

### 3.2.3 What Is Model Uncertainty?

Model uncertainties arise because of gaps in the scientific theory that is required to make predictions on the basis of causal inferences. Common types of model uncertainties in various risk assessment–related activities include:

a) Relationship errors (e.g., incorrectly inferring the basis of correlations between chemical structure and biological activity).

b) Oversimplified representations of reality (e.g., representing a three-dimensional aquifer with a two-dimensional mathematical model).

c) Incompleteness, i.e., exclusion of one or more relevant variables (e.g., relating asbestos to lung cancer without considering the effect of smoking on both those exposed to asbestos and those unexposed).

d) Use of surrogate variables for ones that cannot be measured (e.g., using wind speed at the nearest airport as a proxy for wind speed at the facility site).

e) Failure to account for correlations that cause seemingly unrelated events to occur more frequently than expected by chance (e.g., two separate components of a nuclear plant are both missing a particular washer because the same newly hired assembler put them together).

f) Extent of (dis)aggregation used in the model (e.g., whether to break up fat compartment into subcutaneous and abdominal fat in a physiologically based pharmacokinetic, or PBPK, model).

Model uncertainty is often difficult to quantify. Further, it is inherent in risk assessment that seeks to capture the complex processes impacting release, environmental fate and transport, exposure, and exposure-response. EPA's models are often incomplete and knowledge of specific processes limited. As a result, EPA relies on specific default assumptions as a response to uncertainty. Again, these represent scientifically plausible choices that are intended to more likely guard against underestimating risks.

### 3.2.4 How Does Variability Differ From Uncertainty?

Variability is considered with uncertainty since, like uncertainty, it does not allow for a specific single correct value. However, they differ significantly in terms of their impact and role in risk assessments. Uncertainty — the lack of knowledge — can be reduced through additional

investigation. Variability is inherent heterogeneity across space, in time, or among individuals; it cannot be reduced with additional investigation, only better understood or characterized.

"Human variability" refers to person-to-person differences in biological susceptibility or in exposure. Although both human variability and uncertainty can be characterized as ranges or distributions, they are fundamentally different concepts. Again, uncertainty can be reduced by further research that supports a model or improves a parameter estimate, but human variation is a reality that can be better characterized, but not reduced, by further research. (Note that fields other than risk assessment use "variation" or "variability" to mean dispersion about a central value, including measurement errors and other random errors that risk assessors address as uncertainty.)

EPA addresses variability by assessing the risk to the sensitive portions of the population or ecosystem (see section 3.6 for fuller discussion). Accordingly, EPA typically makes explicit choices to characterize the risks at the upper end of the expected distribution. Typically, EPA focuses on the critical parameter or assumption in any particular aspect of an assessment and based on the particular needs of the decision maker, including the mandates and constraints in statutory language, and selects the appropriate choice — such as the maximum or a specified percentile. In general, EPA's approach to variability has focused on exposure (e.g., characterizing the risks to the *individual most exposed* within the Air Toxics or Hazardous Air Pollutant program, section 2.2.7), but the approach may also be reflected in the toxicity part of the assessment (e.g., UF for human variability).

## 3.3 Characterizing Uncertainty and Variability

One of the major comments on EPA risk assessment practices is that they do not characterize uncertainty and variability transparently enough. This is an issue EPA is attempting to address. The comments also encourage EPA to begin using probabilistic analysis to describe quantitatively the uncertainty in risk estimates where appropriate.

### 3.3.1 Why Is Characterizing Uncertainty and Variability Important?

The very heart of risk assessment is the responsibility to use whatever information is at hand or can be generated to produce an estimate, a range, a probability distribution — whatever best expresses the present state of knowledge about the effects of some hazard in some specific setting. To ignore the uncertainty in any process is almost sure to leave critical parts of the process incompletely examined and hence to increase the probability of generating a risk estimate that is incorrect, incomplete, or misleading (NRC, 1994).

The NRC (1994) further noted that risk assessments that do not pay sufficient attention to uncertainty are vulnerable to four common, potentially serious pitfalls:

a) They do not allow for optimal weighting of the probabilities and consequences of error.

b) They do not permit a reliable comparison of alternative decisions.

c) They fail to communicate the range of control options that would be comparable with different assessments of the true state of nature.

d) They preclude the opportunity for identifying research initiatives.

Further, uncertainty analysis will play a more prominent and formal role in regulatory decision making. For example, OMB's recent revisions to its regulatory analysis guidelines (USOMB, 2003c) state that formal quantitative uncertainty analysis be performed for economic assessment in support of overall regulatory analysis. Whenever possible, appropriate statistical techniques should be used to determine the probability distribution of relevant outcomes. For major rules involving annual economic effects of $1 billion or more, a formal quantitative analysis of uncertainty is required. The OMB guidelines outline analytical approaches of varying levels of complexity, which could be used for uncertainty analysis such as qualitative disclosure, numerical sensitivity analysis, and formal probabilistic analysis (required for rules with impacts greater than $1 billion).

### 3.3.2 When Should EPA Conduct Uncertainty Analysis?

As the NRC notes, EPA's analysis of uncertainty has tended to be piecemeal and highly focused on an assessment's sensitivity to the accuracy of a few specified assumptions rather than a full exploration of the process from data collection to final risk assessments (NRC, 1994). Further, EPA tends to conduct uncertainty analysis *a posteriori,* focusing on specific assumptions or variables. The level of treatment again depends on the decision and the value such an analysis would provide to that decision. EPA believes this is an appropriate approach to balance resources with value added. Some have suggested that EPA should integrate analysis and uncertainty into the basic development of the primary study (see USEPA, 2000e). Until such methods and supporting data are developed, though, it is not feasible to do a full and integrated assessment for every analysis. Further, in most instances EPA is not the data developer: much EPA analysis is based upon third-party literature.

Ideally, an uncertainty analysis would be built into the study design; hopefully, as methods and supporting data continue to be developed, we will move toward such an approach in the future.

### 3.3.3 What Is an Appropriate Level of Uncertainty Analysis?

Over the years, improved computer capabilities have created more opportunities to characterize uncertainty. As a result, advocates promote such characterization in all cases. We need to be judicious in which methods we apply, such as Monte Carlo analysis. Uncertainty analysis is not a panacea, and full formal assessments can still be time- and resource-intensive. Further, the time and resources needed to collect an adequate database for such analyses can be a problem. While uncertainty analysis arguably provides significant information to aid in decision making, its relative value is case-specific and depends on the characteristics of the assessment and the decision being made. In some cases, a full probabilistic assessment may add little value relative to simpler forms. This may occur where more detailed uncertainty analysis (or analysis focused on non-critical uncertainties) does not provide information which has any impact on the overall decision.

Accordingly, EPA's practice is to use a "tiered approach" to conducting uncertainty analysis; that is, EPA starts as simply as possible (e.g., with qualitative description) and sequentially employs more sophisticated analyses (e.g., sensitivity analysis to full probabilistic), but only as warranted by the value added to the analysis and the decision process. Questions regarding the appropriate way to characterize uncertainty include:

a) Will the quantitative analysis improve the risk assessment?

b) What are the major sources of uncertainty?

c) Are there time and resources for a complex analysis?

d) Does this project warrant this level of effort?

e) Will a quantitative estimate of uncertainty improve the decision? How will the uncertainty analysis affect the regulatory decision?

f) How available are the skills and experience needed to perform the analysis?

g) Have the weaknesses and strengths of the methods involved been evaluated?

h) How will the uncertainty analysis be communicated to the public and decision makers?

### 3.3.4 With What Precision Should EPA Results Be Reported?

Uncertainty is inherent in almost all the data EPA reviews. However, at the end of many risk assessments, values are provided that appear to have several significant digits. This implies a certain level of precision that probably is not justified. EPA is currently addressing this practice of characterizing the values.

With the increased use of computational methods and readily available software, estimates can be reported with apparently unlimited precision. Nonetheless, the precision to which a value is reported should be commensurate with the level of confidence in that number. EPA has been criticized in the past for reporting risk estimates with unnecessary and unsupportable degrees of precision. Often, this is the result of automated reporting within commercial software packages, or of incremental stages of analysis that are not reviewed and edited in accordance with their incremental nature.

EPA agrees that results should be reported to their proper level of precision and the value(s) well characterized. EPA guidance on presentation of risk information comes from various documents developed since 1989. The *Risk Assessment Guidance for Superfund: Volume I — Part A* (USEPA, 1989a) provides guidance on the number of significant figures to use in presenting cancer risks in the risk characterization. RAGS Part A's page 8-7 (Exhibit 8-2, footnote b) states that "All cancer risks should be expressed as one significant figure only." In addition, the *Risk Assessment Guidance for Superfund: Volume 1 — Part D* (USEPA, 2001b) provides standardized tables for presenting risk information. The Record of Decision guidance provides standardized tables for presenting risk information (USEPA, 1999a); the same document provides guidance on presenting non-cancer hazard indices.

In a final risk description, all cancer risks and non-cancer hazard indices should be presented with one significant figure only. More figures, however, should be carried along the way to minimize rounding errors and to make it possible for others to verify calculations. The number of significant figures presented during a risk assessment changes to reflect the needs of the assessment. For example, risk information presented in RAGS Part D tables (tables 6 through 9 in RAGS) typically present cancer risk information with two significant figures. This summary information allows risk assessors reviewing the document to check the accuracy of the cancer risk calculations and provide appropriate comments.

## 3.4 Issues in Characterizing Uncertainty and Variability

EPA has been examining its practice of uncertainty analysis over the years. Generally speaking, EPA has applied a qualitative approach to characterizing uncertainty (and variability for that matter). Quantitative characterization has been used more unevenly. Also, EPA

typically uses deterministic approaches to characterize risk — although, increasingly often, EPA applies probabilistic techniques for characterization of risk, usually within exposure assessments.

### 3.4.1 What Is the Issue of Deterministic Versus Probabilistic Approaches?

Risk assessments may consider both non-cancer and cancer endpoints when data are available. A non-cancer assessment includes an oral reference dose (RfD) and an inhalation reference concentration (RfC). A cancer assessment includes a cancer weight-of-evidence narrative statement and, if data permit, a quantitative cancer risk estimate. Many of the risk estimates from these assessments are found in the Agency's Integrated Risk Information System (IRIS) as well as in EPA offices that also develop health assessments..

As an example for the practice of characterizing uncertainty, we examine the traditional approach to dose-response assessment analysis of non-cancer endpoints (RfD or RfC). The approach is to identify the no-observed-adverse-effect level (NOAEL) and lowest-observed-adverse-effect level (LOAEL) from an appropriate study. The NOAEL (or LOAEL, if a NOAEL is not available) is adjusted downward by uncertainty factors (UFs) intended to account for uncertainties in the available data, producing an exposure that is likely to pose no increased risk of adverse effects for chronic exposure. EPA currently also uses the benchmark dose (BMD) approach to deriving a RfD/RfC when appropriate data are available.

While there may be significant statistical analysis and consideration of (both qualitative and quantitative) uncertainty, an estimated reference value or exposure is largely developed via a deterministic process and yields a deterministic estimate. Part of this approach is the use of UFs that are applied to adjust for uncertainties in extrapolating from the type of study serving as the basis for the RfD/RfC to the situation of interest for the risk assessment. UFs are used to account for each of the extrapolations used in the assessment, such as use of animal data (interspecies extrapolation) and protection of susceptible individuals (intraspecies variability).

A deterministic value for "a dose or exposure not expected to cause adverse effects" may convey an inappropriate degree of precision, and some think that a single number is not appropriate. A single number is frequently presented as "the risk assessment"; it is clear this is not an adequate way to characterize any potential risk. Several considerations need to be characterized to increase confidence in the assessment (i.e., to reduce uncertainty):

a) The degree of confidence needed in stating that susceptible/sensitive individuals are affected or not above the resulting exposures.

b) The condition of the database from which the original BMD, NOAEL, or LOAEL was developed (i.e., the quality of the individual study on which the value is based).

c) The numerical values for the UFs employed in calculating the RfC/RfD.

d) The procedure by which the BMD, NOAEL, or LOAEL from the experimental system may have been adjusted to account for known differences in the human population of interest (e.g., dosimetric adjustments).

e) The shape of the dose-response curve.

A probabilistic approach could be advantageous in that it would eliminate the need to define a single value and might be less likely to imply undue precision. In such a framework, a probability distribution would be used to express the belief that any particular value represents the dose or exposure concentration that would pose no appreciable risk of adverse effects. In contrast with the imprecision of a deterministic approach, several researchers still attempt to interpret what is meant by the RfD or RfC and what is the degree of protection afforded (e.g., Swartout et al., 1998; Baird et al., 1996; Gaylor and Kodell, 2000).

### 3.4.2 What Is the Importance of Quantitative Characterization of Uncertainty in Dose-Response?

EPA has stressed the importance of attention to uncertainty in risk assessment. We quote extensively here from the successive versions of EPA's Cancer Guidelines, since they articulate the issue well in terms of cancer risk assessment. Many of the points apply to this issue for risk assessment in general. The 1986 Cancer Guidelines state:

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

At that time EPA recognized the need for further development of methodologies to address uncertainties:

> It is also recognized that there is a need for new methodology that has not been addressed in this document in a number of areas, e.g., the characterization of uncertainty. As this knowledge and assessment methodology are developed, these Guidelines will be revised whenever appropriate.

The 1999 draft revised Cancer Guidelines contain substantial additional discussion about uncertainty. In the key area of dose-response assessment, they state:

> The characterization presents the results of analyses of dose data, of response data, and of dose-response. When alternative approaches are plausible and persuasive in selecting dose data, response data, or extrapolation procedures, the characterization follows the alternative paths of analysis and presents the results.

This discussion goes on to address specific areas where uncertainties, including quantitative uncertainties, are to be considered:"Uncertainty analyses, qualitative or quantitative if possible, are highlighted in the characterization." The guidelines continue with a list of specific factors to include in the characterization.

The 2003 draft final Cancer Guidelines contain a further expansion in the discussion of uncertainty issues:

> Some aspects of model uncertainty that should be addressed in an assessment include the use of animal models as a surrogate for humans, the influence of cross-species differences in metabolism and physiology, the use of effects observed at high doses as an indicator of the potential for effects at lower doses, the effect of using linear or nonlinear extrapolation to estimate risks, the use of small samples and subgroups to make inferences about entire human populations or subpopulations with differential susceptibilities, and the use of experimental exposure regimens to make inferences about different human exposure scenarios (NRC, 2002).
>
> Toxicokinetic and toxicodynamic models are generally premised on site concordance across species … The assessment should discuss the relevant data that bear on this form of model uncertainty. …
>
> Probabilistic risk assessment, informed by expert judgment, has been used in exposure assessment to estimate human variation and uncertainty in lifetime average daily dose. Probabilistic methods can be used in this exposure assessment application because the pertinent variables (for example, concentration, intake rate, exposure duration, and body weight) have been identified, their distributions can be observed, and the formula for combining the variables to estimate the lifetime average daily dose is well defined (see USEPA, 1992a). Similarly, probabilistic methods can be applied in dose-response assessment when there is an understanding of the important parameters and their relationships, such as identification of the key determinants of human variation (for example, metabolic polymorphisms, hormone levels, and cell replication rates), observation of the distributions of these variables, and valid models for combining these variables. With appropriate data and expert judgment, formal approaches to probabilistic risk assessment can be applied to provide insight into the overall extent and dominant sources of human variation and uncertainty. In doing this, it is important to note that analyses that omit or underestimate some principal sources of variation or uncertainty could provide a misleadingly narrow description of the true extent of variation and uncertainty and give decision-makers a false sense of confidence in estimates of risk. Specification of joint probability distributions is appropriate when variables are not independent of each other. In each case, the assessment should carefully consider the questions of uncertainty and human variation and discuss the extent to which there are data to address them.
>
> Probabilistic risk assessment has been used in dose-response assessment to determine and distinguish the degree of uncertainty and variability in toxicokinetic and toxicodynamic modeling. Although this field is

> less advanced than probabilistic exposure assessment, progress is being made and these guidelines are flexible enough to accommodate continuing advances in these approaches.

### 3.4.3 How Does EPA Use Probabilistic Analysis?

EPA cancer and other risk assessments have not included full probabilistic uncertainty analyses to date, primarily due to the need to develop relevant probability distributions in the toxicity part of risk assessment. However, quantitative statistical uncertainty methods are routinely applied in evaluation of fitting of dose-response models to tumor data, and quantitative uncertainty methods have been used to characterize uncertainty in pharmacokinetic and pharmacodynamic modeling.

As stated in EPA's *Policy for Use of Probabilistic Analysis in Risk Assessment* (1997c):

> For human health risk assessments, the application of Monte Carlo and other probabilistic techniques has been limited to exposure assessments in the majority of cases. The current policy and associated guiding principles are not intended to apply to dose-response evaluations for human health risk assessment until this application of probabilistic analysis has been studied further.

The EPA's Science Advisory Board (SAB) suggests that the Agency reconsider this policy (see USEPA, 2000e) for dose-response evaluations for human health risk assessments.

Generally, "probabilistic analysis" is a means for describing the uncertainty in risk estimates by characterizing the uncertainty and population variability in the individual steps by probability distributions. Thus, the likelihood of each risk is quantitatively characterized in the resulting estimates. This is generally implemented by a "Monte Carlo" approach, which performs a computer simulation to produce the estimates.

EPA provided guidance on the use of probabilistic analysis and made significant efforts to use probabilistic techniques to characterize uncertainty. In recent years, as stated above, EPA developed the March 1997 *Guiding Principles for Monte Carlo Analysis,* the May 1997 Policy Statement, and the December 2001 Superfund document *Risk Assessment Guidance for Superfund: Volume III — Part A, Process for Conducting Probabilistic Risk Assessment*. Before performing a probabilistic risk assessment, one should carefully examine the available data to determine if they are adequate for the analysis and to consider the costs associated with the probabilistic risk assessment relative to the proposed action.

There is still more work to be done, particularly in the hazard assessment area and in combining the hazard assessment and exposure assessment results. Probabilistic risk assessment could probably be used more frequently; it could provide useful information for many risk assessments beyond screening-level assessments. Probabilistic risk assessment incorporating the entire process, including the toxicity part of the risk assessment and not only the exposure

component, would be particularly useful. In addition, probabilistic analysis may avoid some potential problems of apparent overestimation of risk estimates from multiplying UFs in a deterministic risk assessment. (Most of the time, there is more of a tendency for the estimate to be higher than the "true" risk rather than lower. In individual cases, however, it is most certainly possible that the final, overall risk estimates will underestimate "true" risk for a particular risk assessment.)

We recognize that probabilistic analysis is no panacea. It will not necessarily result in different outcomes or decisions, nor are its implementation and use simple processes. The accuracy of probabilistic analysis will still depend upon the quality of the data used for the analysis. Moreover, in some situations, a probabilistic analysis may not be appropriate. For example, at cleanup sites (such as Superfund, Resource Conservation and Recovery Act, or Brownfields sites), the resident populations may not have the wide-ranging susceptibilities that would warrant conducting a probabilistic risk assessment.

Although EPA provided guidance for some use of probabilistic techniques, and facilitated its use via the May 1997 policy statement, the Task Force agrees that EPA should attempt, where appropriate, to make more use of probabilistic techniques: they could provide useful information on both variability in exposures and uncertainty in the estimates for many risk assessments that go beyond screening-level assessments. In addition to making variability and uncertainty more transparent (e.g., sometimes allowing disaggregation of the effects of variability and uncertainty by "nested" simulations), probabilistic analysis can provide more consistency in characterizing risks. Further, probabilistic analysis may avoid some potential problems of upwardly biased estimates of risk from combining uncertainty/variability input variables in a deterministic risk assessment. (By "upwardly biased," it is meant that, *on average,* there is more of a *tendency* for the estimate to be higher than the "true" risk rather than lower. In individual cases, however, it is most certainly possible that the final, overall risk estimates will underestimate "true" risk for a particular risk assessment.)

### 3.4.4 How Does Research Help Reduce Uncertainty?

The NRC strongly stated (NRC, 1994) and EPA agrees that risk assessment is intended to be an iterative process that promotes further investigation (research) to reduce uncertainties. Once available, this improved information is applied to the risk assessment process to refine the assessment. The information obtained from research improves confidence in the risk assessment and reduces the associated uncertainty. Further, the use of defaults does not preclude, reduce, or eliminate the push for new science.

Such an approach suggests that the responsibility to push for or conduct research be placed on those imposing the risks (and most likely benefitting from the activity) rather than the

general public or the population at risk, which may not have the resources to promote such research.

## 3.5 Inherent Variability in Biological Response

### 3.5.1 Why Consider Populations and Life-Stages?

Consideration of the variability among humans is a critical aspect of risk assessment. It is the goal of EPA risk assessments to identify all potentially affected populations, including human populations (e.g., gender, nutritional status, genetic predisposition) and life-stages (e.g., childhood, pregnancy, old age) that may be more susceptible to toxic effects or are highly or disproportionately exposed (e.g., children, ethnic groups; see USEPA, 1992a, 2000a, 2002c). While susceptible ecosystems and ecological entities (e.g., endangered species) are also considered in ecological assessments, the following discussion focuses on human populations and life-stages.

The term "susceptible" is also used to describe *sensitive* populations/life-stages, as these two terms are often used interchangeably and no convention for their use is widely accepted. Although "susceptible" has been used to describe susceptibility to toxic effect(s) and disproportionate or unique exposures, it is more transparent to identify the two issues separately in risk assessment.

The NRC recognized that many types of variability enter into risk assessment, including the nature and intensity of exposure and susceptibility to toxic insult as affected by age, lifestyle, genetic background, ethnicity, and other factors (NRC, 1994). The NRC outlined several recommendations for considering such variability in risk assessment. The NRC also made recommendations specific to risk assessments for children and infants as a potentially susceptible group (NRC, 1993).

The consideration of populations and life-stages has also been described through the concept of vulnerability, or the propensity for the system to suffer harm (USEPA, 2003b). Factors included in vulnerability are:

a) *Susceptibility or sensitivity to adverse effect:* increased likelihood of sustaining an adverse effect as a relationship to a factor describing the population (e.g., genetic polymorphisms, prior immune reactions, disease state, or prior damage) and/or life-stage.

b) *Differential exposure:* differences in current exposure, historical body burden, and background exposure.

c) *Differential preparedness to withstand a stressor:* linked to coping systems and resources of an individual, population, or community with respect to prospective mitigation efforts (e.g., immunization).

d) *Differential ability to recover:* linked to coping systems (e.g., differential survival rates).

Legislation has called on EPA to consider potentially susceptible populations and life-stages. The SDWA Amendments mandate that EPA consider risks to groups within the general population that are identified as being at greater risk of adverse health effects, including children, the elderly, and people with serious illness (SDWAA, 1996). Similarly, the Food Quality Protection Act (FQPA) contains special provisions for the consideration of pesticide risks to children (FQPA, 1996). In addition to legislative mandates, EPA has further guidance in considering health and safety risks to children (EO 13045, 1997; USEPA, 1995b).

### 3.5.2 Are Certain Populations and Life-Stages Always at Greater Risk?

When conducting risk assessments, EPA examines populations and life-stages that may be especially sensitive to the stressor(s) being assessed. This does not imply that these examined populations and life-stages are always at greatest risk. For each stressor and exposure scenario, different data are available such that there is not a single or exact method for examining potential susceptibility and associated risk. Risk assessment often uses an iterative approach. Populations and life-stages may be assessed using defaults and assumptions in screening-level assessments, while more detailed analyses of these groups will be performed for more refined assessments. While the question "Are certain populations and life-stages at greater risk?" needs to be asked in all iterations of risk assessment, the answer may not remain the same throughout.

### 3.5.3 How Are Sensitivities to Toxic Effects Considered?

When data are available to describe the toxicological differences for a susceptible population or life-stage, then those data are summarized and analyzed, and the decisions based on this information are presented. It is preferable to have population- and chemical-specific data to describe a susceptibility to toxic effects. For example, the IRIS RfC for beryllium is based on the human subpopulation that is susceptible to chronic beryllium disease (USEPA, 1998b; also see section 3.5.5 below). Similarly, if data are available to indicate that susceptible populations or life-stages are not at risk, those data are also used.

When one is analyzing and describing the available toxicity and effects data for susceptible populations or life-stages, pragmatic considerations include:

a) Are epidemiologic, toxicologic, mechanistic, or other studies available that investigate reproductive effects, developmental effects or windows of susceptibility, multi-generation effects, gender or race differences, genetic predisposition, modulation of existing disease, old age, or other possible susceptibilities? Were qualitative (i.e., different effect) or quantitative (i.e., same effect at lower dose or more severe effect at same dose) differences in susceptibility observed? Both?

b) Have the available studies examined the likely and plausible endpoints?

c) Has a distribution been used in assessing toxicity (e.g., benchmark dose lower confidence level)? Is it feasible that an entirely different distribution exists for susceptible populations or life-stages, such that selecting a central value or even tail of the existing distribution is not necessarily representative of susceptibility?

d) How have the findings of available studies been incorporated in the dose-response assessment? Is it clear how susceptible populations and life-stages have been considered in the quantitative assessment of toxicity?

e) What data are absent/needed? Can information from existing studies be synthesized to identify specific data gaps?

It has been recognized that limited data are currently available for the *a priori* identification of susceptible populations and life-stages for many chemicals and risk assessments (USEPA, 2002c). In these situations, it is important that risk assessments clearly identify and summarize data needs and uncertainties, in addition to the available data. Typically, when data are limited, default practices are used:

a) *Non-cancer effects:* An intraspecies UF is used to account for variations in susceptibility within the human population (USEPA, 2002c). This UF typically has a value of 10-fold, but can be increased or reduced when sufficient data are available. One can apply the same UF to carcinogens using a non-linear dose-response model. For example, the IRIS chloroform oral carcinogenesis assessment considers the non-cancer assessment to be protective against cancer risk, and the same intraspecies UF is applied in the chloroform oral cancer and non-cancer assessments (USEPA, 2001c).

A database UF may also be applied for deficiencies in the available data or when existing data suggest that additional data may yield a lower reference value (USEPA, 2002c). This UF is most often used when developmental or two-generation reproduction studies are not available, but it may be applied in other

situations to account for the lack of data for potentially susceptible populations or life-stages.

b) *Cancer effects:* An evaluation should be made as to whether low-dose linear extrapolation is sufficient to protect susceptible populations (USEPA, 2003c). For example, available data indicate that early life exposure to mutagenic carcinogens may lead to greater incidence of cancer than the same lifetime average daily dose received later in life. This information has been reviewed by EPA and published in draft (USEPA, 2003d).

### 3.5.4 Are Unique or Disproportionate Exposures Considered?

Unique or disproportionate exposures are considered in parallel to the toxicity considerations described in the preceding section. Available data describing potential unique or disproportionate exposure are summarized and decisions based on this information are presented. Again, it is preferable to have population- and chemical-specific data to address specific exposure scenarios. For example, the Integrated Exposure Uptake Biokinetic Model (IEUBK) is a very specialized tool that allows detailed analysis of lead exposures for children ages 6 months to 7 years (USEPA, 2001d).

While some risk assessments consider exposure to the population as a whole, many only consider exposed populations, not unexposed members of the general population. For example, a hazardous waste assessment only considers the people who live near or on the site, or near the incinerator or point source involved.

For exposure assessment, pragmatic considerations for unique and disproportionate exposures include:

a) Are there unique or disproportionate exposures based on age, race, ethnicity, gender, lifestyle, cultural practices, economic status, or other considerations?

b) How well do existing studies or distributions of exposure consider identified populations or life-stages?

c) Have the exposures within these groups been considered in detail? For example: pica soil ingestion by children, consumption of breast milk by infants, traditional or subsistence diets, mouthing of objects by toddlers, different/increased food and water consumption rates for children, restrictions on mobility (e.g., children cannot leave a location or residence by choice as adults can; elderly populations often have reduced mobility).

d) Have unique or disproportionate childhood exposures been integrated for lifetime exposure? Children are not children for a chronic duration, but all adults spent a portion of their lifetime in childhood; this needs to be taken into account as a source of unique and disproportionate exposure as a part of a continuous lifetime.

When exposure data for certain populations and life-stages are limited, exposure assumptions and default values are used to assess plausible current and/or future exposure scenarios. The exposure factors published in EPA's *Exposure Factors Handbook* are widely used in Agency exposure assessments (USEPA, 1997d, 2002d). In some scenarios, additional default data may be used to describe situations the handbook does not cover. For example, exposure parameters used to estimate childhood exposures due to mouthing exposures have been used in pesticide exposure assessments (USEPA, 1997d).

It is essential that risk assessments, and exposure assessment in particular, explicitly address environmental justice concerns. The goal of environmental justice is to ensure that all people, regardless of race, national origin, or income, are protected from disproportionate impacts of environmental hazards (EO 12898, 1994). For example, concerns regarding risks associated with subsistence and the practice of traditional lifeways have been voiced by Native American people across the United States and by the National Environmental Justice Advisory Council (NEJAC, 2002). Executive Order 13175, "Consultation and Coordination With Indian Tribal Governments," requires EPA to consult with tribes regarding protection of their land and health. Tribes have specifically requested that EPA enter into formal consultation with them on environmental issues, including exposures unique to tribes. These exposures are a result of cultural or behavioral preferences of certain native populations. EPA typically bases its decisions on the general U.S. population; statistics derived from the general population's exposure patterns do not necessarily represent exposure patterns for all peoples. For example, the exposure distribution for four Columbia River Treaty Tribes (CRITFC, 1994; also, see table 3-1, below) shows that these peoples' average fish consumption rate is much higher than the rate for the average U.S. person. This disproportionate exposure leaves Native Americans in these communities with exposures that are underestimated or ignored in risk assessments.

| Table 3-1: Fish Consumption Patterns for General U.S. and Native American Populations | | | | |
|---|---|---|---|---|
| | General Public | | Columbia River Treaty Tribes | |
| Exposure Parameter | AFC | HFC | AFC | HFC |
| Tissue chemical concentration | Average | Average | Average | Average |
| Ingestion rate of fish tissue (g/day) | | | | |
| Adults | 7.5[a] | 142.4[b] | 63.2[c] | 389[d] |
| Children under 15 | 2.83[a] | 77.95[b] | — | — |
| Children under 6 | — | — | 24.8[c] | 162[d] |

AFC — average fish consumption; HFC — high fish consumption
[a] Mean U.S. per capita consumption rate of uncooked freshwater and estuarine fish (USEPA, 2000f).
[b] 99th percentile U.S. per capita consumption rate of uncooked freshwater and estuarine fish (USEPA, 2000f).
[c] Mean consumption rate for fish consumers in the Umatilla, Nez Perce, Yakama, and Warm Springs Tribes of the Columbia River Basin (CRITFC, 1994).
[d] 99th percentile consumption rate for fish consumers in the Umatilla, Nez Perce, Yakama, and Warm Springs Tribes of the Columbia River Basin (CRITFC, 1994).

Situations for unique exposures also need to be considered. For example, the Asian Pacific American population typically consumes foods that are not included in the U.S. diet survey (Sechena et al., 2003). Thus, their exposure to contaminants in seaweed or moon snails are not included in routine market basket surveys or the resulting exposure estimates.

Sometimes disproportionate exposures occur because of racial or socioeconomic conditions. For example, exposure to environmental and household agents can trigger asthma attacks (AAP, 2003); the Centers for Disease Control and Prevention reported that asthma mortality rates among African Americans are 2.5-fold higher than among Caucasians (CDC, 2003). Similarly, children of low income families constitute 83% of the children ages 1 to 5 with elevated blood lead levels that may result in health effects.

### 3.5.5 Is Human Variability Considered in Occupational Cohorts?

Occupational epidemiologic cohorts are an important source of human toxicity information for use in risk assessment. Occupational studies are conducted with select populations. Because of their presence in the workforce, these persons are believed to be healthy (this is the "healthy worker effect"), predominately male, and generally adult. These factors obviously limit a population's diversity with respect to susceptibility. As a result, the general approach is to use an intraspecies UF greater than 1 for non-cancer endpoints when the RfD or RfC is based on occupational data. This practice is consistent with the recommendation that the reduction of the intraspecies UF of 10-fold be considered only if available data are sufficiently representative of the exposure/dose-response data for the most sensitive populations and life-stages (USEPA, 2002c).

The 2003 IRIS database was examined for all instances in which human occupational data had been used to set an RfD and/or RfC. In the six instances in which occupational data had been used to set the RfD, an intraspecies factor of 10 was used three times, an intraspecies factor of 3 was used twice, and an intraspecies factor of 1 was used once. In the latter case (that of beryllium and beryllium compounds), the health effect involved is beryllium sensitization and progress to CBD. The intraspecies factor of 1 was chosen because only a small percentage of the population is sensitive to CBD and sensitization: the affected population already represented a sensitive group of people.

Occupational epidemiologic cohorts are also used in cancer assessments. For example, in the health assessment of 1,3-butadiene, the cancer slope factor (CSF) was derived using an occupationally exposed male cohort. In animal studies, mammary gland cancer was the only common tumor in mice and rats (USEPA, 2002e), so a factor of 2 was used to protect the female population. This is an example of observations from the occupational cohort — which did not include women and children — not necessarily representing the risks to the larger, more diverse human population.

### 3.5.6 What Needs To Be Done To Address the Risk Assessment of Populations and Life-Stages?

Risk characterization is the opportunity to bring together all population and life-stage considerations, including the toxicity and exposure concerns discussed in the preceding sections. A risk characterization considers all information, integrated through a logical and transparent approach, and provides a thorough discussion of the findings and related uncertainties (USEPA, 2000a). A risk characterization needs to provide a clear answer to the question, "Are there susceptible populations and life-stages?"

Overarching guidance is needed to improve the ease and consistency of risk assessment for susceptible populations and life-stages. Such guidance needs to provide more detail, methods, and advice for considering the issues discussed in this section. As a part of such efforts, EPA needs to continue developing a strategy and methodologies to address the cumulative risks of unique or disproportionate exposures.

Continued research and the availability of new types of information (e.g., genomics, improved early life animal toxicity testing, cultural practices awareness, demographics) will provide an improved ability to identify and describe susceptible populations and life-stages. It is anticipated that these advances will improve the quality of risk assessments and guidance documents — both for risk assessment in general and for susceptible groups in particular. Until such data are available, the current practice of considering available data and following Agency practice where data are limited is a reasonable approach. The Agency continues to strive to be

clear and transparent in explaining the assessment of populations, uncertainties, and life-stages in risk assessment.

## 3.6 What Is EPA Doing To Improve Assessment of Uncertainty and Variability?

### 3.6.1 How Can EPA Move Forward Toward Quantitative Characterization of Uncertainty in Dose-Response?

In successive versions of its Cancer Guidelines, EPA expressed increasing emphasis on a full examination of uncertainties, with the recognition that both qualitative and quantitative approaches to uncertainty assessment are important and can (applied appropriately) help clarify the nature of assessment findings. The use of sophisticated uncertainty tools also involves substantial issues of science and mathematics, as well as specialized issues such as the appropriate presentation and characterization of probabilistic estimates in the decision making context where appropriate. We note that active research is underway in this field, and that EPA guidelines are intended to be flexible to allow use of advances in this field as they develop. As also discussed in various places in the guidelines, substantial quantitative uncertainty analyses are also appropriate (and at a more advanced stage of development) for a number of important components of cancer risk assessments, including exposure assessment, characterization of statistical uncertainty in fits of both descriptive and biologically based dose-response models, and dosimetry estimates using pharmacokinetic data and modeling. It is not, however, EPA's intent to suggest that full probabilistic models of cancer risks are generally feasible at this time, or that the role of a qualitative presentation of uncertainties should be diminished.

EPA health scientists currently conduct research that both generates necessary quantitative data and integrates data-driven hypotheses on mode of action into biologically based models. These models simulate key biological processes and provide quantitative predictions that improve risk assessment, including cross-route extrapolation and animal-to-human extrapolation. One can use these models to develop better estimates of variability in the human population, to identify susceptible subpopulations, to help determine the applicability of proposed modes of action, and to derive likelihoods of adverse effects in humans.

### 3.6.2 Will EPA Move Toward Integrated Uncertainty Analysis?

EPA recognizes the importance of uncertainty analysis and the value of having an integrated framework for evaluating uncertainty within an overall analysis, rather than as an *a posteriori* exercise that generates deterministic estimates. EPA is working toward integrated probabilistic frameworks within risk models now being developed. Once these are implemented, EPA will be able to conduct probabilistic analyses as part of any original analysis, though it is recognized that probabilistic frameworks will be pointless without adequate data to insert.

Models that include integrated probabilistic frameworks include APEX, TRIM, and CALENDEX.

In addition, EPA is developing a flexible modeling framework, the Multimedia Integrated Modeling System (MIMS), which provides an infrastructure for modeling that supports composing, configuring, applying, and evaluating models. It can support connecting models in a controlled manner, provide graphical user interfaces to help users configure models, and manage the execution of complex systems of models. MIMS has been designed to support a wide range of models. One of the benefits of using MIMS to manage model executions is that it can initiate repetitive executions, including simulation of multiple scenarios and sensitivity, uncertainty, and optimization studies.

MIMS's sensitivity and uncertainty capabilities are designed to work with a broad range of models or modeling systems. The initial sensitivity capability will support local evaluation. Uncertainty capabilities will initially include simple Monte Carlo and Latin Hypercube Sampling. More sophisticated capabilities, such as analysis of variance for sensitivity studies, are planned for the future. In addition to handling parameter sampling and model execution, MIMS will provide data analysis tools to help the user interpret the results of the computations.

# 4. CONSIDERING INFORMATION GAPS IN HEALTH ASSESSMENTS: USE OF DEFAULT AND EXTRAPOLATION ASSUMPTIONS

## 4.1 Default Assumptions

Given the nature of uncertainty and data gaps, it is accepted practice to use default assumptions (sometimes simply called defaults) to address these uncertainties. A default assumption is "the option chosen on the basis of risk assessment policy that appears to be the best choice in the absence of data to the contrary" (NRC, 1983). The NRC, in its review of EPA risk assessment practices titled *Science and Judgment in Risk Assessment* (NRC, 1994), supported EPA's use of defaults as a reasonable way to deal with uncertainty. That report stated that EPA should have principles for choosing default options and for judging when and how to depart from them. It identified a number of criteria that NRC believed ought to be taken into account in formulating such principles, including protecting the public health, ensuring scientific validity, minimizing serious errors in estimating risks, maximizing incentives for research, creating an orderly and predictable process, and fostering openness and trustworthiness.

The first two sections (4.1.1 and 4.1.2) discuss default assumptions in general. The rest of section 4.1 discusses specific examples of defaults that have been highlighted in the comments we received.

### 4.1.1 How Does EPA Use Default Assumptions?

EPA's current practice is to examine all relevant and available data first when performing a risk assessment. When the chemical- and/or site-specific data are unavailable (i.e., when there are data gaps) or insufficient to estimate parameters or resolve paradigms, EPA uses a default assumption in order to continue with the risk assessment. Under this practice EPA invokes defaults only after the data are determined to be not usable at that point in the assessment — this is a different approach from choosing defaults first and then using data to depart from them. The default assumptions are not chemical- or site-specific, but are relevant to the data gap in the risk assessment. They are based on peer reviewed studies and extrapolation to address specific data gaps. These defaults are based on published studies, empirical observations, extrapolation from related observations, and/or scientific theory.

EPA's use of defaults is appropriate. These choices are well within the range of plausible outcomes and often at specific percentiles (for variability) within that range of observation. The aim of the risk assessment is to ensure that EPA does not underestimate risk to a chemical or stressor; the default assumptions used in the risk assessment are used to help pursue this goal.

Conversely, the intent of using default assumptions is not to produce a risk assessment that appreciably overestimates the risk being assessed.

A default is usually based on some data and/or some consensus agreement for the generic situation pertinent to the data gap being assessed. Defaults usually have undergone peer review, in which interested parties are involved or have the opportunity to be involved in those reviews. For example, many of the exposure defaults we use are found in EPA's Exposure Factors Handbook, a resource that was peer reviewed and continues to be examined and re-evaluated. When chemical- or site-specific information becomes available and is adequate to use, our risk assessments attempt to use those data rather than the default(s) — assuming those data are relevant for the specific risk being assessed, and the peer review supports their use.

There is always room for improvement in highlighting where these defaults are used in actual assessments, and one should not assume that people who read an assessment will automatically know that a default position is used (NRC, 1994). Where and when we use default assumptions to address a data gap may not be clear to many. We need to point to the places where the default was developed and explain why it is a reasonable assumption.

### 4.1.2 Are Default Assumptions Science Policy?

One of the major comments EPA received suggests that default assumptions are actually conservative policy decisions embedded in risk assessments. While default assumptions have data and/or scientific consensus supporting them, use of defaults to address data gaps in a risk assessment is a science policy decision. In keeping with the EPA's goal of protecting public health and the environment, the default assumptions are used ensure that we do not underestimate risk to chemicals and stressors. (Again, though, they are not intended to overtly overestimate a risk.) Defaults are ideally developed in a transparent manner, and we believe this is the practice of the Agency. Therefore, while defaults are indeed used in risk assessments, they are not buried (hidden) policy decisions. Recently, the GAO reflected that we use assumptions (defaults) as an unavoidable part of risk assessment because science cannot always provide clear answers to questions at various stages of an assessment (USGAO, 2001). Particular assumptions are chosen through an evaluation of available scientific information or policy decisions related to our regulatory mission and mandates. Due to the complexity of risk assessments, transparency in choice and use of assumptions is critical in risk assessments and communication efforts.

If a default assumption is a policy position that some deem too conservative or not conservative enough, this issue can be addressed during the peer review. Generally speaking, EPA's current default assumptions are positions based on data or scientific consensus and supported through peer review. Some older risk assessments may contain defaults that have not yet been reexamined in light of new data. Thus, in the context of current data, they appear inadequate (i.e., to some, they can be too conservative or not conservative enough). This is not

an inherent problem with the original risk assessments, but a process problem for EPA: EPA needs to determine the need to reexamine older assessments based on current priorities, needs, and resources.

### 4.1.3 Does Any Change Seen in Animals Indicate There Will Be a Problem for Humans?

It is generally accepted that there can be numerous changes to the recipient organism (the animal under study) following exposure to a chemical, some of which may be beneficial, adaptive, early manifestations on a continuum to toxicity, overtly toxic, or several of these things in combination. Unless there are data to indicate otherwise, a change that is considered adverse (i.e., associated with toxicity) is assumed to indicate a problem for humans.

It is recognized that a diversity of opinion exists regarding what is "adverse" versus "adaptive," both within EPA and in the general scientific community. At present, there is no Agency-wide guidance from which all health assessors can draw when making a judgment about adversity. Therefore, various experts may have differing opinions on what constitutes an adverse effect for some changes. Moreover, as the purpose of a risk assessment is to identify risk (harm, adverse effect, etc.), effects that appear to be adaptive, non-adverse, or beneficial may not be mentioned.

As a further look at this issue, an "adaptive" example is used. The human body is capable of adapting to certain toxic insults. When adaptive responses become adverse and irreversible is not yet defined. In cases where data are not available to determine when the capacities of repair mechanisms are exceeded and adaptive responses become toxic, health assessments are based on any adverse response that is deemed biologically significant. As a general principle, our practice is not to base risk assessments on adaptive, non-adverse, or beneficial events.

The IRIS database includes some examples of effects reported at the designated NOAEL level, but their toxicological significance was questioned and not relied upon in setting RfDs/RfCs; for example, mild nasal lesions and hyaline droplet degeneration in the nasal cavity. These effects were not considered of sufficient toxicological significance to warrant designating the levels that produced them as LOAELs. In these instances, the LOAELs were set at the next highest dose level, based on more severe nasal responses.

There are also differences of opinion regarding the severity of effects and their qualitative and quantitative relationship to more overt toxicity. To explore past Agency practice, we performed a series of IRIS database searches on a selected example: "liver weight" as the sole critical effect, expanded slightly to incorporate examples in which liver weight appears to be listed as the predominant critical effect, along with other effects. The results indicate that liver weight change has been listed as a sole critical effect in the IRIS database in seven profiles (and

as a predominant effect in three more; examples dating from 1987 to 1994). More in-depth examination of these IRIS records generally indicates that at least one additional toxicity consideration was incorporated into the Agency decision at that time. This additional toxicity information comes in three forms: (1) listing of co-critical adverse events; (2) supporting information on adverse effects from other studies, at lower, similar, or slightly higher doses, that have not been reported in the IRIS critical events column; and/or (3) a clear progression of toxicity at slightly higher doses. All three sources of additional toxicity information may be discussed in a particular IRIS file, with the predominant information being supportive toxicity data at similar doses in other studies. The extent to which additional toxicity information is available and/or reported varies among the chemical assessments. This examination does not rule out the possibility of using liver weight (or other changes) as a critical effect if it is deemed to contribute to adversity in an assessment.

### 4.1.4 Are Benign (Histologically Non-Malignant) Tumors Presumed To Have Potential To Progress to Malignant Tumors, and Are They Counted as if They Were Malignant?

When evaluating the results of long-term bioassays in rodents for evidence of potential carcinogenicity, we do not always presume that all benign tumor types have the potential to progress to malignant tumors. Although the total incidence of benign tumors is considered in evaluations of carcinogenic potential, the weight placed on evidence of an increase in a benign tumor response is subject to consideration of the potential to progress to a malignant tumor.

For example, a uterine polyp, a benign tumor, is not considered to have the potential to progress to a malignant tumor; in the absence of evidence of an increase in malignant tumors in the same tissue (or in other tissues), evidence of carcinogenic potential would be considered to be weak. In such cases, the chemical would likely be assigned the descriptor “Suggestive Evidence of Carcinogenicity, but Not Sufficient To Assess Human Carcinogenic Potential” or “Not Likely To Be Carcinogenic to Humans” and a quantitative risk assessment would not be recommended (USEPA, 1999b).

On the other hand, hepatocellular adenomas are presumed to have the potential to progress to hepatocellular carcinomas. This is because there is considerable evidence showing that hepatocellular adenomas may progress to carcinomas in both rodents and humans. In addition, there is no clear histological demarcation between hepatocellular adenomas and carcinomas. Thus, when the results of rodent bioassays give evidence of the induction of both adenomas and carcinomas, the combined incidence of the two tumor types is considered in a weight-of-evidence evaluation of potential carcinogenicity. This approach is consistent with that of the National Toxicology Program (NTP), which recommends combining hepatocellular adenomas and hepatocellular carcinomas for an overall analysis of carcinogenicity (in addition to analyzing them separately).

### 4.1.5 Should We Expect Target Organ Concordance Between Effects Observed in Animals and Those Expected in Humans?

Thus far, there is evidence that growth control mechanisms at the level of the cell are homologous among mammals, but there is no evidence that these mechanisms are necessarily target organ concordant. For example, Haseman and Lockhart (1993) examined a database of 379 long-term carcinogenicity studies in rats and mice to evaluate sex and species correlations in site-specific carcinogenic responses. Within a species, most target sites showed a strong correlation between males and females. If a chemical produced a site-specific carcinogenic effect in female rats or mice, there was a 65% probability that it would be carcinogenic at the same site in males. On the other hand, the interspecies correlation was lower. Approximately 36% of the site-specific carcinogenic effects observed in one species were also observed in the other species.

EPA's cancer risk assessment guidelines recommend that site concordance of tumor effects between animals and humans be considered in each case. Site concordance is not assumed *a priori*. On the other hand, in establishing a mode of action that involves certain processes with consequences for particular tissue sites (e.g., disruption of thyroid function, hormonal effects leading to mammary tumors), one would expect site concordance when evaluating the weight of evidence. We address the issue of target organ concordance on a case-by-case basis, considering information on the mode of action. For example, based on a significant body of data, EPA determined that male rat kidney tumors are not relevant to the assessment of human cancer risk where tumors arise as a result of accumulation of a protein (alpha-2u-globulin) unique to the male rat (USEPA, 1991a). Similarly, the reproductive and developmental endpoints need to be judged considering what is known about the mode of action for an individual chemical as well as how the effects in animals may or may not be predictive of effects in humans.

Similarly, the expectation of concordance of fetal effects would depend on knowledge about mode of action. Most fetal effects are exquisitely dependent on time as well as dose. As the developmental process of laboratory animals and humans differs, especially for rodents, slight differences in the pattern of exposure or the toxicokinetics in the mother may greatly affect the observed outcome. In this regard, it is worth noting that the standard metric of daily dosing may have different implications for a laboratory animal than a human. Moreover, certain reproductive endpoints that are useful for evaluating fetal effects in rodents have no human equivalent, e.g., number and percent of live offspring per litter. With regard to fetal effects, for example, we consider both individual effects and total affected implants. Changes in dose may not only change the specific malformation, but also cause fetolethality rather than malformation. Thus, concordance of endpoint may not be the best predictor of developmental effects in people.

## 4.2 Extrapolations

Some federal regulatory agencies can estimate risks based on historical data about those risks. Thus, we have information on risks such as house fires, traffic accidents, and occupational injuries. At EPA, we are often faced with the problem of estimating the risk from chemicals for which we have limited or no historical data, e.g., new chemicals. Even when such a chemical has been in use, most if not all of the available information about its toxicity comes from studies with laboratory animals that use concentrations that are designed to produce effects. Similarly, available human data usually involve either exposures at high concentrations for short durations (e.g., accidents) or occupational exposures. Our concerns often involve situations where people might be exposed to much lower concentrations of the chemical for various periods of time. We, therefore, need to use procedures to extrapolate from responses (or lack thereof) at exposures where data are available to exposures of interest. Frequently, we also need to extrapolate from responses of laboratory animals to those that might reasonably be expected in people.

The rest of section 4.2 discusses how EPA extrapolates from animals to humans.

### 4.2.1 Why Does EPA Use Experiments With Animals To Predict Effects in Humans?

As a general rule, we assume that toxic responses observed in laboratory animals are indicative of toxic responses that are likely to occur in people. To predict responses in people more accurately, we would prefer information from animals that are as similar to people as possible. The use of other mammals, such as dogs, rats, and mice, as models for responses in humans is based on the assumption that there are important similarities among mammals in the way they respond to chemicals. A qualitative similarity has been established in the response of laboratory animals and humans to carcinogenic substances. For example, one analysis of chemicals that have been tested by the NTP for their ability to produce cancer (Huff et al., 1991) indicates that rats and mice predict responses for each better than 80% of the time for liver cancers. Most known human carcinogens have been shown to be positive for tumorigenicity in well-conducted animal studies. Data are not available to determine whether the opposite is true.

Nevertheless, important differences can exist. Metabolic patterns of both activation and deactivation along with differences in pharmacokinetics, for example, can give rise to significant differences in sensitivity between species and within species. Unless we know the rates of the activation-deactivation processes, for example, it is impossible to predict the differences in response among species. Presently, there is a paucity of comparative data on metabolism for specific chemicals and other interspecies differences that can affect toxicity. Important differences require experiments on the toxicokinetics or toxicodynamics within each species. We cannot be certain which results are most like those for humans in the absence of these data.

The increasing use of mode of action information will make it clearer whether a similarity or difference is apparent between test animals and humans.

### 4.2.2 How Does EPA Adjust for the Differences Between Animals and Humans?

If we have sufficient information about how the environmental agent is absorbed, distributed, metabolized, and excreted in the laboratory animal and the human and about the physiologies of the two species, we can develop a model (such as a PBPK model) that will allow us to compare how differences in these factors and processes affect the biologically effective dose to a target tissue. For some chemicals, this process has resulted in a very specific adjustment from the animal to the human. When information is limited, we can only make general adjustments.

Our practice is to interpret the findings of long-term rodent bioassays in conjunction with results of subchronic studies along with toxicokinetic studies and other pertinent information, if available. When data are appropriate, we use metabolic and toxicokinetic data to:

a) Identify and compare the relative activities of metabolic pathways in animals and humans and at different ages.

b) Describe anticipated distribution within the body and possibly identify target organs.

c) Identify changes in toxicokinetics and metabolic pathways with changes in dose.

d) Determine the bioavailability via different routes of exposure by analyzing uptake processes under various exposure conditions.

This issue is addressed on a case-by-case basis where data are available. When species-comparative data are not available, we assume concordance between the animals tested and humans.

### 4.2.3 Does EPA Use the Most Sensitive Animal Results To Predict Effects in Humans?

Use of the most sensitive species, strain, sex, age, etc., was originally justified on the need to be health protective in the absence of data. From the 1970s through the 1990s, the most sensitive responding species (given several data sets to choose from) were frequently selected: very little information was available as to what was scientifically the most representative choice for human risk prediction. Combined with UFs and other upper-bound estimates, basing cancer and non-cancer risks on the most sensitive animal data gives reasonable assurance that the

potential for harm will not be underestimated, most likely even when some toxicity endpoints have not been evaluated. Our early cancer risk assessment guidelines (USEPA, 1986a) comment on the use of the most sensitive animal response in the absence of appropriate human studies. The most recent draft guidelines (USEPA, 2003c) suggest a wide range of considerations based on knowledge and uncertainty for selecting the appropriate data set(s), assuming several sets are available.

Other federal agencies use a similar approach. According to Rhomberg (1997), the Occupational Safety and Health Administration (OSHA) also generally uses the data set demonstrating the highest estimated sensitivity, although it tends to "present several alternatives together or do several analyses and present the median result." The Consumer Product Safety Commission (CPSC) also uses similar methods, but differs in its approach to combining tumors from different sites.

The availability of more biological knowledge or insight potentially enables the risk assessor to make more scientifically informed choices among the available experiments with animals for use in dose-response analysis. The recent draft cancer guidelines (USEPA, 2003c) ask for a full display of all data options as well as the rationale for why one data set is selected and why others are rejected. Data from humans are preferred, although they are not always amenable to quantitative risk assessment. Alternatively, data from species that respond most like humans should be used if information to support this selection exists. If this is not known, all of the available data sets are considered and compared. An informed, scientific judgment is made as to what data best represent the observed data and important biological features such as mode of action. When faced with critical uncertainties or data gaps, however, we will favor a selection that ensures against underestimating risk as a policy choice.

We are continuing our trend toward using relevant data before using a default assumption to address data gaps. Use of toxicokinetic and toxicodynamic data for interspecies extrapolation should better address this source of uncertainty. Using the available data should accommodate some of the differences among species. For example:

a) When data on mode of action indicate that most the sensitive species/strain/sex is not an appropriate model for humans, e.g., alpha-2u-globulin, those data are not used for estimating human risks.

b) In some cases, use of the data that estimate the highest potential for harm can result in the use of less-well-designed studies, a concern that should be attenuated by increased use of the benchmark dose procedure.

c) Additionally, we have begun to investigate the utility of meta-analysis as one way to evaluate an entire database quantitatively.

### 4.2.4 Is an Additional Body Scaling Factor for Children Considered?

The draft final cancer guidelines (USEPA, 2003c) suggest that different adjustments may need to be made for children inhaling carcinogenic substances than for adults, but the guidelines do not specify how this should be done. For example, their language has been incorrectly interpreted as suggesting the use of a scaling factor of body weight to the 0.75 power before adjusting for a child's breathing rate. These draft guidelines recognize that correlations exist "between breathing rate, respiratory tract dimensions, and body weight."

The methods that could be used for interspecies extrapolation with appropriate adjustments for children are the subject of draft guidance (USEPA, 2003d), ongoing discussions, and further assessment by the Agency. The supplemental guidance for early-life exposures, for example, illustrates the importance of information regarding the mode of action of the carcinogen on evaluating the potential risks to children. We understand the need to address this issue further and will do so, in part, based on the comments received on the draft guidance.

### 4.2.5 How Are Effects Observed at High Exposures Used To Predict Responses to Much Lower Exposures?

The role of animal tests is to provide maximum detectability of effects within the narrow constraints of test sensitivity. This is typically achieved through the use of "high" exposures or doses to increase the probability that a possible adverse effect will be observed. This practice is usually used when sample sizes are too small to be sensitive to effects that are not overtly toxic or seen in response to potent stressors. Generally, the methods used to extrapolate from high to lower exposures differ for cancer and non-cancer effects. Both are discussed below. It should be noted that as more mode of action information becomes available, the distinction between cancer and non-cancer endpoints will dissipate and the dose-response evaluation is more likely to distinguish between linear and nonlinear extrapolations from high exposures to lower exposures.

*Cancer*

Carcinogenic effects are typically evaluated using a long-term rodent carcinogenicity bioassay. Current practice for these tests suggests using at least 50 animals per sex per dose group in each of three treatment groups and in a concurrent control group, usually for 18 to 24 months, depending on the rodent species tested. One group is exposed to a high dose, often the maximum tolerated dose, or MTD (discussed further below). Two groups are given lower doses of the test compound. The high dose in long-term studies is generally selected to provide the maximum ability to detect treatment-related carcinogenic effects while not compromising the outcome of the study through excessive toxicity or inducing inappropriate toxicokinetics (e.g., overwhelming absorption or detoxification mechanisms). The purpose of the two or more lower

doses is to provide some information on the shape of the dose-response curve, including possible subtle, precursor, and/or early events that can provide a more meaningful dose-response relationship relative to anticipated human exposures. Similar protocols have been and continue to be used by many laboratories worldwide.

Because of biological variability and the relative low sensitivity of most bioassay designs, positive evidence at the high dose is considered a rebuttable presumption of carcinogenicity. These findings are not altered by negative data at lower exposures unless there are supporting studies that provide a biological basis to conclude that the mode of action responsible for the high-dose effects is not relevant at lower doses and therefore for lower levels of exposure experienced by humans.

There is often a need to estimate risks for exposures significantly below exposures for which we have data. Two methods are used for this process. For mutagenic chemicals and those for which insufficient data exist to determine a mode of action, a point of departure (POD) at the low end of the exposures for which we have data is selected and a straight line is drawn from (the lower confidence limit on) that dose to zero. A nonlinear approach is selected when there are sufficient data to ascertain the mode of action and conclude that it is not linear at low doses and the agent does not demonstrate mutagenic or other activity consistent with linearity at low doses.

The use of low-dose linear assumptions for mutagenic carcinogens has been the peer reviewed, common practice for decades for many federal regulatory agencies including EPA. Modifications for non-mutagenic carcinogens are currently under review.

*Non-Cancer*

For non-cancer effects, we routinely use the absence of an adverse effect (a NOAEL) in a relevant animal study as the basis for estimating a negligible risk exposure level for humans. Adjustments are made to the NOAEL to account for limitations in knowledge, i.e., uncertainty, or limitations inherent in the test system, e.g., the use of relatively small numbers of genetically relatively homogeneous animals. Similarly, mode of action, physiological differences, or other factors have been used to improve the accuracy of estimation of the RfD or RfC for non-cancer endpoints.

The Agency has also been using and is continuing to refine the BMD methodology (USEPA, 2000g) for estimating non-cancer risks. In this procedure, the POD is derived from the dose-response curve, and further adjustments analogous to those for the NOAEL may be necessary to account for similar limitations of the data.

The EPA RfD/RfC review (USEPA, 2002c) delineates several options that have been used as the Agency developed more accurate methods for evaluating these parameters from

responses in animals. It also provides a series of examples, as well as lists of recommendations that address the predictivity of results from animal studies.

### 4.2.6 Why Is the Maximum Tolerated Dose Used?

Because tests for cancer are long and expensive, they are designed to increase the likelihood of detecting a positive response with a limited number of animals. The highest dose that does not produce other overt toxicity (i.e., the MTD) and fractions of that dose are selected for the dosing regimen. The MTD is usually determined from a shorter, e.g., 90-day, experiment with the same animal species and chemical.

The use of such high doses in animal cancer tests has been the subject of considerable debate. Limitations inherent in using the MTD approach and suggestions for improvement have been controversial since its use became standard. Testing at the MTD has been criticized because it may cause damage which may lead to cellular proliferation, increased mitosis, and eventually carcinogenicity. It has been argued that biochemical and physiological distortions occurring at high doses may lead to toxicity-induced carcinogenic effects that might not be expected at lower doses (Ames and Gold, 1990). Haseman (1985) has shown that more than two-thirds of the carcinogenic effects detected in feeding studies conducted under the NTP would have been missed if the highest dose had been restricted to doses below the MTD, if the chemicals were, in fact, animal carcinogens. Additionally, Bickis and Krewski (1989) found the upper confidence limit on the linear term ($q_1^*$) from the linearized, multi-stage model based on 263 data sets to be highly correlated with the maximum dose tested.

The draft final cancer guidelines (USEPA, 2003c) cautions the assessor in the use of results from the MTD and recommends that this issue be addressed on a case-by-case basis in the context of other study results and other lines of evidence. The guidelines state that the results of such studies would not be considered suitable for dose-response extrapolation if it is determined that the mode of action underlying the tumorigenic response at high doses is not operative at lower doses. For example, EPA developed a science policy position for agents that cause thyroid follicular cell tumors as a result of high-dose effects that are not present at low doses (USEPA, 1998c). In addition, the recent chloroform risk assessment was predicated on the recognition of a high-dose effect leading to cell proliferation and then carcinogenicity. The MCL was therefore developed based on an assumption of nonlinearity for carcinogenicity.

### 4.2.7 Does the Use of a Maximum Tolerated Dose Affect Cancer Potency Estimates?

The draft final cancer guidelines (USEPA, 2003c) include an extensive discussion of the role of dose selection in the conduct and interpretation of data from cancer tests carried out in animals. Section 2.2.2.1 deals with “Long-Term Carcinogenicity Studies,” and section 2.2.2.1.1,

"Dosing Issues," is in essence a discussion of the role of the MTD in cancer testing in animals and interpretation of data from such tests.

Section 2.2.2.1.1 makes the following critical points, among others, to justify the use of MTDs: "...failure to reach a sufficient dose reduces the sensitivity of the studies." Because animal tests are conducted with a relatively small number of animals, dosing at the MTD is necessary to provide the best chance that cancer will be seen in test animals if a chemical is carcinogenic. Further, the document recognizes that "...overt toxicity or inappropriate toxicokinetics due to excessively high doses may result in tumor effects that are secondary to the toxicity rather than directly attributable to the agent."

The document also stresses the importance of establishing that the MTD has not been exceeded. Criteria for determining this include whether "...an adequate high dose would generally be one that produces some toxic effects without unduly affecting mortality from effects other than cancer or producing significant adverse effects on the nutrition and health of the test animal." Thus, our practice is that the experimental conditions are scrutinized to ensure that the MTD has been reached but not exceeded. We further believe that effects seen at the MTD may be appropriate for use in risk assessment when the data have been critically evaluated.

### 4.2.8 Are the Cancer Risks Estimated by EPA the Expected Incidence of Cancer for a Given Exposure?

Our risk estimates are designed to ensure that risks are not underestimated, which means that a risk estimate is the upper bound on the estimated risk. In past guidelines (USEPA, 1986a), we have explicitly stated that the true cancer potency "could be as low as zero." Since the CSF is generally multiplied by the exposure estimate to generate an estimated risk, a CSF of zero would result in a risk estimate of zero.

If other procedures, e.g., odds ratios from epidemiologic studies, result in risk estimates that should be interpreted differently, these should be clearly delineated. In particular, if the method either requires or results in maximum likelihood estimates (MLEs) rather than bounding estimates being used, this should be clearly stated.

According to Rhomberg (1997), OSHA uses the same methodology as EPA, but it presents and focuses on the MLE dose-response curve. CPSC also uses the same model as EPA, but modifies its constraints so that, with the limited data usually available, the MLE of the dose-response function is linear at low doses. Rhomberg asserts that this procedure also results in an upper-bound estimate, although no proof is given in the cited document.

### 4.2.9 How Does EPA Evaluate Incidence Versus Mortality for Cancer Rates?

Incidence of a disease is a measure of disease occurrence. Incidence rate measures the frequency of a given disease in a population. A concept closely related to incidence rate is that of the risk or probability of developing a disease over a period of time. The incidence rate for a disease is also referred to as the hazard rate for developing that disease (Kalbfleisch and Prentice, 1980). Mortality rate, on the other hand, measures the frequency of death rate in a population from a given disease. Both incidence and mortality rates are measures of risk and have been used to calculate risk in different risk assessments. The appropriate measure depends upon the particular risk of interest and the type of data available. It is important that this be generally made clear in risk assessments.

Incidence rates give a more accurate indication of risk of a disease in a population. However, obtaining information about incidence rates in a population is usually difficult. One place to look for incidence rates of site-specific cancers in a population is the Surveillance, Epidemiology, and End Results (SEER) data. Obtaining mortality data and computing rates is comparatively easy, because every death is recorded. Thus, over the years, the most commonly used data for risk assessments have been the mortality rates. An adjustment between mortality and incidence can be carried out using information from databases such as SEER.

In diseases for which survival is poor and mortality is 95% to 100% within a short period of time (e.g., lung cancer, pancreatic cancer), the mortality rates are good surrogates for incidence rates. A risk assessment based on these mortality rates provides a good estimation of true risk in a given population. In diseases for which survival is higher and mortality is lower (e.g., non-melanoma skin cancer, urinary bladder cancer), the mortality rates are poor surrogates for incidence rates. A risk assessment based on these mortality rates therefore provides an underestimation of risk in a given population.

EPA's draft final Guidelines for Carcinogen Risk Assessment (USEPA, 2003c) recognize this issue. Section 3.2.1 of the guidelines states that "Analysis of epidemiologic studies depends on the type of study and quality of data, particularly the availability of quantitative measures of exposure. The objective is a dose-response curve that estimates the incidence of cancer attributable to exposure to the agent." It further notes that "The analysis is tailored to the nature of each study, with due consideration of the consequences of study design. For example, many studies collect information from death certificates, which leads to estimates of mortality rather than incidence. Because survival rates vary for different cancers, the analysis can be improved by adjusting mortality figures to reflect the relationship between incidence and mortality." This was the approach used in EPA's assessment of 1,3-butadiene, found in IRIS (USEPA, 2002e). A similar approach was used in the NRC's assessment of the cancer risk from arsenic ingestion (NRC, 2001).

It should be noted that when the CSF is derived from animal data, the extrapolation uses the tumor incidence data.

### 4.2.10 If Data Are Only Available for One Route of Exposure, Does EPA Evaluate Other Routes of Exposure?

Our general assumption is that chemicals that cause internal toxicity by one route of exposure will do so by any other route if that route also produces an internal dose. This position is based on the scientific principle that the internal dose to the tissue of interest is the ultimate determinant of toxicity. We advocate development and use of agent-specific pharmacokinetic information to inform this position.

In its 1994 report, the NRC (NRC, 1994) states that "... the target-site dose is the ultimate determinant of risk ..." Thus, the manner or route of systemic exposure is secondary (except as it affects internal doses, e.g., by first-pass metabolism) to the internal dose. Others in government also reflect this fundamental principle (e.g., Corish and Fitzpatrick, 2002; ACGIH, 1991).

Some chemicals have demonstrated the same (or very close to the same) target organs in response to chemical exposure, regardless of route. For example, for cumene, a kidney effect is observed after oral exposure, implying an internal dose of cumene or its metabolites to the kidney via the oral route; a kidney effect is also seen after an inhalation exposure to cumene, implying an internal dose of cumene or its metabolites to the kidney via the inhalation route. This demonstrates a target organ dosing independent of route of administration and a fundamental role for establishing an internal dose leading to toxicity. Three other examples from IRIS of chemical assessments having a common critical target organ via inhalation (RfC) and oral (RfD) routes of administration are shown below, in table 4-1.

**Table 4-1: Route Comparison**

| | Target Organ Affected via the | |
|---|---|---|
| **Agent** | **Oral Route** | **Inhalation Route** |
| Cumene (CAS # 98-82-8) | Kidney | Kidney |
| 1,1-dichloroethylene (CAS # 75-35-4) | Liver | Liver |
| Vinyl chloride (CAS # 75-01-4) | Liver | Liver |

Several examples from studies measuring kinetic endpoints of compounds administered by various routes of exposure found results to be mostly parallel across routes. This indicates similar kinetic processes regardless of route.

a) After exposing rats to benzo(*a*)pyrene via either oral, inhalation, or intravenous routes, Ramesh et al. (2002) observed data indicating metabolism to all principal metabolites across all routes.

b) Kaneko et al. (1995) reported metabolism in rats of xylene by inhalation, oral, and intraperitoneal (ip) routes of exposure, with a significant "first-pass" effect via the oral route. This latter effect was observed only when the liver enzymes were altered by compounds known to induce activating enzymes.

c) After administration of alkyl ketones to rats via oral or inhalation route, Duguay and Plaa (1995) observed dose-related increases in parent and metabolite in all tissues examined (plasma, liver, and lung).

d) Gospe and Albayati (1994) demonstrated that oral administration of toluene produces blood toluene concentrations that can simulate blood levels achieved after inhalation exposures of this agent.

e) Timchalk et al. (1991) observed that radiolabeled dichloropropane administered to rats was readily absorbed, similarly metabolized, and excreted after both oral and inhalation exposures.

Studies measuring both toxicokinetics and toxicity administered by various routes of exposure found, in general, parallel results regardless of route.

a) In a study measuring toxic in addition to kinetic endpoints, Gansewendt et al. (1991) reported that both oral and inhalation methyl bromide administration in rats resulted in the same levels and species of DNA adducts, with the highest levels found in the stomach and forestomach for both routes.

b) After administering rats hepatotoxins that were either poorly (carbon tetrachloride) or highly (chloroform) metabolized, Wang et al. (1997) observed and reported that toxicodynamics were parallel and toxicity was present for both compounds via oral, inhalation, or ip routes.

c) McEuen et al. (1995) noted that oral and ip routes of administration of dinitrobenzene produced significantly higher levels of active metabolite via the oral route but only subtle differences in target organ (testes) toxicity in rats.

d) Using a PBPK model for cumene to examine measures of dose associated with renal toxicity across an oral and an inhalation study in rats, Foureman and Clewell

(1999) noted that measures of dose generated from either study were associated with the toxicity in a dose-related manner, regardless of route.

*What Data May Set Aside the Route-to-Route Extrapolation?*

The draft final cancer guidelines (USEPA, 2003c) discuss some evidence that could modify the application of this extrapolation, stating that an exception would occur when there is convincing toxicokinetic evidence that absorption for a given chemical does not occur by another route of interest. The basis for this exception is establishing the absence of an internal dose such that the extrapolation may not be applied and such a chemical could be characterized as not being toxic via this particular route of administration.

In the absence of data to the contrary, the route-to-route extrapolation assumption used is that if a chemical is absorbed via one route, it will also be absorbed by all other routes. Demonstration of any degree of uptake for each of the routes of interest is sufficient to allow the qualitative judgment to apply the route-to-route extrapolation. The level of information necessary to set aside the extrapolation (i.e., "contrary" evidence ) may largely be a function of the individual stressor being tested. The similarities of processes and structures involved in absorptive processes across biological systems may be regarded as an unstated principle that underlies the extrapolation. Absorption obligates passing through a tissue and, fundamentally, tissues have more similarities than differences. The presumption is that if the stressor can penetrate tissue via one route, it can penetrate the tissue if it arrives there via another route. Thus it is anticipated that most agents are absorbed to some degree into biological systems regardless of route (e.g., see results summarized in Owen, 1990).

Chemicals causing point-of-entry toxicity are often very reactive, and thus are less likely than other chemicals to produce an internal dose. Chemicals that do not establish an internal dose by a given route would not be presumed to give internal toxicity by that route. Evidence of robust point-of-entry toxicity by an agent may thus be considered contrary evidence for the contention that toxicity by one route of exposure by that agent is presumed to do so by another route. IRIS lists assessments for a number of highly reactive agents that elicit only clear and robust portal-of-entry effects — in the context of this issue, contrary evidence. Nearly all of these assessments explicitly consider that these agents would not be anticipated to elicit internal toxicity and discount the need for information on systemic toxicity. This discounting is done by decreasing or even eliminating the "database" UF typically evoked at a value of 10 when studies such as developmental and reproductive toxicity (toxicities typically associated with an internal dose) are absent. Table 4-2 gives some examples. The practice of omitting the database UF when portal-of-entry effects occur may not be followed in some older IRIS assessments that have not yet been updated. For example, a database UF of 3 is applied specifically for missing developmental data in the assessment of one other of the isocyanates, TDI (CAS # 26471-62-5).

A database UF of 10 is also applied in the RfC for 2-chloroacetophenone (CAS # 532-27-4), a highly reactive tearing agent.

| Table 4-2: Point-of-Entry Toxicity Examples | |
|---|---|
| **Agent (CAS#)** | **Value[a] and Commentary of "Database" UF (Online IRIS)** |
| Acrolein (107-02-8) | UF = 1, not applied due to lack of systemic distribution (2002) |
| Ammonia (7664-41-7) | UF = not greater than 3, no significant distribution likely to occur (1991) |
| Epoxybutane (106-88-7) | UF = 1, extrarespiratory circulation of EBU minimal (1992) |
| HCl (7647-01-0) | UF = 1, expected portal-of-entry effect, UF not applied (1995) |
| HDI (822-06-0) | UF = 3, unlikely that HDI is distributed in significant amounts (1994) |
| $H_3PO_4$ (7664-38-2) | UF = 1, no UF to be applied (1995) |
| MDI (101-68-8) | UF = 3 (1998) |
| Methacrylate (80-62-6) | UF = 1, reactive, potential for systemic effects is remote (1998) |
| [a] This UF has a maximum value of 10. | |

## 4.3 Use of Uncertainty Factors

### 4.3.1 How Are NOAELs or LOAELs Used for Extrapolations?

When developing a non-cancer reference value (i.e., a RfD or RfC) for a chemical substance, EPA surveys the scientific literature and selects a critical study and a critical effect to serve as the point of departure for the assessment. The critical effect is defined as the first adverse effect, or its known precursor, that occurs in the most sensitive species as the dose rate of an agent increases (USEPA, 2002c). Such a study, whether an occupational human study, a deliberately dosed animal study, or some other study, typically involves exposure at a range of doses. The highest exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control is — as mentioned above — called the "no-observed-adverse-effect level" (NOAEL). When a NOAEL can be identified in a critical study, it becomes the basis of the reference value derivation. The NOAEL is divided by appropriate UFs (e.g., for intraspecies

variation or study duration) to derive the final reference value. If NOAEL cannot be identified, then a "lowest-observed-adverse-effect level" (LOAEL) — also mentioned above — is identified instead. A LOAEL is the lowest exposure level at which there are biologically significant increases (with or without statistical significance) in frequency or severity of adverse effects between the exposed population and its appropriate control group. The NOAEL is generally presumed to lie between zero and the LOAEL, so an UF (generally 10 but sometimes 3 or 1) is applied to the LOAEL to derive a nominal NOAEL. Other factors are then applied to derive the reference value. More recently, some assessments have used a BMD approach instead of the traditional NOAEL/LOAEL approach (USEPA, 2000g).

For each non-cancer reference value (i.e., RfC or RfD) for a chemical substance in IRIS, the following table shows whether that value was based on a NOAEL, a LOAEL, or a benchmark dose lower confidence level. The table clearly shows that the majority of non-cancer reference values in IRIS are based on NOAELs.

**Table 4-3: Effect Levels Used To Derive Non-Cancer Reference Values in IRIS**

| | RfCs | RfDs |
|---|---|---|
| **NOAEL** | 35 | 294 |
| **LOAEL** | 21 | 51 |
| **BMD** | 13 | 10 |
| **Total** | 69 | 355 |

### 4.3.2 What Uncertainty Factors Does EPA Use To Reduce the Experimental NOAEL in Health Assessments?

The Risk Assessment Forum RfD/RfC Technical Panel report (USEPA, 2002c) defines an uncertainty factor as:

> one of several, generally 10-fold, default factors used in operationally deriving the RfD and the RfC from animal experimental data. The factors are intended to account for (1) variation in sensitivity among the members of the human population; (2) the uncertainty in extrapolating animal data to humans; (3) the uncertainty in extrapolating from data obtained in a study with less-than-lifetime exposure to lifetime 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. The exact value of the UFs chosen depend on the quality of the studies available, the extent of the database, and scientific judgment.

Some investigators evaluated the accuracy and limitations of allocating a value of 10 for each area of uncertainty (Dourson and Stara, 1983; Woutersen et al., 1984; Kadry et al., 1995;

Dourson et al., 1996). For example, the subchronic-to-chronic UF, some conclude, is probably closer to a two- to threefold difference of uncertainty (95% of the time) and a UF of 10 should be considered as a loose upper-bound estimate of the overall uncertainty. Others (see Dourson and Stara, 1983, for survey of arguments representing various points of view) conclude that a 10-fold interspecies UF could account for all animal to human differences and likewise for interhuman variability. There are opposing arguments that EPA's application of UFs may not be conservative enough (Dourson and Stara, 1983).

EPA applies UFs in health assessments based on the available data and the scientific judgment of EPA risk assessors and internal and external peer reviewers. In cases where chemical-specific data are lacking, each UF is typically no greater than 10. For example, the majority of IRIS Toxicological Reviews provide justifications for the individual UFs selected for a particular chemical substance. The choices of critical health endpoint, principal study, and UFs undergo rigorous internal and independent, external scientific peer review before being presented in the IRIS database.

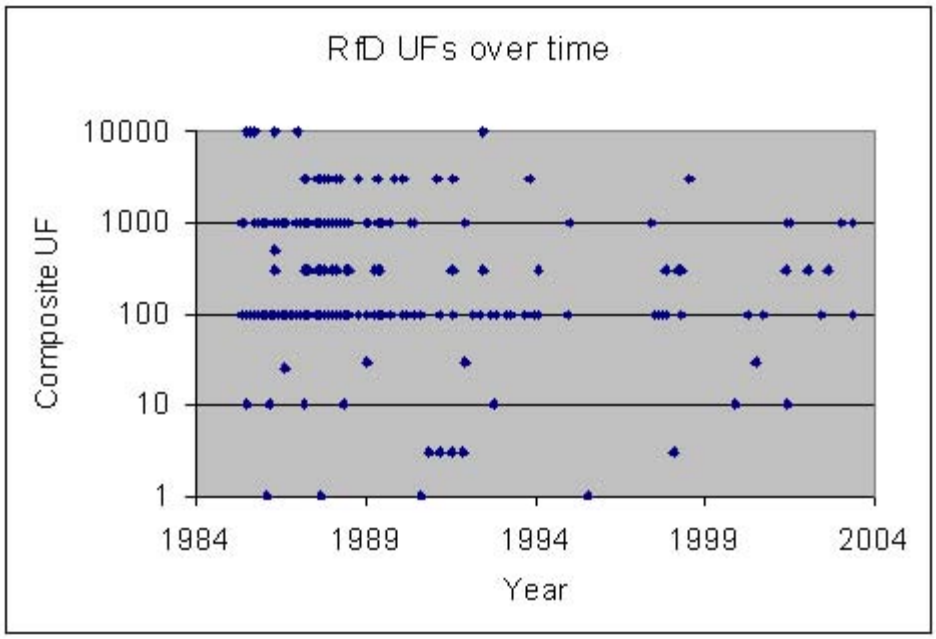
For several years, EPA has used a more qualitative approach to modify the usual 10-fold default values. For example, in deriving inhalation RfC values, the interspecies variability UF of 10 is used in the absence of data, where the distribution is assumed to be log-normally distributed. While EPA has not yet established guidance for the use of chemical-specific data for derived UFs, the reference concentration methodology guidance (USEPA, 1994a) provides opportunities for using data-derived interspecies UFs by subdividing the factor of 10 to allow for separate evaluations of toxicokinetics and toxicodynamics. The advantage to such subdivision is a default UF of 10 for interspecies variability that can now be reduced to 3 when animal data are dosimetrically adjusted to account for toxicokinetics (USEPA, 2002c).

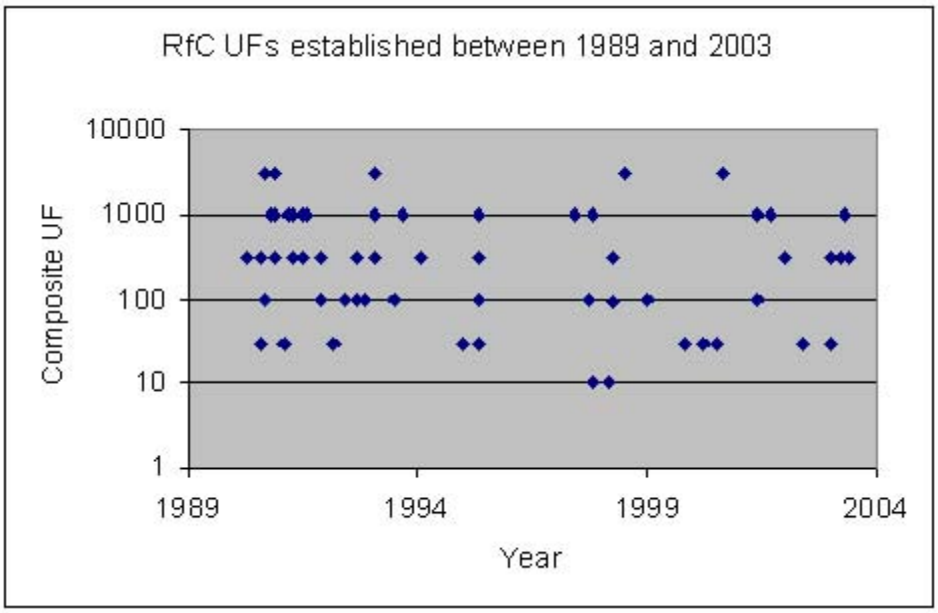
### 4.3.3 Does EPA Consider the Effects of Combining Several UFs?

In its review of the RfD/RfC practice in the Agency, the RfD/RfC Technical Panel reviewed the application of UFs in health assessments (USEPA, 2002c). The Panel recognized the potential for overlap in the individual UFs and concluded that the application of five UFs of 10 for the RfD/RfC is inappropriate. Therefore, due the uncertainty inherent in values when so many UFs are applied, the Panel recommended that no reference value (RfD/RfC) for any particular chemical substance be derived if the composite UF is greater than 3,000. It further recommended avoiding the derivation of a reference value that involves application of the full 10-fold in four or more areas of extrapolation. The report also recommended to discontinue the use of the modifying factor as a UF.

The two graphs below show EPA's historical practice for combining UFs to derive a RfD or RfC. As can be seen, when reference values were originally being generated, some combinations of UFs reached as high as 10,000. As we gained greater and greater experience and obtained more usable data, the trend has shown a decrease in the combination of UFs to such high levels. The more recent RfDs and RfCs are more in line with the recommendation from the Technical Panel: no derived reference value exceeds 3,000.

**Figure 4-1: Use of Uncertainty Factors Over Time (to Set RfDs and RfCs)**






### 4.3.4 Does EPA Use UFs Derived for Specific Chemicals?

With respect to departing from the default UFs with data-derived or chemical-specific adjustment factors, the RfD/RfC Technical Panel indicated that guidance on how to use chemical-specific adjustment factors in risk assessment will be developed. Toward that end, EPA has recently begun to consider data-derived approaches for chemical-specific adjustment factors in risk assessments.

## 4.4 Weight of Evidence

### 4.4.1 How Does EPA Consider Weight of Evidence?

Risk assessment involves consideration of the weight of evidence provided by all available scientific data. In other words, "weight of evidence evaluation is a collective evaluation of all pertinent information so that the full impact of biological plausibility and coherence is adequately considered" (USEPA, 1999b). Judgment on the weight of evidence involves consideration of the quality and adequacy of data and consistency of responses induced by the stressor. The weight-of-evidence judgment requires combined input of relevant disciplines: toxicology, biology, chemistry, epidemiology, statistics, etc. Initial views of the database may change significantly when other data are brought into consideration. For example,

the impact of a positive animal carcinogenicity study may be diminished by high-quality negative studies; or a weak association in human epidemiologic studies may be bolstered by consideration of other key data from animal or other assays. Generally, no single study, whether positive or negative, drives the overall weight-of-evidence judgment. And study findings are not scored by any mathematical algorithm; rather, they are based on professional scientific judgment.

### 4.4.2 Does EPA Give More Weight to Positive Animal Studies Than to Negative Animal Studies?

EPA evaluates all data, including both "positive" and "negative" studies, in assessing the weight of evidence for toxicity of a chemical. Assessment decisions are, by nature, subject to professional judgment and require input from experts in a variety of disciplines including toxicology, pathology, and statistics. Generally, EPA does consider positive animal studies even in the presence of negative studies. Negative studies do not "outweigh" positive studies, though, unless they present a robust, compelling outcome relative to the positive studies.

To examine the comment that EPA always gives more weight to positive animal studies than to negative studies, 108 chemicals classified as Class D (not classifiable as to human carcinogenicity) in IRIS were searched for evidence as to whether they included any positive animal studies. Seventeen of those substances were found to have *statistically significant positive* animal data, but it was the professional judgment of those performing the assessment and those who scientifically peer reviewed it that these substances were indeed nonclassifiable as to human carcinogenicity. Thus, we do not automatically translate positive animal carcinogenicity data into classification as a human carcinogen. Rather, it is EPA's practice to base decisions on the best available science.

### 4.4.3 Does EPA Sometimes Weight Animal Data More Than Human Data?

EPA prefers high-quality human studies over animal studies because they provide the most relevant kind of information for human hazard identification. In the absence of usable human data, the default assumption is that positive effects in animal cancer studies indicate carcinogenic potential in humans. Risk analysis takes all studies into account, whether they show positive associations, null results, or even protective effects. In weighing positive versus null studies, those judged to be of high quality are given more weight than those judged methodologically less sound, and possible reasons for inconsistent results should be sought. Null results from a single epidemiologic study cannot prove the absence of health effects because of the limitations that may be associated with the study — inadequate statistical power, inadequate design, imprecise estimates, confounding factors, and others. However, null results from a well-designed and well-conducted epidemiologic study that contains usable exposure data can help define upper limits for the estimated dose of concern for human exposure if the overall weight of evidence indicates the agent is a potential human carcinogen.

EPA tries to use human studies whenever feasible and appropriate. As the NRC pointed out (NRC, 1994), "Epidemiologic studies clearly provide the most relevant kind of information for hazard identification, simply because they involve observations of human beings, not laboratory animals." As a result, human data form the basis of a number of highly visible assessments such as those of arsenic, chromium, nickel, benzene, and vinyl chloride. It is true, however, that we have relied more heavily on animal data for some assessments even when human data were available. The case of epichlorohydrin is an example of such an instance. Despite the fact that two retrospective cohort mortality studies of exposed workers found no evidence of an increased cancer risk, epichlorohydrin is classified as a "probable human carcinogen." The apparent discounting of the human data is explained as follows:

> The first study is inadequate for valid carcinogenicity assessment because of low exposure, short exposure duration, short study period, and the young age of the cohort. The second study suffers from some of the same limitations, as well as that of a small cohort size with at least 10 years exposure (274 individuals).

While pointing out the benefits of epidemiologic evidence, the NRC noted that the obvious and substantial advantage (of human studies) is offset to various degrees by the difficulties associated with obtaining and interpreting epidemiologic information (NRC, 1994). As the Office of Science and Technology Policy indicated (OSTP, 1985), the epidemiologic method (in the assessment of cancer risk) is often hampered by the long latent period between exposure to a carcinogenic agent and the development of cancer, by the inability to control for the confounding influences of unknown risk factors, by problems in assessing specific agents when the human exposures are to mixtures, by the frequent absence of appropriate groups for study, and by a variety of difficulties associated with accurate and unbiased historical exposure assessment or disease ascertainment. In short, the Agency, while recognizing the clear advantages of epidemiologic studies, also recognizes the limitations of these studies; and it is these limitations which sometimes constrain the use of human data in an assessment.

For some comparisons of the use of animal versus human studies in risk assessment decisions within the IRIS database, please see the tables below.

*Animal Versus Human Studies Tables*

All substances in the IRIS database were searched for the keyword "human," which provided 541 matches (essentially the whole database). From these matches, the first 60 were examined to identify compounds for which human data are available, yet animal studies serve as the basis for assigning RfCs, RfDs, or cancer risk estimates.

Within the first 60, there were 15 with RfC listings. Of those 15, 8 list specific human data in the inhalation exposure section. Four of those eight use human data as the principal study guiding determination of the RfC, and four did not. Table 4-4 outlines the explanations provided for why the available human data were not a principal study for the latter four substances. The

table also presents brief statements of positive and negative results from listed human studies and interspecies UFs.

Among the first 60, there were 42 with RfD listings. Of these 42, nine list specific human data in the oral exposure section. Five use the human data as the principal study in guiding determination of the RfD; four do not. Table 4-5 outlines the explanations provided for these four studies as to why they were not used to set the RfD.

There were 27 among the first 60 with classifications of carcinogenicity. Of these 27, 10 identify human data in the assessment. Four of these assessments use human data to make the classification of carcinogenicity; six do not. Table 4-6 provides the explanations as to why, for the latter six substances, human data were not used to make the carcinogenic classification.

**Table 4-4: Explanations of Why Human Results Were Not Used To Estimate the RfC for Four Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | HD = Principal Study | Explanation |
|---|---|---|---|
| **Bromomethane** | Yes | No | Several studies have been conducted on the longer-term effects of occupational exposure bromomethane. **None of these studies can serve as the basis for the derivation of an RfC for bromomethane because of concurrent exposures to other chemicals, inadequate quantitation of exposure levels and/or durations, and other deficits in study design.**<br><br>[UF] of 3 for interspecies extrapolation because dosimetric adjustments have been applied.<br><br>• *Positive human occupational study: neurological effects (Anger et al., 1986).*<br>• *Four cases of reported occupational toxicity (Herzstein and Cullen, 1990).* |

| | | | |
|---|---|---|---|
| **Dichlorvos** | Yes | No | Justification not provided. A UF of 3 is used for interspecies extrapolation.<br><br>• *Limited reporting of mixed results in study of residential fumigators and residents (Gold et al., 1984).*<br>• *Plasma cholinesterase activity and vitamin A decrease in five of six pesticide manufacturing plant workers (Ember et al., 1972).*<br>• *Case studies of poisonings report clinical and self-reported symptoms (Low et al., 1980; Reeves et al., 1981).*<br>• *Decrease in plasma cholinesterase activity, but no significant difference in reported clinical symptoms between test and control subjects in volunteer human studies (Rider et al., 1967; Slomka and Hine, 1981).* |

**Table 4-4: Explanations of Why Human Results Were Not Used To Estimate the RfC for Four Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | HD = Principal Study | Explanation |
|---|---|---|---|
| **Acrylonitrile** | Yes | No | Justification not provided.<br><br>An uncertainty factor of 3 for interspecies variability is used because the use of the dosimetric adjustments account for part of this area of uncertainty.<br><br>• *No excess mortality for non-cancer endpoints in human occupational epi studies (O'Berg, 1980; Ott et al., 1980; Werner and Carter, 1981; O'Berg et al., 1985; Chen et al., 1987).*<br>• *Self-reported and clinical symptoms apparent in occupational study (Wilson et al., 1948).*<br>• *No effects in human volunteer study (Jakubowski et al., 1987).* |
| **Propylene glycol monomethyl ether (PGME)** | Yes | No | Comment on use of human data: A human exposure study (Stewart et al., 1970) indicates that the RfC should prevent irritation. Additional justification not provided.<br><br>[An uncertainty] factor of 3 is used for interspecies extrapolation given the dosimetric adjustment and the zero-order pharmacokinetic rate of elimination of PGME that is suggestive of an adaptive metabolic response (Morgott and Nolan, 1987).<br><br>• *Human volunteer study was positive for self-reported symptoms but negative for clinical effects (Stewart et al., 1970).* |

[a] These four substances were identified from the first 60 matches in the IRIS data file for which the word "human" was identified. RfCs were estimated for 15 of these 60 substances. Human data relevant to the RfC were available for 8 of the 15 substances. The human data were used as the principal study for the RfC for four of those eight substances. This table describes the reasoning provided as to why the human data were not used in the other four instances. Interspecies UFs are listed. Results of human studies are included in italics.

**Table 4-5: Explanations of Why Human Results Were Not Used To Estimate the RfD for Four Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | HD = Principal Study | Explanation |
|---|---|---|---|
| **Chromium (VI)** | Yes | No | In 1965, a study of 155 subjects exposed to drinking water at concentrations of approximately 20 mg/L was conducted outside Jinzhou, China … **Precise exposure concentrations, exposure durations, and confounding factors were not discussed, and this study does not provide a NOAEL for the observed effects**. However, the study suggests that gastrointestinal effects may occur in humans following exposures to hexavalent chromium at levels of 20 ppm in drinking water (Zhang and Li, 1987).<br><br>[UF] — two 10-fold decreases in dose to account for both the expected interhuman and interspecies variability in the toxicity of the chemical in lieu of specific data.<br><br>• *Positive drinking water study (Zhang and Li, 1987).*<br>• *Studies of positive allergic response (Bruynzeel et al., 1988; Polak, 1983; Cronin, 1980; Hunter, 1974).* |

**Table 4-5: Explanations of Why Human Results Were Not Used To Estimate the RfD for Four Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | HD = Principal Study | Explanation |
|---|---|---|---|
| **Uranium, soluble salts** | Yes | No | Uranium is a classical nephrotoxic. The toxicity of this chemical to humans has been of interest since the 1800's when uranium was used as a homeopathic cure for diabetes mellitus (Hodge, 1973). These early reports demonstrate the susceptibility of humans to the nephrotoxicity of ingested uranium, but **provide inadequate basis for estimating the threshold dose for toxic effects.**<br><br>[UF of] 10 for both intraspecies and interspecies variability to the toxicity of the chemical in lieu of specific data.<br><br>• *Reports of nephrotoxicity with therapeutic use (Hodge, 1973).*<br>• *Mixed results from experiments on uranium excretion and toxicity (Hursh and Spoor, 1973; Lussenhop et al., 1958; Hursh et al., 1969).* |
| **Naphthalene** | Yes | No | Humans exposed via inhalation, combined inhalation and dermal exposure, and combined inhalation and oral exposure have developed hemolytic anemia. Hemolytic anemia is characterized by findings of lowered hemoglobin, hematocrit, and erythrocyte values, elevated reticulocyte counts, Heinz bodies, elevated serum bilirubin, and fragmentation of erythrocytes. In severe cases, the hemolytic anemia was accompanied by jaundice, high serum levels of bilirubin, cyanosis, and kernicterus with pronounced neurological signs. **Neither oral nor inhalation exposure levels were available in human studies reporting anemia** (Melzer-Lange and Walsh-Kelly, 1989; Owa, 1989; Owa et al., 1993).<br><br>[UF of] 10 to extrapolate from rats to humans.<br><br>• *Human studies positive for hemolytic anemia (Melzer-Lange and Walsh-Kelly, 1989; Owa, 1989; Owa et al., 1993).* |

**Table 4-5: Explanations of Why Human Results Were Not Used To Estimate the RfD for Four Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | HD = Principal Study | Explanation |
|---|---|---|---|
| **1,3-Dichloro-propene** | Yes | No | There are **no chronic human studies suitable for dose-response assessment**. Chronic feeding studies in rats (Stott et al., 1995) and mice (Redmond et al., 1995) and chronic gavage studies (NTP, 1985) using both species are available. The feeding studies are favored over the gavage studies because the route of administration is more relevant to human exposure.<br><br>The **toxicokinetics in humans are similar to those observed in rats**. Waechter et al. (1992) showed that the absorption of 1,3-dichloropropene from inhalation exposure of humans (72%–82%) was similar to absorption in rats (82%; Stott and Kastl, 1986).<br><br>The default uncertainty factor of 10 for interspecies extrapolation is applied because there are no data on the relative sensitivity of rats and humans to stomach irritation.<br><br>• *Human studies demonstrate contact dermatitis (Bousema et al., 1991; Nater and Gooskens, 1976).*<br>• *Human poisonings result in neurotoxic symptoms (Hernandez et al., 1994).* |

[a] These four substances were identified from the first 60 matches in the IRIS data file for which the word "human" was identified. RfDs were estimated for 42 of the 60 substances. Human data relevant to the RfD were available for 9 of the 42 substances. The human data were used as the principal study for the RfC for five of the nine substances. This table describes the reasoning provided as to why the human study were not used in the other four instances. Interspecies uncertainty factors are listed. Results of human studies are included in italics.

**Table 4-6: Explanations of Why Human Results Were Not Used for the Carcinogenic Classification for Six Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | Explanation |
|---|---|---|
| **Bromomethane** | Yes | D (Not classifiable as to human carcinogenicity)<br><br>Inadequate human data. A prospective mortality study was reported for a population of 3579 white male chemical workers. The authors noted that it was difficult to draw definitive conclusions as to causality because of the lack of exposure information and the likelihood that exposure was to many brominated compounds.<br><br>• *(Wong et al., 1984) study looked at cancer mortality in the above mentioned occupational cohort. Only increased cancer mortality was from testicular cancer.* |
| **p,p'-Dichlorodiphenyl-trichloroethane (DDT)** | Yes | B2 (Probable human carcinogen — based on sufficient evidence of carcinogenicity in animals)<br><br>Inadequate human data. The existing epidemiological data are inadequate.<br><br>• *Autopsy studies relating tissue levels of DDT to cancer incidence have yielded conflicting results. Three studies reported that tissue levels of DDT and DDE were higher in cancer victims than in those dying of other diseases (Casarett et al., 1968; Dacre and Jennings, 1970; Wasserman et al., 1976). In other studies no such relationship was seen (Maier-Bode, 1960; Robinson et al., 1965; Hoffman et al., 1967). Studies of occupationally exposed workers and volunteers have been of insufficient duration to be useful in assessment of the carcinogenicity of DDT to humans.* |

**Table 4-6: Explanations of Why Human Results Were Not Used for the Carcinogenic Classification for Six Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | Explanation |
|---|---|---|
| **Polychlorinated biphenyls (PCBs)** | Yes | B2 (Probable human carcinogen — based on sufficient evidence of carcinogenicity in animals)<br><br>Inadequate human data. The human studies are being updated; currently available evidence is inadequate, but suggestive.<br><br>• *Significant cancer excesses in human cohort studies (Bertazzi et al., 1987; Brown, 1987; Sinks et al., 1992).*<br>• *Inconclusive results of occupational studies due to design limitations (NIOSH, 1977; Gustavsson et al., 1986; Shalat et al., 1989).*<br>• *Liver cancer and skin disorders in populations exposed to PCB-contaminated rice oil in Japan and Taiwan (ATSDR, 1993; Safe, 1994).* |
| **Dieldrin** | Yes | B2 (Probable human carcinogen — based on sufficient evidence of carcinogenicity in animals)<br><br>Inadequate human data. Two studies of workers exposed to aldrin and to dieldrin reported no increased incidence of cancer. Both studies were limited in their ability to detect an excess of cancer deaths … Exposure was not quantified, and workers were also exposed to other organochlorine pesticides (endrin and telodrin). The number of workers studied was small, the mean age of the cohort (47.7 years) was young, the number of expected deaths was not calculated, and the duration of exposure and of latency was relatively short.<br><br>• *No statistically significant excess in cancer deaths among occupational cohort (Van Raalte, 1977; Ditraglia et al., 1981).*<br>*~ same studies as cited in the aldrin assessment* |

**Table 4-6: Explanations of Why Human Results Were Not Used for the Carcinogenic Classification for Six Substances in the IRIS Database[a]**

| Chemical | Human Data (HD) Listed | Explanation |
|---|---|---|
| **Acrylamide** | Yes | B2 (Probable human carcinogen — based on sufficient evidence of carcinogenicity in animals)<br><br>Inadequate human data. There are two studies on the relationship of workers exposed to acrylamide and cancer mortality. A basic limitation of both studies is that the design is insufficient to derive an inference of relative risk.<br><br>Based on inadequate human data and sufficient evidence of carcinogenicity in animals; significantly increased incidences of benign and/or malignant tumors at multiple sites in both sexes of rats, and carcinogenic effects in a series of one-year limited bioassays in mice by several routes of exposures. The classification is supported by positive genotoxicity data, adduct formation activity, and structure-activity relationships to vinyl carbamate and acrylonitrile.<br><br>• *No statistically significant increase in cancer mortality in human study (Collins, 1984).*<br>• *Confounded results in human study (Sobel et al., 1986).* |
| **Aldrin** | Yes | B2 (Probable human carcinogen — based on sufficient evidence of carcinogenicity in animals)<br><br>Inadequate human data. Two studies of workers exposed to aldrin and dieldrin (a metabolite of aldrin) did not find these workers to have an excess risk of cancer. Both studies, however, were limited in their ability to detect an excess of deaths from cancer … Exposure was not quantified, and workers were also exposed to other organochlorine pesticides (endrin and telodrin). A small number of workers was studied, the mean age of the cohort (47.7 years) was low, the number of expected deaths was not calculated, and the duration of exposure and of latency was relatively short.<br><br>• *No statistically significant excess in cancer deaths among occupational cohort (Van Raalte, 1977; Ditraglia et al., 1981). ~ same studies as cited in the dieldrin assessment* |

[a] These six substances were identified from the first 60 matches in the IRIS data file in which the word "human" was identified. Of these 60, 27 were classified as to their carcinogenicity. Of the 27, 10 identify human data in the assessment, and 4 use human data for identifying a carcinogenic classification. This table describes the reasoning provided as to why the human data were not used in the other six instances. Results of human studies are included in italics.

## 4.5 Toxicity Equivalency Factor Approach

Several comments stated that the use of Toxicity Equivalency Factors (TEFs) is inappropriate because TEFs are based on the unproven assumption of additive toxic effects, compounded by the uncertainties inherent in the generation of TEFs, their application, and the heterogeneity of effects for which they are being used. However, some comments acknowledge the value of TEFs in certain circumstances.

### 4.5.1 What Are TEFs?

In 1986, EPA published the *Guidelines for the Health Risk Assessment of Chemical Mixtures* (USEPA, 1986b). These provide the bases for evaluating human risk from exposure to combinations of two or more chemical substances, regardless of source or of spatial or temporal proximity. They were extensively peer reviewed, submitted for public comments, and reviewed again by EPA's Science Advisory Board. The guidelines have been supplemented by a technical support document (USEPA, 1988a) and the *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures* (USEPA, 2000h). These guidance documents refer the assessor to several different methods for evaluating mixtures data of varying types: epidemiologic or toxicological data on the environmental mixture (the whole mixture) to which exposures occur; test data on a tested mixture judged to be sufficiently similar to the environmental mixture; data from a group of similar mixtures; and data on the mixture components. TEFs are used in evaluating mixture toxicity from data on components. Their use is based on an assumption of additivity among similarly acting components of a mixture.

In the supplementary mixtures guidance (USEPA, 2000h), the Agency provided procedures for developing the relative potency factor (RPF) method and described TEFs as a special case of RPFs. The RPF method is component-based and relies on two types of information: toxicological dose-response data for at least one component of the mixture being addressed (referred to as the index compound or IC) and scientific judgment as to the toxicity of the other individual compounds in the mixture and of the mixture as a whole. If data are limited, the applicability of a set of RPFs may be restricted to certain effects, a specific route of exposure, or exposure duration. Application may also be limited to a certain portion of the dose-response curve. TEFs are a special case of RPFs, because they can be applied to any effect, route, or exposure duration, and across the entire dose-response curve (USEPA, 2000h). They are thus used only in relatively data-rich situations. Both RPFs and the special case of TEFs are intended to serve as interim approaches for addressing any mixture, pending the development of new data on the mixture's toxicity. The supplementary mixtures guidance (USEPA, 2000h) noted an early Agency application of an RPF-like approach for polycyclic aromatic hydrocarbons (PAHs) (USEPA, 1993a; Schoeny et al., 1998). EPA has since applied the RPF method to organophosphate pesticides in the *Guidance on Cumulative Risk Assessment of Pesticide*

*Chemicals That Have a Common Mechanism of Toxicity* (USEPA, 2002f). This was followed with the publication of biostatistical criteria for developing RPFs (USEPA, 2003e).

The RPF method (including TEFs) is based on dose addition and assumes that the chemicals in a mixture share a common toxic mode of action that is relevant to the health endpoints being assessed (see section 5.7.2 for a complete discussion of dose addition). Operationally, this means that mixture components tested in the same bioassay should have dose-response curves of similar shape between the toxicity thresholds and the maximum response. The components are assumed to be true toxicological representations of each other, although their relative toxic potencies may differ.

When one uses the RPF method, it is necessary to identify the constraints of its application. For example, a set of RPFs may be restricted to oral exposures and not be usable for exposures to the same mixture through the inhalation route; this was the recommendation for RPFs applied to PAHs. TEFs for chlorinated dibenzo-*p*-dioxins, dibenzo-*p*-furans, and coplanar polychlorinated biphenyls (PCBs) have no identified constraints.

To apply the method (see the following formula), one estimates an RPF for each mixture component; that is, one estimates component toxicity *relative* to that of the IC.

$$R_m = f_1\left(\sum_{i=1}^{n} RPF_i x D_i\right) \qquad \text{Equation 4.5-1}$$

Where:

| | | |
|---|---|---|
| $R_m$ | = | risk posed by chemical mixture |
| $f_1$ (*) | = | dose-response function of index chemical |
| $D_i$ | = | dose of the i$^{th}$ mixture component ($i$ = 1, …, $n$) |
| $RPF_i$ | = | toxicity proportionality constant relative to index chemical for the $i^{th}$ mixture component ($i$ = 1, …, $n$) |
| $n$ | = | number of chemical components in the mixture |

One estimates the RPF by choosing a response level and calculating the ratio of the dose causing that response for the component and the IC. For example, say that the $ED_{10}$ (the effective dose at which 10% of a test population exhibits an effect) for component A is 500 micrograms (µg) and that of the index chemical is 5 µg. If the index chemical RPF is set at 1, the component RPF would be 5/500 or 0.01. The next step is to calculate the IC equivalent dose for the mixture. The equivalent dose of an individual component is the product of the amount of that component in the mixture and the RPF of the component. In our example, if there is 100 µg of component A in the mixture, then its IC equivalent dose is 100 µg × 0.01, or 1 µg. These equivalent doses are summed across all components to determine the mixture total equivalent dose. Lastly, one

estimates the risk posed by the mixture by comparing its summed IC equivalent dose to the dose-response function (potency) of the IC.

TEFs have been developed for three structurally related groups of compounds which have been demonstrated to exert similar biochemical and toxic endpoints: the polychlorinated dibenzo-*p*-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), and (co-planar) PCBs. Of these, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is the most extensively researched and has served as the IC. Many studies support the TEF approach for these compounds because they share a common mechanism of action involving binding to a cellular receptor (AhR). TEFs were derived for 29 dioxin-like agents, which express each congener's toxicity as a fraction of that of TCDD, the IC (van den Berg et al., 1998).

### 4.5.2 How Did the TEF Approach Evolve?

In the 1970s and 1980s, human health risk assessments of complex mixtures of PCDDs and PCDFs were generally done only by considering 2,3,7,8-TCDD or by assuming that all congeners were equally potent to 2,3,7,8-TCDD (USEPA, 1987a, 1989b). It appears that PCDD and PCDF congeners (and co-planar PCBs) act through a common mechanism of action and induce similar biochemical and toxicological effects. However, the toxicity of individual congeners was shown to vary in different bioassays, leading to a recognition of the uncertainty associated with the earlier approach.

Eadon et al. (1986) described the first TEF-like method as a means to estimate potential human health risks associated with a PCB transformer fire in Binghamton, New York. Subsequently, EPA concluded that a TEF approach was the best available interim science policy for dealing with complex emissions of dioxins and furans from waste incineration. In 1987, EPA adopted an interim procedure, based on TEFs, for estimating the hazard and dose-response of complex mixtures containing PCDDs and PCDFs in addition to 2,3,7,8-TCDD (USEPA, 1987a).

Following adoption of the TEF methodology in the United States and Canada, the North Atlantic Treaty Organization Committee on the Challenges of Modern Society (NATO/CCMS) examined the methodology and concluded that it was the best available interim method for PCDD/PCDF human health risk assessment (NATO, 1988a, b). NATO/CCMS refined the TEFs by including more recent data sets and more *in vivo* data. The NATO/CCMS panel assigned TEFs to OCDD and OCDF, and removed TEFs for all congeners lacking chlorine in the 2,3,7,8-positions. Although it is theoretically possible to detect nearly all of the 210 PCDD/PCDF isomers in the environment, only the seventeen 2,3,7,8-substituted congeners are known to bioaccumulate. EPA officially adopted the revised TEFs in 1989 (TEFs-$NATO_{89}$), with the caveat that the methodology remain interim and continued revisions be made (USEPA, 1989b; Kutz et al., 1990). The use of the TEF methodology for human health risk assessment and risk

management purposes has since been formally adopted by a number of other countries (e.g., Canada, Germany, Italy, the Netherlands, Sweden, and the United Kingdom) (Yrjänheiki, 1992).

The dioxin-like activity of co-planar PCBs has also been established (Safe, 1990, 1994). Studies show that only PCB congeners substituted in the meta and para positions are approximate stereoisomers of 2,3,7,8-TCDD and induce dioxin-like biochemical and toxicological effects (Leece et al., 1985). In 1991, EPA convened a workshop which concluded that a small subset of the PCBs displayed dioxin-like activity and met the criteria for TEFs (Barnes et al., 1991; USEPA, 1991b).

In the years since initial adoption of the TEF methodology, additional data have been developed on the toxicological potency of individual PCDDs, PCDFs, and PCBs relative to 2,3,7,8-TCDD. A joint project conducted by the World Health Organization European Centre for Environmental Health (WHO-ECEH) and the International Programme on Chemical Safety (IPCS) resulted in development of a database of all available relevant toxicological data for dioxin-like compounds available through 1993. Following a review of almost 1,200 peer reviewed publications, 146 were selected and used to calculate TEFs for PCBs (TEFs-$WHO_{94}$). A panel of experts from eight different countries recommended interim TEFs for 13 dioxin-like PCBs, based on the reported results for 14 different biological and toxicological parameters, from a total of 60 articles (Ahlborg et al., 1994). EPA reaffirmed application of this methodology in human health risk assessment in its dioxin reassessment (USEPA, 2000i).

In 1997, a second WHO-ECEH group expanded the TEF methodology to include class-specific TEFs for mammals, birds, and fish. The resulting report (van den Berg et al., 1998) included TEFs for 7 PCDD, 10 PCDF, and 12 PCB congeners for mammals, birds, and fish (TEFs-$WHO_{98}$).

In summary, the TEF concept and associated TEF factors for PCDDs, PCDFs, and PCBs were developed through an international consensus process that has

a) Included input from numerous international experts, agencies, and stakeholders.

b) Re-examined the assumptions and limitations of the concept and the TEFs.

c) Revised TEFs as new scientific information and data became available.

This method of adding weighted doses for related compounds is considered by the international scientific community to be an improvement over other options, such as (1) ignoring risks for non-TCDD congeners, (2) treating each congener as a separate chemical evaluation (i.e., ignoring the risk posed by the mixture), or (3) assigning the full toxicity of TCDD to all similar congeners.

### 4.5.3 Is the Additivity Assumption for TEF Justified?

A substantial effort has been made to test the assumptions of additivity and the ability of the TEF approach to predict the effects of mixtures of dioxin-like chemicals. Current experimental evidence shows that for PCDDs, PCDFs, co-planar dioxin-like PCBs, and strictly AhR-mediated events, the concept of TEF additivity adequately estimates the dioxin-like toxicity of either synthetic mixtures or environmental extracts, despite the variations in relative contributions of each congener (see chapter 9 of the Dioxin Reassessment, USEPA, 2000a — particularly van Birgelen et al. [1994a, b] and van der Plas et al. [1999, 2000]; also Hamm et al., 2003).

Studies in fish and wildlife species of mixtures of PCDDs, PCDFs, and PCBs (including binary mixtures, synthetic complex mixtures, and environmental mixtures) support the additivity assumption (Zabel et al., 1995; Walker et al., 1996; Tillitt and Wright, 1997). Further, numerous studies that have examined the effects of environmental mixtures in marine mammals and avian species show a correlation between toxic effects and dietary concentrations (toxic equivalency quotient; TEQ) (Ross et al., 1996; Summer et al., 1996a, b; Giesy and Kannan, 1998; Restum et al., 1998; Shipp et al., 1998a, b; Ross, 2000).

In summary, current experimental evidence shows that for PCDDs, PCDFs, and co-planar PCBs, additivity adequately estimates the dioxin-like toxicity of both synthetic mixtures and mixtures found in the environment. Interactions other than additivity have been observed with a variety of effects in both binary combinations and complex synthetic mixtures of dioxin-like and non-dioxin-like chemicals (commercial PCBs, PCB153). However, it appears that at these high-dose exposures, multiple mechanisms of action not under the direct control of the AhR are responsible for these non-additive effects.

### 4.5.4 What Is EPA's Experience With TEFs?

EPA and other organizations have estimated the risks of exposure to PCDD and PCDF mixtures using TEFs for more than a decade. The TEF concept and supporting literature are covered in detail in chapter 9 of the draft dioxin reassessment ("Toxic Equivalency Factors [TEF] for Dioxin and Related Compounds") and also in the Part III, Risk Characterization. Both of these documents have undergone interagency review and other external peer review. The RPF method was also extensively reviewed outside the Agency. The Regions used the "RPF-like" approach developed for PAHs (USEPA, 1993a) to estimate risks posed by this group of compounds. The Agency used the RPF method in its assessment of the cumulative risks associated with exposures to organophosphorus pesticides (USEPA, 2002f, g); this effort underwent extensive internal and external peer review.

In 1998, EPA and the Department of the Interior sponsored a workshop that examined application of the TEF concept and supporting assumptions using ecological risk assessment case studies. The resulting document, *Workshop Report on the Application of 2,3,7,8-TCDD Toxicity Equivalence Factors to Fish and Wildlife* (USEPA, 2001e), provides general conclusions regarding application of and uncertainties associated with the TEF methodology in ecological risk assessment, as follows:

a) The TEF methodology is technically appropriate for evaluating risks to fish, birds, and mammals associated with AhR agonists, and it can support risk analyses beyond screening-level assessments.

b) The methodology entails less uncertainty and is less likely to underestimate risks than are methods based on single compounds (i.e., 2,3,7,8-TCDD).

c) Because total PCBs in the environment can consist of many compounds that vary in concentration and potency as AhR agonists, the TEF methodology provides a means for accounting for their variable potency.

d) The uncertainties associated with using the methodology are not thought to be larger than other sources of uncertainty within the ecological risk assessment process.

For human health assessment, issues concerning TEFs are being addressed through the dioxin reassessment, its review by the NAS, and the policy implementation document planned for after the release of the reassessment. In the interim, the Agency is obliged in many instances to use the best available science and risk assessment methods in regulatory decisions: hence the movement to use the internationally accepted practice of TEQ calculations for PCBs. From an international scientific perspective, the current TEQ concept and values are a necessary and appropriate interim procedure in advance of the development of more refined methods (van den Berg et al., 1998; USEPA, 2001e).

## 4.6 Model Use

### 4.6.1 Why Does EPA Use Environmental Models?

A wide variety of models have been used over the years to support environmental decision making by the Agency. These models often involve environmental characterization, environmental fate and transport, contact between stressors and receptor populations, relationships between exposure and adverse human health and ecological effects, quantitative estimation of risk, and economic impact models, all of which affect the manner in which the Agency chooses to address a multitude of environmental questions. Frequently, modeled

estimates become the basis for environmental cleanup, protection, or regulation. Environmental models are vital to the Agency's mission because they allow us to (1) predict the effects of future actions, (2) estimate exposures and risks where no measured data are available, and (3) attribute risks to individual sources. Models, therefore, underlie how the Agency chooses to address a multitude of environmental questions. At issue is whether the underlying assumptions, algorithms, etc., used in the modeling — and potentially the analytic structure of the models themselves — are articulated to the public well enough for stakeholders to assess the validity of modeling assessments. Further, analysis of the uncertainties in modeled estimates should be clear and transparent to those unfamiliar with the models.

### 4.6.2 How Does EPA Approach Environmental Models?

As the Agency's scientific understanding of environmental issues has improved over the years, the models we use to understand them have become increasingly complex, requiring the involvement of a wide range of scientific disciplines in the development of the models as well as in the interpretation of the results. As a consequence, the Agency and others have become increasingly focused on consistency, quality, and duplication of effort in model development, selection, and application. Such issues can, unfortunately, lead to risk characterizations that are not "transparent." Transparency requires explicitness in the presentation of each step in the risk assessment process. It ensures that any reader understands all the steps, logic, key assumptions, limitations, and decisions in the risk assessment, and comprehends the supporting rationale that lead to the outcome, including any modeling that is done for fate/transport and exposures.

Generally, EPA risk assessments address several categories of uncertainty. Each of these merits consideration. For example, for models, there are uncertainties associated with the selection of specific scientific models for each parameter in risk assessment, e.g., dose-response models, models of environmental fate and transport, and exposure models. The selection of parameters for these models creates even more sources of uncertainty. Ultimately, EPA guidance recommends that risk assessors identify those uncertainties that, if changed, would substantially affect the modeled outcomes.

### 4.6.3 What Is EPA Doing To Improve the Use of Environmental Models?

As early as 1989, EPA's SAB expressed its concerns on EPA's use of environmental models in *Resolution on Use of Mathematical Models by EPA for Regulatory Assessment and Decision Making* (USEPA, 1989c). Since that time, the SAB has continued to recommend that EPA "establish a general model evaluation protocol, provide sufficient resources to test and confirm models with appropriate field and laboratory data, and establish an Agency-wide task group to assess and guide model use by EPA."

In response to these concerns, the Deputy Administrator established the Agency Task Force on Environmental Regulatory Modeling (ATFERM) in 1992. In 1994, the ATFERM published the *Report of the Agency Task Force on Environmental Regulatory Modeling — Guidance, Support Needs, Draft Criteria and Charter* (USEPA, 1994b). The report concludes that (1) there is a need for training and technical support, (2) there is a need for model use acceptability criteria, (3) there is a need for Agency guidance for conducting external peer review of environmental regulatory modeling, and (4) there is a need for a Committee on Regulatory Environmental Modeling (CREM). To this end, the report included a section entitled "Guidance for Conducting External Peer Review of Environmental Regulatory Modeling."

In 1994, the EPA Risk Assessment Forum developed a document entitled *Model Validation for Predictive Exposure Assessments* (USEPA, 1994c), a draft protocol for model validation that defined a set of procedures for evaluating models for exposure assessments. In 1997, EPA's Officer of Research and Development (ORD) and program offices conducted an Agency-wide conference (called the Models-2000 Workshop) to facilitate Agency adherence to existing guidance on modeling, to define and implement improvements to the way in which the Agency develops and uses models, and to recommend an implementation and improvement plan for enhancing modeling within the Agency. As a follow-on to the workshop, 10 Action Teams and a Steering/Implementation Team were established to develop further action plans for improving the Agency's use of models.

In 1999, a *White Paper on the Nature and Scope of Issues on Adoption of Model Use Acceptability Guidance* was developed (USEPA, 1999c). The paper, by discussing current practices and case studies, reviewed progress in and identified issues associated with developing model use acceptability guidance.

In February 2000, the Deputy Administrator established the CREM and developed the framework of guiding principles for its activities. The CREM was established to promote consistency and consensus within the Agency on mathematical modeling issues (including modeling guidance, development, and application) and to enhance both internal and external communications on modeling activities. The CREM supports and enhances existing modeling activities by Agency program offices. The CREM provides the Agency with consistent yet flexible modeling tools to support environmental decision making, in particular as they relate to the development and implementation of programs with cross-Agency implications. Further, the CREM provides EPA staff and the public with a central point for inquiring about EPA's use of modeling.

In May 2000, the CREM initiated several cross-Agency activities that were designed to enhance the Agency's development, use, and selection of regulatory environmental models. One of these activities involved the ultimate development of modeling guidance. The CREM determined that a workshop to facilitate discussion of good modeling practices among

participants would constitute a viable starting point for developing modeling guidance. To organize the workshop and develop issue papers for discussion, a cross-Agency Model Evaluation Action Team was formed in June 2000. Over the next several months, the Action Team (1) reviewed the *White Paper on the Nature and Scope of Issues on Adoption of Model Use Acceptability Guidance* to identify model evaluation issues; (2) developed a list of components, or elements, to follow in evaluating a regulatory environmental model; (3) identified several general principles to support the use of model evaluation guidance; (4) developed an organizational plan for the workshop that covered the identified model evaluation elements; (5) held the workshop; and (6) developed a summary report of the workshop.

In February 2003, EPA's Administrator directed the CREM to draft Agency guidance on environmental models and create a Web-accessible inventory of EPA's most frequently used models. Both these efforts address the transparency and uncertainty with which environmental models — including fate and transport models — are used to make decisions. First, in its guidance on environmental models, EPA intends to recommend that the modeling community document all steps taken to develop and evaluate models; the guidance will also recommend best practices for evaluating models. Second, EPA intends to house this documentation of model development and evaluation within the Web-accessible inventory of models.

In addition to the activities noted above, EPA issued Order 5360.1 A2 (USEPA, 2000c), which covers collection and use of environmental data, including information produced from models in the areas of (1) use of a systematic planning approach to develop acceptance or performance criteria; (2) approved Quality Assurance Project Plans (QAPPs) or equivalent documents; and (3) assessment of existing data used to support Agency decisions or other secondary purposes (to verify that they are of sufficient quantity and adequate quality for their intended use). Requirements for QAPPs are provided in EPA QA Manual 5360 A1 (USEPA, 2000d) for EPA personnel and EPA QA/R-5 (USEPA, 2001f) for extramural personnel.

The atmospheric dispersion models we use (e.g., the Industrial Source Complex model, or ISC, and AERMOD) have been through rigorous scientific peer review. To improve our tools for modeling fate and transport from air into other media, EPA developed the Fate, Transport and Ecological Exposure Module of the Total Risk Integrated Methodology, or the TRIM.FaTE model. TRIM.FaTE not only handles the partitioning of pollutants between environmental media, but does so in an interactive way, allowing pollutants to move back and forth between media throughout the simulation according to the simulation conditions. The input values, any assumptions, and transfer algorithms are completely transparent and accessible to the user because they are stored in an electronic "library" accessed by the model, not "hardwired" into the model itself. This feature allows the Agency to easily replace algorithms or values as new findings indicate they are appropriate, or as appropriate for different applications. A variety of documents have been prepared on the model and algorithms compiled in the library for applications to date (e.g., USEPA, 2002h, i, j); these documents are available along with the

model on the Agency's Web site. TRIM.FaTE also includes features for sensitivity and Monte Carlo analyses to assist in developing analyses of uncertainty and variability for multimedia risk assessments.

With regard to inhalation exposure models, EPA has developed methods that use human time activity patterns and stochastic sampling to derive probabilistic activity-based estimates of exposure. The current example of this is the Human Exposure Event module of TRIM (the TRIM.Expo inhalation model, also known as the air pollutant exposure model or APEX). This model is derived from the probabilistic National Ambient Air Quality Standards exposure model (pNEM) and has been developed for both criteria and HAP applications. TRIM.Expo is available on the EPA Web site and by its design it relies on data from input files that are completely transparent to the user.

Regarding uncertainty analysis, each Agency risk assessment needs to include an analysis of important uncertainties that, if reduced, might change the outcome of the assessment (see the Agency's 1997 *Policy for Use of Probabilistic Analysis in Risk Assessment* and *Guiding Principles for Monte Carlo Analysis*). The Agency believes that it is generally preferred that quantitative uncertainty analyses be presented in each risk characterization, but there is no single recognized guidance on how to conduct such uncertainty analyses. Even if the uncertainty analysis is conducted qualitatively, the Agency still considers it to be of great value to a risk manager.

### 4.6.4 How Does EPA Approach Fate and Transport Modeling and Account for Uncertainty?

Once chemicals are released to the environment, they are subject to a wide range of environmental processes that influence their fate and transport. These processes may affect their location (e.g., meteorology, wind, water currents), their chemical structure (e.g., transformation, reactivity), and the media in which they partition or reside. These processes' level of influence varies significantly, often depending on the chemicals' physical and chemical properties as well as the conditions they encounter in the environment.

This section describes the modeling conducted under OPP to illustrate EPA's approach to fate and transport modeling. (Fate and transport modeling in OPP efforts typically encounters multiple media and therefore is a good example of broad modeling practices used by EPA.) Fate and transport models are used in OPP to estimate ambient concentrations for ecological risk assessment and drinking water concentrations for human risk assessments. For ecological risk assessments, estimated environmental concentrations are divided by a toxicity reference value to yield an index that is compared to a level established to constitute an acceptable risk. For human risk assessments, estimated drinking water concentrations are incorporated into a total dietary assessment, which is then evaluated for risk.

EPA responds to uncertainty in these assessments depending on the type of decision to be made. In estimating pesticide concentrations in water resources, OPP uses a tiered screening approach (consistent with the general approach discussed in chapter 2) that involves a combination of modeling and analysis of monitoring data. In OPP's tiered modeling system, the program begins with a public-health-protective, national, highly vulnerable estimate and then, if concerns exist, progresses to regional and eventually local, probabilistic, more resource-intensive estimates. This triage approach is designed to provide a thorough analysis of each pesticide, while focusing OPP's efforts on those pesticides that pose the greatest potential risk.

The intent of the screening approach is to estimate pesticide concentrations in water from sites that are vulnerable to runoff or leaching so that the program can be confident that any pesticide that passes the screening tiers poses a low possibility of harming human health, wildlife, or the environment. Therefore, during these screening assessments, the intent is to address the uncertainty regarding the most sensitive scenario that may exist in areas of pesticide application.

These initial screening assessments establish the environmental setting and input parameters that would result in a high-end exposure of a pesticide after application. Half-life data from laboratory studies serve as inputs, as described in the publicly available input parameter guidance. The initial screening levels models used by OPP are: for the surface water pathway, GENEEC2 (GENeric Estimated Exposure Concentration) and FIRST (FQPA Index Reservoir Screening Tool), and for the groundwater pathway, SCI-GROW (Screening Concentration in GROund Water).

If a pesticide raises a concern after the initial screening tier, a set of higher-tiered surface water models called PRZM/EXAMS are used to conduct a more refined assessment. These models, developed by ORD, allow the Agency to focus on specific sites (by accounting for site-specific properties including soil type, weather, and agricultural practices) or watersheds in its analysis. At the present time, OPP does not have a higher-tiered groundwater model.

A tier II or PRZM/EXAMS refined water assessment uses a region-specific site with an agriculturally relevant soil that is vulnerable to runoff and is of a large extent in the watershed (i.e., a benchmark soil). The assessment also uses 30 years of monitored weather data for the site and maximum pesticide use rates. The combination of site characteristics and fate property inputs is designed to approximate a "high-end" exposure scenario. A tier II drinking water assessments uses an index reservoir drinking water source (USEPA, 1998d), and ecological exposures assessments are conducted using a small fixed-volume water body, the "farm pond." The drinking water reservoir modeled is Shipman City Lake in Illinois; the farm pond is located in Georgia. Drinking water estimates are adjusted to reflect the amount that a major crop, specifically corn, wheat, cotton, or soybeans (USEPA, 1999d), may be grown in a watershed using a national percent crop area factor. Crop area factors for other crops or combinations of

crops to which a pesticide may be applied in the same watershed are assigned a national default value representing the highest percent of a watershed in agricultural production.

Recently, region-specific default percent crop areas were developed to reflect different cropping practices across the United States. Estimates may be further refined by focusing on additional region-specific use sites and typical usage and application practices. Estimated concentrations from PRZM/EXAMS in the initial tiers are a 1 in 10 year occurrence frequency and, for cancer risk assessments, a 30-year overall mean. Additional refinements are pesticide-specific and may be performed at the request of risk managers seeking to focus on risk mitigation. These refinements may include a probabilistic dietary assessment using the total PRZM/EXAMS output of 30 years of daily values.

If degradates are observed in the laboratory studies that are of toxicological significance, have or are predicted to have pesticidal activity based on structure, or are a significant component of the residues in an environmental compartment, modeling is conducted on these degradates. The screening-level models are bypassed in favor of PRZM EXAMS to model degradates and are assessed in separate modeling runs. The amount of data available to serve as inputs varies. Half-lives can sometimes be obtained from the registrant-submitted studies if those studies were conducted for a period greater than the half-life of the parent. Defaults are used in the absence of data, as described in the input parameter guidance.

### 4.6.5 Does EPA Consider or Review Existing Data on Environmental Concentrations?

Monitoring data are preferred when estimating concentrations; however, monitoring data are generally either not available or inadequate. Monitoring data are not available for new, never-registered chemicals. No or little monitoring data will be available for a chemical that has little use in the United States or has a "niche" market. In some cases, after reviewing the monitoring data, modeling outputs are selected because monitoring was conducted in areas in which the chemical was not used or had little use. Generally, monitoring data in targeted pesticide use areas are most useful for estimating long-term time-weighted concentrations and potential chronic impacts. Low sample frequency and lack of statistically representative use sites limits the usefulness of monitoring data for estimating peak concentrations and its potential acute impacts. Sufficient data are sometimes available for widely used pesticides, pesticides that are in multi-residue monitoring methods, or pesticides that have the interest of the public. In all risk assessments, available monitoring results are compared to screening-level modeling estimates.

### 4.6.6 Does EPA Consider and Use Case-Specific Data in Models? (Or: How Does EPA Decide To Modify Default Assumptions in Later Tiers Using Models?)

As in all aspects of risk assessment, EPA prefers to use specific data rather than default assumptions when working with models. Default assumptions are used when there are data gaps. When data of sufficient quality are available, EPA uses such data. For example, the pesticide industry is responsible for submitting the data that serve as the inputs into these models. Data requirements are described in 40 CFR Part 158, and the protocols for conducting these studies are discussed in the Pesticide Assessment Guidelines. Registrants may use Organisation for Economic Cooperation and Development (OECD) guidelines or other guidance provided the data meets the purpose of the study. In selecting model parameters, OPP established publicly available guidance for selecting inputs (Water Models, Guidance for Selecting Input Parameters). The amount of data available determines how confident an input will become.

### 4.6.7 How Will EPA Improve Pesticide Fate and Transport Modeling in the Future?

The Agency's goal is to develop realistic estimates of pesticide concentrations in ambient and drinking water. The following efforts are underway to improve its tools and technology:

a) *Ground water*: A tier II model that generates daily or short-time-step concentrations for use in a probabilistic dietary assessment will be in use in late 2004. Currently available groundwater models will be assessed jointly with the Pest Management Regulatory Agency/Health Canada. In addition, a refined version of SCI/GROW is being investigated that uses all of the data from prospective groundwater studies submitted by the pesticide industry.

b) *Surface water*: Modifications are being made to PRZM/EXAMS to allow the size of the receiving body to be adjusted to reflect the size of a typical pond or reservoir in the area being modeled. A geographic information system is being developed that maps all of the drinking water intakes in the United States and their associated watersheds. Additional data layers will be added to allow the Agency to model specific sites and their cropping patterns. Spray drift is a fixed variable in PRZM/EXAMS. Spray drift models will be linked to PRZM/EXAMS to generate realistic drift inputs (USEPA, 1997e, 1999e). Watershed-scale regression models are being developed in cooperation with the U.S. Geological Survey to estimate a distribution of concentrations according to watershed properties and pesticide usage areas (Larson et al., in press; Larson and Gilliom, 2001).

c) *Drinking water treatment*: A workshop planned for January 2004 will address the types of data needed to be able to factor drinking water treatment into estimates of pesticide concentrations. Additional workshops and a Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel meeting are expected to finalize any new data requirements that may be imposed on industry to provide data.

d) *Tiered water exposure assessment process*: Interaction is conducted with the human health risk assessors and risk managers in OPP to standardize the modeling tiers and generate the data needed for conducting probabilistic exposure assessments earlier in the process and to identify the tools and technologies to routinely conduct probabilistic analyses. The program's goal is to account for daily and seasonal variations in residues over time due to the time of application and runoff/leaching events, year-to-year variations in weather patterns, and place-to-place variability in residues caused by the water source and regional/local factors (such as soil, geology, hydrology, climate, crops, pest pressures, and usage) that affect the vulnerability of a source.

### 4.6.8 Does EPA Use Worst-Case Assumptions in Modeling That Are Unreasonable and Do Not Account for Degradation, Partitioning, and "Sinks" in the Environment?

Default assumptions are frequently used in modeling, specifically in exposure assessment for each medium or cumulative assessment. Model inputs for fate and transport use the most sensitive soil type, combined with the lowest rates of degradation, air exchange, or water dispersion, and 100% of mass used in each scenario. These default assumptions are used to address data gaps and to ensure that risks will not be underestimated given those uncertainties.

The degree to which defaults are incorporated into any modeling exercise depends on the purpose of the assessment and the availability of data. In a screening assessment with little or no available data, several default assumptions may be used, such as the assumption that the chemical is stable and not subject to transformation or degradation. In fact, the potential for degradation of pesticides in the environment is and has been incorporated into many aspects of OPP risk assessments, as illustrated in the examples below.

a) *Food*: In dietary exposure and risk assessments, the use of monitoring data (e.g., the U.S. Department of Agriculture's Pesticide Data Program or market basket surveys) incorporates potential degradation of pesticides from "farm gate" residues to "dinner plate" residues. Dietary exposure estimates include consideration of the impact of typical consumer practices such as washing and peeling, and commercial practices during processing (juicing, blanching, freezing,

etc.), that may reduce or concentrate residues. For pesticides found largely on the surfaces of fruits and vegetables, considerable reduction of residues has been observed in processing studies, as well as in monitoring data when results for fresh and processed commodities can be compared. Concentration of residues is observed more often in dried commodities.

b) *Water*: In determining the potential concern for pesticide residues in drinking water, OPP risk assessments incorporate modeled surface and groundwater values generated using chemical-specific environmental fate inputs (e.g., hydrolysis, anaerobic and aerobic metabolism, and microbial degradation). If sufficient data are available, assessments also consider the impact of treatment on potential residues in drinking water. Finally, monitoring data are evaluated for use in the risk assessment. (If used quantitatively, then degradation in the environment is inherently considered; if used qualitatively, a comparison between modeled and monitored values is provided as characterization.) Currently, the impact of water treatment on drinking water concentrations is being assessed qualitatively. Following a recommendation from the FIFRA Scientific Advisory Panel meeting during September 2000 (USEPA, 2000j), ambient concentrations are used in a drinking water assessment unless there are data on the impact of water treatment.

c) *Residential*: Exposure and risk estimates for pesticides in residential settings are generated using both default assumptions and chemical-specific data. The default assumptions are used largely in screening-level assessments. Chemical-specific or activity-based information is incorporated as available. The timing and duration of exposure are the most critical factors in estimating risk. For example, in estimating toddlers' dermal exposure to pets, the highest residues on the pets' fur (and the highest exposure to the child) are expected to occur on the day of application. For longer-term exposure durations, the decline in residues on fur is taken into consideration. In assessing dermal exposure to golfers from treated turf, short-term non-cancer risks are estimated on the day of application at the label rate, but cancer risks are estimated based on average application rates, and take into account the frequency of golfing over a lifetime. Exposure assessments that consider "sinks" (e.g., carpeting or soil) can be problematic due to data deficiencies, but there is an assumption that less pesticide would be available for dermal exposure, and an adjustment can be made in calculating the risks using, for example, a "soft surface" exposure factor. In estimating indoor inhalation exposures, the registered label is used as the basis for determining the re-entry time and ambient concentrations, but breathing rate and level of activity considerations can be included in the final risk estimates.

d) *Occupational*: For occupational handlers (those who mix, load, and/or apply pesticides), estimated exposures are driven by pesticide formulation type, the application method (e.g., aerial or ground boom), and the number of acres treated per day. Exposure and risk estimates are calculated assuming different levels of protective clothing and using the geometric mean of exposures (unit exposures based on real data conducted with a variety of formulation types and application equipment). Dermal risks estimated for different clothing types adjust for potential protection afforded by a second layer of clothing or gloves, and consider how much of the pesticide may actually get on the skin. Inhalation risks calculated using unit exposure data already take into account activity levels, breathing rate, and the potential ambient air concentrations for various pesticide formulations during application. For harvesting or other foliar-contact activities undertaken following application, the degradation of pesticide residues on leaves over time is considered and used in the calculations. For estimating cancer risks, exposures are calculated based on typical application rates, then amortized over a lifetime.

In all of the above, the best available data are used to estimate exposure and risk and, to the extent possible, reflect the use pattern for specific pesticides. If risks exceed the level of concern, more specific information about a particular pesticide/crop/pest combination may be obtained to provide more refined estimates of exposure and risk.

# 5. SITE- AND CHEMICAL-SPECIFIC ASSESSMENTS

## 5.1 Overview

This chapter addresses EPA's use of site- or chemical-specific information in risk assessments and the default assumptions that are often used in the absence of such information. Many comments that EPA received regarding site- and chemical-specific information pertain to assessments concerning waste sites. Therefore, this chapter largely focuses on the comments in the context of human health risk assessment guidance for Superfund, which is also followed by the Resource Conservation and Recovery Act Corrective Action program, as appropriate. EPA also received a number of comments directed at risk management decisions made at specific Superfund sites. Indeed, many of the site-specific comments repeat comments filed with the Agency either during the development of the proposed plans or during the public comment period for proposed plans at those specific sites. The EPA officials responsible for implementing the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) at those sites either have already responded to those comments or will be responding to them through the CERCLA process. Therefore, this document does not respond to comments on specific sites. Further information about the role of site- and chemical-specific data in other regulatory programs may be available from individual program offices (refer to the EPA home page at http://www.epa.gov/epahome/programs.htm) and by other examples found elsewhere in this document (e.g., the "porch potato").

Due to the uniqueness of individual waste sites, site-specific information plays an important role in risk assessments and management decisions on a regional level. In general, EPA considers site- and chemical-specific information in risk assessments when it is available and appropriate, then uses default assumptions when data gaps exist. Default assumptions are typically based on scientific, peer reviewed data and aid in the evaluation of aspects such as various age groups and various populations (e.g., residents, workers, Native American tribes, subsistence fishers).

In terms of program goals and protecting public health, it is appropriate for EPA to use default assumptions when data are not available to use in a risk assessment. Note that the default exposure values are not all high-end, are not "worst-case," and have data supporting their derivation (in addition to being peer reviewed). The following table presents examples of commonly used default exposure factors and their associated percentile distribution, which range from median to 95$^{th}$ percentile.

| Table 5-1: Examples of Default Exposure Values With Percentiles | | |
|---|---|---|
| **Exposure Pathway** | **Percentile** | **Source of Data** |
| Drinking water consumption: 2 liters/day | 90th | Approximately a 90th percentile value (USEPA, 2000k). |
| Soil ingestion rate for children: 200 mg/day | 65th | Analyses and distributions constructed by Stanek and Calabrese (1995a, b, 2000) places the 200 mg ingestion rate around the 65th percentile of average daily intakes throughout the year. The Stanek and Calabrese analyses suggests that ingestion rates for children in the top 10% (i.e., the high end) of the distribution would be greater than 1,000 mg/day. |
| Residence duration: 30 years | 90th<br>80th<br>90th–95th | For home owners, farms, and rural populations; 30 years is greater than the 95th percentile residence time for renters and urban populations. |
| Body weight: 70 kg | 50th | For males and females 18 to 75 years old (NCHS, 1987) |

Risk assessments and the collection of data and other information are conducted on a site-specific basis by the EPA regional offices, within the regulatory requirements of the waste programs. Public participation is a critical component of the regional Superfund process. For example, public knowledge about the history of disposal practices at an abandoned waste site often informs the conceptual site model design and development of sampling plans around the site. Another example of site-specific information involves people's interaction with their environment. Some populations may eat more fish than would be estimated using EPA default exposure values based on differences in consumption patterns. If resources are available, and considering the population and other factors, a site-specific survey may be conducted to assess the consumption patterns for the specific population.

Often, when EPA actively seeks public participation, the Agency staff also need to clarify what constitutes useful data for risk assessment. For example, a soil sample placed in a household container does not provide useful or adequate data for a risk assessment. EPA would explain that information about the extent of contamination is collected following the quality assurance/quality control guidance (USEPA, 2000l) using a standard protocol with quality control and quality assurance procedures.

To address the concerns and comments about use of site- and chemical-specific data and to fully understand assessments of cancer risks and non-cancer health hazards, it is important to

understand their underpinnings. Among these are the regulatory requirements and purpose of the programs for which the assessments are conducted, whom the programs seek to protect, the role of default assumptions and site- and chemical-specific information in risk assessments, the types of risk assessment performed, and factors considered in risk management decisions. The following sections present information about these aspects of the Superfund program.

### 5.1.1 What Are the Superfund Program's Purpose and Regulatory Requirements?

CERCLA, 42 U.S.C. 9605 (CERCLA, 1980), as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA), P.L. 99-499 (SARA, 1986), requires that actions selected to remedy hazardous waste sites be protective of human health and the environment. The National Oil and Hazardous Substances Pollution Contingency Plan, or NCP, is the regulation that implements CERCLA (NCP, 1990a, b). The NCP establishes the overall approach for determining appropriate remedial action at Superfund sites (NCP, 1990a, b).

The national goal of the Superfund program is to select remedies " ... that are protective of human health and the environment, that maintains protection over time, and that minimizes untreated waste"(NCP, 1990a). This means that the Superfund program needs to protect human health and the environment from current and potential future threats posed by uncontrolled hazardous substance releases. Pursuant to the NCP, decisions at Superfund sites involve consideration of site-specific information and cancer risks and non-cancer health hazards associated with both current and future land use conditions.

### 5.1.2 Whom Does the Superfund Program Seek To Protect? The Reasonable Maximum Exposure Scenario

One of the policy goals of the Superfund program is to protect a high-end, but not worst-case, individual exposure: the reasonable maximum exposure (RME). The RME is the highest exposure that is reasonably expected to occur at a Superfund site. As described in the preamble to the NCP (NCP, 1990a, b), the RME will

> ... result in an overall exposure estimate that is conservative but within a realistic range of exposure. Under this policy, EPA defines "reasonable maximum" such that only potential exposures that are likely to occur will be included in the assessment of exposures. The Superfund program has always designed its remedies to be protective of all individuals and environmental receptors that may be exposed at a site; consequently, EPA believes it is important to include all reasonably expected exposures in its risk assessments ...

In addition to evaluating the risks to the RME individual, EPA evaluates risks for the central tendency exposure (CTE) estimate, or average exposed individual. This approach is consistent with the Risk Characterization Policy and Handbook (USEPA, 1995a, 2000a). CTE estimates give the risk manager additional information to consider while making decisions at a

site. Pursuant to the NCP, decisions at Superfund sites are based on cancer risks and non-cancer health hazards associated with RME estimates under both current and future land use conditions.

### 5.1.3 Is the RME Overly Conservative?

EPA received comments implying that the RME represents a "worst-case" or "overly conservative" exposure estimate. As described above, the RME is not a worst-case estimate — the latter would be based on an assumption that the person is exposed for his or her entire lifetime at the site. As the preamble to the NCP (NCP, 1990a, b) states:

> … The reasonable maximum exposure scenario is "reasonable" because it is a product of factors, such as concentration and exposure frequency and duration, that are an appropriate mix of values that reflect averages and 95$^{th}$ percentile distribution …

The RME represents an exposure scenario within the realistic range of exposure, since the goal of the Superfund program is to protect against high-end, not average, exposures. The "high end" is defined as that part of the exposure distribution that is above the 90$^{th}$ percentile, but below the 99.9$^{th}$ percentile. The approach was developed with technical support from ORD and is consistent with the peer reviewed EPA Exposure Assessment Guidelines (USEPA, 1992a).

Further, for some individual activity patterns, the RME may be modified based on site-specific considerations (e.g., workers who may remain in one location, individuals who live in a residence for their lifetimes, Native Americans remaining on tribal lands, children exhibiting pica behavior). The Superfund program supports the development of reasonable risk assessments that address the exposure and risk to all segments of the community, not only the "average" individual. EPA seeks to protect "sensitive populations," segments of the general population that are at greater risk, either because of particular sensitivity to the toxic effects of certain chemicals or because they experience higher exposures than the general population, as children do. Under the NCP (NCP, 1990a, b), the Superfund program achieves this goal of protecting public health by using the RME approach.

## 5.2 Superfund Risk Assessment Guidance

### 5.2.1 What Are Some Types of Risk Assessment Under Superfund?

EPA conducts risk assessments that vary in length and degree of detail, from screening-level assessments to comprehensive site-specific baseline assessments of hazardous waste sites. Risk assessments may focus on chronic, long-term exposures and/or evaluation of acute exposures that may require an emergency response. They are often conducted using a triage approach, beginning with a screening-level assessment to determine if a more comprehensive assessment is necessary (USEPA, 1988b, 1996a, 1999a, 2001b, 2002k).

### 5.2.2 How Does Superfund Use Site-Specific Concentration and Maximum Concentration Data in Risk Assessment?

When evaluating the information used in risk assessment, it is important to consider the level of risk assessment being conducted. In screening-level assessments, risk assessors may assume that receptors are in contact with the maximum concentration for the entire exposure duration to place an upper bound on exposure estimates. This approach is used to ascertain which exposure pathways or contaminants are most important, so that resources may be more efficiently used in developing a comprehensive risk assessment for the most important contaminants and exposure pathways.

The concentration term is a sensitive parameter in an exposure assessment. In most risk assessments, an estimate of the average chemical concentration is used to represent the exposure concentration or exposure point concentration (EPC) for assessing long-term or chronic exposure scenarios. Because of the uncertainty associated with estimating the true average concentration, the 95% upper confidence limit of the arithmetic mean ($UCL_{95}$) is used to account for uncertainties (USEPA, 1992b, 2002l).

The $UCL_{95}$ is highly dependent on the number of samples collected. The greater the number of environmental samples, the closer the $UCL_{95}$ will be to the true mean. Sometimes, environmental data are limited or there is extreme variability in measured or modeled data. When this is true, the upper confidence limit can be greater than the highest measured or modeled concentration and neither the average nor $UCL_{95}$ is a reliable estimate of the exposure point concentration. If additional data cannot practicably be obtained, it is more appropriate to use the highest measured or modeled value as the EPC. In cases where the true mean is actually higher than this maximum value, especially if sampling data are very limited, the maximum concentration is used in the calculation (USEPA, 1992b).

Site-specific contaminant data are collected for all sites. In some cases, the maximum concentration is used when data are limited or the calculated $UCL_{95}$ exceeds the maximum concentration. However, the maximum concentration is rarely used for hazardous waste sites in which a comprehensive baseline risk assessment is being conducted, since reliable site-specific data are typically collected as part of the remedial investigation (RI) process. During the planning and scoping phase, the minimum number of samples that is sufficient to characterize environmental media is determined using EPA's data quality objectives process (USEPA, 2000l). The reality is that the number of samples collected from each medium, and the resulting certainty achieved, need to be balanced by the resources available and the purpose of the assessment.

Lastly, it is important to note that it is appropriate to assume individuals are exposed to the maximum concentration when acute exposure scenarios are being evaluated, such as those conducted for emergency removal actions. Another example: when "hot spots" with relatively

high concentrations are present in surface soil, then acute toxicity is of most concern and the focus is on measuring or estimating short-term, peak environmental concentrations.

### 5.2.3 Has the Superfund Guidance Been Updated and Externally Peer Reviewed Since It Was First Released in 1989?

EPA continues to develop and peer review guidance to assist risk assessors conducting various types of risk assessments in Superfund. Since the development of the Superfund Public Health Evaluation Manual (USEPA, 1986c), and later in 1989 the RAGS Part A, Human Health Evaluation Manual (USEPA, 1989a), the Superfund program developed appropriate guidance to help risk assessors evaluate risks to the RME individual (USEPA, 1996a, 1999a, f, 2001a, b, g, 2002k). The guidance consistently included internal review by risk assessors across the Agency, external comment by stakeholders, and external peer review where appropriate based on our peer review guidance.

External peer reviews have been conducted for the majority of guidance documents developed since 1989. For example, as early as 1989, the RAGS Part A was reviewed by the Agency's Science Advisory Board. In addition, since the issuance of the Agency policy on peer review in 1994 (USEPA, 1994d) and subsequent handbook (USEPA, 2000b), where appropriate and consistent with peer review policies, Superfund Risk Assessment guidance has undergone external peer review. The external peer review meetings have included opportunities for the stakeholders to participate through public comment. All guidance documents are made available on the Internet for public comment and addressed in the final documents.

The probabilistic risk assessment guidance is one example of the process used in the Superfund program to develop guidance. In that case, Superfund identified the emerging science, developed an EPA workgroup to evaluate the available science and its application within the Superfund program, released the draft guidance document for public comment, and conducted an external peer review. The guidance document (USEPA, 2001a) provides program-specific information regarding the conduct of probabilistic risk assessments and supplements the earlier policy on this issue (USEPA, 1997c). In addition, EPA has developed training courses on the application of this methodology within the Superfund program. To date, probabilistic risk assessment methods have been used or are being developed at several sites to evaluate cancer risks and non-cancer health hazards (USEPA, 2000m, 2003f).

The risk assessment processes used at individual Superfund sites have evolved over time based on new science and EPA's understanding of new potential exposure pathways. For example, in the early days of the program, dermal exposure was not fully evaluated based on a lack of dermal exposure information; this guidance was updated including an external peer review (USEPA, 2001g). The Superfund program has updated other guidance documents, including external peer review where appropriate to address the current scientific knowledge.

## 5.3 Default Assumptions and Site-Specific Information

### 5.3.1 What Is the Role of Default Assumptions and Site-Specific Information in Superfund Risk Assessments?

Risk assessments incorporate both *default* assumptions and *site-specific* information. The supplemental guidance document, "Standard Default Exposure Factors" (OSWER Directive 9285.6-03, March 25, 1991), presents the Superfund program's default exposure factors for calculating RME exposure estimates (USEPA, 1991c). This guidance was developed in response to requests that EPA make Superfund risk assessments more transparent and their assumptions more consistent. However, the guidance clearly states that the defaults should be used where "there is a lack of site-specific data or consensus on which parameter to choose, given a range of possibilities."

The table in section 5.1 presented examples of default exposure values and the percentile of the population the values represent, as well as the peer reviewed studies supporting these assumptions. Again, the RME approach uses default values designed to estimate the exposure of a high-end individual in the 90$^{th}$ percentile of exposure or above (USEPA, 1992a). Consistent with this guidance, relevant default assumptions for various activity levels and age groups are used for drinking water consumption rates, soil ingestion rates, residence times, body weight, and inhalation rates. The table illustrates the range of percentiles — some defaults included the 50$^{th}$ percentile (e.g., body weight), 80$^{th}$, 90$^{th}$, and 95$^{th}$ percentiles.

Although the Superfund program routinely uses default assumptions to assess the risk to the RME individual at many sites, the characteristics of the surrounding population change from site to site. For example, the distributions of individual residence times will vary depending on whether the site is located in a rural or an urban area. Individuals in rural communities are likely to have longer residence times than individuals in urban communities. Thus, a default value of 30 years may fall at the 80$^{th}$ percentile for farmers but above the 95$^{th}$ percentile for renters in an urban setting. The extent to which a single default value will impact the final exposure estimate depends on the values and variabilities of all the parameters used to estimate exposure. The goal is to estimate an individual exposure that actually occurs and is above the 90$^{th}$ percentile. In some cases, use of default assumptions may produce an estimate near the 90$^{th}$ percentile; in others, the estimate may be higher in the range.

In general, Superfund's default factors are designed to be reasonably protective of the majority of the exposed population. The assumptions used in Superfund's risk assessments are consistent with the 90$^{th}$ percentile or above and the Agency's exposure assessment guidelines (USEPA, 1992a). Default exposure factors used to assess the RME are a mix of average and high-end estimates (see table 5-1). The use of these default exposure assumptions does not

automatically result in an overestimation of exposures. These following sections illustrate this with several examples.

### 5.3.2 What Is the Drinking Water Consumption Rate?

EPA uses site-specific data on drinking water consumption patterns that have met appropriate review criteria at individual sites, but these site-specific studies are generally limited or absent. In these cases, to evaluate exposures to contaminants in drinking water, the Superfund program applies a default ingestion rate of 2 liters per day for the RME individual (USEPA, 1989a, 1991c, 2000k). The 2 liters per day assumption represents a reasonable default assumption. The results of the 1994 - 1996 Continuing Survey of Food Intake by Individuals (CSFII) analysis indicate that the arithmetic mean, 75th, and 90th percentile values for adults 20 years and older are 1.1, 1.5, and 2.2 liters per day, respectively (USEPA, 2000k). The 2 liters per day value represents the 86th percentile for adults.

In certain parts of the United States — or, during summer seasons, throughout the country — people may drink 4 to 4.5 liters per day (USEPA, 1997d). Similarly, a report from OW on the analysis of the CSFII data shows that the 99$^{th}$ percentile is 4.2 liters per day (the report provides no maximum) (USEPA, 2000k). In some circumstances, then, this default value may underestimate cancer risks and non-cancer health hazards for the reasonable high-end exposure scenario.

### 5.3.3 What Is the Inhalation Rate?

The 1989 RAGS Part A describes 20 cubic meters per day as an average inhalation rate and 30 cubic meters per day as an upper bound (USEPA, 1989a, 1997d). The inhalation rate varies depending upon the activity levels of the exposed individual (USEPA, 1997d). The default value of 20 cubic meters per day reflects a typical mix of activity levels; however, more accurate, site-specific activity level data may be used when available. The default value is consistent with the recommendations from the International Commission on Radiological Protection (ICRP) data set for the "reference man" (ICRP, 1981; cited in USEPA, 1997d). The ICRP used values of 20 to 23 cubic meters per day. These values assumed 16 hours of light activity and 8 hours of resting. Daily inhalation rates for individuals performing activities at levels other than resting or light activities are not presented. Thus, the values or the exposure period may need to be adjusted for individuals (e.g., construction workers) with moderate and/or heavy activity levels where inhalation rates would be higher.

### 5.3.4 What Is the Exposure Duration for Residences?

When site-specific data are not available or appropriate, the default residential exposure durations are 30 years. This value represents the 90$^{th}$ percentile for home owners and the 90$^{th}$ to

95th percentile for rural populations (USEPA, 1991c, 1997d). It is not reflective of the residence times for farms, where 30 years represent an 80th percentile for this specific population (USEPA, 1997d).

Some site-specific exposures may continue even if the individual changes residences, provided that he or she remains in the general area of the site. For example, Connelly et al. reported that individuals are willing to travel up to 34 miles to continue fishing in a favorite area (Connelly et al., 1992). At an existing Superfund site where more than 100 river miles have been contaminated, it is thus possible for an angler to change his residence within the surrounding counties and continue to fish from the impacted river. An assumption of an exposure duration based only on residence time in this instance would underestimate the risks to the angler.

### 5.3.5 What Is the Exposure Duration for Workers?

It is Superfund practice to assume that workers are generally exposed for a duration of 25 years, unless site-specific data are available (USEPA, 1991c). The exposure duration is assumed to be equivalent to job tenure for receptors in an occupational scenario. An analysis of Bureau of Labor Statistics data shows that the projected job tenure varies from a few months to 50 years, depending on the specific occupational category (Burmaster, 2000).

More specifically, the projected 95th percentile values for job tenure for men and women in the manufacturing sector are 25 years and 19 years, respectively. The projected 95th percentile job tenure values for workers in the transportation/utility and wholesale sectors are only somewhat less than the values for manufacturing workers — 22 years and 18 years for men and women, respectively. Values are lower for other non-industrial sectors — approximately 13 years for workers in the finance and service sectors, and 7 years for retail workers. Note that the 95th percentile is within the range for the high-end distribution, which is defined as that part of the exposure distribution that is between the 90th and 99.9th percentiles (USEPA, 1997d).

Thus, the 25-year default value is an estimate of the 95th percentile exposure duration for workers across a wide spectrum of industrial and commercial sectors. This value may be an overestimate or underestimate of the actual exposure duration, depending on the particular circumstances of the employment, e.g., for future workers on a cleaned-up Superfund site. EPA supports the use of alternative exposure durations, if they are based on adequate data on job tenure and the anticipated industrial/commercial site use. For example, in the evaluation of a construction worker scenario, the exposure duration is typically 1 to 2 years, representing the time an individual is actually engaged in construction activities on the site (USEPA, 2002k).

As mentioned above, EPA also typically evaluates CTE scenarios, which is consistent with EPA's guidance on risk characterization (USEPA, 1995a, 2000a). For example, 6.6 years is often used as a CTE duration for workers. This value is based on the median occupational tenure

of all workers in the United States (USEPA, 1997d). Evaluation of CTE is usually limited to comprehensive baseline risk assessments, whereas 25 years is the default exposure duration for screening-level assessments in EPA's waste programs.

### 5.3.6 What Is the Ingestion Rate for Construction Workers?

Comments submitted implied that the Superfund program continues to use an ingestion rate of 480 milligrams per day for construction workers. This ingestion rate was updated in 2000 with EPA's externally peer reviewed soil screening level guidance to 330 milligrams per day (USEPA, 2002k).

### 5.3.7 What Are the Incidental Soil Ingestion Rates for Children?

A common question asked of EPA is why Superfund risk assessments evaluate "dirt-eating kids": Why should Superfund sites be cleaned up to levels such that children can safely "eat" the soil there? Actually, EPA does not typically assume that children are eating the dirt; rather, EPA assumes that they are exposed to contaminants through the course of normal activities of play on the ground, exposure to dust in the home, and incidentally through mouthing behavior (USEPA 1996b, 2001d).

It is commonly observed that young children suck their thumbs or put toys and other objects in their mouths. This behavior occurs especially among children from 1 to 3 years old (Behrman and Vaughan, 1983; Charney et al., 1980). This "hand-to-mouth" exposure is well documented in the scientific literature for children under 6, and is especially prevalent among children 1½ to 3 years old, a critical period for brain development. This time period is of special concern regarding potential exposure, since children may be at special risk of exposure to specific chemicals, e.g., lead (ATSDR, 1991). Superfund experience has taught us that children do incur exposures to contaminated soil, as is evident at lead-contaminated sites in which elevated blood levels occur in children residing at those sites (USEPA 1996b, 2001d).

Scientists agree that because of this behavior, children may incidentally or accidentally take in soil and dust (Calabrese et al., 1989; Davis et al., 1990; Van Wijnen et al., 1990). Where children are likely to be exposed to contaminated soils (in residential areas, for example), it is appropriate for EPA to evaluate potential risks and set cleanup levels that will protect children for this widely recognized pathway of exposure, especially during this sensitive developmental period in the child's lifetime.

The basis of EPA's default soil ingestion rate is generally a point of contention. EPA has developed soil ingestion rates that are used as "default exposure assumptions" for adults and children. For young children (6 years or younger), the Superfund program default value is 200 milligrams of soil and dust ingested per day (USEPA, 1991c, 1996a). EPA's risk estimates

address the "incidental" ingestion that might occur when a child puts a hand or toy in his or her mouth, or eats food that has touched a dusty surface. Although this default assumption is often presented as an overly conservative value, the amount (200 milligrams per day) represents a small amount of soil ingested. It is less than 1/100 of an ounce (or one-fifth of the contents of a single-serving packet of sugar, e.g., a small package of Equal sweetener) a day. This peer reviewed value is applied in estimates of RME exposures (USEPA, 1989a, 1991c, 1997d).

In Superfund risk assessments, this soil ingestion rate for young children is combined with site-specific assumptions about exposure frequency (days per year) to estimate an average intake over the 6-year exposure period. Exposure frequency varies depending on site-specific current and future land uses. Soil ingestion studies report daily averages; the amount of soil ingested cannot be prorated on an hourly basis. Also, soil ingestion is episodic in nature and dependent upon a child's activity patterns, so prorating by time is not always appropriate. This is a common misapplication of soil ingestion rates in risk assessment.

Some children deliberately eat soil and other non-food items (a behavior known as pica). Pica behavior has been identified in children at rates of up to 5,000 milligrams per day (ATSDR, 1996, 2000; Calabrese et al., 1991). The Agency for Toxic Substances and Disease Registry uses this pica ingestion rate when calculating Environmental Media Evaluation Guides, which are used to select contaminants of concern at hazardous waste sites (ATSDR, 1996). EPA itself does not routinely address this form of exposure unless site-specific information is available. The default soil ingestion rate of 200 milligrams per day applied in Superfund risk assessments is intended to ensure reasonable protection of children in cases where they are likely to become exposed to contaminated soils and dust associated with a Superfund site.

## 5.4 Site-Specific Information

### 5.4.1 When Does EPA Use Site-Specific Information?

Again, site-specific information is considered in risk assessment where available. EPA risk assessment guidance supports the use of site-specific data, where feasible and appropriate according to a careful evaluation of the information including the specifics of the study design, numbers of individuals evaluated, representativeness of the exposures, etc. Examples of useful information are location of residences (receptor location), current and future land uses, behavior patterns and preferences (drinking water use, fish preparation), unique exposure pathways (hunting of specific game), and residence time within specific geographic areas.

The nature of the data submitted by stakeholders varies from anecdotal information to knowledge about creel surveys or other surveys conducted by universities or other groups. The nature and amount of data submitted will vary based on the site, ranging from no submission to

data on specific work practices at a facility. When such information is submitted, EPA critically evaluates it to determine its use in the risk assessment portion of the remedial investigation (RI).

The most accurate way to characterize potential site-specific exposures to populations around Superfund sites would be to conduct a detailed census at each site considering both current and future land uses. Theoretically, this would involve interviewing all potentially exposed individuals regarding their lifestyles, daily patterns, water usage, cleanliness, consumption of local fish and game and procedures, working locations, and exposure conditions, while collecting environmental samples (e.g., indoor air, water, soil, food, and possibly more). Although site-specific data are collected on environmental media (e.g., soil, groundwater, air, etc.) as appropriate during the RI, such collection has significant limitations. The three almost insurmountable difficulties are time, expense, and a major intrusion on privacy.

Discussed below are two examples of exposure parameters for which EPA has collected and applied site-specific data in risk assessment: bioavailability and fish consumption rates.

### 5.4.2 How Is Bioavailability Addressed?

The Agency addresses bioavailability by using default values and, in some cases, developing site-specific values supported by laboratory studies. In the absence of reliable information, RAGS Part A recommends assuming that the relative absorption efficiency between food or soil and water is 1. In other words, the assumption is that the bioavailability of the contaminant on the site, regardless of exposure medium, is the same as the bioavailability in the toxicity study used to derive the RfD or CSF. Such a default value is used to ensure that we do not underestimate risk in the face of uncertainty. The bioavailability of a compound in the exposure medium of concern at the site may actually be greater than in the exposure medium used in the critical toxicity study that formed the basis of the RfD or CSF. If this is the case, assuming a relative bioavailability of 1 for the medium of concern may result in an underestimate of risk at the site, depending on how all the other parameters are evaluated.

Appendix A of RAGS Part A (USEPA, 1989a) specifically addresses the consideration of site-specific bioavailability information in human health risk assessments. This information is referred to as "adjustments for absorption efficiency." In particular, Appendix A states that site-specific bioavailability adjustments may be appropriate:

> Adjustments also may be necessary for different absorption efficiencies depending on the medium of exposure (e.g., contaminants ingested with food or soil might be less completely absorbed than contaminants ingested with water).
>
> If the medium of exposure in the site exposure assessment differs from the medium of exposure assumed by the toxicity value (e.g., RfD values usually are based on or have been adjusted to reflect exposure via drinking water, while the site medium of concern may be soil), an absorption adjustment may, on occasion, be appropriate. For example, a substance might be more completely absorbed following exposure to

> contaminated drinking water than following exposure to contaminated food or soil (e.g., if the substance does not desorb from soil in the gastrointestinal tract).

For example, juvenile swine have been used as a model to derive site-specific bioavailability adjustments at several mining and smelter sites across the country (Casteel et al., 1997; USEPA, 1999g). These adjustments have been primarily limited to lead and arsenic.

Default values have been developed for some metals. For example, lead risks in children are typically assessed by predicting blood lead levels using the Integrated Exposure Uptake Biokinetic Model for Lead in Children (IEUBK) developed by EPA (USEPA, 2001d). The IEUBK model assigns default absolute bioavailability factors to all lead exposure media. The default values for air, water, and soil are 32%, 50%, and 30%, respectively (USEPA, 1999g, 2001d). For nonresidential scenarios, EPA developed the Adult Lead Methodology, which assumes the absolute bioavailability of lead in soil is 12% (USEPA, 1996b).

EPA finds that the site-specific application of quantitative bioavailability adjustments in risk assessments is not frequently supported by available scientific data. This occurs principally because of a lack of validated data and models to assess bioavailability for common site contaminants.

### 5.4.3 How Are Fish Ingestion Rates Evaluated?

Fish ingestion is evaluated at many Superfund sites with contaminated river systems, streams, and lakes. These contaminated sites range in size from small creeks to large river systems covering hundreds of miles. The contaminants of concern found at these sites include PCBs, dioxins and furans, mercury, pesticides, and other chemicals identified as persistent organic pollutants. State health departments have established fish consumption advisories at many of these water bodies.

The 1989 RAGS Part A guidance (USEPA, 1989a) recognized the importance of evaluating the potential cancer risks and non-cancer health hazards associated with consuming contaminated fish associated with Superfund sites. Page 6-6 of RAGS Part A specifically states:

> ... Be sure to include potentially exposed distant populations, such as public water supply consumers and distant consumers of fish or shellfish or agricultural products from the site area. ...

Further, on page 6-7, the guidance states:

> ... Identify any site-specific population characteristics that might influence exposure. For example, if the site is located near major commercial or recreational fisheries or shell fisheries, the potentially exposed population is likely to eat more locally-caught fish and shellfish than populations located inland. ...

A variety of data sources are available to evaluate fish ingestion rates. Fish ingestion studies can be either creel surveys, in which anglers are interviewed in person while fishing, or mail surveys, in which anglers (often identified as people with fishing licenses) are sent questionnaires in the mail (reviewed in USEPA, 1992c). Creel surveys typically involve interviews with anglers at the dockside requesting information about their fishing activities (fish preference, consumption rates, cooking methods, age, gender, frequency of fishing the specific water body, etc.). This survey method can provide information on both licensed and unlicensed anglers, depending upon who is interviewed. Mail surveys, meanwhile, typically involve sending questionnaires to licensed anglers requesting information on fishing practices; preferred rivers, lakes, or streams; fish consumption; and other information. If mailing addresses are obtained from lists of licensed anglers, unlicenced anglers are not represented. A third type of survey, diary surveys — in which participants are asked to record the frequency of fish ingestion, the types of fish eaten, and the meal size — require more effort from survey participants, but are generally assumed to yield more accurate results because they minimize the potential recall bias found in the other survey methods.

Any of these methods may be appropriate to evaluate the impact of Superfund site contamination on individuals consuming fish from contaminated water bodies impacted by a site. The approaches used in assessing fish consumption rates at sites vary based on the availability of creel surveys, the size of the impacted water body (e.g., a creek as opposed to a large river system), the available information on fishing practices within the specific water body, and the time and resources necessary to conduct a site-specific survey of fish consumption patterns.

The risk assessments include an evaluation of the RME and CTE individuals. In the absence of site-specific fish consumption rates, the default values from the 1997 Exposure Factors Handbook (USEPA, 1997d) may be used. The handbook makes a number of recommendations based on $95^{th}$ and $50^{th}$ percentile fish ingestion rates for specific water bodies, geographic locations, and fish species (USEPA, 1997d). Table 5-2, below, provides some of those recommendations.

| Table 5-2: Comparison of Default Fish Ingestion Rates | | | |
|---|---|---|---|
| **Default** | **Ingestion Rate (grams/day)** | **Percentile** | **Data Source** |
| 1997 EPA Exposure Factors Handbook<br>General population<br>Long-term intake distribution | 63 | 95th | USEPA, 1997d, based on TRI (Javitz, 1980; Ruffle et al., 1994) |
| Total fish | 20.1 | 50th | USEPA Analysis of CSFII, 1989–1991 |
| Marine fish | 14.1 | 50th | |
| Freshwater/estuarine | 6.0 | 50th | |
| 1997 EPA Exposure Factors Handbook (recreational marine anglers) | | | |
| Atlantic | 18.0<br>5.6 | 95th<br>50th | USEPA, 1997d; NMFS, 1993 |
| Gulf | 26.0<br>7.2 | 95th<br>50th | |
| Pacific | 6.8<br>2.0 | 95th<br>50th | |
| 1997 EPA Exposure Factors Handbook (freshwater) | | | |
| Maine | 13<br>5 | 95th<br>50th | Ebert et al., 1993 |
| New York | 18<br>5 | 95th<br>50th | Connelly et al., 1996 |
| Michigan | 39<br>12 | 95th<br>50th | West et al., 1989 |
| Michigan | —<br>17 | 95th<br>50th | West et al., 1993 |

Species-specific intake information may be derived from a state-wide angler survey (Connelly et al., 1992) to calculate the concentration of PCBs ingested in fish. For example, at a PCB site, EPA determined that each species of fish has a characteristic PCB concentration and the average concentration an angler consumes will, in part, be based on the relative percentages of the different fish species consumed. This analysis yields both central tendency and RME estimates for each of the parameters needed to calculate the cancer risks and non-cancer health hazards. A site-specific analysis using this database found that, for adults, the central tendency fish ingestion rate was 4.0 grams per day, or about six half-pound meals per year; the RME fish ingestion rate was 31.9 grams per day, or about 51 half-pound meals per year for adults. Fish ingestion rates for adolescents and young children were reduced based on the ratio of adolescent or child body weight to that of an adult.

Other risk assessments have used studies from the published literature that are representative of the relevant site-specific consumption patterns. Risk assessors review the literature to determine how well it represents the site they are evaluating. In determining the literature's appropriate use, they may evaluate the study design, hypotheses being tested, number of individuals interviewed within the survey, rates of consumption, and populations studied.

In several other examples, site-specific surveys have been conducted to evaluate the consumption patterns for specific populations that the published surveys do not capture. These surveys found considerably higher consumption rates among these populations than if the standard default assumptions from the 1997 Exposure Factors Handbook were used (USEPA, 1997d). For example, a 3½-year site-specific creel survey (Toy et al., 1996) included information on whether or not adults harvested fish and shellfish from Puget Sound. The survey included 190 adults and 69 children between the ages of 0 and 6. The study found that tribal seafood consumption rates were considerably higher than Exposure Factors Handbook values. Among the Squaxin, the average consumption rate was 72.8 grams per day and the 90$^{th}$ percentile ingestion rate was 201.6 grams per day. Among the Tulalips, the average consumption rate was 72.7 grams per day and the 90$^{th}$ percentile was 192.3 grams per day. Other site-specific consumption surveys found similar differences in consumption rates (APEN, 1998; USEPA, 2001h; Sechena et al., 2003).

In cases where EPA has conducted individual surveys to identify fish consumption rates, EPA has found it important to include the community in the process (USEPA, 1999f). EPA and other agencies (both private and governmental) have spent considerable resources and time to plan and implement these studies. The surveys (APEN, 1998; USEPA 2001h; Sechena et al., 2003) were all conducted using one-on-one interviews, as opposed to creel or mail surveys. The people conducting the interviews were always specially trained members of the ethnic group or community being surveyed. (EPA realized early on that creel and mail surveys do not necessarily work for tribes and other such groups. One-on-one interviews involving the affected communities are more appropriate.)

The determination of the need to conduct a creel survey or use site-specific information is dependent on a number of site-specific factors. These include available information on consumption patterns, whether an existing consumption advisory may bias responses when surveys are conducted, time, and resources.

## 5.5 Other Factors in Superfund Assessments

### 5.5.1 What Is the Role of Stakeholders in Superfund Assessments?

EPA works with members of the community to understand their concerns at the national and site-specific levels. In 1996, EPA conducted meetings with the communities to discuss the Superfund process and their concerns (USEPA, 1996c). Under the 1996 RAGS Reform process, EPA met with community members at several meetings to discuss their concerns and identify priorities for new guidance documents. At these meetings, the communities identified ideas for improving communication between the regulators and the community. This resulted in guidance for involving communities in the risk assessment process (USEPA, 1999f).

Within the Superfund program, EPA conducts extensive outreach with stakeholders at the site. As described in the guidance on community involvement (USEPA, 1999f, 2000n), at each site the remedial project manager, the community involvement coordinator, and other members of the team including the risk assessor regularly communicate with the members of the community, including the potentially responsible parties. Each site has a Community Involvement Plan, developed following interviews with the community, that outlines mechanisms by which the community will be involved in the process.

EPA guidance and educational materials help illustrate the ways that citizens can be involved in the risk assessment process (USEPA, 1999f, h). For example: Community-specific information on fishing preferences helped to identify exposure areas for sampling and fish species consumed by people who fish in a contaminated bay. Information from farmers on pesticide applications helped EPA determine why certain contaminants were present in an aquifer. Discussions with farmers about certain harvesting practices helped EPA refine exposure models and assumptions at another site (USEPA, 1999h).

EPA uses a range of communication tools to include the community in the Superfund process. These include newsletters, fact sheets, site-specific home pages, public meetings, public availability sessions, and 1-800- numbers to contact EPA staff. EPA strives to communicate information about the RI, the results of the risk assessment, proposed actions at the site, and the proposed and final decisions for remedial actions. The Record of Decision (ROD) includes a responsiveness summary that addresses comments including those from the community. During the period of the remedial action, communication with the community continues, including updates during the 5-year review process.

EPA conducts peer reviews of site-specific risk assessments following appropriate peer review guidance (USEPA, 2000b).

Consistent with EPA Superfund goals of improving the transparency of the process, the methods for summarizing risk information are found in the RAGS Part D (USEPA, 2001b). Superfund continues to update guidance documents to improve the transparency of risk information.

### 5.5.2 How Are Superfund Assessments Factored in Superfund Risk Management Decisions?

Risk assessment is one of several tools used to inform risk management decisions. Risk managers weigh a number of factors, including uncertainties in exposure and risk estimates, when developing health and environmental protective decisions. EPA considers a variety of alternatives to protect human health and the environment at sites and evaluates them by considering long-term effectiveness, use of treatment, short-term effectiveness, implementability, and cost. EPA then proposes a protective, cost-effective, ARAR-compliant (that is, compliant with the Applicable or Relevant and Appropriate Requirements (ARAR)) remedy, which it may modify based on state and public comments (see also CERCLA § 121, 42 U.S.C. § 9621 and 40 CFR § 300.430[e][9]) (USEPA, 1988b, 1999a, f, 2000n; NCP 1990a, b).

Following the completion of the RI during which the risk assessment is conducted, EPA develops a feasibility study (FS) that evaluates remedial actions at the site (USEPA, 1988b). The FS evaluates the risks in the absence of remedial actions or institutional controls. This provides a baseline for comparison with other remedial alternatives. The FS includes the development of Remedial Action Objectives, including Preliminary Remediation Goals (PRGs) that are developed based on the RME exposure assumptions used in the risk assessment. The PRGs provide concentration levels that are protective of the RME individual who is currently exposed or may be exposed in the future.

During the FS, remedial alternatives are developed to achieve the program goals through a variety of different methods, generally including containment and treatment alternatives. The alternatives reflect the scope and complexity of the site problem. The Superfund program evaluates these alternatives using nine criteria described by the NCP. The criteria address protectiveness, effectiveness, implementability, and acceptability issues. The criteria were derived from remedy selection criteria provided by Congress in SARA 121 (SARA, 1986). The detailed analysis consists of an assessment of the individual alternatives against each of the nine evaluation criteria and a comparative analysis focusing upon the relative performance of each alternative against those criteria. The nine criteria are presented in table 5-3, below. In addition to viable remedial alternatives, EPA evaluates a no-action remedial alternative at all sites. The no-action alternative provides a baseline for comparison of the various alternatives that are

appropriate for a specific site. All of this information is provided in a Proposed Plan, which is released with the RI/FS for public review and comment.

EPA provides opportunities for public review of this information. A public meeting is held to discuss the proposed remedial alternatives and to obtain comments. Public comments are addressed at the meeting and in the Response to Comments that is developed as part of the ROD. The ROD identifies remedial actions that have been selected for the site.

Following the ROD, EPA begins the remedial design process and the implementation of construction. Depending on the nature of the remedial actions and the amount of time required to complete the construction, EPA may conduct 5-year reviews to determine the protectiveness of the remedy (USEPA, 2001i). Throughout this process, information is shared with the affected community regarding the progress of the remedial actions.

**Table 5-3: Nine Evaluation Criteria for Superfund Remedial Alternatives**

| |
|---|
| ***Threshold Criteria*** |
| ***Overall protection of human health and the environment*** determines whether an alternative eliminates, reduces, or controls threats to public health and the environment through institutional controls, engineering controls, or treatment.<br><br>***Compliance with ARARs*** evaluates whether the alternative meets federal and state environmental statutes, regulations, and other requirements that pertain to the site, or whether a waiver is justified. |
| ***Primary Balancing Criteria*** |
| ***Long-term effectiveness and permanence*** considers the ability of an alternative to maintain protection of human health and the environment over time.<br><br>***Reduction of toxicity, mobility, or volume of contaminants through treatment*** evaluates an alternative's use of treatment to reduce the harmful effects of principal contaminants, their ability to move in the environment, and the amount of contamination present.<br><br>***Short-term effectiveness*** considers the length of time needed to implement an alternative and the risks the alternative poses to workers, residents, and the environment during implementation.<br><br>***Implementability*** considers the technical and administrative feasibility of implementing the alternative, including factors such as the relative availability of goods and services.<br><br>***Cost*** includes estimated capital and annual operation and maintenance costs, as well as present worth cost. Present worth cost is the total cost of an alternative over time in terms of today's dollar value. Cost estimates are expected to be accurate within a range of +50% to -30%. |
| ***Modifying Criteria*** |
| ***State acceptance*** considers whether the state agrees with the EPA's analyses and recommendations, as described in the RI/FS and Proposed Plan.<br><br>***Community acceptance*** considers whether the local community agrees with EPA's analyses and preferred alternative. Comments received on the Proposed Plan are an important indicator of community acceptance. |

## 5.6 Superfund Site-Specific Conclusions

In general, the information applied in EPA Superfund risk assessments and the level of protectiveness are considered in the context of the regulatory program. From the previous discussion, the following conclusions are highlighted.

a) *Reasonable maximum exposures*. For the Superfund program, EPA bases decisions on current and future risks associated with reasonable high-end exposures or RME, not only the average exposures. By definition, the RME *is* within a realistic range of exposures, not a "worst-case" estimate.

b) *Default assumptions.* Risk estimates are based on a combination of site-specific information, when available and appropriate, and general default exposure factors developed for Superfund baseline risk assessments. The default factors are based on published scientific data and are designed to be reasonably protective of the majority of those exposed. The use of defaults also streamlines risk assessments, reduces unwarranted variability in the exposure assumptions used to characterize exposures, and provides reasonable risk estimates where there is a lack of site-specific data or consensus on which value to choose. For some exposure parameters, default values are used because the costs of collecting data for specific behaviors in a population may be prohibitive.

c) *Site-specific information.* Information submitted to the EPA regions or collected by EPA is critically evaluated to determine how to apply it in the risk assessment. Anecdotal information is valuable but not sufficient for quantifying exposures. It may, however, be useful to consider such information when developing a conceptual site model and characterizing uncertainties. In some cases, site-specific data would be preferable to using default exposure data that represent the majority of the population at the high end of the distribution for specific activity patterns. However, the costs of collecting data for specific behaviors in a population may be prohibitive. For example, a site-specific creel survey may cost $100,000 or more and take several years to conduct (Toy et al., 1996; USEPA 2001h; Sechena et al., 2003).

Specialized studies may reduce uncertainties in a site-specific risk assessment; however, the region considers the sensitivity of the parameter in question and the added value of the information in light of uncertainties of other parameters, including the toxicity information. In some cases, a research study of exposure will not significantly reduce the uncertainty of the risk assessment when the toxicity information is very limited (USEPA, 1992a).

d) *Peer review.* Consistent with Agency guidance on peer review, external peer reviews are conducted for specific risk assessments and guidance documents.

e) *Probabilistic risk assessment.* Superfund has developed specific guidance for conducting site-specific probabilistic risk assessments, and this guidance was externally peer reviewed. At the present time, site-specific probabilistic risk assessments have been conducted at several sites to examine exposure assessments.

f) *Guidance documents.* The Superfund program continues to update guidance to include new science where appropriate. The guidance documents are made available for public review and external peer review.

g) *Transparency and clarity in risk assessments.* Superfund developed specific guidance for improving the clarity and transparency of risk assessments. This guidance covers community involvement, standardization of reporting of risk information, and standardization of ROD information (USEPA, 1999f, 2000n).

EPA's risk assessment practices and policies apply a mix of site-specific information and science-based default assumptions that are consistent with the goals of the Superfund program. Site-specific considerations are addressed at the EPA site-specific level. In turn, the regions inform EPA program offices of the research needed to reduce uncertainties in risk assessments. There will always be a need for additional research on exposure and toxicity parameters to improve understanding of how people contact their environments, chemical toxicity, and chemical fate and transport.

## 5.7 Chemical Mixtures Risk Assessment Methods and Practice

### 5.7.1 What Are the Issues Regarding Chemical Mixture Site Assessments?

EPA published guidance on the health risk assessment of chemical mixtures (USEPA, 1986b; 2000h) and its use in site assessments (USEPA, 1989a; 1999a). As part of the implementation of the 1996 Food Quality Protection Act, EPA also developed guidance for conducting cumulative risk assessments of chemicals that appear to act by a common mechanism of toxicity (USEPA, 2002f). EPA (USEPA, 2000h) guides the risk assessor through an examination of data relevant to a mixture of concern to facilitate selection of an appropriate mixture risk assessment method. Suggested methods include both whole-mixture approaches and component-based algorithms that include interaction data. A comment regarding EPA's practice is that the Agency "assumes the toxicity of a chemical mixture is equal to the sum of the toxicity of each individual chemical, regardless of the toxicity type, competition or antagonism among chemicals." This comment apparently assumes that:

a) EPA uses only models of mixture toxicity based on dose addition,

b) EPA inappropriately applies dose addition by ignoring the requirement that the mixture's chemical components share a common toxic mode of action (MOA),

c) EPA ignores interactions data that may reduce the toxicity of a mixture through competition or antagonism, and

d) EPA never determines the toxicity of a whole mixture directly.

These assumptions do not reflect current guidance, as shown by the Agency's mixtures guidelines (USEPA, 1986b) and supplementary guidance (USEPA, 2000h). These documents offer detailed descriptions of a variety of component-based and whole-mixture methods far beyond dose addition, provide instructions on appropriate application of the methods, and recommend consideration of interaction data whenever possible. Further, program guidance (USEPA, 1989a) specifies the use of component-based approaches to evaluate chemical mixtures. The following sections explain the theory behind these procedures and discuss current practice.

### 5.7.2 What Are Dose Addition and Response Addition?

To address low exposure levels when no interaction information or whole-mixture toxicity data are available, it is EPA's practice to use response addition and dose addition as the recommended default methods when the component chemicals in a mixture show dissimilar toxicity and similar toxicity, respectively (USEPA, 2000h). Dose addition and response addition are fundamentally different methods, based on different assumptions about toxicity. Because both methods are relatively easy to apply and use single chemical toxicity and exposure information, they are the methods most commonly used in site-specific assessments (USEPA, 1989a). The two additivity assumptions are further described as follows:

a) *Dose addition* sums the doses of the components in a mixture after they have been scaled for toxic potency relative to each other. The risk of a toxic effect is determined from this summed dose. Dose addition requires the components to share a common toxic MOA.

   Superfund site assessments have applied dose addition in the form of a hazard index (HI) to evaluate sites for indications of health risk (USEPA, 1989a). The HI is calculated as the sum of hazard quotients (HQs) for the chemical components of the mixture; thus, it indicates risk, but is not an explicit risk estimate. A HQ is typically calculated as the ratio of a chemical's exposure level to its safe level, such that values larger than 1 are of concern. For a group of $n$ chemicals in a

mixture with exposure levels *(E)* and using the RfD as a safe level, the HI is calculated as:

$$HI = \sum_{i=1}^{n} E_i / RfD_i \qquad \text{Equation 5.7-1}$$

The HI is usually calculated for groups of chemicals whose effects are observed in a common target organ. As an initial screening step, though, risk assessors calculate HIs by combining across health endpoints, contradictory to the assumption of common MOA. If the HI across all effects indicates no risk (i.e., is less than 1), then the assessment can stop. If risk is indicated (i.e., the HI is greater than 1), the risk assessor repeats the analysis by breaking the mixture into groups of chemical components with a common target organ and recalculating the HI for each smaller group.

b) *Response addition* first estimates the probabilistic risk of observing a toxic response for each chemical component in the mixture. Then the component risks are summed to estimate total risk from exposure to the mixture, assuming independence of toxic action (i.e., that the toxicity of one chemical in the body does not affect the toxicity of another chemical). This can be thought of as an organism receiving two (or more) independent insults to the body, so that the risks are added under the statistical law of independent events.

Superfund site assessments calculate total cancer risk by summing the individual cancer risks for the carcinogens in the mixture (USEPA, 1989a). For example, the mixture risk *(Rm)* for two chemicals is the sum of the risks for chemical one *($r_1$)* and chemical two *($r_2$)* minus the probability that the toxic event from exposure to chemical one would overlap in time with the toxic event from exposure to chemical two, as follows:

$$Rm = r_1 + r_2 - (r_1 \times r_2) \qquad \text{Equation 5.7-2}$$

Because cancer risks are typically very small (i.e., in the $10^{-4}$ to $10^{-6}$ range), the amount of risk subtracted in Equation 5.7-2 (e.g., an amount in the $10^{-8}$ to $10^{-12}$ range) is insignificant and is therefore usually ignored. Risks are appropriately aggregated for cancers across various target organs because the result is interpreted as the risk of any cancer, and the cancers from each chemical component are considered to be independent events in the body.

### 5.7.3 How Are Dose Addition and Response Addition Applied in Practice?

To get a general idea of how dose addition and response addition are actually practiced in EPA's site assessment work, a search of EPA's RODs in the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) database was conducted in August 2003 (http://www.epa.gov/superfund/sites/siteinfo.htm), keying on the term "hazard index." From that list, 10 RODs were selected from across the Agency. These RODs were examined for tables and text relevant to the application of response addition and the Hazard Index (HI). RODs were used as a surrogate for the complete risk assessments since only a limited number of risk assessments are available on the internet. It is noted that, as stated in the guidance for preparing RODs, the full risk assessment need not be included in the ROD document (USEPA, 1999a). The risk assessments are available in the Administrative Record for each site and, based on the time limits of this analysis, it was not possible to evaluate each of the individual site records.

To estimate cancer risks using response addition, the procedure is to sum risk estimates across all carcinogenic modes of action, assuming independence of action. There is no requirement to demonstrate independence empirically; however, when component exposures are low, such biological independence is likely. Most of the examples appear to have appropriately applied response addition, although in several cases, this conclusion is inferred from the available text. In one instance, the technique was not used because there was only one carcinogen found at the site. In two other cases, additional information from the risk assessment would be needed to determine exactly how the evaluations were completed.

To apply the HI, the procedure in the guidance is to first perform a screening-level analysis, summing Hazard Quotients (HQs) across all target organs. If the HI is below 1 across all target organs, then the analysis is complete; this was the case for tap water risks in one example. If the HI is above 1 across all target organs, then the analysis is to be repeated by summing the HQs for only those mixture components that affect the same target organ; this occurred in the three of the examples. The purpose of the re-analysis is to comply with the dose addition assumption of a common toxic mode of action. In one case, the HI was shown in the ROD to change from greater than 1 in the screening analysis to less than 1 under the re-analysis. In two instances, the HI was greater than 1 in the screening analysis, but the ROD did not show the re-analysis by target organ; in these cases, however, the individual chemical HQs were much larger than 1, making all HI calculations in these cases moot (i.e., no HI below 1 could be calculated under any constraints). Finally, in four of the examples, additional information would be needed to determine exactly how the evaluations were completed.

In conclusion, the results from this limited analysis are generally positive, illustrating that EPA practice is consistent with its own guidance regarding chemical mixtures risk assessment. The clarity of EPA's risk communications could be improved by transferring additional

information from the risk assessment in the Administrative Record for each site to its associated ROD.

### 5.7.4 Are Interaction Data Used in Mixture Risk Assessment?

EPA defines interactions as observed effects greater than (synergism) or less than (antagonism) those expected under a specified form of additivity. EPA (USEPA, 1986b, 2000h) also states that mixture risk assessment should reflect known toxicological interactions. In practice, sufficient data are seldom available to model interactions. Data on binary combinations of chemicals are most prevalent, so for mixtures of more than two chemicals, the true nature of joint toxic action may be speculative at best. For exposures at low doses with low component risks, the likelihood of a significant interaction is usually considered to be low. Interaction arguments based on saturation of metabolic pathways or competition for cellular sites usually imply an increasing interaction effect with dose, so that the importance for most low environmental exposures is probably small.

EPA mixture risk assessments have usually considered the information on interactions in a qualitative sense. For example, a Superfund site may receive more scrutiny or its remediation may proceed faster if there are indications of potential synergism among the detected chemicals. The cleanup goals and the estimated risk would not change, though, because the synergism could not be quantified. No standard methods or mathematical models are yet in common use in regulatory agencies to incorporate interactions and to serve as defaults. There is one newly developed procedure, the interaction-based HI (USEPA, 2000h), that allows for numerical adjustments to the HI based on evidence of synergism or antagonism for pairs of chemicals in the mixture. This method is currently undergoing verification using laboratory data.

Despite the difficulties in using interaction information, some regional assessments have considered relevant data. An interesting assessment by Region III explicitly evaluated synergy and antagonism of multiple metals in soil (USEPA, 2002m). Text from that ROD reads as follows:

> EPA (1994b; *[this reference is in the ROD]*) prepared an extensive review of interactions among cadmium, lead, and zinc reported in the open literature. Available evidence does not support a change in assumed absorption of metals due to interactions among metals following ingestion of contaminated soils. The quantitative risk assessment does not alter the estimates of exposure based on co-exposure to cadmium, lead, and zinc. Further, available information suggests that any potential impact of cadmium on neurotoxicity of lead in young children is largely speculative. No modification to the assessment of lead toxicity is thus justified based on co-exposure to cadmium.

### 5.7.5 Are Data on Whole Mixtures Used in Risk Assessment?

Risk assessments based on tests of whole mixtures or on epidemiologic data determine combined effects empirically. Examples of these (USEPA, 2003g) include (1) RfDs on

commercial PCB mixtures (Aroclors 1016 and 1254) based on primate data, and (2) a CSF for coke oven emissions based on human occupational exposures. Drinking water disinfection byproducts represent a complex mixture for which epidemiologic data suggest potential health risks (USEPA, 2003b). An EPA study, called the 4 Lab project, is currently underway to toxicologically and chemically characterize this complex mixture to produce data on reproductive and developmental effects in rats exposed to concentrations of this complex mixture for use in risk assessment (Simmons et al., 2002).

Whole-mixture studies are routinely used in ecological risk assessments. The Agency has developed subchronic toxicity tests for whole aqueous effluents and for contaminated ambient waters, sediments, and soils (USEPA, 1989d, 1991d, 1994e). Further, the effects of mixtures in aquatic ecosystems are evaluated using bioassessment techniques that are equivalent to epidemiology, but more readily employed (USEPA, 1999i). Similar bioassessment methods are sometimes used at Superfund sites (USEPA, 1994f). These empirical approaches to assessing ecological risks from mixtures are employed in National Pollutant Discharge Elimination System permitting and the development of Total Maximum Daily Loads, and are often used in Superfund baseline ecological risk assessments.

### 5.7.6 What Other EPA Applications Exist for Mixtures Risk Assessment?

EPA program offices and regional risk assessors have a great need for both assessment information and risk assessment methods to evaluate human health and ecological risks from exposure to chemical mixtures. Further, several pieces of legislation exist that require EPA to consider the evaluation of chemical mixtures (see below). The Agency conducted some specific mixture risk assessments in the past and continues to pursue additional evaluations. Currently, an IRIS assessment is being developed for the PAHs (USEPA, 2002n). In addition, EPA used a dose additive approach to assess PCBs (USEPA, 1996d), organophosphorus pesticides (USEPA, 2002g), and dioxins (USEPA, 1989b; 2000i). (See section 4.5 for a discussion of this approach.) Examples of other mixtures of interest to the Agency include creosote, particulate matter, drinking water disinfection byproducts (DBPs), brominated flame retardants, mixtures of metals, and pesticide mixtures in soils. Current Agency interests and supporting legislation include:

a) OW is concerned with contaminant mixtures in drinking water in response to requirements of the Safe Drinking Water Act Amendments of 1996, including mixtures of DBPs and of Contaminant Candidate List chemicals (e.g., organotins, pesticides, metals, pharmaceuticals). Information and methods are being developed to better evaluate the toxic mode of action, the risk posed by drinking water mixtures, exposure estimates for mixtures via multiple routes, and the relative effectiveness of advanced treatment technologies (USEPA, 2003e, h).

b) The Office of Air Quality Planning and Standards used a mixture risk assessment approach for multiple route exposures in conducting the National Air Toxics Assessment of 33 air pollutants (a subset of 32 air toxics from the Clean Air Act's list of 188 air toxics plus diesel particulate matter) (USEPA, 2001j).

c) The Office of Air and Radiation issued a proposed rule to evaluate the national emission standards for HAPs for stationary combustion turbines, suggesting the application of HIs to evaluate potential health risks (USEPA, 2003i).

d) The Officer of Pesticide Programs conducted a multiple-route assessment of organophosphorus pesticide mixtures (USEPA, 2002g), in response to the FQPA of 1996. Future assessments may be performed on additional pesticide classes and other co-occurring substances for which a common toxic mode of action can be identified.

e) The National Homeland Security Research Center is considering the potential toxicological interactions with co-exposure to respiratory toxicants, dust, and smoke.

f) The Office of Solid Waste and Emergency Response evaluates contaminant mixtures at Superfund sites (USEPA, 1989a) under CERCLA (see sections 5.7.2 and 5.7.3 above). Examples of current needs and research areas are the assessment of mixtures of metals and guidance on how to group the environmental contaminants most commonly found together at sites into common mode of action classes for use in developing HIs or assessments based on RPFs.

# 6. ECOLOGICAL ASSESSMENT

## 6.1 Overview

The early 1980s saw both the emergence of risk assessment as a regulatory paradigm and the first widespread use of ecological impacts to inform regulatory and policy decisions. The use of ecological information in decision making has expanded slowly through the 1980s, as illustrated by the cancellation of the pesticide diazinon based on impacts to birds (Bascietto, 1998), the reductions in sulfur dioxide emissions to lessen the adverse impacts of acid deposition on lakes and forests and the setting of ozone standards to reduce the damaging effects of tropospheric ozone on crops. In the middle to late 1980s, tools and methods for conducting ecological risk assessments began to be standardized with the publication of EPA's Ambient Water Quality Criteria (WQC) methodology, the pesticide program's Standard Evaluation Procedures, and Superfund's Environmental Evaluation Manual (these are discussed in Bascietto et al., 1990). More recently, EPA published the Ecological Risk Assessment Guidance for Superfund (USEPA, 1997f) and the *Guidelines for Ecological Risk Assessment* (USEPA, 1998a).

In 1992, the Agency published the *Framework for Ecological Risk Assessment* (USEPA, 1992d; Norton et al., 1992) as the first statement of principles for ecological risk assessment. That document and the 1998 first draft *Guidelines for Ecological Risk Assessment* not only describe methods for conducting the more conventional single-species, chemical-based risk assessments, but also describe techniques for assessing risks to ecosystems from multiple stressors and multiple endpoints. With the publication of these important documents came the need to enhance EPA's ability to do better ecological assessments. These important publications recognized the need to advance the science of multiple-scale, multiple-stressor, and multiple-endpoint ecological assessments and, more important, called for the consideration of human dimensions in the planning and conduct of ecological risk assessments (Bachmann et al., 1998; Barton and Sergeant, 1998; Cooper, 1998; Linthurst et al., 2000).

Following EPA's paradigm, an ecological risk assessment evaluates the potential for adverse ecological effects resulting from human activities. The risk assessment process provides a way to develop, organize, and present scientific information so that it is relevant to environmental decisions. When conducted for a particular place such as a watershed, the ecological risk assessment process can be used to identify vulnerable and valued resources, prioritize data collection activity, and link human activities with their potential effects. Risk assessments can also provide a focal point for cooperation between local communities and state and federal government agencies. Risk assessment results provide a basis for comparing different management options, enabling decision makers and the public to make more informed decisions about the management of ecological resources.

Ecological risk assessments have become a common subject of discourse at EPA. A summary of the evolution of this discourse is described by Bascietto et al. (1990), Maki and Slimak (1990), Kutz et al. (1992), Norton et al. (1992), Bachmann et al. (1998), and Linthurst et al. (2000). Beginning in 1984, the concepts of risk assessment and risk management became a common language for justifying regulatory proposals throughout EPA. In 1987 a major Agency report stated that the "fundamental mission of EPA is to reduce risks," and another in 1990 recommended that EPA "target its environmental protection efforts on opportunities for the greatest risk reduction" (USEPA 1987b, 1990). And, as described previously, the rhetoric of risk became the Agency's primary language for justifying its decisions (Russell and Gruber, 1987).

### 6.1.1 What Is the EPA Ecological Risk Assessment Approach?

Ecological risk assessment evolved out of the human health risk assessment process, an approach to understanding and protecting humans from the threats — predominately cancer — from exposure to chemicals. The National Academy of Sciences' National Research Council's seminal document on risk assessment (NRC, 1983) defined risk assessment as assessment of the "probability that an adverse effect may occur as a result of some human activity." This report, and the articulation by EPA Administrator Ruckelshaus (Ruckelshaus, 1983) that risk was to be EPA's defining operational paradigm, ushered the Agency into a period in which risk assessment and risk management became the primary discourse among its staff. Recognizing that its definition of risk is human-oriented, the report recommended that the EPA develop a counterpart for non-human or ecological endpoints. Thus began a lengthy process within EPA to complement the human risk assessment process with procedures for wildlife, ecosystems, and endangered species, all of which fall into the category of ecological risk assessment. The EPA relied on scientific and technical deliberations among experts in the field to complete the *Framework for Ecological Risk Assessment* (USEPA, 1992d; Norton et al., 1992). The Framework defined ecological risk assessment as "a process that evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors." EPA subsequently published case studies (USEPA, 1993b, 1994g) and a companion document on issues relating to the elements of an ecological risk assessment (USEPA, 1994h). Then (as mentioned above) EPA published an Ecological Risk Assessment Guidance for Superfund in 1997 (USEPA, 1997f) and *Guidelines for Ecological Risk Assessment* in 1998 (USEPA, 1998a).

According to EPA (USEPA, 1998a) and others (Suter, 1993; Lackey, 1994), an ecological risk assessment is a flexible process for organizing and analyzing data, information, assumptions, and uncertainties to evaluate the likelihood of adverse ecological effects. Ecological risk assessment provides a critical element for environmental decision making by giving risk managers an approach for considering available scientific information along with the other factors they need to consider (e.g., social, legal, or economic factors) in selecting a course of action. An ecological risk assessment includes an initial planning step, problem formulation,

analysis of stressors and effects, and risk characterization. Problem formulation, the initial step in analysis, involves defining and specifying the issue under consideration. Analysis is separated into exposure (stressor) and effects (hazard) analysis. Risk characterization includes both quantitative estimation of risk and a qualitative description of risk. The final step is communicating the results to a manager and stakeholders. Important properties of the assessment include data gathering and review during and after each step and the possibility of repeating steps if new data become available.

Ecological risk assessments at EPA are generally characterized by a particular emphasis at the problem formulation phase. The interface among risk assessors, risk managers, and interested parties during planning at the beginning of the risk assessment, as well as communication of risk at its end, is critical to ensure that the results of the assessment can be used to support a management decision. Because of the diverse expertise required, especially in complex ecological risk assessments, risk assessors and risk managers frequently work in multi-disciplinary teams during problem formulation. This approach has important implications for how an ecological risk assessment is formulated, since it requires an up-front discussion on what is at risk, what the assessment endpoints are, how they are measured, and what unacceptable risk levels are. This discussion and its outcome are shaped by the values, beliefs, and attitudes of the risk assessors, risk managers, and stakeholders. Those conducting ecological risk assessments have learned that risk managers, risk assessors, and stakeholders bring valuable perspectives to the initial planning activities for an assessment. Risk managers charged with protecting the environment can identify information they need to develop their decisions, risk assessors can ensure that science is effectively used to address ecological concerns, and stakeholders bring a sense of realism and purpose. Together, all can evaluate whether a risk assessment can address the identified problems.

## 6.2 Organism-Level Versus Population-Level Ecological Risk Assessments

The Agency received comments expressing concern about the use of organism (individual) level attributes to assess ecological risk. An example comment follows:

> EPA (1997) ecological risk assessment guidance recommends that potential ecological risks should be assessed at the population-level for all but threatened and endangered species. [*note added here*: This reference and accompanying citation appear to be incorrect, in that that document does not make such a recommendation.] Although no explicit guidance is provided, this is typically accomplished through the use of measurement endpoints that are related to population effects (e.g., using Toxicity Reference Values based on growth or reproductive effects). However, in many EPA ecological risk assessments, the agency has defaulted to assessing effects on individual animals. This metric has no significance scientifically and is entirely useless as a basis for making risk management decisions.

The above statement cites text from EPA documents emphasizing the importance of protecting populations or communities. The Office of Solid Waste and Emergency Response's guidance for Superfund sites (USEPA, 1999j) states:

> Ecological risk assessments incorporate a wide range of tests and studies to either directly estimate community effects (e.g., benthic species diversity) or indirectly predict local population-level effects (e.g., toxicity tests on individual species), both of which can contribute to estimating ecological risk. Superfund remedial actions generally should not be designed to protect organisms on an individual basis (the exception being designated protected status resources, such as listed or candidate threatened and endangered species or treaty-protected species that could be exposed to site releases), but to protect local populations and communities of biota. Levels that are expected to protect local populations and communities can be estimated by extrapolating from effects on individuals and groups of individuals using a lines-of-evidence approach.

As an initial matter, the comment may be more properly considered as a risk management decision to be made in the problem formulation stage of an ecological risk assessment. The cited text notes the important role that measurements of organism-level effects can play in extrapolations to population-level effects. Where possible, toxicity tests are performed using measurement endpoints that are related to population effects (e.g., measurements of growth rate or reproduction). We agree that it would be desirable to define the full range of ecological benefits by considering effects on populations, communities, and ecosystem processes as well as organisms. But it is also important to recognize that the use of organism-level attributes such as mortality as endpoints for ecological risk assessments is well-supported by policy and precedent (USEPA, 1994i, 2003j) and by the courts — e.g., in the diazinon decision (USEPA, 1988c; *Ciba Geigy v. EPA*, 874 F.2d 277 [5th Cir. 1989])[1]. In addition, some laws including the Marine Mammal Protection Act, the Migratory Bird Treaty Act, The Bald Eagle Protection Act, and the Endangered Species Act call for protection of organism properties and even individual organisms. Nonetheless, the protection of organism-level attributes is generally interpreted as occurring in a population or community context (USEPA, 2003j). That is, increased mortality or decreased fecundity or growth of organisms in an assessment population or community is assumed to be significant, even with no demonstration that a population- or community-level property is affected.

That assumption is necessitated by the extreme difficulty of predicting effects at higher ecological levels. Most population-level attributes, including abundance attributes, are determined by the vital rates (births, deaths) of individuals within the population, as well as the rates of migration into and out of the population. Linkages between effects on vital rates and those on population dynamics, or between effects on the biochemical and physiological processes that determine vital rates and effects on populations, can either be established empirically — by correlating responses at different levels of biological organization — or by determining causal

[1] On March 29, 1988, the Agency cancelled registrations for the pesticide diazinon unless amended to prohibit use on golf courses and sod farms (USEPA, 1988c). As noted in the remand decision (*Ciba Geigy v. EPA*, 874 F.2d 277 (5th Cir. 1989)), the circuit court rejected the industry's argument that a risk is unreasonable only if it endangers bird populations, stating: "FIFRA gives the Administrator sufficient discretion to determine that recurring bird kills, even if they do not significantly reduce bird population, are themselves an unreasonable environmental effect."

relationships and constructing mechanistic models (see Maltby et al., 2001, for a conceptualization and discussion of these linkages). Thus, with some degree of understanding of how changes in vital rates manifest into population consequences (i.e., mechanistic understanding), effects measured at the organism level can be *extrapolated* to expected population-level responses. Extrapolation of this latter type is usually accomplished using models that integrate the effects of stressors on survivorship and fecundity. However, the practicality of this approach is cast into question by experience which indicates that so many unverifiable assumptions are involved that in real cases the very different outcomes of dueling models cannot be resolved (Barnthouse et al., 1984). Also, it is impractical to estimate density dependent responses in real populations, particularly when modeling the effects of toxicants.

The bottom line is that, although methods exist for predicting the effects of chemicals at levels of organization higher than the organism, they are still in the development phase and have not been shown to be reliable. On the other hand, assessment of ecological risks using measures of organism-level effects is justified by experience, policy, and judicial decisions.

## 6.3 Conservatism and Ecological Risk Assessments

### 6.3.1 How Is Conservatism Addressed in Ecological Risk Assessments?

A number of ecological concerns expressed to the Agency revolve around the concept of being overly conservative. In early tiers of a risk assessment (e.g., for screening), a high degree of conservatism is sought. Definitive assessments strive to be as realistic as possible, replacing conservative assumptions with best estimates of exposures and effects and associated uncertainties. Three examples in which conservatism is addressed in ecological risk assessments are discussed in the following sections.

### 6.3.2 How Are Toxicity Reference Values Developed in Ecological Risk Assessments?

A number of comments have been received relating to the development of toxicity reference values (TRVs) for use in ecological assessments. One of these notes that EPA guidance recommends use of a "weight-of-evidence" method that does not generate TRVs appropriate for individual sites. The guidance that this comment refers to is the Ecological Soil Screening Levels (Eco-SSLs) document (USEPA, 2003k). The intent of this document is precisely to develop generic TRVs that are *not* site-specific: the TRVs are meant to be conservative and to be used for screening purposes.

In the aquatic arena, the use of conservative threshold concentrations for ecological protection is not new. EPA Water Quality Criteria (WQC), designed to protect aquatic organisms and their uses — specifically, 95% of the taxa from an appropriate variety of

taxonomic and functional groups (USEPA, 1985) — have been employed in a regulatory context for many years. The degree to which Eco-SSLs (derived from generic TRVs) may be more conservative than chronic WQC can not be assessed, because of differences in their derivation.

Another comment was received relating to inconsistency in use of UFs in developing avian TRVs for dioxin. At issue are the conditions under which a subchronic-to-chronic UF should be used. Use of an UF where unwarranted would result in an overly conservative TRV. Based on guidance in Sample et al. (1996), avian studies where exposure duration is 10 weeks or less are generally considered to be subchronic. Sample et al., however, considered 10-week exposure of ring-necked pheasant hens to TCDD in a study by Nosek et al. (1992a) to represent chronic conditions, in part because exposure occurred through a critical life-stage (reproduction). It must be noted, though, that TCDD's persistence and bioaccumulative properties make it an exception to the rule that exposure during a critical life-stage is sufficient for consideration as chronic exposure. Another study by Nosek et al. (1992b) found a half-life for whole-body elimination of TCDD in non egg-laying pheasant hens of approximately 1 year. Based on this information, only 13% of steady-state accumulation would be achieved from a 10-week exposure. A truly chronic exposure could presumably have had nearly an order of magnitude lower concentration in the food and still elicited the same tissue levels and effects (USEPA, 1993c). Thus, the 10-week exposure in the Nosek et al. study (1992a) can be considered to be subchronic, and a UF of 10 used to derive a TRV. In summary, all available information is taken into account when deciding on the use of an UF.

### 6.3.3 How Do Screening-Level and Definitive (Baseline) Assessments Differ?

The Agency has been criticized for using screening-level ecological risk assessments rather than definitive "baseline" assessments in making remedial decisions. The former are designed to be conservative, and are generally meant to be followed by more detailed assessment. A screening-level ecological risk assessment provides a comparison of abiotic media concentrations to ecotoxicological benchmarks, and does not include extensive site-specific information. In a definitive assessment, exposure concentrations are derived for wildlife receptors as well as for abiotic media. Definitive assessments require analysis of exposure-response relationships for the chemicals of concern. The use of site-specific information to derive appropriate exposure concentrations is an important element of a definitive assessment and is discussed below.

### 6.3.4 How Is Site-Specific Information Used To Derive Exposure Concentrations in Ecological Risk Assessments?

The Agency has been criticized for ignoring site-specific habitat information in a number of ecological risk assessments for Superfund sites. The degree to which spatial and temporal use of habitats by receptor species should be considered in the conceptual site model is an important question. Distinct management areas, known as operable units or OUs, are often evaluated in the

RI/FS process. It is important to have an understanding of the degree to which various habitats within an OU are, or will be, used by receptor species. For example, if we know that fish move throughout the area of an OU, exposure concentrations should be derived using data from the entire OU. Exposure estimates are generated using point estimates of exposure parameters (e.g., upper 95 % confidence limits of the mean chemical concentrations in media). More realistic estimates of exposure may be generated using distributions of exposure parameters and uncertainty analysis (e.g., Monte Carlo simulation; see USEPA, 1997b, and the references in chapter 7 of Suter et al., 2000). In the case of a large river system with tributaries that may be "relatively" clean compared to the main channel, the decision as to whether or not to include tributaries in the assessment typically depends on how contaminated they are and how they are used by the receptor species.

## 6.4 Water Quality Criteria

### 6.4.1 What Is the Agency Standard Methodology for Deriving Aquatic Life Water Quality Criteria?

Criteria are derived according on the principles set forth in the 1985 *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses* (USEPA, 1985). These guidelines are ordinarily applied as follows:

a) Acute toxicity test data need to be available for species from a minimum of eight diverse taxonomic groups. The diversity of tested species is intended to ensure protection of various components of an aquatic ecosystem.

b) The Final Acute Value (FAV) is derived by extrapolation or interpolation to a hypothetical genus more sensitive than 95% of all tested genera. The FAV, which typically represents an $LC_{50}$ or $EC_{50}$, is divided by two in order to obtain an acute criterion protective of nearly all individuals in such a genus.

c) Chronic toxicity test data (representing longer-term survival, growth, or reproduction end points) need to be available for at least three taxa to be calculated directly. When such data are not available, which is often the case, the chronic criterion is set by determining an appropriate acute-chronic ratio (the ratio of acutely toxic concentrations to chronically toxic concentrations) and applying that ratio to the acute value of the hypothetical genus more sensitive than 95% of all tested genera.

d) When necessary, the acute and/or chronic criterion may be lowered to protect recreationally or commercially important species.

### 6.4.2 What Is the GLI Tier II Methodology?

The Tier II methodology for aquatic life is contained in the final Water Quality Guidance for the Great Lakes System, or Great Lakes Initiative (GLI). The methodology was developed for use in deriving translators for narrative criteria to protect aquatic life in waters of the Great Lakes system when insufficient data are available to derive criteria. Tier II values may be based on toxicity data from as little as a single taxonomic family, provided the data are acceptable. The values are generally more stringent than criteria to reflect greater uncertainty in the absence of additional toxicity data. As more data become available, the Tier II values tend to become less conservative and more closely approximate the criteria. The general procedure is:

a) The lowest Genus Mean Acute Value in the database is divided by the Secondary Acute Factor (an empirically derived adjustment factor that varies according to the number of satisfied minimum data requirements listed in the criteria methodology) to derive the Secondary Acute Value (SAV).

b) A Secondary Chronic Value (SCV) is calculated from either the FAV (from the criteria) or the SAV divided by either the Secondary Acute to Chronic Ratio or Final Acute to Chronic Ratio.

c) The Secondary Maximum Concentration is equal to half of the SAV, and the Secondary Continuous Concentration is equal to the SCV or the Final Plant Value, if available — whichever is lower.

d) When necessary, the acute and chronic values can be adjusted to protect recreationally or commercially important species.

The final rule allows and encourages dischargers, states, and authorized tribes to develop additional data that could reduce the Tier II adjustment factors or allow development of WQC (Tier I criteria) using the methodology described above. Additionally, in situations when dischargers generate data to derive Tier I criteria or Tier II values, the final rule specifies that permit authorities may grant a reasonable period of time for data generation before the criteria or values become effective for the discharge permit.

### 6.4.3 Was the GLI Tier II Methodology Made Available for Public Comment or Vetted Through Other Public Processes?

The Tier II methodology was first developed by a Steering Committee, consisting of directors of water programs from EPA's national and regional offices and the Great Lakes states' environmental agencies (as co-regulators of Clean Water Act water quality programs), and a Technical Work Group consisting of technical staff from the Great Lakes states' environmental agencies, EPA, the U.S. Fish and Wildlife Service, and the U.S. National Park Service. All of

the work was reviewed and commented upon by a Public Participation Group consisting of representatives from environmental groups, municipalities, industry, and academia who observed the deliberations of the other two groups, advised them of the public's concerns, and kept their various constituencies informed.

EPA's SAB reviewed the draft Tier II methodology in 1992. (See its final report, EPA-SAB-EPEC/DWC-93-005, completed December 16, 1992.) The SAB commended EPA for the interactions among the states, EPA, the private sector, and the scientific community during the development of this proposed guidance and provided substantial comments on many elements of the submitted draft documents. The Board provided specific comments on the Tier II methodology. EPA responded to each of these comments in proposing and finalizing the methodology. The preamble to the proposed rule (see below) describes these comments and EPA's response to each.

The Tier II methodology was part of a 1995 rulemaking (USEPA, 1995c, d) and as such received full public comment, response, and revision based on the comments, in accordance with the Administrative Procedure Act. Federal Register notices requesting public comment on the guidance were published on April 16, 1993, August 9, 1993, September 13, 1993, and August 30, 1994 (e.g., see USEPA,1993d). Over 26,500 pages of comments, data, and information from over 6,000 respondents were received in response to these Federal Register notices and meetings held with the public.

## 6.5 Uncertainty and Pesticide Ecological Assessment

### 6.5.1 How Does EPA Address Uncertainty Analysis in Pesticide Ecological Assessments?

As a result of a dialogue with internal and external experts in the field of ecological risk assessment and probabilistic methods, the Environmental Fate and Effects Division of the Office of Pesticide Programs is currently developing methodology to routinely conduct probabilistic risk assessments (http://www.epa.gov/oppefed1/ecorisk/index.htm). The models were initially presented to the FIFRA Scientific Advisory Panel in March 2001 (http://www.epa.gov/scipoly/sap/2001/index.htm#march). The Panel recommended a number of improvements prior to implementation; these improvements will be discussed with the Panel in a SAP meeting originally scheduled for February 2004 followed by implementation in fiscal year 2004. The models were used in an assessment of the risks of the use of carbofuran flowable which is currently scheduled for completion in fiscal year 2004. In the interim, pesticide registrants have submitted probabilistic exposure assessments for chlorpyrifos, chlorfenapyr, diazinon, and atrazine. The chlorfenapyr assessment was reviewed by the FIFRA SAP, which concluded that the registrants' conclusion of negligible risk is not supported for any geographic scale. The scaling up of field exposures to the Cotton Belt scale was done incorrectly. The probabilistic exposure assessment had several other flaws that require remediation (see

http://www.epa.gov/scipoly/sap/1999/september/finalrpt.pdf). The Environmental Fate and Effects Division conducted a modified probabilistic assessment for endosulfan in which a distribution of effects was overlaid on a distribution of exposure to generate a joint probability curve.

Registrants have indicated that probabilistic risk assessments are planned for a rodenticide, while the Division may conduct one for a registered chemical going through the re-registration process in early fiscal 2004.

### 6.5.2 How Does OPP Address Rodenticide Use Patterns?

EPA issued the document *Potential Risks of Nine Rodenticides to Birds and Nontarget Mammals: A Comparative Approach* in January 2003 for public comment (USEPA, 2003l). The Agency is currently reviewing the comments submitted to OPP's docket. Similar comments were submitted in response to OMB's February 3, 2003, request for public comments on its *Draft 2003 Report to Congress on the Costs and Benefits of Federal Regulations* (USOMB, 2003a).

The document presented the Agency's preliminary assessment of potential risks to birds and non-target mammals from nine rodenticides, predominantly used to control commensal rats and mice in and around buildings and transport vehicles, and in sewers. Some also have products registered for other outdoor uses against other rodent and small mammalian pests. The assessment focused on the potential primary and secondary risks to birds and non-target mammals posed by applications of these nine rodenticides (11 baits) to control rats and mice in and around buildings (commensal use) and in field and other outdoor settings to control various rodent and other small mammalian pests. Typical use information used to estimate non-target organism exposure, such as amount of rodenticide active ingredient or formulated product applied per unit area, was not available. Thus, exposure estimates are largely based on the amount of active ingredient available per kilogram of the grain bait formulation. In preliminary risk assessments, it is assumed that birds and non-target mammals are likely to be exposed to the pesticide without attempting to establish a quantitative measure of likelihood. The existence of incident data along with liver residues provides important support for the assumption that birds and non-target mammals are exposed and adversely affected by applications of rodenticide baits. The fact that numerous species, including predators and scavengers, have been found exposed to these baits indicates that both primary and secondary exposures are occurring.

The risk conclusions were based both on the weight of evidence of the available data and comparative analysis modeling. Each rodenticide is ranked or categorized and compared to the other rodenticides according to the following criteria:

a) Overall potential risk.

b) Potential primary risk to birds.

c) Potential primary risk to non-target mammals.

d) Potential secondary risk to avian predators and scavengers.

e) Potential secondary risk to mammalian predators and scavengers.

A comment has been received that suggests the rodenticide comparative assessment fails to evaluate the likelihood of ecological exposure for the many different use patterns in the preliminary comparative "ecological risk" assessment of rodenticides.

A number of factors contributed uncertainty to the assessment. These were discussed in the document released to the public. Those that appeared to contributed the greatest uncertainty were:

a) Missing data on acute, chronic, and secondary effects, as well as retention of some active ingredients in the liver and blood.

b) The variable quality and quantity of data on metabolism and retention times in rodents and non-target species.

c) Specific use information by formulation, including typical amounts applied, distances applied from buildings, amounts used in rural versus urban areas, and so forth.

d) Information on the number and species of birds and non-target mammals likely to find and consume bait in the various use areas.

e) Methods to determine liver concentration(s) that would corroborate death from anticoagulant exposure, or even if such a cause-effect relationship is appropriate, e.g., the "threshold of toxicity" concentration in liver tissue.

f) Not accounting for the impacts of sub-lethal effects on non-target mortality, e.g., clotting abnormalities, hemorrhaging, and stress factors (including environmental stressors such as adverse weather conditions, food shortages, and predation).

g) Comparing rodenticides with different modes of action, i.e., vitamin K antagonists that disrupt normal blood-clotting (anticoagulants), a diphenylamine that is a neurotoxicant, an inorganic compound that kills by liberation of phosgene gas, and a sterol that kill by inducing hypercalcemia.

A comparative analysis model was used to rank and compare potential primary and secondary risks. The underlying methodology is a simple multi-attribute rating technique, or SMART (Goodwin and Wright, 1998). SMART was adapted for comparing potential risks among rodenticides based on a number of measure-of-effect values for primary and secondary risk to birds and mammals. It is similar to the technique used in the Agency's *Comparative Analysis of Acute Avian Risk From Granular Pesticides* (USEPA, 1992e) and *A Comparative Analysis of Ecological Risks From Pesticides and Their Uses: Background, Methodology, and Case Study* (USEPA, 1998e); both were reviewed by a FIFRA Scientific Review Panel. Concerning the latter analysis, the Panel noted the many scientific uncertainties in the method, yet agreed that it was a useful screening tool that provides a rough estimate of relative risk. The Panel made a number of helpful suggestions to improve the utility of the method, most of which are included in this comparative assessment. In this analysis, a risk quotient (RQ), calculated as the ratio of toxicant potentially ingested to the inherent toxicity of the rodenticide, was used to compare potential primary risks to birds and non-target mammals. RQs are compared among rodenticide baits based on the amount of bait and number of bait pellets that birds or non-target mammals of various sizes would need to eat to ingest an acute oral ($LD_{50}$) dose. Dietary data ($LC_{50}$) also are available for birds (but not for mammals), and RQs based on bait concentration and avian dietary toxicity are compared among the rodenticides. As noted by the Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM, 1999), RQs do not quantify risk but are useful for comparisons among alternative compounds. EPA's *Guidelines for Ecological Risk Assessment* (USEPA, 1998a) also notes that quotients provide an efficient, inexpensive means of identifying high- or low-risk situations that can allow risk management decisions to be made without the need for further information.

A major uncertainty has been how and where these pesticides are used. In estimating pesticide exposure, risk assessors rely upon the label instructions, since the label prescribes the level and manner of use. Use other than that specific on the label is illegal. The commensal use is common to all nine rodenticides. The terminology "in and around buildings" appears on product labels registered for commensal use. This statement does not limit bait placements to any specified distance from buildings, and in many non-urban areas bait applications might pose an exposure scenario comparable to some field uses. Only two of the nine rodenticide labels limit the "in and around buildings" use to urban areas; applications in non-urban areas must be indoors. Because of this label language, exposure to non-targets has been difficult to estimate: the phrase "in and around buildings" — i.e., how far from a building bait may be applied — is subject to interpretation. Further, many of these products are sold to the general public and available in outlets ranging from grocery to hardware stores. Homeowners may or may not follow the label instructions. How and where the public and pest control operators use these products is largely unknown. Currently, general marketing data are available. At best, these data report sales on a regional basis, and thus may be used to estimate volume of use. California does require that reports be maintained of pesticide use, but homeowners are exempt.

For agricultural uses, the Agency obtains information on how the pesticide is typically used: application rates, number of applications, areas in the United States where it is used, etc. The registrants, the Rodenticide Registrants Task Force, have volunteered to provide usage information for the rodenticide of most concern which will allow the Agency to further refine its risk assessment.

## 6.6 Summary and Conclusions

This chapter examines ecological risk assessment approaches in a few areas within the Superfund, water, and pesticides programs in order to address issues raised by comments received by EPA. Within the Superfund program, we discuss EPA's approach regarding organism- versus population-level ecological risk assessments and provide three examples in which conservatism is addressed. Within the water program, we discuss the standard methodology for deriving aquatic life WQC and the Great Lakes Initiative Tier II Methodology. In the pesticides program we discuss how EPA addresses uncertainty analysis in pesticide ecological assessments and rodenticide use patterns. To paraphrase a statement made earlier in this document, we attempt to use all information available in an objective, realistic, scientifically balanced way to make decisions.

Ecological risk assessment "evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors." It is a flexible process for organizing and analyzing data, information, assumptions, and uncertainties to evaluate the likelihood of adverse ecological effects. Ecological risk assessment provides a critical element for environmental decision making by giving risk managers an approach for considering available scientific information along with the other factors they need to consider (e.g., social, legal, political, or economic) in selecting a course of action.

Ecological risk assessment includes three primary phases: problem formulation, analysis, and risk characterization. In problem formulation, risk assessors evaluate goals and select assessment endpoints, prepare the conceptual model, and develop an analysis plan. During the analysis phase, assessors evaluate exposure to stressors and the relationship between stressor levels and ecological effects. In the third phase, risk characterization, assessors estimate risk through integration of exposure and stressor-response profiles, describe risk by discussing lines of evidence and determining ecological adversity, and prepare a report.

The interface among risk assessors, risk managers, and interested parties during planning at the beginning of the risk assessment, and communication of risk at its end, is critical to ensure that the results of the assessment can be used to support a management decision. Because of the diverse expertise required (especially in complex ecological risk assessments), risk assessors and risk managers frequently work in multidisciplinary teams. Both risk managers and risk assessors bring valuable perspectives to the initial planning activities for an ecological risk assessment. Risk managers charged with protecting the environment can identify information they need to

develop their decisions, risk assessors can ensure that science is effectively used to address ecological concerns, and together they can evaluate whether a risk assessment can address identified problems. However, this planning process is distinct from the scientific conduct of an ecological risk assessment. This distinction helps ensure that political and social issues, while helping to define the objectives for the risk assessment, do not introduce undue bias.

Problem formulation, which follows these planning discussions, provides a foundation upon which the entire risk assessment depends. Successful completion of problem formulation depends on the quality of three products: assessment endpoints, conceptual models, and an analysis plan. Since problem formulation is an interactive, nonlinear process, substantial reevaluation is expected to occur during the development of all problem formulation products. The analysis phase includes two principal activities: characterization of exposure and characterization of ecological effects. The process is flexible, and interaction between the two evaluations is essential. Both activities evaluate available data for scientific credibility and relevance to assessment endpoints and the conceptual model. Exposure characterization describes sources of stressors, their distribution in the environment, and their contact or co-occurrence with ecological receptors. Ecological effects characterization evaluates stressor-response relationships or evidence that exposure to stressors causes an observed response. The bulk of quantitative uncertainty analysis is performed in the analysis phase, although uncertainty is an important consideration throughout the entire risk assessment. The analysis phase products are summary profiles that describe exposure and the stressor-response relationships. During risk characterization, the final phase of an ecological risk assessment, risk assessors estimate ecological risks, indicate the overall degree of confidence in the risk estimates, cite evidence supporting the risk estimates, and interpret the adversity of ecological effects. To ensure mutual understanding between risk assessors and managers, a good risk characterization will express results clearly, articulate major assumptions and uncertainties, identify reasonable alternative interpretations, and separate scientific conclusions from policy judgments.

The Agency's Ecological Risk Assessment Guidelines are just that — guidelines, meant not to be overly prescriptive. Each program office administers different statutes that in many cases are silent on assessment endpoints or that stipulate endpoints that can be difficult to quantify. As such, programs have the responsibility to develop assessment procedures that are consistent with their enabling legislation. For the most part, the various programs' practice of ecological risk assessment is consistent across the Agency, but there are enough differences to raise questions of consistency, as illustrated by the comments this document seeks to address. EPA's goal is to apply ecological risk assessment methods in as consistent and transparent a process as possible, enabling our stakeholders to engage in a fully open debate on the risks and the management of those risks for the overall protection of the environment.

# 7. SUMMARY AND RECOMMENDATIONS

## 7.1 Summary

EPA's mission and statutes require that EPA protect human health and the environment, and ensure that EPA's risk management decisions provide some "margin of safety" when risk assessment answers are uncertain or variable. The risk estimates provided in risk assessments and ultimately addressed in EPA decisions are driven by the uncertainty and variability inherent in practically all the information and methodologies EPA uses. This uncertainty stems to a large extent from the complex nature of the environment and the processes that interact with the stressor(s) to cause risk. It is clear that risk assessors best understand the basic elements of a risk assessment, in particular how to analyze the uncertainty and variability found in individual risk assessments. Risk managers, by taking into consideration the characterization of all the elements of a risk assessment — including the uncertainty and variability, provided by the risk assessors — should determine the appropriate degree of protection to ensure that risk is not underestimated and conversely, is not appreciably overestimated.

Comments on EPA risk assessment principles and practices have been received over the years, more recently in response to the OMB Federal Register notice asking for such comments. An overarching position in the comments suggests that EPA inappropriately mixes policy and science: that, in generating risk assessments, risk assessors "decide" the degree of protection in a way that is not transparently characterized. Further, in more than a few instances, it is suggested that EPA's current risk assessment practice compounds risk estimates to a protective position that is beyond reasonable. Also, in many instances, risk managers apparently are not aware of the degree of protection inherent in each risk assessment. Some other comments suggest that EPA risk estimates tend to underestimate risk, since a majority of the assessments tend to be single-stressor assessments and do not adequately take into account multiple stressors, multiple exposures, and susceptible populations.

EPA risk assessment practices are meant to produce credible, science-based risk assessments that provide reasonable risk estimates. The assessments are designed to ensure that risk is not underestimated or grossly overestimated. EPA views a critical analysis of the available data specific to the relevant chemical (stressor) and/or site as the *starting* point for a risk assessment. The derivation of risk estimates improves continually with the addition of newer techniques and relevant data. In the absence of relevant data, EPA uses assumptions in developing estimates of risk. The assumptions (defaults and extrapolations) used in the assessments are necessary components of the assessment, as they help bridge the data gaps encountered during the assessment. These assumptions are also focused on providing a reasonable default position and usually have scientific data and scientific support. They are also publicly vetted and peer reviewed. For the conduct of risk assessment, EPA uses a triage approach (from screening to "in-depth" assessments) to determining how much time and

resources need to be expended. This is a practical approach for dealing with the generation of different "depths" of assessment.

Another major theme in the comments regarding EPA risk assessment principles and practices suggests that transparency and clarity are still absent in characterizing the elements of the risk assessment, particularly the uncertainty and variability as well as the choices made in the assessment. This may be an aspect of EPA's practices that needs strengthening. It should be clear that the development of risk assessments for risk management purposes is a mix of science data, science policy, judgments, guidelines, and best professional judgement. Transparency about all these matters is crucial so that nothing appears "hidden" or "buried" in an assessment — so that nothing keeps one from understanding the impact of the elements that go into estimating and characterizing a risk.

## 7.2 Several Current EPA Activities That Will Enhance EPA Risk Assessment Principles and Practices

EPA works on many fronts to review our practices and methodologies as part of an ongoing evaluation of risk assessment principles and practices. Several current examples demonstrate our commitment to enhancing EPA risk assessment principles and practices:

a) EPA is currently finalizing its proposed Guidelines for Cancer Risk Assessment. Included in this effort is the generation of a supplemental guidance focused on cancer risk to children, particularly from mutagenic carcinogens. This is a major effort that explores many newer issues, such as examining data before invoking defaults, extensive use of mode of action to inform the assessment, more clearly defining default positions that can be used when data are unavailable, and making the cancer risk assessment practice more transparent and clear.

b) EPA is engaged in a long-term examination to "harmonize" cancer and non-cancer health assessments into a unified approach more geared toward examining how health endpoints, via a mode of action, present linear or nonlinear dose-response curves that can be used for quantitation. Other major aspects of this effort are the use of chemical-defined uncertainty factors, guidance for selecting a point of departure from observed data to extrapolate to lower doses of expected exposures, and clarification of dosing and scaling factors to use in health assessments.

c) The use of environmental models to support risk assessments is constantly increasing. EPA promotes consistency and consensus on mathematical modeling issues including modeling guidance, development, and application, and enhances both internal and external communications on modeling activities through organizations such as the EPA Council on Regulatory Environmental Modeling.

d) The better the data and information are, the better the risk assessment will be in terms with being more reasonable with more certainty. To that end, EPA initiated an effort to examine the usefulness of the toxicity tests needed, and required under several statutes, to support EPA assessments and decisions. As the field of toxicology evolves, particularly recently with the Human Genome Project and the burgeoning area of genomics and related technologies, it is appropriate to examine what will be the best tests we can use to generate data useful for risk assessments. EPA is working with the NAS to examine these toxicity testing needs as projected into the next two decades.

e) EPA is undertaking a major effort to upgrade and improve its process for creating assessments to enter onto IRIS. As a result of an increase in IRIS staff and funding, the number of new and updated assessments completed each year will be increased. A more open external peer review component will provide more transparency and interaction for external stakeholders.

f) EPA focuses much attention on its research program to help reduce major uncertainties in risk assessment by understanding and elucidating the fundamental determinants of exposure and dose and the basic biological changes that follow exposure to stressors leading to a toxic response. For example, in EPA's recent *Human Health Research Strategy* (USEPA, 2003m), EPA will focus on ways to improve the information supporting its risk assessments. These include approaches to harmonizing the use of mechanistic data in human health assessments, predicting the effects of aggregate and cumulative exposure, and protecting susceptible populations such as children, older adults, and those with preexisting disease or genetic predispositions for different responsiveness to environmental stressors. Many of these areas are discussed in this document.

## 7.3 Recommendations for Further Improvements to Risk Assessment Principles and Practices

Based on the evaluation in this document, the staff provides several recommendations for EPA to consider as part of an agenda for further improvements to risk assessment principles and practices.

a) Confidence in our risk assessments is critical. This confidence is increased by the use of scientifically sound, chemical- and site-specific information. In the absence of such data, EPA should continue to use default assumptions to fill data gaps.

1) We need to continue to encourage the development of the specific data necessary to more accurately assess potential risks, particularly with those responsible for generating appropriate data (e.g., pesticide registrants). In the absence of strong data requirements, EPA should continue to work with industry and other partners to implement voluntary programs, such as the High Production Volume Challenge, to generate such data, thus lessening our reliance on defaults in developing risk assessments.

2) Since defaults assumptions are used in risk assessments to address uncertainty when chemical- and/or site-specific data are not available or useful, we need to ensure that the defaults themselves are supported by the best available data. For example, the draft supplemental children's guidance accompanying the EPA draft cancer guidelines suggests a default position for children exposed to mutagenic carcinogens that is based on data, albeit with a small sample number, that supports the default assumption (USEPA, 2003d). We should continually look for opportunities to increase our certainty and confidence in the defaults and extrapolations we use. As more data become available, we develop and/or refine the default assumptions and extrapolations to use; for example, defaults might be developed for different life-stages or for chemicals with a specific mode of action. EPA should encourage and support research needed to develop and refine assumptions for use in risk assessments. Further, to help increase our confidence in risk assessments, EPA might consider the use of sensitivity analysis to examine what data points and/or defaults involved in the assessment are the most critical.

b) Transparency in our risk assessment practice and our risk management process is an issue that needs to be addressed. Initiatives for EPA to improve our transparency in risk assessments and risk management include:

1) Continued concerted and conscious use of planning and scoping with risk assessors and risk managers before a risk assessment is started. The intent is *not* to point to a specific decision choice, but to outline what needs to be addressed and how to do so. Just as important is to be clear about what will not be addressed (i.e., the scope of the assessment effort needs to be defined).

2) Continued use of the triage approach to decide how much time and resources are necessary for a risk assessment. However, we need to be more clear about why the risk assessment is being performed. This would help reduce potential confusion about why EPA decides upon particular risk estimates in one decision and others in another decision (e.g., a

screening assessment versus a detailed assessment; different circumstances at different sites).

3) We need to focus on better communication of the data and assumptions and choices used in our risk assessments. We already have the guidance in place in the form of our Risk Assessment Guidelines and Agency Risk Characterization Policy and Handbook. Also, we have Planning and Scoping Guidance as the first part of our Cumulative Risk initiative, and this is incorporated into the new Cumulative Risk Framework. Close attention to our guidances will ensure transparency and clarity in our risk assessments. This will also help ensure that risk assessments will appropriately support and inform decision making and risk management.

4) Regarding transparency in the risk management process: since we do not have much guidance here, we should encourage work on a decision making framework. This framework would help everyone, from those inside EPA to all interested parties outside EPA, understand the process risk managers (decision makers) go through to arrive at a decision. It should not be a checklist of steps, but a set of actions (e.g., peer review) and factors (e.g., cost/benefit analysis) risk managers should consider before making a decision. While it is certainly feasible to produce such a framework explaining our methods of procedure in general terms, it cannot explain relevant "weights" given in risk management decisions to the various factors (costs, benefits, treatment options, etc.) as each decision will be made on a case-by-case basis. So most importantly, this transparency will clarify how a risk assessment is incorporated into a decision and how it informs the decision. Places to look for information to construct such a framework include the NAS reports and the Commission on Risk Assessment and Risk Management.

c) A major method to help address uncertainty in risk assessment is probabilistic analysis. We can use that technique for exposure assessments, but even for these assessments, we should encourage greater use and reliance on it when appropriate. Further, we need to explore the feasibility of using probabilistic analyses in any phase of the risk assessment. For example, we need to encourage the development of Monte Carlo and other probabilistic analyses in the dose-response assessment. If probabilistic analyses are identified for particular uses and it is generally agreed they are appropriate, we should start incorporating the relevant analyses where the risk assessor feels it is appropriate to do so.

1) It should be noted that the seemingly slow implementation of probabilistic analysis is not strictly an internal EPA issue. Much environmental risk

assessment is performed by affected parties rather than EPA itself, and outside parties' pace in implementing and/or accepting probabilistic/quantitative analysis of uncertainty is seemingly slow as well. Notwithstanding, we recommend that EPA examine a few recent or ongoing risk assessments to determine what the obstacles are to use of probabilistic analysis to describe uncertainty and how to overcome those obstacles more frequently in the future. A major area that needs work is the development of probability distributions for many endpoints in the toxicity part of risk assessment, as well as putting the toxicity and exposure components together in an overall probabilistic risk assessment. Preliminary work in this area has been published, with some key papers concerning the development of probability distributions for uncertainty factors (Dourson et al., 1992; Baird et al., 1996; Swartout et al., 1998).

d) The staff feels greater harmonization with our partners — states, other national governments (as in the case of OPP's interactions with Canada), OECD — as well as greater cross-program interaction inside the Agency to ensure consistency and scientific consensus for practices and methodologies (e.g., probabilistic analyses) will help avoid a duplication of effort or differences in approaches and interpretations.

e) EPA has historically assessed risks based on individual chemicals/stressors and often focused on one source, pathway, or adverse effect. However, people do not live in a "single chemical" world: in reality, they are exposed to multiple chemicals/stressors from a variety of sources. Because people are exposed to multiple stressors, this suggests that the total risk they may experience in their daily lives may be underestimated by assessments of single stressors. Tools are needed to understand combined risks. While the Agency has moved forward with emphasis on evaluating cumulative risks (e.g., planning and scoping activities for cumulative risk, the Cumulative Risk Framework, OPP examination of aggregate risk), EPA needs to continue moving aggressively to flesh out approaches for cumulative risk and to produce the most scientifically rigorous and realistic evaluation of cumulative risk that the state-of-the-science can accommodate.

f) Overarching guidance is needed to improve the ease and consistency of risk assessment for susceptible populations and life-stages. Such guidance should provide more detail, methods, and advice for considering the issues regarding susceptibility. As a part of such efforts, EPA needs to continue developing a strategy and methodologies to address the cumulative risks of unique or disproportionate exposures.

g) Peer review is fundamental to ensuring the quality and integrity of all assessments. We need to continue with constant vigilance to peer review all our major scientific and technical work products.

h) Many practices EPA is criticized for are practices we no longer follow (e.g., compound uncertainty factors up to 10,000). However, there are older risk assessments supporting standing decisions that may need reexamination. This would need to be done in light of current resources and priorities, e.g., looking at older IRIS values for chemicals that are still EPA priorities.

# LIST OF USEFUL ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| ACC | American Chemistry Council |
| ADI | Acceptable Daily Intake |
| AEGL | acute exposure guideline level |
| ARAR | applicable or relevant and appropriate requirement |
| ARE | acute reference exposure |
| AQI | Air Quality Index |
| ATFERM | Agency Task Force on Environmental Regulatory Modeling |
| ATSDR | Agency for Toxic Substances and Disease Registry |
| AUC | area under the curve |
| BAF | bioaccumulation factor |
| BAT | best available technology |
| BMC | benchmark concentration |
| BMCL | benchmark concentration lower confidence limit |
| BMD | benchmark dose |
| BMDL | benchmark dose lower confidence level |
| BMR | benchmark response |
| CAA | Clean Air Act |
| CAAA | Clean Air Act Amendments |
| CBD | chronic beryllium disease |
| CDC | Centers for Disease Control and Prevention |
| CEPPO | Chemical Emergency Preparedness and Prevention Office |
| CERCLA | Comprehensive Environmental Response, Compensation, and Liability Act |
| CERCLIS | Comprehensive Environmental Response, Compensation, and Liability Information System |
| CFSAN | Center for Food Safety and Nutrition |
| CNS | central nervous system |
| CPSC | Consumer Product Safety Commission |
| CRARM | Congressional/Presidential Commission on Risk Assessment and Risk Management |
| CREM | Committee on Regulatory Environmental Modeling |
| CSAF | chemical-specific adjustment factor |
| CSF | cancer slope factor |
| CSFII | Continuing Survey of Food Intakes by Individuals |
| CTE | central tendency exposure |
| CWA | Clean Water Act |
| DAF | dosimetric adjustment factor |
| DBP | disinfection byproduct |
| DENR | Department of Environment and Natural Resources |
| DHHS | Department of Health and Human Services |

| | |
|---|---|
| DHS | Department of Homeland Security |
| DNA | deoxyribonucleic acid |
| DNT | developmental neurotoxicity |
| DOD | Department of Defense |
| DOE | Department of Energy |
| DOJ | Department of Justice |
| DOT | Department of Transportation |
| ECC | Environmental Clearance Committee |
| Eco-SSLs | Ecological Soil Screening Levels |
| ED | effective dose |
| EDC | endocrine disrupting chemical |
| EMAP | Environmental Monitoring and Assessment Program |
| EPA | Environmental Protection Agency |
| EPCRA | Emergency Planning and Community Right-To-Know Act |
| ETV | Environmental Technology Verification |
| FAV | Final Acute Value |
| FDA | Food and Drug Administration |
| FDAMA | Food and Drug Administration Modernization Act |
| FFDCA | Federal Food, Drug, and Cosmetic Act |
| FIELDS | Field Environmental Decision Support |
| FIFRA | Federal Insecticide, Fungicide, and Rodenticide Act |
| FQPA | Food Quality Protection Act |
| FR | Federal Register |
| FS | feasibility study |
| GAO | General Accounting Office |
| GHS | globally harmonized system |
| GIS | geographic information system |
| GLI | Great Lakes Initiative |
| GLNPO | Great Lakes National Program Office |
| GPRA | Government Performance and Results Act |
| GPS | Global Positioning System |
| GSA | General Services Administration |
| HA | Health Advisory |
| HAP | hazardous air pollutant |
| HAZMAT | hazardous materials |
| HEC | human equivalent concentration |
| HED | human equivalent dose |
| HEPA | high-efficiency particulate air |
| HI | hazard index |
| HMIS | Hazardous Materials Information System |
| HQ | hazard quotient |
| HVAC | heating, ventilation, and air conditioning |

| | |
|---|---|
| IBI | Index of Biological Integrity |
| IC | index compound |
| ICRP | International Commission on Radiological Protection |
| IEUBK | Integrated Exposure Uptake Biokinetic Model for Lead in Children |
| IME | individuals most exposed |
| IPCS | International Programme on Chemical Safety |
| IRIS | Integrated Risk Information System |
| ISO | International Organization for Standardization |
| LED | lowest effective dose |
| LOAEL | lowest-observed-adverse-effect level |
| LOEL | lowest-observed-effect level |
| MACT | maximum achievable control technology |
| MCL | maximum contaminant level |
| MCLG | maximum contaminant level goal |
| MEI | maximally exposed individual |
| MF | modifying factor |
| MIMS | Multimedia Integrated Modeling System |
| MLE | maximum likelihood estimate |
| MOA | mode of action |
| MOE | margin of exposure |
| MOU | Memorandum of Understanding |
| MRL | minimal risk level |
| MTD | maximum tolerated dose |
| NAAQS | National Ambient Air Quality Standards |
| NAS | National Academy of Sciences |
| NATO/CCMS | North Atlantic Treaty Organization Committee on the Challenges of Modern Society |
| NCEA | National Center for Environmental Assessment |
| NCER | National Center for Environmental Research |
| NCP | National Oil and Hazardous Substances Pollution Contingency Plan |
| NCTR | National Center for Toxicological Research |
| NEIC | National Enforcement Investigation Center |
| NERL | National Environmental Research Laboratory |
| NESHAP | National Emissions Standard for Hazardous Air Pollutants |
| NHEERL | National Health and Environmental Effects Research Laboratory |
| NHSRC | National Homeland Security Research Center |
| NIEHS | National Institute for Environmental Health Sciences |
| NIOSH | National Institute for Occupational Safety and Health |
| NNI | National Nanotechnology Initiative |
| NOAA | National Oceanic and Atmospheric Administration |
| NOAEL | no-observed-adverse-effect level |
| NOEL | no-observed-effect level |

| | |
|---|---|
| NPDES | National Pollutant Discharge Elimination System |
| NRC | National Research Council |
| NRMRL | National Risk Management Research Laboratory |
| NSF | National Science Foundation |
| NTP | National Toxicology Program |
| OAQPS | Office of Air Quality, Planning, and Standards |
| OAR | Office of Air and Radiation |
| OECD | Organisation for Economic Cooperation and Development |
| OERR | Office of Emergency and Remedial Response |
| OMB | Office of Management and Budget |
| OPEI | Office of Policy, Economics, and Innovation |
| OPP | Office of Pesticide Programs |
| OPPT | Office of Pollution Prevention and Toxics |
| OPPTS | Office of Prevention, Pesticides and Toxic Substances |
| ORD | Office of Research and Development |
| OSHA | Occupational Safety and Health Administration |
| OSTP | Office of Science and Technology Policy |
| OSW | Office of Solid Waste |
| OSWER | Office of Solid Waste and Emergency Response |
| OU | operable unit |
| OW | Office of Water |
| OWOW | Office of Wetlands, Oceans, and Watersheds |
| PAD | population-adjusted dose |
| PAH | polycyclic aromatic hydrocarbon |
| PBPK | physiologically based pharmacokinetic |
| PCB | polychlorinated biphenyl |
| PCDD | polychlorinated dibenzo-*p*-dioxin |
| PCDF | polychlorinated dibenzofurans |
| PEL | permissible exposure limit |
| PM | particulate matter |
| PMN | premanufacture notification |
| POD | point of departure |
| PPE | personal protective equipment |
| ppm | parts per million |
| PRG | Preliminary Remediation Goal |
| QA | quality assurance |
| QAPP | Quality Assurance Project Plan |
| QA/QC | quality assurance/quality control |
| QSAR | quantitative structure-activity relationship |
| RAF | Risk Assessment Forum |
| RAGS | Risk Assessment Guidance for Superfund |
| RCRA | Resource Conservation and Recovery Act |

| | |
|---|---|
| RDDR | regional deposited dose ratio |
| ReVA | Regional Vulnerability Assessment |
| RGDR | regional gas dose ratio |
| RfC | reference concentration |
| RfD | reference dose |
| RI | remedial investigation |
| RME | reasonable maximum exposure |
| RNA | ribonucleic acid |
| ROD | Record of Decision |
| RPF | relative potency factor |
| RQ | risk quotient |
| RSC | relative source contribution |
| SAB | Science Advisory Board |
| SAP | Scientific Advisory Panel |
| SAR | structure-activity relationship |
| SAV | Secondary Acute Value |
| SARA | Superfund Amendments and Reauthorization Act |
| SCV | Secondary Chronic Value |
| SDWA | Safe Drinking Water Act |
| SDWAA | Safe Drinking Water Act Amendments |
| SEER | Surveillance, Epidemiology, and End Results |
| SEQL | Sustainable Environment for Quality of Life |
| SIP | State Implementation Plan |
| SMART | simple multi-attribute rating technique |
| SMCL | secondary maximum contaminant level |
| SOPs | standard operating procedures |
| STEL | short-term exposure limit |
| STAR | Science To Achieve Results |
| TCCR | transparency, clarity, consistency, reasonableness |
| TCDD | tetrachlorodibenzo-*p*-dioxin |
| TEF | Toxicity Equivalency Factor |
| TEQ | Toxic Equivalency Quotient |
| TI | tolerable intake |
| TIO | Technology Innovation Office |
| TRC | Toxicogenomics Research Consortium |
| TRI | Toxics Release Inventory |
| TRV | toxicity reference value |
| TSCA | Toxic Substances Control Act |
| TWA | time-weighted average |
| UCL | upper confidence limit |
| $UCL_{95}$ | 95% upper confidence limit of the arithmetic mean |

| | |
|---|---|
| UF | uncertainty factor |
| USDA | United States Department of Agriculture |
| WHO-ECEH | World Health Organization European Centre for Environmental Health |
| WQC | water quality criteria |

# GENERAL REFERENCES

Agency for Toxic Substances and Disease Registry (ATSDR). (1991). Preventing lead poisoning in young children. Atlanta, GA.

Agency for Toxic Substances and Disease Registry (ATSDR). (1996). ATSDR public health assessment guidance manual. Atlanta, GA.

Agency for Toxic Substances and Disease Registry (ATSDR). (2000). Summary report for the ATSDR Soil-Pica Workshop, June 2000. Atlanta, GA.

Ahlborg, M.G., G.C. Becking, L.S. Birnbaum, A. Brouwer, H.J.G.M. Derks, M. Feeley, G. Golog, S. Hanberg, J.C. Larsen, A.K.D. Liem, S. Safe, C. Schlatter, F. Waern, M. Younes, and E. Yrjänheikki. (1994). Toxic equivalency factors for dioxin-like PCBs: Report on a WHO-ECEH and IPCS consultation, December 1993. *Chemosphere* 26(6):1049-1067.

American Academy of Pediatrics (AAP). (2003). Pediatric environmental health, 2nd ed. Ruth A. Etzel, ed. Elk Grove Village, IL.

American Conference of Governmental Industrial Hygienists (ACGIH). (1991). Guidelines for the classification of occupational carcinogens. In: Documentation of the threshold limit values and biological exposure indices, 6th ed. Cincinnati, OH.

Ames, B.N., and L.S Gold. (1990). Chemical carcinogenesis: too many rodent carcinogens. *Proc Natl Acad Sci USA* 87:7772-7776.

Asian Pacific Environmental Network (APEN). (1998). A seafood consumption survey of the Laotian community of West Contra Costa County, CA. Oakland, CA.

Bachmann, R., A.L. Barton, J.R. Clark, P.L. deFur, S.J. Ells, C.A. Pittinger, M.W. Slimak, R.G. Stahl, and R.S. Wentsel. (1998). A multi-stakeholder framework for ecological risk management: Summary of a SETAC technical workshop. Supplement to *Environ Toxicol Chem* 18(2). 24 pp.

Baird, S., J. Cohen, J. Graham, A. Shlyakhter, and J. Evans. (1996). Noncancer risk assessment: A probabilistic alternative to current practice. *Hum Ecol Risk Assess* 2(1):79-102.

Barnes, D., A. Alford-Stevens, L. Birnbaum, F.W. Kutz, W. Wood, and D. Patton. (1991). Toxicity equivalency factors for PCBs? *Qual Assur* 1(1):70-81.

Barnthouse, L.W., J. Boreman, S.W. Christensen, C.P. Goodyear, D.S. Vaughan, and W. van Winkle. (1984). Population biology in the courtroom: The Hudson River controversy. *Bioscience* 34:14-19.

Barton, A., and A. Sergeant. (1998). Policy before the ecological risk assessment: What are we trying to protect? *Hum Ecol Risk Assess* 4(4):787-795.

Bascietto, J.J. (1998). A framework for ecological risk assessment: Beyond the quotient method. In: M.C. Newman and C.L. Strojan, eds. Risk assessment: Logic and measurement. Ann Arbor Press. 352 pp.

Bascietto, J., D. Hinckley, J. Plafkin, and M. Slimak. (1990). Ecotoxicity and ecological risk assessment. *Environ Sci Technol* 24(1):10-15.

Behrman, L.E., and V.C. Vaughan, III. (1983). *Textbook of Pediatrics.* Philadelphia, PA: W.B. Saunders Co.

Bickis, M., and D. Krewski. (1989). Statistical issues in the analysis of the long-term carcinogenicity bioassay in small rodents: an empirical evaluation of statistical decision rules. *Fundam Appl Toxicol* 12:202-221.

Burmaster, D.E. (2000). Distributions of projected job tenure for men and women in selected industries and occupations in the United States, February 1996. *Risk Anal* 20(2):205-224.

Calabrese, E.J., R. Barnes, E.J. Stanek 3rd, H. Pastides, C.E. Gilbert, P. Veneman, X.R. Wang, A. Lasztity, and P. T. Kostecki. (1989). How much soil do young children ingest: An epidemiologic study. *Regul Toxicol Pharmacol* 10:123-137.

Calabrese E.J., E.J. Stanek, and C.E. Gilbert. (1991). Evidence of soil-pica behavior and quantification of soil ingestion. *Hum Exp Toxicol* 10:245-249.

Casteel, S., R. Coward, C. Weis, G. Hermingsen, E. Hoffman, W. Brattin, R. Guzman, M. Starost, J. Payne, S. Stockharm, S. Becker, J. Drexler, and J. Turk. (1997). Bioavailability of lead to juvenile swine dosed with soil from the Smuggler Mountain NPL site of Aspen, Colorado. *Fundam Appl Toxicol* 36(2):177-187.

Centers for Disease Control and Prevention (CDC). (2003). Second national report on human exposure to environmental chemicals. NCEH Publication No. 02-0716. National Center for Environmental Health (NCEH), Centers for Disease Control and Prevention, Atlanta, GA. (http://www.cdc.gov/exposurereport/)

Charney, E., J. Sayer, and M. Coulter. (1980). Increased lead absorption in inner city children: Where does the lead come from? *Pediatrics* 65:226.

Cogliano, V.J. (1997). Plausible upper bounds: Are their sums plausible? *Risk Anal* 17:77-84.

Columbia River Inter-Tribal Fish Commission (CRITFC). (1994). A fish consumption survey of the Umatilla, Nez Perce, Yakama, and Warm Springs Tribes of the Columbia River Basin. Technical Report 94-3. CRITFC, Portland, OR.

Commission on Risk Assessment and Risk Management (CRARM). (1997a). Framework for environmental health risk management, Final report, Volume 1. The Presidential/Congressional Commission on Risk Assessment and Risk Management, Washington, DC.

Commission on Risk Assessment and Risk Management (CRARM). (1997b). Risk assessment and risk management in regulatory decision-making, Final report, Volume 2. The Presidential/Congressional Commission on Risk Assessment and Risk Management, Washington, DC.

Comprehensive Environmental Response, Compensation, and Liability Act (Superfund) (CERCLA). (1980). 42 U.S.C. 9601-9627.

Connelly, N.A., B.A. Knuth, and C.A. Bisogni. (1992). Effects of the health advisory changes on fishing habits and fish consumption in New York sport fisheries. Report for the New York Sea Grant Institute, Project No. R/FHD-2-PD. Dept. of Natural Resources, Cornell University, Ithaca, NY. (Raw survey data also received electronically from study authors.)

Connelly, N.A., B.A. Knuth, and T.L. Brown. (1996). Sportfish consumption patterns of Lake Ontario anglers and the relationship to health advisories. *N Amer J Fish Man* 16:90-101.

Cooper, W.E. (1998). Risk assessment and risk management: An essential integration. *Hum Ecol Risk Assess* 4(4):931-937.

Corish, J., and D. Fitzpatrick. (2002). Theoretical models of percutaneous absorption. In: Proceedings of the International Conference on Occupational and Environmental Exposures of Skin to Chemicals, Crystal City, VA. Sept. 8-11. National Institute for Occupational Safety and Health, Washington, DC.

Davis, S., P. Waller, R. Buschbon, J. Ballou, and P. White. (1990). Quantitative estimates of soil ingestion in normal children between the ages of 2 and 7 years: Population-based estimates using aluminum, silicon, and titanium as oil tracer elements. *Arch Environ Health* 45:112-122.

Dourson, M., S. Felter, and D. Robinson. (1996). Evolution of science-based uncertainty factors in noncancer risk assessment. *Regul Toxicol Pharmacol* 24:108-120.

Dourson, M., L. Knauf, and J. Swartout. (1992). On reference dose (RfD) and its underlying toxicity data base. *Toxicol Ind Health* 8:171-189.

Dourson, M., and J. Stara. (1983). Regulatory history and experimental support of uncertainty (safety) factors. *Regul Toxicol Pharmacol* 3:224-238.

Duguay, A.B., and G.L. Plaa. (1995). Tissue concentrations of methyl isobutyl ketone, methyl n-butyl ketone and their metabolites after oral or inhalation exposure. *Toxicol Lett* 75(1-3):51-58.

Eadon, G., L. Kaminsky, J. Silkworth, K. Aldous, D. Hilker, P. O'Keefe, R. Smith, J. Gierthy, J. Hawley, N. Kin, and A. DeCaprio. (1986). Calculation of 2,3,7,8-TCDD equivalent concentrations of complex environmental contaminant mixtures. *Environ Health Perspect* 70:221-227.

Ebert, E., N. Harrington, K. Boyle, J. Knight, and R. Keenan. (1993). Estimating consumption of freshwater fish among Maine anglers. *N Amer J Fish Man* 13:727-745.

Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM). (1999). Terrestrial draft report. Report by ECOFRAM Terrestrial Workgroup. May 10. 677 pp.

Executive Office of the President of the United States (EOP). (1991). Regulatory Program of the United States Government. Current Regulatory Issues in Risk Assessment and Risk Management, 1 April 1990 - 31 March 1991.

Executive Order (EO) 12898. (1994). Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations. *Fed Reg* 59(32):7629.

Executive Order (EO) 13045. (1997). Protection of children from environmental health and safety risks. *Fed Reg* 62(78):19883-19888. April 23.

Food Quality Protection Act (FQPA). (1996). U.S. Public Law 104-170. 104th Cong., 3 August 1996. Section 408(b)(2)(C).

Foureman, G.L and H.J. Clewell. (1999). Route-to-route extrapolation with a physiologically-based pharmacokinetic (PBPK) model for cumene. *Toxicol Sci* 48(1-S):395 (abstract).

Gansewendt, B., U. Foest, D. Xu, E. Hallier, H.M. Bolt, and H. Peter. (1991). Formation of DNA adducts in F-344 rats after oral administration or inhalation of [14C]methyl bromide. *Food Chem Toxicol* 29:557-563.

Gaylor, D.W., and R.L. Kodell. (2000). Percentiles of the product of uncertainty factors for establishing probabilistic reference doses. *Risk Anal* 20:245-250.

Giesy, J.P., and K. Kannan. (1998). Dioxin-like and non-dioxin-like toxic effects of polychlorinated biphenyls (PCBs): Implications for risk assessment. *Crit Rev Toxicol* 28:511-569.

Gillies, M.E., and H.V. Paulin. (1983). Variability of mineral intakes from drinking water: A possible explanation for the controversy over the relationship of water quality to cardiovascular disease. *Int J Epid* 12(1):45-50.

Goodwin, P., and G. Wright. (1998). Decision analysis for management judgement, 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc. 454 pp.

Gospe, S.M., and M.A.S. Albayati. (1994). Comparison of oral and inhalation exposures to toluene. *J Am Coll Toxicol* 13:21-32.

Hamm, J.T., C.-Y. Chen, and L.S. Birnbaum. (2003). A mixture of dioxins, furans and non-ortho PCBs based upon consensus toxic equivalency factors produces dioxin-like reproductive effects. *Toxicol Sci* 74:182-191.

Haseman, J.K. (1985) Issues in carcinogenicity testing: Dose selection. *Fundam Appl Toxicol* 5:66-78.

Haseman, J.K., and A.M. Lockhart. (1993). Correlations between chemically related site-specific carcinogenic effects in long-term studies in rats and mice. *Environ Health Perspect* 101:50-54.

Hill, A.B. (1965). The environment and disease: Association or causation? *Proc R Soc Med* 58:295-300.

Huff, J., J. Cirvello, J. Haseman, and J. Bucher. (1991). Chemicals associated with site-specific neoplasia in 1394 long-term carcinogenesis experiments in laboratory rodents. *Environ Health Perspect* 93:247-270.

International Commission on Radiological Protection (ICRP). (1981). Report of the task group on reference man. New York: Pergammon Press.

Javitz, H. (1980). Seafood consumption data analysis. SRI International. Final report prepared for EPA Office of Water Regulations and Standards. EPA Contract 68-01-3887.

Kadry, A.M., G.A. Skowronski, and M.S. Abdel-Rahman. (1995). Evaluation of the use of uncertainty factors in deriving RfDs for some chlorinated compounds. *J Toxicol Environ Health* 45:83-95.

Kalbfleisch, J.D., and R.L. Prentice. (1980). The statistical analysis of failure time data. New York: Wiley.

Kaneko, T., P.Y. Wang, H. Tsukasa, and A. Sato. (1995). m-xylene toxicokinetics in phenobarbital-treated rats — comparison among inhalation exposure, oral administration, and intraperitoneal administration. *Toxicol Appl Pharmacol,* 131(1):13-20.

Kutz, F.W., D.G. Barnes, E.W. Bretthauer, D.P. Bottimore, and H. Greim. (1990). The international toxicity equivalency factor (I-TEF) method for estimating risks associated with exposures to complex mixtures of dioxins and related compounds. *Toxicol Environ Chem* 26:99-109.

Kutz, F.W., R.A. Linthurst, C. Riordan, M.W. Slimak, and R. Frederick. (1992). Ecological research at EPA: New directions. *Environ Sci Technol* 26:861-866.

Kuzmack, A.M., and R.E. McGaughy. (1975). Quantitative risk assessment for community exposure to vinyl chloride. EPA Office of Planning and Management and Office of Health and Ecological Effects, December 1975.

Lackey, R.L. (1994). Ecological risk assessment. *Fisheries* 19(9):14-18.

Larson, S.J., C.G. Crawford, and R.J. Gilliom. (In press). Development and application of watershed regression for pesticides (WARP) methods to estimate atrazine concentration distributions in streams. U.S. Geological Survey Water Resources Investigations Report 03-4047. In press. (Draft version: (http://in.water.usgs.gov/nawqa/mmwg/warp.atrazine.110502.pdf)

Larson, S.J., and R.J. Gilliom. (2001). Regression models for estimating herbicide concentrations in U.S. streams from watershed characteristics. *J Amer Wat Res Assoc* 37:1349-1367.

Leece, B., M.A. Denomme, R. Towner, S.M.A. Li, and S. Safe. (1985). Polychlorinated biphenyls: Correlation between in vivo and in vitro quantitative structural-activity relationships (QSARs). *J Toxicol Environ Health* 16:379-388.

Linthurst, R.A., L.A. Mulkey, M.W. Slimak, G.D. Veith, and B.M. Levinson. (2000). Ecological research in the Office of Research and Development at the U.S. Environmental Protection Agency: An overview of new directions. *Environ Toxicol Chem* 19(4):1222-1229.

Maki, A.W., and M.W. Slimak. (1990). The role of ecological risk assessment in environmental decision making. In: Grodzinski, W., ed. Ecological risks: Perspectives from Poland and the United States. National Academy Press: Washington, DC. pp. 77-91.

Maltby, L., T.J. Kedwards, V.E. Forbes, K. Grasman, J.E. Kammenga, W.R. Munns, Jr., A.H. Ringwood, J.S. Weis, and S.N. Wood. (2001). Linking individual-level responses and population-level consequences. In: Baird, D.J., and H.A. Burton, Jr., eds. Ecological complexity: New directions for assessing responses to stress. SETAC Special Publication Series, Society of Environmental Toxicology and Chemistry, Pensacola, FL. pp. 27-82.

McEuen, S.F., C.F. Jacobson, C.D. Brown, and M.G. Miller. (1995). Metabolism and testicular toxicity of 1,3-dinitrobenzene in the rat: Effect of route of administration. *Fundam Appl Toxicol* 28(1):94-9.

National Center for Health Statistics (NCHS). (1987). Anthropometric reference data and prevalence of overweight, United States, 1976-1980. Data from the National Health and Nutrition Examination Survey. Series 11, No. 238. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, Hyattsville, MD. DHHS Publication No. (PHS) 87-1688.

National Contingency Plan (NCP). (1990a). National Oil and Hazardous Substances Pollution Contingency Plan. 40 CFR Part 300, *Fed Reg* 55:8666. March 8.

National Contingency Plan (NCP). (1990b). Preamble to the National Oil and Hazardous Substances Pollution Contingency Plan. *Fed Reg* 53:51394. March 8.

National Environmental Justice Advisory Committee (NEJAC). (2002). Fish consumption and environmental justice: A report on the public meeting convened by the National Environmental Justice Advisory Council, December 3-6, 2001. Seattle. (http://www.epa.gov/compliance/resources/publications/ej/fish_consump_report_1102.pdf)

National Marine Fisheries Service (NMFS). (1993). Data tapes for the 1993 NMFS provided to U.S. EPA National Center for Environmental Assessments. Cited in USEPA, 1997f.

National Research Council (NRC). (1983). Risk Assessment in the federal government: Managing the process. National Academies Press: Washington, DC.

National Research Council (NRC). (1993). Pesticides in the diets of infants and children. National Academies Press: Washington, DC.

National Research Council (NRC). (1994). Science and judgment in risk assessment. National Academies Press: Washington, DC.

National Research Council (NRC). (1996). Understanding risk: Informing decisions in a democratic society. Stern, P.C., and H.V. Fineberg, eds. National Academies Press: Washington, DC.

National Research Council (NRC). (2001). Arsenic in drinking water: 2001 update. National Academies Press: Washington, DC.

North Atlantic Treaty Organization (NATO). (1988a). International toxicity equivalency factor (I-TEF), method of risk assessment for complex mixtures of dioxins and related compounds. Report No. 176. NATO.

North Atlantic Treaty Organization (NATO). (1988b). Scientific basis for the development of international toxicity equivalency factor (I-TEF), method of risk assessment for complex mixtures of dioxins and related compounds. Report No. 178. NATO Committee on the Challenges of Modern Society.

Norton, S.B., D.J. Rodier, J.H. Gentile, W.H. van der Schalie, W.P. Wood, and M.W. Slimak. (1992). A framework for ecological risk assessment at the EPA. *Environ Toxicol Chem* 11:1663-1672.

Nosek, J.A., S.R. Craven, J.R. Sullivan, S.S. Hurley, and R.E. Peterson. (1992a). Toxicity and reproductive effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens. *J Toxicol Environ Health* 35(3):187-198.

Nosek, J.A., S.R. Craven, J.R. Sullivan, J.R. Olson, and R.E. Peterson. (1992b). Metabolism and disposition of 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens, chicks, and eggs. *J Toxicol Environ Health* 35(3):153-164.

Office of Science and Technology Policy (OSTP). (1985). Chemical carcinogens: A review of the science and its associated principles. OSTP, Executive Office of the President. *Fed Reg* 50:10371-10442. March 14.

Office of Science and Technology Policy (OSTP). (1995). Science, risk, and public policy. OSTP, Executive Office of the President, Washington, DC.

Owen, B.A. (1990). Literature-derived absorption coefficients for 39 chemicals via oral and inhalation routes of exposure. *Regul Toxicol Pharmacol* 11(3):237-252.

Pacala, S.W., E. Bulte, J.A. List, and S.A. Levin. (2003). False alarm over environmental false alarms. *Science* 301:1187-1188.

Ramesh, A., D.B. Hood, F. Inyang, M. Greenwood, A.M. Nyanda, A.E. Archibong, and M.E. Knuckles. (2002). Comparative metabolism, bioavailability, and toxicokinetics of benzo[a]pyrene in rats after acute oral, inhalation, and intravenous administration. *Polycyclic Aromatic Hydrocarbons* 22(3-4):969-980.

Restum, J., S. Bursian, J. Giesy, J. Render, W. Helferich, E. Shipp, D. Verbrugge, and R. Aulerich. (1998). Multigenerational study of the effects of consumption of PCB-contaminated carp from Saginaw Bay, Lake Huron, on mink. 1. Effects on mink reproduction, kit growth and survival, and selected biological parameters. *J Toxicol Environ Health* 54:343-375.

Rhomberg, L.R. (1997). A survey of methods for chemical health risk assessment among federal regulatory agencies. Supporting document for Framework for Environmental Health Risk Management Report, National Commission on Risk Assessment and Risk Management. (http://www.riskworld.com/nreports/1996/risk_rpt/)

Ross, P. (2000). Marine mammals as sentinels in ecological risk assessment. *Hum Ecol Risk Assess* 6:29-46.

Ross, P., R. De Swart, R. Addison, H. Van Loveren, J. Vos, and A. Osterhaus. (1996). Contaminant-induced immunotoxicity in harbour seals: Wildlife at risk? *Toxicology* 112:157-169.

Ruckelshaus, W.D. (1983). Science, risk, and public policy. *Science* 221:1026-1028.

Ruffle, B., D. Burmaster, P. Anderson, and D. Gordon. (1994). Lognormal distributions for fish consumption by the general U.S. population. *Risk Anal* 14(4):395-404.

Russell, M. and M. Gruber. (1987). Risk assessment in environmental policy-making. *Science* 236:286-290.

Safe Drinking Water Act Amendments (SDWAA). (1996). U.S. Public Law 104-182. 104th Cong. 6 August 1996. Sections 102(b)(1)(C) and 103(b)(3)(C)(i)(V).

Safe, S. (1990). Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: Environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). *Crit Rev Toxicol* 21(1):51-88.

Safe, S. (1994). Polychlorinated biphenyls (PCBs): Environmental impact, biochemical and toxic responses, and implications for risk assessment. *Crit Rev Toxicol* 24(2):87-149.

Sample, B.E., D.M. Opresko, and G.W. Suter, II. (1996). Toxicological benchmarks for wildlife. 1996 Revision. Risk Assessment Program, Health Science Research Division, Oak Ridge National Laboratory. Oak Ridge, TN. ES/ER/TM-86/R3.

Schoeny, R., P. Muller, and J. Mumford. (1998). Risk assessment for human health protection — Applications to environmental mixtures. In: Douben, P., ed. Pollution risk assessment and management. New York: John Wiley & Sons. pp. 205-234.

Sechena, R., S. Liao, R.M. Lorenzana, C. Nakano, N. Polissar, and R.A. Fenske. (2003). Asian American and Pacific Islander seafood consumption — A community-based study in King County, Washington. *J Expo Anal Environ Epidemiol* 13(4):256-66.

Sexton, K., J.L.Adgate, G. Ramachandran, G.C. Pratt, S.J. Mongin, T.H. Stock, and M.T. Morandi. (2004). Comparison of personal, indoor and outdoor exposures to hazardous air pollutants in three urban communities. *Environ Sci Technol* 38:423-430.

Shipp, E., J. Restum, J. Giesy, S. Bursian, R. Aulerich, and W. Helferich. (1998a). Multigenerational study of the effects of consumption of PCB-contaminated carp from Saginaw Bay, Lake Huron, on mink. 2. Liver PCB concentration and induction of hepatic cytochrome P-450 activity as a potential biomarker for PCB exposure. *J Toxicol Environ Health* 54:377-401.

Shipp, E., J. Restum, J. Giesy, S. Bursian, R. Aulerich, and W. Helferich. (1998b). Multigenerational study of the effects of consumption of PCB-contaminated carp from Saginaw Bay, Lake Huron, on mink. 3. Estrogen receptor and progesterone receptor concentrations, and potential correlation with dietary PCB consumption. *J Toxicol Environ Health* 54:403-420.

Simmons, J.E., S.D. Richardson, T.F. Speth, R.J. Miltner, G. Rice, K. Schenck, E.S. Hunter, III, and L.K. Teuschler. (2002). Development of a research strategy for integrated technology-based toxicological and chemical evaluation of complex mixtures of drinking water disinfection byproducts. *Environ Health Perspect* 110(6):1013-1024.

Stanek, E.J., and E.J. Calabrese. (1995a). Daily estimates of soil ingestion in children. *Environ Health Perspect* 103(3):276-285.

Stanek, E.J., and E.J. Calabrese. (1995b). Soil ingestion estimates for use in site evaluations based on the best tracer method. *Hum Ecol Risk Assess* 1:133-156.

Stanek, E.J., III, and E.J. Calabrese. (2000). Daily soil ingestion estimates for children at a Superfund site. *Risk Anal* 20:627-635.

Summer, C., J. Giesy, S. Bursian, J. Render, T. Kubiak, P. Jones, D. Verbrugge, R. Aulerich. (1996a). Effects induced by feeding organochlorine-contaminated carp from Saginaw Bay, Lake Huron, to laying white leghorn hens. I. Effects on health of adult hens, egg production, and fertility. *J Toxicol Environ Health* 49:389-407.

Summer, C., J. Giesy, S. Bursian, J. Render, T. Kubiak, P. Jones, D. Verbrugge, R. Aulerich. (1996b). Effects induced by feeding organochlorine-contaminated carp from Saginaw Bay, Lake Huron, to laying white leghorn hens. II. Embryotoxic and teratogenic effects. *J Toxicol Environ Health* 49:409-438.

Superfund Amendments and Reauthorization Act (SARA). (1986). 42 U.S.C. 9605.

Suter, G. (1993). Ecological risk assessment. Boca Raton, FL: Lewis Publishers, Inc.

Suter, G.W., II, R.A. Efroymson, B.E. Sample, and D.S. Jones. (2000). Ecological risk assessment for contaminated sites. Boca Raton, FL: CRC Press LLC.

Swartout, J.C., P.S. Price, M.L. Dourson, H.L. Carlson-Lynak, and R.E. Keenan. (1998). A probabilistic framework for the reference dose (probabilistic RfD). *Risk Anal* 18:271-282.

Tillitt, D., and P. Wright. (1997). Dioxin-like embryotoxicity of a Lake Michigan lake trout extract to developing lake trout. *Organohalogen Cmpds* 34:221-225.

Timchalk, C., M.D. Dryzga, F.A. Smith, and M. Bartels. (1991). Disposition and metabolism of [14C]1,2-dichloropropane following oral and inhalation exposure in Fischer 344 rats. *Toxicology* 68(3):291-306.

Toy, K., G. Gawne-Mittelstaedt, N. Polissar, and S. Liao. (1996). A fish consumption survey of the Tulalip and Squaxin Island tribes of the Puget Sound region. Tulalip Tribes, National Resources Department, Marysville, WA.

Train, R. (1976). Interim procedures and guidelines for health risk and economic impact assessments of suspected carcinogens. EPA Office of the Administrator, May 1976.

U.S. Environmental Protection Agency (USEPA). (1980). Water quality criteria documents: availability. *Fed Reg* 45: 79318-79379.

U.S. Environmental Protection Agency (USEPA). (1984). Risk assessment and management: Framework for decision making, EPA 600/9-85-002, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1985). Guidelines for deriving numerical National Water Quality Criteria for the protection of aquatic organisms and their uses. EPA 822/R-85-100. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1986a). Guidelines for carcinogen risk assessment. EPA/600/8-87/045. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1986b). Guidelines for the health risk assessment of chemical mixtures. EPA/600/8-87/045. Risk Assessment Forum, Washington DC.

U.S. Environmental Protection Agency (USEPA). (1986c). Superfund public health evaluation manual. EPA/540/1-86/060. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1987a). Interim procedures for estimating risks associated with exposures to mixtures of chlorinated dibenzo-p-dioxins and -dibenzofurans (CDDs and CDFs). EPA/625/3-87/012. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1987b). Unfinished business: A comparative assessment of environmental problems. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1988a). Technical support document on risk assessment of chemical mixtures. EPA/600/8-90/064. Risk Assessment Forum, Washington DC.

U.S. Environmental Protection Agency (USEPA). (1988b). Guidance for conducting remedial investigations and feasibility studies. OSWER Directive 9355.3-01. EPA/540/G-89/004. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1988c). Special review final decision relating to diazinon in the matter of Ciba-Geigy Corporation et al. FIFRA Docket Nos. 562 et seq. March 29.

U.S. Environmental Protection Agency (USEPA). (1989a). Risk assessment guidance for Superfund, Volume I: Human health evaluation manual (Part A), Interim Final. EPA/540/1-89/002. Office of Emergency and Remedial Response, Washington, DC. (http://www.epa.gov/superfund/programs/risk/ragsa)

U.S. Environmental Protection Agency (USEPA). (1989b). Interim procedures for estimating risks associated with exposures to mixtures of chlorinated dibenzo-p-dioxins and -dibenzofurans (CDDs and CDFs) and 1989 update. EPA/625/3-89/016. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1989c). Resolution on use of mathematical models by EPA for regulatory assessment and decision making. SAB-EEC-89-012. Science Advisory Board, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1989d). Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA 600/4-89/001. Cincinnati, OH.

U.S. Environmental Protection Agency (USEPA). (1990). Reducing risk: Setting priorities and strategies for environmental protection. EPA-SAB-EC-90-021. Science Advisory Board, Washington, DC. 26 pp.

U.S. Environmental Protection Agency (USEPA). (1991a). Alpha-2u-globulin: association with chemically induced renal toxicity and neoplasia in the male rat. EPA/625/3- 91/019F. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1991b). Workshop report on toxicity equivalency factors for polychlorinated biphenyl congeners. EPA/625/3-91/020. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1991c). Risk assessment guidance for Superfund, Volume I: Human health evaluation manual, Supplemental Guidance, "Standard default exposure factors." OSWER Directive 9285.6-03. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1991d). Methods for measuring the acute toxicity of effluents to freshwater and marine organisms, 4th ed. EPA-600/4-90/027. Cincinnati, OH.

U.S. Environmental Protection Agency (USEPA). (1992a). Guidelines for exposure assessment. EPA 600Z-92/001. Risk Assessment Forum, Washington, DC. 170 pp. (http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=559070)

U.S. Environmental Protection Agency (USEPA). (1992b). Supplemental guidance to RAGS: Calculating the concentration term. OSWER Directive 9285.7-081. Office of Solid Waste and Emergency Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1992c). Consumption surveys for fish and shellfish: A review and analysis of survey methods. EPA 822/R-92-001. Office of Water, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1992d). Framework for ecological risk assessment. EPA/630/R-92/001. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1992e). Comparative analysis of acute avian risk from granular pesticides. Office of Pesticide Programs, Washington, DC. 71 pp.

U.S. Environmental Protection Agency (USEPA). (1993a). Provisional guidance for quantitative risk assessment of polycyclic aromatic hydrocarbons. EPA/600/R-93/089. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1993b). A review of ecological assessment case studies from a risk assessment perspective. EPA/630/R-92/005. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1993c). Interim report on data and methods for assessment of 2,3,7,8-tetrachlorodibenzo-p-dioxin risks to aquatic life and associated wildlife. EPA/600/R-93/055. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1993d). Water quality guidance for the great lakes system, Proposed rule. *Fed Reg* 58:20802.ff April 16.

U.S. Environmental Protection Agency (USEPA). (1994a). Methods for Derivation of Inhalation reference concentrations and application of inhalation dosimetry. EPA/600/8-90/066F. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1994b). Report of the Agency Task Force on environmental regulatory modeling — Guidance, support needs, draft criteria and charter. EPA 500-R-94-001. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1994c). Model validation for predictive exposure assessments. Risk Assessment Forum, Washington, DC. (http://cfpub.epa.gov/crem/cremlib.cfm)

U.S. Environmental Protection Agency (USEPA). (1994d). Peer review program. Memorandum from Administrator Carol Browner to Assistant Administrators, General Counsel, Inspector General, Associate Administrators, Regional Administrators, and Staff Office Directors. Science Policy Council, Washington, DC. June 7. (http://www.epa.gov/osp/spc/memo0607.htm)

U.S. Environmental Protection Agency (USEPA). (1994e). Catalog of standard toxicity tests for ecological risk assessment. EPA 540-F-94-013. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1994f). Field studies for ecological risk assessment. EPA 540-F-94-014. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1994g). A review of ecological assessment case studies from a risk assessment perspective, Vol. II. EPA 630-R-94-003. Washington, D.C.

U.S. Environmental Protection Agency (USEPA). (1994h). Ecological risk assessment issue papers. EPA/630/R-94/009. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1994i). Managing ecological risks at EPA: Issues and recommendations for progress. EPA/600/R-94/183. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1995a). Policy for risk characterization. Science Policy Council, Washington, DC. (http://www.epa.gov/osp/spc/rcpolicy.htm)

U.S. Environmental Protection Agency (USEPA). (1995b). Policy on evaluating health risks to children. Science Policy Council, Washington, DC. (http://www.epa.gov/osp/spc/memohlth.htm)

U.S. Environmental Protection Agency (USEPA). (1995c). Water quality guidance for the Great Lakes system, Final Rule. *Fed Reg* 60(56): 15366-15425.

U.S. Environmental Protection Agency (USEPA). (1995d). Water quality guidance for the Great Lakes system: Supplementary information document (SID). EPA-820-B-95-001. Office of Water, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1996a). Soil screening guidance: User's guide. 2nd ed. OSWER Pub. 9355.4-23. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1996b). Recommendations of the technical review workgroup for lead for an approach to assessing risks associated with adult exposures to lead in soil. EPA-540-R-03-001. Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1996c). Risk assessment guidance reforms. Office of Emergency and Remedial Response, Washington, DC. (http://www.epa.gov/superfund/programs/reforms/reforms/3-5a.htm#des)

U.S. Environmental Protection Agency (USEPA). (1996d). PCBs: Cancer dose-response assessment and application to environmental mixtures. EPA/600/P-96/001F. National Center for Environmental Assessment, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1997a). Guidance on cumulative risk assessment. Part 1: Planning and scoping. Science Policy Council, Washington, DC. (http://www.epa.gpv/osp/spc/cumrisk2.htm)

U.S. Environmental Protection Agency (USEPA). (1997b). Guiding principles for Monte Carlo analysis. EPA/630/R-97/001. Risk Assessment Forum, Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1997c). Policy for use of probabilistic analysis in risk assessment at the U.S. Environmental Protection Agency. Fred Hansen, Deputy Administrator. Science Policy Council, Washington, DC. (http://www.epa.gov/osp/spc/probpol.htm)

U.S. Environmental Protection Agency (USEPA). (1997d). Exposure factors handbook. EPA/600/P-95/002F(a,b,c). Office of Research and Development, Washington, DC. (http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=12464)

U.S. Environmental Protection Agency (USEPA). (1997e). Aerial spray drift review. FIFRA Scientific Advisory Panel, Washington, DC. December 10-11. (http://www.epa.gov/scipoly/sap/1997/december/spraydrift.htm)

U.S. Environmental Protection Agency (USEPA). (1997f). Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments - Interim Final. Office of Solid Waste and Emergency Response, OSWER 9285.7-25, EPA 540-R-97-006, PB97-963211, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1998a). Guidelines for ecological risk assessment. EPA/630/R-95/002F. Risk Assessment Forum, Washington, DC. 171 pp. (http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=12460)

U.S. Environmental Protection Agency (USEPA). (1998b). Beryllium and compounds. Integrated Risk Information System, Office of Research and Development, Washington, DC. (http://www.epa.gov/iris/subst/0012.htm)

U.S. Environmental Protection Agency (USEPA). (1998c). Assessment of thyroid follicular cell tumors. EPA/630/R-97/002. Science Policy Council, Washington, DC. (http://www.epa.gov/osp/spc/2thyfoll.htm)

U.S. Environmental Protection Agency (USEPA). (1998d). Proposed methods for basin-scale estimation of pesticide concentrations in flowing water and reservoirs for tolerance reassessment, Chapter 4: An index reservoir for use in assessing drinking water exposure. FIFRA Scientific Advisory Panel, Washington, DC. July. (http://www.epa.gov/oscpmont/sap/1998/index.htm)

U.S. Environmental Protection Agency (USEPA). (1998e). A comparative analysis of ecological risks from pesticides and their uses: Background, methodology & case study. Office of Pesticide Programs, Washington, DC. (http://www.epa.gov/scipoly/sap/1998/index.htm#december8)

U.S. Environmental Protection Agency (USEPA). (1999a). Guide to preparing Superfund proposed plans, records of decision, and other remedy selection decision documents. EPA 540/R-98/031. Office of Solid Waste and Emergency Response, Washington, DC. (http://www.epa.gov/superfund/resources/remedy/rods)

U.S. Environmental Protection Agency (USEPA). (1999b). Guidelines for carcinogen risk assessment, Revised draft. NCEA-F-0644. Office of Research and Development, Washington, DC. (http://cfpub.epa.gov/ncea/raf/cancer_gls.pdf)

U.S. Environmental Protection Agency (USEPA). (1999c). White paper on the nature and scope of issues on adoption of model use acceptability guidance. Science Policy Council., Washington, DC. (http://cfpub.epa.gov/crem/cremlib.cfm)

U.S. Environmental Protection Agency (USEPA). (1999d). Proposed methods for determining watershed-derived percent crop areas and considerations for applying crop area adjustments to surface water screening models. FIFRA Scientific Advisory Panel, Washington, DC. May 25-27. (http://www.epa.gov/oscpmont/sap/1999/index.htm)

U.S. Environmental Protection Agency (USEPA). (1999e). Two documents: 1) Downwind deposition tolerance bounds for ground hydraulic boom sprayers; 2) Downwind deposition tolerance bounds for orchards. FIFRA Scientific Advisory Panel, Washington, DC. July 20-23. (http://www.epa.gov/scipoly/sap/1999/index.htm)

U.S. Environmental Protection Agency (USEPA). (1999f). Risk assessment guidance for Superfund, Volume 1, Human health evaluation manual (Supplement to Part A): Community involvement in Superfund risk assessments. EPA 540-R-98-042. Office of Solid Waste and Emergency Response, Washington, DC. March.

U.S. Environmental Protection Agency (USEPA). (1999g). Short Sheet: IEUBK Model Bioavailability Variable. EPA 540-F-00-006. Office of Solid Waste and Emergency Response, Washington, DC. October.

U.S. Environmental Protection Agency (USEPA). (1999h). Superfund risk assessment and how you can help (videotape). EPA-540-V-99-002. Office of Solid Waste and Emergency Response, Washington, DC. September.

U.S. Environmental Protection Agency (USEPA). (1999i). Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish, 2nd edition. EPA 841-B-99-002. Office of Water, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (1999j). Issuance of final guidance: Ecological risk assessment and risk management principles for Superfund Sites. OSWER Directive 9285.7-28 P. Office of Solid Waste and Emergency Response, Washington, DC. October 7.

U.S. Environmental Protection Agency (USEPA). (2000a). Science Policy Council Handbook: Risk Characterization Handbook. EPA 100-B00-002. Science Policy Council, Washington, DC. December. (http://www.epa.gov/osp/spc/rchandbk.pdf)

U.S. Environmental Protection Agency (USEPA). (2000b). Science Policy Council Handbook: Peer Review Handbook, 2nd edition. EPA 100-B00-001. Science Policy Council, Washington, DC. December. (http://www.epa.gov/osp/spc/prhandbk.pdf)

U.S. Environmental Protection Agency (USEPA). (2000c). Policy and program requirements for the mandatory agency-wide quality system. EPA Order 5360.1 A2. Office of Environmental Information, Washington, DC. (http://www.epa.gov/quality/qs-docs/5360-1.pdf)

U.S. Environmental Protection Agency (USEPA). (2000d). EPA quality manual for environmental programs. EPA Order 5360 A1. Office of Environmental Information, Washington, DC. (http://www.epa.gov/quality/qs-docs/5360.pdf)

U.S. Environmental Protection Agency (USEPA). (2000e). An SAB advisory on the USEPA's draft case study analysis of the residual risk of secondary lead smelters, EPA Science Advisory Board, May 2000, pp. 20-21. (http://www.epa.gov/sab/pdf/ecadv05.pdf)

U.S. Environmental Protection Agency (USEPA). (2000f). Estimated per capita fish consumption in the United States. EPA-821-R-00-025.

U.S. Environmental Protection Agency (USEPA). (2000g). Benchmark dose technical guidance document. EPA/630/R-00/001. Risk Assessment Forum, Washington, DC. (http://cfpub1.epa.gov/ncea/cfm/recordisplay.cfm?deid=22506)

U.S. Environmental Protection Agency (USEPA). (2000h). Supplementary guidance for conducting health risk assessment of chemical mixtures. EPA/630/R-00/002. Risk Assessment Forum, Washington DC.

U.S. Environmental Protection Agency (USEPA). (2000i). Exposure and human health reassessment of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and related compounds (External Review Draft). EPA/600/P-00/001B(a-f). National Center for Environmental Assessment, Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2000j). Session VI — Progress report on estimating pesticide concentrations in drinking water and assessing water treatment effects on pesticide removal and transformation: A consultation. FIFRA Scientific Advisory Panel, Washington, DC. September, 26-29. (http://www.epa.gov/oscpmont/sap/2000/index.htm)

U.S. Environmental Protection Agency (USEPA). (2000k). Methodology for deriving ambient water quality criteria for the protection of human health. EPA-822-B-00-004. Office of Water, Office of Science and Technology, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2000l). Data quality objectives process for hazardous waste site investigations — Data quality objectives guidance. EPA/600/R-00/007. Office of Environmental Information, Washington, DC. January.

U.S. Environmental Protection Agency (USEPA). (2000m). Phase 2 report: Further site characterization and analysis, Volume 2F. Revised human health risk assessment Hudson River PCBs reassessment RI/FS. Region 2, New York, NY. November. (http://www.epa.gov/hudson/revisedhhra-text.pdf)

U.S. Environmental Protection Agency (USEPA). (2000n). Superfund community involvement guidance. EPA 540-K-01-003. Office of Solid Waste and Emergency Response, Washington, DC. April.

U.S. Environmental Protection Agency (USEPA). (2001a). Risk assessment guidance for Superfund: Volume III - Part A, Process for conducting probabilistic risk assessment. EPA 540-R-02-002. Office of Emergency and Remedial Response, Washington, DC. December. (http://www.epa.gov/superfund/programs/risk/rags3a/index.htm)

U.S. Environmental Protection Agency (USEPA). (2001b). Risk assessment guidance for Superfund, Volume I, Human health evaluation manual (Part D, Standardized planning, reporting and review of Superfund risk assessments) Final. Office of Solid Waste and Emergency Response, Washington, DC. December. (http://www.epa.gov/superfund/programs/risk/ragsd)

U.S. Environmental Protection Agency (USEPA). (2001c). Chloroform. Integrated Risk Information System, Office of Research and Development, Washington, DC. (http://www.epa.gov/iris/subst/0025.htm)

U.S. Environmental Protection Agency (USEPA). (2001d). Integrated exposure uptake biokinetic model for lead in children (IEUBKwin v1.0). Washington, DC. (http://www.epa.gov/superfund/programs/lead/ieubk.htm)

U.S. Environmental Protection Agency (USEPA). (2001e). Workshop report on the application of 2,3,7,8-TCDD toxicity equivalence factors to fish and wildlife. EPA/630/R-01/002. Risk Assessment Forum, Washington, DC. (http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=23763)

U.S. Environmental Protection Agency (USEPA). (2001f). EPA requirements for quality assurance project plans, EPA QA/R-5. EPA/240/B-01/003. Office of Environmental Information, Washington, DC. (http://www.epa.gov/QUALITY/qs-docs/r5-final.pdf)

U.S. Environmental Protection Agency (USEPA). (2001g). Risk assessment guidance for Superfund, Volume I, Human health evaluation manual (Part E): Supplemental guidance for dermal risk assessment, Interim. PB99-963312. Office of Solid Waste and Emergency Response, Washington, DC. May.

U.S. Environmental Protection Agency (USEPA). (2001h). Record of decision: Alcoa (Point Comfort)/Lavaca Bay Site Point Comfort, TX. CERCLIS #TXD008123168. Region VI, Dallas, TX. December.

U.S. Environmental Protection Agency (USEPA). (2001i). Comprehensive five-year review guidance. EPA 40-R-01-007. Office of Emergency and Remedial Response, Washington, DC. June.

U.S. Environmental Protection Agency (USEPA). (2001j). National-scale air toxics assessment for 1996. Science Advisory Board Preliminary Draft. EPA-453/R-01-003, Office of Air Quality Standards and Planning, Research Triangle Park, NC. (http://www.epa.gov/ttn/atw/nata/natsaov.html)

U.S. Environmental Protection Agency (USEPA). (2002a). Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity, of information disseminated by the Environmental Protection Agency. EPA/260R-02-008. Office of Environmental Information, Washington, DC. October. (http://www.epa.gov/oei/qualityguidelines)

U.S. Environmental Protection Agency (USEPA). (2002b). Lessons learned on planning and scoping for environmental risk assessments. Science Policy Council. January. (http://www.epa.gov/osp/spc/2cumrisk.htm)

U. S. Environmental Protection Agency (USEPA). (2002c). A review of the reference dose and reference concentration processes. EPA/630/P-02/002F. Risk Assessment Forum, Washington, DC. December. 192 pp. (http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55365)

U.S. Environmental Protection Agency (USEPA). (2002d). Child-specific exposure factors handbook (Interim report). EPA-600-P-00-002B. Office of Research and Development, Washington, DC. September.

U.S. Environmental Protection Agency (USEPA). (2002e). 1,3-Butadiene Integrated Risk Information System, Office of Research and Development, Washington, DC. (http://www.epa.gov/iris/subst/0139.htm)

U.S. Environmental Protection Agency (USEPA). (2002f). Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity. Office of Pesticide Programs, Washington, DC. (http://www.epa.gov/oppfead1/trac/science/cumulative_guidance.pdf)

U.S. Environmental Protection Agency (USEPA). (2002g). Organophosphate pesticides: Revised OP cumulative risk assessment. Office of Pesticide Programs, Washington, DC. (http://www.epa.gov/pesticides/cumulative/rra-op)

U.S. Environmental Protection Agency (USEPA). (2002h). Total risk integrated methodology: TRIM.FaTE technical support document, Volume I: Description of module. EPA-453/R-02-011a. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. Environmental Protection Agency (USEPA). (2002i). Total risk integrated methodology, TRIM.FaTE technical support document, Volume II: Description of chemical transport and transformation algorithms. EPA-453/R-02-011b. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. Environmental Protection Agency (USEPA). (2002j). Evaluation of TRIM.FaTE, Volume I: Region III, Philadelphia, PA. Approach and initial findings. EPA-453/R-02-012. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. Environmental Protection Agency (USEPA). (2002k). Supplemental guidance for developing soil screening levels for Superfund sites. OSWER 9355.4-24. Office of Solid Waste and Emergency Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2002l). Calculating upper confidence limits for exposure point concentrations at hazardous waste sites. OSWER 9285.6-10. Office of Emergency and Remedial Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2002m). Record of decision: Palmerton Zinc Pile, Palmerton, PA. EPD ID: PAD002395887.OU 03. EPA/ROD/R03-02/007.

U.S. Environmental Protection Agency (USEPA). (2002n). Peer consultation workshop on approaches to polycyclic aromatic hydrocarbon (PAH) health assessment. EPA/635/R-02/005. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003a). A summary of general assessment factors for evaluating the quality of scientific and technical information. EPA 100/B-03/001. Science Policy Council, Washington, DC. June. (http://www.epa.gov/oei/qualityguidelines/af_home.htm)

U.S. Environmental Protection Agency (USEPA). (2003b). Framework for cumulative risk assessment. EPA/630/P-02/001F. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003c). Draft final guidelines for carcinogen risk assessment. EPA/630/P-03/001A. NCEA-F-0644A. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003d). Supplemental guidance for assessing cancer susceptibility from early-life exposure to carcinogens (external review draft). EPA/630/R-03/003. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003e). Developing relative potency factors for pesticide mixtures: Biostatistical analyses of joint dose-response. EPA/600/R-03/052. Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH.

U.S. Environmental Protection Agency (USEPA). (2003f). Human health risk assessment for the Housatonic River — Rest of the river. Region I, Boston, MA. June. (http://www.epa.gov/NE/ge/thesite/restofriver-reports.html)

U.S. Environmental Protection Agency (USEPA). (2003g). Integrated risk information system (IRIS). Office of Research and Development, National Center for Environmental Assessment, Washington, DC. (http://www.epa.gov/iris)

U.S. Environmental Protection Agency (USEPA). (2003h). The feasibility of performing cumulative risk assessments for mixtures of disinfection by-products in drinking water. EPA/600/R-03/051. Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH.

U.S. Environmental Protection Agency (USEPA). (2003i). National emission standards for hazardous air pollutants for stationary combustion turbines; Proposed rule. 40 CFR Part 63. *Fed Reg* 68(9) 1888-1929. January 3.

U.S. Environmental Protection Agency (USEPA). (2003j). Generic endpoints for ecological risk assessment. EPA/630/P-02/004A. Risk Assessment Forum, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003k). Guidance for developing ecological soil screening levels (Eco-SSLs). OSWER Directive 92857-55. Office of Solid Waste and Emergency Response, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2003l). Potential risks of nine rodenticides to birds and nontarget mammals: A comparative approach. January. (http://www.epa.gov/pesticides/rodenticidecluster/)

U.S. Environmental Protection Agency (USEPA). (2003m). Human health research strategy. EPA 600/R-02/050. Office of Research and Development, Washington, DC. July.

U.S. General Accounting Office (USGAO). (2001). Chemical risk assessment: Selected federal agencies' procedures, assumptions, and policies. GAO-01-810. Washington, DC. August. (http://www.gao.gov/new.items/d01810.pdf)

U.S. Office of Management and Budget (USOMB). (2002). Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. *Fed Reg* 67:8452-8460. February 22. (http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf)

U.S. Office of Management and Budget (USOMB). (2003a). Draft 2003 report to congress on the costs and benefits of federal regulations. *Fed Reg* 68:5492-5527. February 3. (http://www.whitehouse.gov/omb/fedreg/2003draft_cost-benefit_rpt.pdf)

U.S. Office of Management and Budget (USOMB). (2003b). Informing regulatory decisions: 2003 report to congress on the costs and benefits of federal regulations and unfunded mandates on state, local, and tribal entities. Office of Information and Regulatory Affairs, Washington, DC. (http://www.whitehouse.gov/omb/inforeg/2003_cost-ben_final_rpt.pdf)

U.S. Office of Management and Budget (USOMB). (2003c). Informing regulatory decisions: 2003 report to congress on the costs and benefits of federal regulations and unfunded mandates on state, local, and tribal entities, Appendix D: OMB Circular A-4. Regulatory Analysis. Office of Information and Regulatory Affairs, Washington, DC.

van Birgelen, A.P., J. van der Kolk, K.M. Fase, I. Bol, H. Poiger, A. Brouwer, and M. van den Berg. (1994a). Toxic potency of 3,3',4,4',5-pentachlorobiphenyl relative to and in combination with 2,3,7,8-tetrachlorodibenzo-p-dioxin in a subchronic feeding study in the rat. *Toxicol Appl Pharmacol,* 127(2):209-221.

van Birgelen, A.P., J. van der Kolk, K.M. Fase, I. Bol, H. Poiger, A. Brouwer, and M. van den Berg. (1994b). Toxic potency of 2,3,3',4,4',5-hexachlorobiphenyl relative to and in combination with 2,3,7,8-tetrachlorodibenzo-p-dioxin in a subchronic feeding study in the rat. *Toxicol Appl Pharmacol* 126(2):202-213.

van den Berg, M., L. Birnbaum, A.T.C. Bosveld, B. Brunstrom, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, J.C. Larsen, F. X. van Leeuwen, A. K. Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tysklind, M. Younes, F. Waern, and T. Zacharewski. (1998). Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environ Health Perspect* 106 (12):775-792.

van der Plas, S.A., M. Haag-Grönlund, G. Scheu, L. Wärngard, M. van der Berg, P. Wester, J.H. Koeman, and A. Brouwer. (1999). Induction of altered hepatic foci by a mixture of dioxin-like compounds with and without 2,2',4,4',5,5'-hexachlorobiphenyl in female Sprague-Dawley rats. *Toxicol Appl Pharmacol* 156(1):30-39.

van der Plas, S.A., J. Sundberg, H. van den Berg, G. Scheu, P. Wester, S. Jensen, A. Bergman, J. de Boer, J.H. Koeman, A. Brouwer. (2000). Contribution of planar (0-1 ortho) and nonplanar (2-4 ortho) fractions of Aroclor 1260 to the induction of altered hepatic foci in female Sprague-Dawley rats. *Toxicol Appl Pharmacol* 169(3):255-68.

van Wijnen, J.H., P. Clausing, and B. Brunekreff. (1990). Estimated soil ingestion by children. *Environ Res* 51:147-162.

Walker, M.K., P.M. Cook, B.C. Butterworth, E.W. Zabel, and R.E. Peterson. (1996). Potency of a complex mixture of polychlorinated dibenzo-*p*-dioxin, dibenzofuran, and biphenyl congeners compared to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in causing fish early life stage mortality. *Fundam Appl Toxicol* 30:178-186.

Wang, P.Y., T. Kaneko, H. Tsukada, M. Nakano, and A. Sato. (1997). Dose- and route-dependent alterations in metabolism and toxicity of chemical compounds in ethanol-treated rats: Difference between highly (chloroform) and poorly (carbon tetrachloride) metabolized hepatotoxic compounds. *Toxicol Appl Pharmacol* 142 (1):13-21.

West, P.C., M.J. Fly, R. Marans, and F. Larkin. (1989). Michigan sport anglers fish consumption survey. A report to the Michigan Toxic Substance Control Commission. Contract No. 87-20141. Michigan Department of Management and Budget, Lansing, MI.

West, P.C., M.J. Fly, R. Marans, F. Larkin, and D. Roesnblatt. (1993). 1991-92 Michigan sport anglers fish consumption study. Technical Report No. 6. Michigan Department of Natural Resources, Lansing, MI. May.

Woutersen, R.A., H.P. Til, and V.J. Feron. (1984). Sub-acute versus sub-chronic oral toxicity study in rats: comparative study of 82 compounds. *J Appl Toxicol* 4:277-280.

Yrjänheiki, E.J. (1992). Review of the models for TEFs in assessing health risks of PCDDs and PCDFs. *Toxic Sub J* 12:283-288.

Zabel, E.W., P.M. Cook, and R.E. Peterson. (1995). Toxic equivalency factors of polychlorinated dibenzo-*p*-dioxins, dibenzofuran and biphenyl congeners based on early life stage mortality in rainbow trout (Oncorhynchus mykiss). *Aquat Toxicol* 31:315-328.

# REFERENCES OF EPA RISK ASSESSMENT GUIDELINES

*Guidelines for Carcinogen Risk Assessment*. Federal Register 51: 33992-34003, 24 September 1986; also EPA Publication No. EPA/600/8-87/045, August 1987.

*Proposed Guidelines for Carcinogen Risk Assessment; Notice*. Federal Register 61: 17960-18011, 23 April 1996.

*Draft Revised Guidelines for Carcinogen Risk Assessment*. EPA Publication No. NCEA-F-0644, July 1999; *Notice* in Federal Register 66: 59593-59594, 29 November 2001.

*Draft Final Guidelines for Carcinogen Risk Assessment (External Review Draft, February 2003)*. EPA Publication No. NCEA-F-0644A, March 2003; *Notice* in Federal Register 68: 10012-10015, 3 March 2003.

*Supplemental Guidance for Assessing Cancer Susceptibility from Early-life Exposure to Carcinogens (External Review Draft)*. EPA Publication No. EPA/630/R-03/003, February 2003; *Notice* in Federal Register 68: 10012-10015, 3 March 2003.

*Guidelines for Mutagenicity Risk Assessment*. Federal Register 51: 34006-34012, 24 September 1986; also EPA Publication No. EPA/600/8-87/045, August 1987.

*Guidelines for the Health Risk Assessment of Chemical Mixtures*. Federal Register 51: 34014-34025, 24 September 1986; also EPA Publication No. EPA/600/8-87/045, August 1987.

*Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures*. EPA Publication No. EPA/630/R-00/002, August 2000.

*Guidelines for Developmental Toxicity Risk Assessment*. Federal Register 56: 63798-63826, 5 December 1991.

*Guidelines for Exposure Assessment*. Federal Register 57: 22888-22938, 29 May 1992; also EPA Publication No. EPA/600/Z-92/001.

*Guidelines for Reproductive Toxicity Risk Assessment; Notice*. Federal Register 61: 56274-56322, 31 October 1996.

*Assessment of Thyroid Follicular Cell Tumors*. EPA Publication No. EPA/630/R-97/002, March 1998.

*Guidelines for Ecological Risk Assessment*. Federal Register 63: 26846-26924, 14 May 1998; also EPA Publication No. EPA/630/R-95/002F, April 1998.

*Guidelines for Neurotoxicity Risk Assessment; Notice*. Federal Register 60: 26926-26954, 14 May 1998.

*Guiding Principles for Monte Carlo Analysis* (contains: Policy for Use of Probabilistic Analysis in Risk Assessment at the U.S. Environmental Protection Agency). EPA Publication No. EPA/630/R-97/001, March 1997.

# ADDITIONAL USEFUL WEB SITES

EPA Risk Assessment Guidelines Web site:
http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=55907

EPA Quality System Web site: http://www.epa.gov/quality

EPA Science Policy Council Web site: http://www.epa.gov/osp/spc

EPA Information Quality Guidelines Web site: http://www.epa.gov/oei/qualityguidelines