Risk Assessment Guidance for Superfund:

Volume III - Part A, Process for Conducting Probabilistic Risk Assessment

EPA 540-R-02-002
OSWER 9285.7-45
PB2002 963302
www.epa.gov/superfund/RAGS3A/index.htm
December 2001

Superfund

Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment

Office of Emergency and Remedial Response
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document provides guidance to EPA Regions concerning how the Agency intends to exercise its discretion in implementing one aspect of the CERCLA remedy selection process. The guidance is designed to implement national policy on these issues.

Some of the statutory provisions described in this document contain legally binding requirements. However, this document does not substitute for those provisions or regulations, nor is it a regulation itself. Thus, it cannot impose legally-binding requirements on EPA, States, or the regulated community, and may not apply to a particular situation based upon the circumstances. Any decisions regarding a particular remedy selection decision will be made based on the statute and regulations, and EPA decision makers retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate.

Interested parties are free to raise questions and objection about the substance of this guidance and the appropriateness of the application of this guidance to a particular situation, and the Agency welcomes public input on this document at any time. EPA may change this guidance in the future.

ABOUT THE REVISION

WHAT IT IS

EPA's *Process for Conducting Probabilistic Risk Assessment* is an update of the 1989 *Risk Assessment Guidance for Superfund (RAGS)*. It is Volume III, an update to the existing two-volume set of RAGS. Volume III: Part A provides policy and guidance on conducting probabilistic risk assessment for both human and ecological receptors.

WHO IT'S FOR

RAGS Volume III: Part A is written primarily for risk assessors. Risk assessment reviewers, remedial project managers, and risk managers involved in Superfund site cleanup activities will also benefit from this addition to RAGS.

WHAT'S NEW

RAGS Volume III: Part A provides guidance on applying probabilistic analysis to both human health and ecological risk assessment. New information and techniques are presented that reflect the views of EPA Superfund program. A tiered approach is described for determining the extent and scope of the modeling effort that is consistent with the risk assessment objectives, the data available, and the information that may be used to support remedial action decisions at Superfund hazardous waste sites.

RAGS Volume III: Part A contains the following information:

- For the risk assessor— updated policies and guidance; discussion and examples of Monte Carlo modeling techniques for estimating exposure and risk.

- For the risk manager and the remedial project manager—an introduction to PRA, a chapter on communicating methods and results of PRA with the public, and a chapter on the role of PRA in decision making.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

1-D MCA	One-dimensional Monte Carlo analysis
2-D MCA	Two-dimensional Monte Carlo analysis
95% UCL	95% upper confidence limit
AM	Arithmetic mean
ARARs	Applicable or relevant and appropriate requirements
AT	Averaging time
AWQC	Ambient water quality criterion
BCa	Bias correction acceleration method
BMD	Benchmark dose
BMDS	Benchmark dose software
BMR	Benchmark Response
BTAG	Biological Technical Assistance Group
BW	Body weight
C	Concentration
CAG	Community advisory group
CDF	Cumulative distribution function
CI	Confidence interval
CIC	Community involvement coordinator
CIP	Community involvement plan
CLT	Central limit theorem
COC	Chemical of concern
CQR	Continuous quadratic regression
CSF	Cancer slope factor
CTE	Central tendency exposure
CV	Coefficient of variation
DI	Daily intake
DQO	Data quality objectives
EC_0	Exposure concentration that produces zero effect
EC_{20}	Concentration that causes a 20% effect
ECDF	Empirical cumulative distribution function
ED	Exposure duration
ED_{10}	Dose that causes a 10% effect
EDF	Empirical distribution function
EF	Exposure frequency
EPA	U.S. Environmental Protection Agency
EPC	Exposure point concentration
ERA	Ecological risk assessment
ERAF	Risk Assessment Forum
ERAGS	Ecological Risk Assessment Guidance for Superfund
EU	Exposure unit
EVIU	Expected value of including uncertainty
EVOI	Expected value of information
EVPI	Expected value of perfect information
EVSI	Expected value of sample information
GIS	Geographical Information Systems
GM	Geometric mean
GoF	Goodness-of-Fit
GSD	Geometric standard deviation
HEAST	Health effects assessment summary table
HHEM	Human Health Evaluation Manual
HI	Hazard Index

HQ	Hazard Quotient
IR	Iterative reduction
Irsd	Soil and dust ingestion rate
IRIS	Integrated Risk Information System
LADD	Life-time average daily intake
LCL	Lower confidence limit
LED_{10}	Lowest effect dose - lower confidence bound for dose that causes a 10% effect
LHS	Latin hypercube sampling
LOAEL	Lowest-observed-adverse-effect level
LOD	Limit of detection
LOEC	Lowest-observed-effect-concentration
MCA	Monte Carlo analysis
MCL	Maximum contaminant levels
MDC	Maximum detected concentration
MEE	Microexposure Event Analysis
MLE	Maximum Likelihood Estimation
MoMM	Method of Matching Moments
NCP	National Oil and Hazardous Substances Pollution Contingency Plan
NOAEL	No-observed-adverse-effect level
NOEC	No-observed-effect-concentration
OLS	Ordinary least squares
PBPK	Physiologically-based pharmacokinetic
PCBs	Polychlorinated biphenyls
pCi/g	Picocuries/gram
PDF	Probability density function
PDFu	Probability distribution for variability
PDFv	Probability distribution for uncertainty
PMF	Probability mass function
PPT	Parts per trillion
PRA	Probabilistic risk assessment
PRG	Preliminary remediation goal
PRP	Potentially responsible party
QAPP	Quality Assurance Project Plan
RAGS	Risk Assessment guidance for Superfund
RAL	Remedial action level
RBC	Risk based concentration
RCRA	Resource Conservation and Recovery Act
RfC	Reference concentration
RfD	Reference dose
RG	Remediation goal
RI/FS	Remedial Investigation/Feasibility Study
RME	Reasonable maximum exposure
RMSE	Root mean squared error
ROD	Record of decision
ROS	Rank order statistic
RPSS	Relative partial sum of squares
RPM	Remedial project manager
RSS	Regression sum of squares
SCM	Site conceptual model
SD	Standard deviation

SE	Standard error
SMDP	Scientific/Management Decision Point
SOW	Statement of Work
SR	Sensitivity ratio
SSD	Species sensitivity distribution
SSE	Sum of squares due to error
SSR	Sum of squares due to regression
SST	Sum of squares for total (regression plus error)
TAB	Technical Assistance to Brownfields Community
TAG	Technical assistance grant
TOSC	Technical outreach services for communities
TRV	Toxicity reference value
TSS	Total sum of squares
UCL	Upper confidence limit
VOI	Value of information

AUTHORS, CONTRIBUTORS, AND REVIEWERS

This manual was developed by EPA's Office of Emergency and Remedial Response. A number of individuals have reviewed and/or have been contributing authors of this document. Members of the EPA RAGS Volume III Workgroup, which was responsible for developing this document, included the following EPA headquarters and regional office staff.

RAGS VOLUME III WORKGROUP PARTICIPANTS

EPA HEADQUARTERS

Office of Emergency and Remedial Response	David A. Bennett S. Steven Chang David E. Cooper Janine Dinan Elizabeth Lee Hofmann
Office of Policy Economics and Innovation	Timothy M. Barry

EPA REGIONAL OFFICES

Region 1	Ann-Marie Burke	Region 5	Amy Mucha James Chapman
Region 2	Audrey Galizia Marian Olsen	Region 6	Maria L. Martinez
Region 3	Nancy Rios Jafolla	Region 8	Susan Griffin Gerry Henningsen Dale Hoff
Region 4	Ted W. Simon Sharon R. Thoms	Region 10	Joe Goulet

Technical assistance and production support was provided to EPA in the development of this guidance under Contract Numbers GS-10F-0137K and GS-35F-0555K.

An earlier draft of this document was peer reviewed by a panel of experts at a peer-review workshop held in November 2000. In addition, individuals in EPA and from the public provided valuable comments on earlier drafts of this guidance during the peer review process.

PREFACE

Risk Assessment Guidance for Superfund (RAGS) Volume III: Part A (hereafter referred to as RAGS Volume 3: Part A) provides technical guidance on the application of probabilistic risk assessment (PRA) methods to human health and ecological risk assessment in the U.S. Environmental Protection Agency (EPA) Superfund program. *RAGS Volume 3: Part A* supplements existing human health and ecological assessment guidance provided in the RAGS series. This guidance focuses on Monte Carlo analysis (MCA) as a method of quantifying variability and uncertainty in risk. Primarily geared toward the risk assessor, it is intended, both in content and format, to be most accessible to those readers who are familiar with risk assessment and basic statistical concepts. Chapters 1, 2, 6, and 7 are also directed towards risk managers. The term risk manager is used in this guidance to refer to individuals or entities that serve as the decision makers at hazardous waste sites. The term is used to emphasize the separation between risk assessment and risk management activities. Risk managers may include individual remedial project mangers (RPMs), site partnering teams, senior EPA managers (Section Chiefs, Branch Chiefs, or Division Directors), or other decision makers.

An attempt has been made in this document to define all relevant technical terms using plain language and to illustrate concepts with examples. An exhibit at the beginning of each chapter provides definitions of terms used in that chapter. In addition, a comprehensive definition of terms is provided in Appendix E. Other useful information has been presented in exhibits placed throughout each chapter. Bullets are used throughout the text to emphasize important concepts and policy statements related to the use of PRA. References are listed at the end of each chapter.

RAGS Volume 3: Part A was developed by the Superfund Probabilistic Risk Assessment Workgroup and the Ecological Risk Assessment Forum (ERAF); both are intra-Agency workgroups that have focused on improving the Risk Assessment Guidance for Superfund and implementing Superfund Reform activities. The guidance has undergone extensive review by Superfund and other programs within the Agency. In February 2000, a draft of the guidance was announced in the *Federal Register* to provide an opportunity for public comment (U.S. EPA, 2000a). In August 2000, a notice of peer review was announced in the *Federal Register* (U.S. EPA, 2000b), and in November 2000, *RAGS Volume 3: Part A* received a formal peer review from panelists outside the Agency.

The Agency may incorporate PRA under fund-lead and Potentially Responsible Party (PRP)-lead risk assessments. Implementation of successful PRAs requires careful planning. EPA strongly recommends that PRPs involve the Agency in all decisions regarding the planning, submittal, and technical details of any PRA. Coordinating with EPA early in the process will help ensure that PRAs conform to the recommended guidelines as part of the Superfund risk assessment process for protecting human and ecological health. PRPs should submit workplans for Agency review before initiating any PRA. Similarly, when EPA chooses to use PRA for an EPA-lead risk assessment, a PRA workplan will assist in directing site investigation and risk assessment activities, whether conducted by EPA or an EPA contractor. A workplan specifies contractor activities in the risk assessment and provides risk assessors and risk managers with an opportunity to obtain internal feedback from knowledgeable EPA staff, prior to initiating work on the assessment.

A tiered approach to PRA is advocated, which begins with a point estimate risk assessment. Important considerations include the time required to perform the PRA, the additional resources involved in developing the PRA, the quality and extent of data on exposure that will be used in the assessment, and

the value added by conducting the PRA. Project scoping is an essential component of all risk assessments and is especially important in PRA.

Implementation of a PRA usually requires special computer software that may be commercially available or that may need to be custom-designed for a specific application. Although commercial software packages are noted in this guidance, any mention or use of a particular product in *RAGS Volume 3: Part A* does not constitute an endorsement of that product by the Agency.

1.0 WHAT IS THE PURPOSE OF RAGS VOLUME 3 PART A?

RAGS Volume 3: Part A addresses the technical and policy issues associated with the use of PRA in EPA Superfund program. This guidance builds upon basic concepts of risk assessment outlined in *RAGS Volume I* (U.S. EPA, 1989a; 2001), recent guidance for ecological risk assessment (U.S. EPA, 1992, 1994, 1997a, 1998a; 1999), and the Agency Probabilistic Analysis Policy document (U.S. EPA, 1997b). *RAGS Volume 3: Part A* addresses the use of PRA for both human health and ecological risk assessments. *RAGS Volume 3: Part A* was developed to provide risk assessors and risk managers with basic guidelines for incorporating PRA into Superfund site-specific risk assessments. It is not intended to be a detailed technical reference on PRA methods, however, it does direct the reader to appropriate literature on important technical subjects. A primary purpose of *RAGS Volume 3: Part A* is to help prevent misuse and misinterpretation of PRA.

2.0 WHAT IS PROBABILISTIC RISK ASSESSMENT AND HOW IS IT USED IN RISK CHARACTERIZATION?

PRA is a risk assessment that uses probability distributions to characterize variability or uncertainty in risk estimates. In a PRA, one or more variables in the risk equation is defined as a probability distribution rather than a single number. Similarly, the output of a PRA is a range or probability distribution of risks experienced by the receptors. The evaluation of variability and uncertainty is an important component of the risk characterization of all risk assessments. As stated in the 1995 Risk Characterization memorandum from Administrator Carol Browner (U.S. EPA, 1995),

> *... we must fully, openly, and clearly characterize risks. In doing so, we will disclose the scientific analyses, uncertainties, assumptions, and science policies which underlie our decisions... There is value in sharing with others the complexities and challenges we face in making decisions in the face of uncertainty.*

In addition, the 1997 EPA Policy for Use of Probabilistic Analysis in Risk Assessment (U.S. EPA, 1997b) states:

> ☞ *It is the policy of the U.S. Environmental Protection Agency that such probabilistic analysis techniques as Monte Carlo analysis, given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments.*

A more extensive general discussion of PRA can be found in Chapter 1 of the guidance. The use of PRA in Superfund remedial decision making is presented in Chapter 7 of the guidance.

3.0 WHAT ARE THE ADVANTAGES AND DISADVANTAGES OF PRA FOR REMEDIAL DECISIONS?

The primary advantage of PRA within the Superfund program is that it can provide a quantitative description of the degree of variability or uncertainty (or both) in risk estimates for both cancer and non-cancer health effects and ecological hazards. The quantitative analysis of uncertainty and variability can provide a more comprehensive characterization of risk than is possible in the point estimate approach.

Another significant advantage of PRA is the additional information and potential flexibility it affords the risk manager. Superfund remedy decisions are often based on an evaluation of the risk to the individual at the reasonable maximum exposure (RME) level (U.S. EPA, 1990). The RME represents the highest exposure reasonably likely to occur (U.S. EPA, 1989a). When using PRA, the risk manager can select the RME from the high-end range of percentiles of risk, generally between the 90^{th} and 99.9^{th} percentiles, referred to in this guidance as the *RME range*.

However, PRA may not be appropriate for every site. Disadvantages of PRA are that it generally requires more time, resources, and expertise on the part of the assessor, reviewer, and risk manager than a point estimate approach.

4.0 HOW IS *RAGS VOLUME 3, PART A* ORGANIZED?

Although the primary audience of this guidance is the risk assessor, Chapter 1 provides a basic overview of PRA for risk assessors and risk managers. The centerpiece of *RAGS Volume 3: Part A* is the tiered approach described in Chapter 2. The tiered approach is a framework that enables the risk manager to decide if and when to undertake a PRA and to determine the appropriate level of complexity for the PRA. Chapter 3 provides a description of using PRA for human health risk assessment. Chapter 4 discusses the issues of using PRA for ecological risk assessment. Chapter 5 presents a discussion of using PRA to determine preliminary remediation goals. Chapter 6 details issues associated with communicating risk estimates developed with PRA. Chapter 7 provides information for risk managers choosing to base remedial decisions on the results of a PRA.

Eight appendices to this guidance expand on technical aspects of topics important to PRA, such as sensitivity analysis and selecting and fitting probability distributions.

5.0 WHAT ARE THE KEY GUIDING CONCEPTS IN *RAGS VOLUME 3: PART A*?

(*1) Use a tiered approach to incorporating PRA into site risk assessments.*

(2) Submit a workplan for Agency review prior to initiating work on a PRA.

(3) Perform a point estimate assessment prior to considering a PRA.

(4) While PRA can provide a useful tool to characterize and quantify variability and uncertainty in risk assessments, it is not appropriate for every site.

(5) PRA generally requires more time, resources, and expertise on the part of the assessor, reviewer, and risk manager than a point estimate risk assessment.

(6) The decision to use PRA is site-specific and is based on the complexity of the problems at the site, the quality and extent of site-specific data, and the likely utility of the result.

(7) If the additional information provided from a PRA is unlikely to affect the risk management decision, then it may not be prudent to proceed with a PRA. However, if there is a clear value added from performing a PRA, then the use of PRA as a risk assessment tool generally should be considered despite the additional resources that may be needed.

(8) Communicating the results of a PRA will be more challenging than communicating the results of a point estimate risk assessment because PRA and its perspective will be new to most participants.

(9) If the decision is made to conduct a PRA, it is important to include community in the planning process. Communication on PRA may involve: providing the community with a basic understanding of the principles of PRA, discussing the proposed workplan and inviting comments on the proposed approach, discussing site-specific data, and communicating the final results and how they impact decisions for the site.

REFERENCES FOR PREFACE

U.S. EPA. 1989a. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment)*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1992. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. *Federal Register,* 22888-22938. May 29.

U.S. EPA. 1994. *Role of Ecological Risk Assessment in the Baseline Risk Assessment.* Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-17.

U.S. EPA. 1995. *Memorandum from Carol Browner on Risk Characterization*. Office of the Administrator, Washington, DC. February 22.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments*. Interim Final. Environmental Response Team, Edison, NJ. EPA/540/R-97/006, OSWER Directive No. 9285.7-25. June.

U.S. EPA. 1997b. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May 15.

U.S. EPA. 1998a. *Guidelines for Ecological Risk Assessment.* Risk Assessment Forum. Environmental Protection Agency, Washington DC. EPA/630/R-95/002F. April. *Federal Register* 63(93): 26846-26924. May 14.

U.S. EPA. 1999. *Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-28P.

U.S. EPA. 2000a. *Superfund Probabilistic Risk Assessment to Characterize Uncertainty and Variability*. Washington, DC. *Federal Register* [FR Doc. 06-3492] 65(31): 7550-7552. February 15.

U.S. EPA. 2000b. *Peer Review for Superfund Probabilistic Risk Guidance*. Washington, DC. *Federal Register* [FR Doc. 00-21197] 65(162): 50694. August 21.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund*: *Volume I. Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments).* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-47. December.

CHAPTER 1

OVERVIEW OF PROBABILISTIC APPROACH TO RISK ASSESSMENT

1.0 INTRODUCTION

This chapter is intended for risk managers and risk assessors as an overview of the probabilistic approach to risk assessment in the context of the Superfund program at the U.S. Environmental Protection Agency (EPA). The goals of this chapter are to provide the reader with information about (1) the role of risk assessment in the Superfund program; (2) the basic concepts of probabilistic risk assessment (PRA); (3) important policies and guiding principles for PRA, as outlined throughout this guidance; and (4) the next steps that will be undertaken in the Superfund program to provide guidance on PRA.

Section 1.1 (1.1.1–1.1.3) describes the role of risk assessment from three perspectives, including the role of risk assessment in areas external to EPA, Agency-wide, and within Superfund. Section 1.1 (1.1.4) also introduces PRA and identifies its place in the Superfund program. Section 1.2 introduces the basic concepts of PRA, including the key terms of variability, uncertainty, Monte Carlo analysis (MCA), and reasonable maximum exposure (RME). PRA concepts are presented using a comparison between PRA and the traditional point estimate approach. Sections 1.2.4 and 1.3 summarize the advantages and disadvantages of PRA and point estimate risk assessment. Section 1.4 provides a summary of policies and guiding principles for using PRA in the Superfund program. EPA's policies on conducting PRA are highlighted throughout the guidance using pointers and are linked to more detailed policy discussions in other chapters in the guidance. Section 1.5 outlines the organization of this document and provides a brief summary of the content of each subsequent chapter and appendix. Section 1.6 presents EPA's next steps for PRA implementation in the Superfund program.

Key terms used throughout this guidance include: Probabilistic Risk Assessment (PRA), Monte Carlo Analysis (MCA), Probability Density Function (PDF), Cumulative Distribution Function (CDF), Reasonable Maximum Exposure (RME), Sensitivity Analysis, Tiered Approach, Variability, Uncertainty, and Preliminary Remediation Goal (PRG). Terms and their definitions are identified in an exhibit at the beginning of each chapter. Terms and definitions relevant to Chapter 1 are presented in Exhibit 1-1. In addition, a glossary of terms used throughout the guidance is given in Appendix E.

EXHIBIT 1-1

DEFINITIONS FOR CHAPTER 1

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in a population, usually considered to be the mean or median of the distribution.

Confidence Interval - A range of values that are likely to include a population parameter. Confidence intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). When used to characterize uncertainty in a risk estimate, it is assumed that methods used to quantify uncertainty in the model inputs are based on statistical principles such as sampling distributions or Bayesian approaches. For example, given a randomly sampled data set, a 95% confidence interval for the mean can be estimated by deriving a sampling distribution from a Student's t distribution.

Confidence Limit - The upper or lower value of a confidence interval.

Countably Infinite - Used to describe some discrete random variables, this term refers to a set of numbers that can be counted with integers (e.g., one, two, three) and that has no upper limit. Examples include the number of tosses required for a coin to show a head—we can count each toss, but it is possible that at least one more toss is needed. The number of dust particles in a volume of air is another example. Countably finite implies there is an upper limit (e.g., days of work per year).

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature - using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

CTE Risk - The estimated risk corresponding to the central tendency exposure.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF, gives the cumulative probability of occurrence for a random independent variable. Each value c of the function is the probability that a random observation x will be less than or equal to c.

Expected Value of Information (EVOI) - The expected increase in the value (or decrease in the loss) associated with obtaining more information about quantities relevant to the decision process. EVOI is a measure of the importance of uncertainty in risk and the potential for changing a risk management decision if uncertainty is reduced (see Appendix D).

Frequency Distribution or Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - A technique for characterizing the uncertainty and variability in risk estimates by repeatedly sampling the probability distributions of the risk equation inputs and using these inputs to calculate a range of risk values.

Numeric Stability - Stochastic variability, or "wobble" associated with random sampling, calculated as the average percent change in the model output after rerunning Monte Carlo simulations with the same set of input assumptions. Used as a metric for evaluating the adequacy of the number of iterations in a MCA.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Point Estimate - In statistical theory, a quantity calculated from values in a sample to estimate a fixed but unknown population parameter. Point estimates typically represent a central tendency or upper bound estimate of variability.

EXHIBIT 1-1

DEFINITIONS FOR CHAPTER 1—*Continued*

Point Estimate Risk Assessment - A risk assessment in which a point estimate of risk is calculated from a set of point estimates for exposure and toxicity. Such point estimates of risk can reflect the CTE, RME, or bounding risk estimate depending on the choice of inputs.

Probabilistic Risk Assessment (PRA) - A risk assessment that yields a probability distribution for risk, generally by assigning a probability distribution to represent variability or uncertainty in one or more inputs to the risk equation.

Probability Density Function (PDF) - A function representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Probability Distribution - A mathematical representation of the function that relates probabilities with specified intervals of values for a random variable. Also called a *probability model*.

Probability Mass Function (PMF) - A function representing the probability distribution for a discrete random variable. The mass at a point refers to the probability that the variable will have a value at that point.

Random Variable - A variable that may assume any value from a set of values according to chance. Discrete random variables can assume only a finite or countably infinite number of values (e.g., number of rainfall events per year). A random value is continuous if its set of possible values is an entire interval of numbers (e.g., quantity of rain in a year).

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989a). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Remedial Investigation/Feasibility Study (RI/FS) - Studies undertaken by EPA to delineate the nature and extent of contamination, to evaluate potential risk, and to develop alternatives for cleanup.

RME Risk - The estimated risk corresponding to the reasonable maximum exposure.

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic r that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Stochastic Dominance - Implies no intersection between two or more CDFs. For example, if the CDF for A and B do not overlap and the CDF for A is greater than the CDF for B, then at every cumulative percentile, the value of A is greater than that of B. Therefore, it can be stated that distribution A stochastically dominates distribution B. It should be noted that even when the CDFs for A and B do not overlap, the PDFs for A and B can overlap.

Uncertainty -Lack of knowledge about specific variables, parameters, models, or other factors. Examples include limited data regarding the concentration of a contaminant in an environmental medium and lack of information on local fish consumption practices. Uncertainty may be reduced through further study.

Variability - True heterogeneity or diversity that characterizes an exposure variable or response in a population. Further study (e.g., increasing sample size, n) will not reduce variability, but it can provide greater confidence (e.g., lower uncertainty) in quantitative characterizations of variability).

1.1 THE ROLE OF RISK ASSESSMENT IN SUPERFUND

The role of risk assessment in the Superfund program today is built upon a foundation of scientific and management principles, policies, and laws that have been established over the past two decades. Since the enactment of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) in 1980 the risk assessment policies and guidance documents have evolved to reflect advances in science and changes in federal regulations.

1.1.1 RISK ASSESSMENT IN THE UNITED STATES

Risk assessment has a long history beginning in 1940. In 1983, the National Research Council published *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983) which outlines the four steps of risk assessment (hazard identification, dose-response, exposure assessment, and risk characterization) that are used today.

The NRC addressed three main objectives in risk assessment: (1) assessment of the benefits of separating the analytical process of risk assessment from the regulatory process of risk management; (2) consideration of the feasibility of creating a single regulatory agency for the purpose of conducting all government risk assessments; and (3) consideration of the feasibility of creating uniform guidelines for risk assessment (NRC, 1983).

The Committee concluded that regulatory agencies should maintain a conceptual distinction between risk assessment and risk management, and develop uniform inference guidelines in risk assessment for use by all federal regulatory agencies. The Committee also recommended that Congress establish a Board on Risk Assessment Methods in order to ensure that risk assessment procedures be continuously reviewed and modified as the science advances. The Committee rejected the proposal for a single federal risk assessment agency based on inadequate evidence to show that one administrative structure would be more advantageous (NRC, 1983).

Since 1983, there have been ongoing advancements in the field of risk assessment. These include: (1) a continued increasing role for risk assessment in the decision-making process of many regulatory agencies, as exemplified by several bills introduced by the 103rd and 104th Congresses in 1994-1995; (2) an increased awareness of the need for uncertainty analysis and for quantifying and communicating uncertainties in risk estimates (*Science and Judgement in Risk Assessment*, NRC, 1994); (3) guidance about more inclusive approaches to risk assessment, as exemplified by environmental health legislation such as the Food Quality Protection Act (FQPA) of 1996 and the Presidential/Congressional Commission on Risk Assessment and Risk Management (1997); and (4) setting the stage for a more open decision-making process through stakeholder involvement in the risk management process, as outlined in *Improving Risk Communication* (NRC, 1989).

1.1.2 RISK ASSESSMENT AT EPA

EPA has refined the risk paradigm through deliberations of the Risk Assessment Forum, Science Policy Council, and other Agency-wide bodies. Such deliberations have led to consensus in guidance, policies, and memoranda that respond to the requirements set out by various environmental statutes. Individual offices have also developed regulations, guidance, and other supporting documents to aid in the implementation of particular environmental statutes.

In 1986, EPA issued final guidelines relating to risk assessment for cancer, mutagenic effects, developmental effects, exposure assessment, and chemical mixtures. Since 1986, EPA has updated or issued revised final guidelines for developmental toxicity, exposure assessment, reproductive toxicity, neurotoxicity, and ecological risk assessment; and is now revising carcinogen risk assessment guidelines. (See http://www.epa.gov/ncea/raf/rafguid.htm for details on *guidelines*.)

Other notable documents that guide risk assessment at EPA include:

- *Framework for Ecological Risk Assessment* (U.S. EPA, 1992b)
- *Guidelines for Ecological Risk Assessment* (U.S. EPA, 1998)
- *Guidance for Risk Characterization* (U.S. EPA, 1995a)
- *Policy for Risk Characterization* (U.S. EPA, 1995c)
- *Policy on Evaluating Health Risks to Children* (U.S. EPA, 1995d)
- *Policy for Use of Probabilistic Analysis in Risk Assessment* (U.S. EPA, 1997g)
- *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment* (U.S. EPA, 1997g)
- *Guidance on Cumulative Risk Assessment. Part 1. Planning and Scoping* (U.S. EPA, 1997e)
- *Risk Characterization Handbook* (U.S. EPA, 2000)

1.1.3 RISK ASSESSMENT IN SUPERFUND

The activities and publications described above have provided a strong foundation for the development of risk assessment guidance on conducting human health—and ecological risk assessments in the Superfund program. EPA uses risk assessment (NRC, 1983, 1994) to carry out CERCLA, as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA). Under CERCLA/SARA, EPA's Superfund program is authorized to protect human health and the environment from current and potential threats posed by releases of hazardous substances, pollutants, or contaminants. The blueprint for the Superfund program is the National Oil and Hazardous Substances Pollution Contingency Plan (NCP) (U.S. EPA, 1990). Among other things, the NCP calls for the identification and mitigation of environmental impacts at hazardous waste sites, and for the selection of remedial actions to protect human health and the environment. An important part of the NCP is the implementation of a Remedial Investigation and Feasibility Study (RI/FS), which is designed to support risk management decisions within the Superfund program. A risk assessment is an integral part of the RI/FS, and is

generally conducted at a site to determine the need for action and to ensure that a selected remedy will be protective. The NCP also establishes some benchmarks for protectiveness and lays out nine criteria (some risk-based) against which each cleanup option should be evaluated (see Exhibit 1-2).

EXHIBIT 1-2

NINE CRITERIA FOR EVALUATION OF CLEANUP ALTERNATIVES (U.S. EPA, 1990)

Threshold Criteria

1. Overall protection of human health and the environment
2. Compliance with ARARs

Balancing Criteria

3. Long-term effectiveness and permanence
4. Reduction in toxicity, mobility, or volume through treatment
5. Short-term effectiveness
6. Implementability
7. Cost

Modifying Criteria

8. State acceptance
9. Community acceptance

Guidance for risk assessment in the Superfund program has been developed to facilitate consistent site-specific responses. Early major guidance documents developed by EPA included*: Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment)* (U.S. EPA, 1989a) and *Risk Assessment Guidance for Superfund.* (RAGS)*: Volume II. Environmental Evaluation Manual* (U.S. EPA, 1989b). *RAGS Volume I: Part A* provides an approach for conducting site-specific baseline (i.e., without remediation or institutional controls) human health risk assessments. *RAGS Volume II*, aimed at site managers, provides a framework for considering environmental effects at sites. More recently, EPA developed guidance for conducting ecological risk assessments within the Superfund program. This guidance, *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments* (U.S. EPA, 1997a), discusses scientific methods and stakeholder input.

Over the years, the Superfund program has expanded RAGS to include the following documents relating to human health:

- *RAGS Volume I, Part B: Development of Risk-based Preliminary Remediation Goals (Risk Equations and Parameters)* (U.S. EPA, 1991b)
- *RAGS Volume I, Part C: Risk Evaluation of Remedial Alternatives* (U.S. EPA, 1991c)
- *RAGS Volume I, Part D: Standardized Planning, Reporting, and Review of Superfund Risk Assessments* (U.S. EPA, 2001a)
- *RAGS Volume I, Part E: Supplemental Guidance for Dermal Risk Assessment (*U.S. EPA, 2001b*)*

Additional ecological guidance documents include:

- *Role of the Ecological Risk Assessment in the Baseline Risk Assessment*. OSWER Directive No. 9285.7-17 (U.S. EPA, 1994a)
- *Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. OSWER Directive 9285.7-28 P (U.S. EPA, 1999)
- *The Role of Screening-Level Risk Assessments and Refining Contaminants of Concern in Baseline Risk Assessments*. 12th Intermittent Bulletin, ECO Update Series. (U.S. EPA, 2001d)

This document (*RAGS Volume 3: Part A*) provides guidance for probabilistic approaches for both human health and ecological risk assessment.

The Superfund program has also issued supplementary documents, including:

- *Human Health Evaluation Manual, Supplemental Guidance: "Standard Default Exposure Factors"* (U.S. EPA, 1991a)
- *Supplemental Guidance to RAGS: Calculating the Concentration Term* (U.S. EPA, 1992d)
- *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions* (U.S. EPA, 1991d)
- *Use of IRIS (*Integrated Risk Information System) *Values in Superfund Risk Assessment* (U.S. EPA, 1993)
- *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide* (U.S. EPA, 1996)
- *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites* (U.S. EPA, 2001c).

EPA will continue to develop Superfund guidance and tools to improve the practice of risk assessment. Superfund guidance documents are available from EPA's Superfund publications web site (http://www.epa.gov/superfund/pubs.htm).

The role of risk assessment in Superfund, described above, can be summarized by a number of principles that are followed and developed in *RAGS Volume 3: Part A*, including:

- The Superfund risk assessment process should rely on early problem formulation, planning, and scoping for improved remedial investigations and feasibility studies, risk assessments, and risk management decisions.

- The use of a tiered process in Superfund risk assessment and management is beneficial in that it promotes an efficient allocation of resources and improved decision-making.

- Early and continuing involvement of stakeholders throughout the Superfund risk assessment process provides an opportunity to build stakeholder trust and meet stakeholder needs, which can result in improved risk assessments and faster, more-informed risk management decisions.

1.1.4 PROBABILISTIC RISK ASSESSMENT AND ITS ROLE IN SUPERFUND

RAGS Volume I (U.S. EPA, 1989a*)* and supporting guidance describe a point estimate approach to risk assessments in the Superfund program. Point estimate risk assessments use single values (point estimates) to represent variables in a risk equation. The output of the risk equation in a point estimate risk assessment is, therefore, a point estimate of risk, which can be a central tendency exposure (CTE) estimate of risk (e.g., the average expected risk) or reasonable maximum exposure (RME) estimate of risk (e.g., the risk expected if the RME was to occur), depending on the input values used in the risk equation. *RAGS Volume 3: Part A* describes a probabilistic approach to risk assessment. Probabilistic risk assessment uses probability distributions for one or more variables in a risk equation in order to quantitatively characterize variability and/or uncertainty. The output of a PRA is a probability distribution of risks that reflects the combination of the input probability distributions. If the input

distributions represent variability, then the output risk distribution can provide information on variability in risk in the population of concern. If the input distributions reflect uncertainty, then the output risk distribution can provide information about uncertainty in the risk estimate. Information from a PRA can be used to make statements about the likelihood of exceeding a risk level of concern, given the estimated variability in elements of the risk equation. Since the results of point estimate methods generally do not lend themselves to this level of risk characterization (e.g., quantitative uncertainty assessment), PRA can provide unique and important supplemental information that can be used in making Superfund risk management decisions at Superfund sites.

Monte Carlo Analysis (MCA) is perhaps the most widely used probabilistic method in PRA. MCA is a specific probabilistic method that uses computer simulation to combine multiple probability distributions in a risk equation (see Section 1.2.2 for further discussion of Monte Carlo simulation). Monte Carlo methods have been in used in modeling since 1946 when Stanislaw Ulam used MCA to conduct uncertainty analysis at Los Alamos during the conceptual stage of the hydrogen bomb project. The history of the use of MCA (from the 1940s to the present) can be found in Rugen and Callahan, 1996.

The application of probabilistic analysis to human health and ecological risk assessment is a relatively recent development that was facilitated by development of statistical sampling techniques to obtain a probabilistic approximation to the solution of a mathematical equation and/or model, and increased speed and capacity of modern computers which can support the intensive computational requirements of MCA. Desktop computers and commercial software are currently available which enable risk assessors to make, in minutes, PRA calculations that only a few years ago would have required days.

The potential value of PRA to support risk-based decisions has become increasingly apparent over the last several years. This has prompted the need for appropriate policy and guidance documents that define the role of PRA in the Superfund program and that promote and facilitate the highest quality and consistent application of PRA in the Program where appropriate. EPA previously issued guidance that addresses the use of quantitative uncertainty analysis in risk assessment. *RAGS Volume I* (U.S. EPA, 1989a) and the *Final Guidelines for Exposure Assessment Guidelines* (U.S. EPA, 1992a) emphasize the importance of assessing variability and uncertainty in risk estimates conducted in the Superfund program. Guidance is also available for characterizing the 95% upper confidence limit (UCL) for the mean exposure concentration (U.S. EPA, 1992d, 1997f). At the regional level, EPA Regions 3 and 8 issued guidance on the appropriate use of probabilistic methods in risk assessment (U.S. EPA, 1994b, 1995e). The importance of adequately characterizing variability and uncertainty is addressed in the 1995 memorandum on *Risk Characterization Policy and Guidance* (U.S. EPA, 1995b). In the spring of 1997, EPA released the memorandum, *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment* (U.S. EPA, 1997g). According to the Policy Statement of the memorandum, probabilistic analysis techniques, "given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments." As such, a PRA, "will be evaluated and utilized in a manner that is consistent with other risk assessments submitted to the Agency." Along with this Policy Statement, the Agency released a set of guiding principles for use and review of probabilistic analyses (U.S. EPA, 1997g). Hence, both RAGS and Agency-wide guidance emphasize the importance of review of the scientific and technical merit of a probabilistic analysis to determine whether or not the assessment is of sufficient quality to support a remedial decision.

Currently, EPA's Office of Emergency and Remedial Response (OERR) is implementing PRA as part of its Superfund reform activities. This guidance, *RAGS Volume 3: Part A*, provides risk assessors with comprehensive guidance on when and how it may be appropriate to conduct PRAs using Monte

Carlo analysis within the Superfund program. It describes basic concepts in PRA, an approach for conducting MCA, and EPA's policy for implementing PRA in the Superfund program. The Agency also intends to supplement this guidance with additional examples and case studies in PRA (see Section 1.6).

1.2 BASIC CONCEPTS OF PROBABILISTIC RISK ASSESSMENT

This section describes what a PRA is and compares and contrasts it to the more familiar point estimate methods for human health risk assessment (U.S. EPA, 1989a) and ecological risk assessment (U.S. EPA, 1997a). A risk assessment performed using probabilistic methods is very similar in concept and approach to the point estimate method, with the main difference being the methods used to incorporate variability and uncertainty into the risk estimate. A variety of modeling techniques can be used to characterize variability and uncertainty in risk. This guidance focuses on MCA, perhaps the most common probabilistic method that risk assessors will encounter. Basic concepts on how to use MCA to propagate variability and uncertainty in exposure through a risk model are presented. Many of the concepts presented in this guidance are applicable to other probabilistic approaches to risk assessment.

At some sites, probabilistic analysis can provide a more complete and transparent characterization of the risks and uncertainties in risk estimates than would otherwise be possible with a point estimate approach. However, a PRA is not necessary or desirable for every site. The tiered approach presented in Chapter 2 highlights important scientific and management decisions for determining if PRA is appropriate at a specific site. The decision to perform PRA is appropriate only after the risk assessor and the remedial project manager (RPM) at the site determine whether a PRA will enhance decision making at the site. If a PRA is conducted, the assumptions and inputs to the probabilistic model should be sufficiently documented so that the results can be independently reproduced.

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

Uncertainty occurs because of a lack of knowledge. For example, we can be very certain that different people drink different amounts of water, but we may be uncertain about how much variability there is in water intakes among the population. Uncertainty can often be reduced by collecting more and better data, while variability is an inherent property of the population being evaluated. Variability can be better characterized with more data, but it cannot be reduced or eliminated.

Sometimes there can be confusion about whether data are representative of variability or uncertainty, especially when the distinction depends on how the problem is framed. For example, one of the exposure variables that may be considered in a risk assessment of workers exposed via inhalation to an indoor air contaminant is the fraction of time spent indoors on site. Assume that time-activity information is available from surveys of a representative population of workers. This data set may be used to define a probability distribution (e.g., empirical, normal) that characterizes inter-individual

variability in exposure times among workers. Sources of uncertainty would include the choice of the probability distribution used to characterize variability, as well as the parameter estimates that are based on a finite data set. Using the same data set, uncertainty in a parameter, such as the arithmetic mean exposure time, may also be defined by a probability distribution. Efforts to clearly distinguish between variability and uncertainty are important for both risk assessment and risk communication. Section 1.2.4 and Chapter 3, Section 3.4 present an overview of the different sources of uncertainty. Guidance on selecting and fitting probability distributions is given in Appendices B and C, and advanced methods for characterizing both variability and uncertainty are discussed in Appendix D.

1.2.1 WHAT IS PRA?

Probabilistic risk assessment is a general term for risk assessments that use probability models to represent the likelihood of different risk levels in a population (i.e., variability) or to characterize uncertainty in risk estimates.

A risk assessment performed using probabilistic methods would rely on the same fundamental exposure and risk equations as do point estimate approaches. U.S. EPA guidance, including *RAGS Volume I: Part A* (U.S. EPA, 1989a), the *Standard Default Exposure Factors Guidance* (U.S. EPA, 1991a), *Supplemental Guidance for Developing Soil Screening Levels* (U.S. EPA, 2001c), and *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments* (U.S. EPA, 1997a) present methods for estimating risk using standardized exposure and risk models. Examples of typical exposure and risk equations that would be used in risk calculations, in this case, for a drinking water exposure scenario, are provided in Exhibit 1-3:

EXHIBIT 1-3

CANCER AND NONCANCER RISK MODELS

Exposure Model:

$$CDI = \frac{C \times IR \times EF \times ED}{BW \times AT}$$

Cancer Risk Model:

$$Risk = CDI \times CSF$$

Noncancer Risk Model:

$$HQ = \frac{CDI}{RfD}$$

CDI	=	chronic daily intake of the chemical (mg/kg-day)
C	=	concentration of the chemical in an exposure medium (e.g., mg/L)
IR	=	ingestion rate (e.g., L/day for water, mg/day for soil, etc.)
EF	=	exposure frequency (days/year)
ED	=	exposure duration (years)
BW	=	body weight (kg)
HQ	=	hazard quotient
AT	=	averaging time (equal to ED x 365 days/year for noncarcinogens and 70 years x 365 days/year for carcinogens)
CSF	=	cancer slope factor (linear low-dose cancer potency factor) for the chemical $(mg/kg\text{-}day)^{-1}$
RfD	=	reference dose for the chemical for assessing noncancer health effects (mg/kg-day)

In the point estimate approach, a single numerical value (i.e., point estimate) is chosen for each variable shown in Exhibit 1-3. For example, point estimates may include a drinking water ingestion rate of 2 L/day and a body weight of 70 kg for an adult. Based on the choices that are made for each individual variable, a single estimate of risk is calculated. In the probabilistic approach, inputs to the risk equation are described as *random variables* (i.e., variables that can assume different values for different receptors in the population) that can be defined mathematically by a probability distribution. For continuous random variables, such as those in Figure 1-1 (body weight), the distribution may be described by a PDF, whereas for discrete random variables (e.g., number of fish meals per month), the distribution may be described by a probability mass function (PMF). The key feature of PDFs and PMFs is that they describe the range of values that a variable may assume, and indicate the relative likelihood (i.e., probability) of each value occurring within that range for the exposed population. For example, the distribution of tap water ingestion (mL/day) among the general U.S. population might be characterized by a lognormal distribution with a log-mean of 6.86 and a log-standard deviation of 0.575 (Table 3-11 of U.S. EPA 1997b). One might use a PDF to show how approximately half the population drinks more than 1 L/day of tap water, but only 10% of the population drinks more than 2 L/day. After determining appropriate PDF types and parameter values for selected variables, the set of PDFs is combined with the toxicity value in the exposure and risk equations given in Exhibit 1-3 to estimate a distribution of risks. Guidance on selecting and fitting distributions for variables in risk equations is provided in Appendix B.

In human health risk assessments, probability distributions for risk should reflect variability or uncertainty in exposure. In ecological risk assessments, risk distributions may reflect variability or uncertainty in exposure and/or toxicity (see Sections 1.4 and 1.4.1, Item 3).

A continuous probability distribution can be displayed in a graph in the form of either a PDF or corresponding CDF; however, for clarity, it is recommended that both representations be presented in adjacent (rather than overlaid) plots. Figure 1-1 illustrates a PDF and CDF for a normal probability distribution for adult body weight. Both displays represent the same distribution, but are useful for conveying different information. Note that it is helpful to include a text box with summary statistics relevant to the distribution (e.g., mean, standard deviation). The types of information that PDFs and CDFs are most useful for displaying are presented in Exhibit 1-4.

EXHIBIT 1-4

USE A PDF AND CDF TO DISPLAY:

PDF

- The relative probability of values
- The most likely values (e.g., modes)
- The shape of the distribution (e.g., skewness, kurtosis, multimodality)
- Small changes in probability density

CDF

- Percentiles, including the median
- High-end risk range (e.g., 90th to 99th percentiles)
- Confidence intervals for selected percentiles
- Stochastic dominance (i.e., for any percentile, the value for one variable exceeds that of any other variable)

Source: U.S. EPA, 1997g

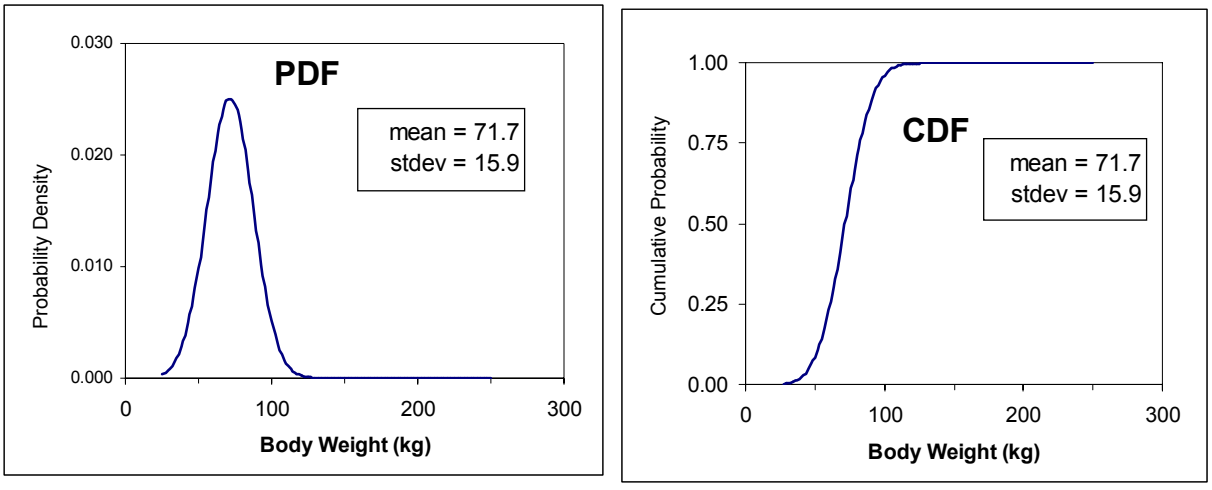


Figure 1-1. Example of a normal distribution that characterizes variability in adult body weight (males and females combined). Arithmetic mean=71.7 kg, standard deviation=15.9 kg (Finley and Paustenbach, 1994). Body weight may be considered a continuous random variable. The left panel shows a bell-shaped curve and represents the PDF, while the right panel shows an S-shaped curve and represents the CDF. Both displays represent the same distribution (including summary statistics), but are useful for conveying different information.

The CDF for risk can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., 95th percentile=1E-06). A text box may also be included on the graph to highlight important summary statistics, such as the parameters of the input distribution, or selected percentiles of the output distribution for risk. For example, a clear description of the parameters for the probability distribution should be given, as well as an indication of whether the distribution represents variability or uncertainty.

1.2.2 WHAT IS A MONTE CARLO SIMULATION?

Perhaps the most common numerical technique for PRA is Monte Carlo simulation. Monte Carlo simulation has been widely used to explore problems in many disciplines of science as well as engineering, finance, and insurance (Rugen and Callahan, 1996). The process for a Monte Carlo simulation is illustrated in Figure 1-2. In its general form, the risk equation can be expressed as a function of multiple exposure variables (V_i) and a toxicity term: Risk=f(V_1, V_2, ...V_n) x Toxicity. Solutions for equations with PDFs are typically too complex for even an expert mathematician to calculate the risk distribution analytically. However, numerical techniques applied with the aid of computers can provide very close approximations of the solution. This is illustrated here for the simplified case in which the assessment variables are statistically independent, that is, the value of one variable has no relationship to the value of any other variable. In this case, the computer selects a value for each variable (V_i) at random from a specified PDF and calculates the corresponding risk. This process is repeated many times (e.g., 10,000), each time saving the set of input values and corresponding estimate of risk. For example, the first risk estimate might represent a hypothetical individual who drinks 2 L/day of water and weighs 65 kg, the second estimate might represent someone who drinks 1 L/day and weighs 72 kg, and so forth. Each calculation is referred to as an iteration, and a set of iterations is called a simulation.

> ☞ *A convenient aid to understanding the Monte Carlo approach for quantifying variability is to visualize each iteration as representing a single individual and the collection of all iterations as representing a population.*

Each iteration of a Monte Carlo simulation should represent a plausible combination of input values (i.e., exposure and toxicity variables), which may require using bounded or truncated probability distributions (see Appendix B). However, risk estimates are not intended to correspond to any one person. The "individuals" represented by Monte Carlo iterations are virtual and the risk distributions derived from a PRA allow for inferences to be made about the likelihood or probability of risks occurring within a specified range for an exposed human or ecological population. A simulation yields a set of risk estimates that can be summarized with selected statistics (e.g., arithmetic mean, percentiles) and displayed graphically using the PDF and CDF for the estimated risk distribution. Often the input distributions are assumed to be independent, as shown in Figure 1-2. More complex Monte Carlo simulations can be developed that quantify a dependence between one or more input distributions by using conditional distributions or correlation coefficients (see Appendix B, Section B.5.5 for a discussion of correlated input distributions).

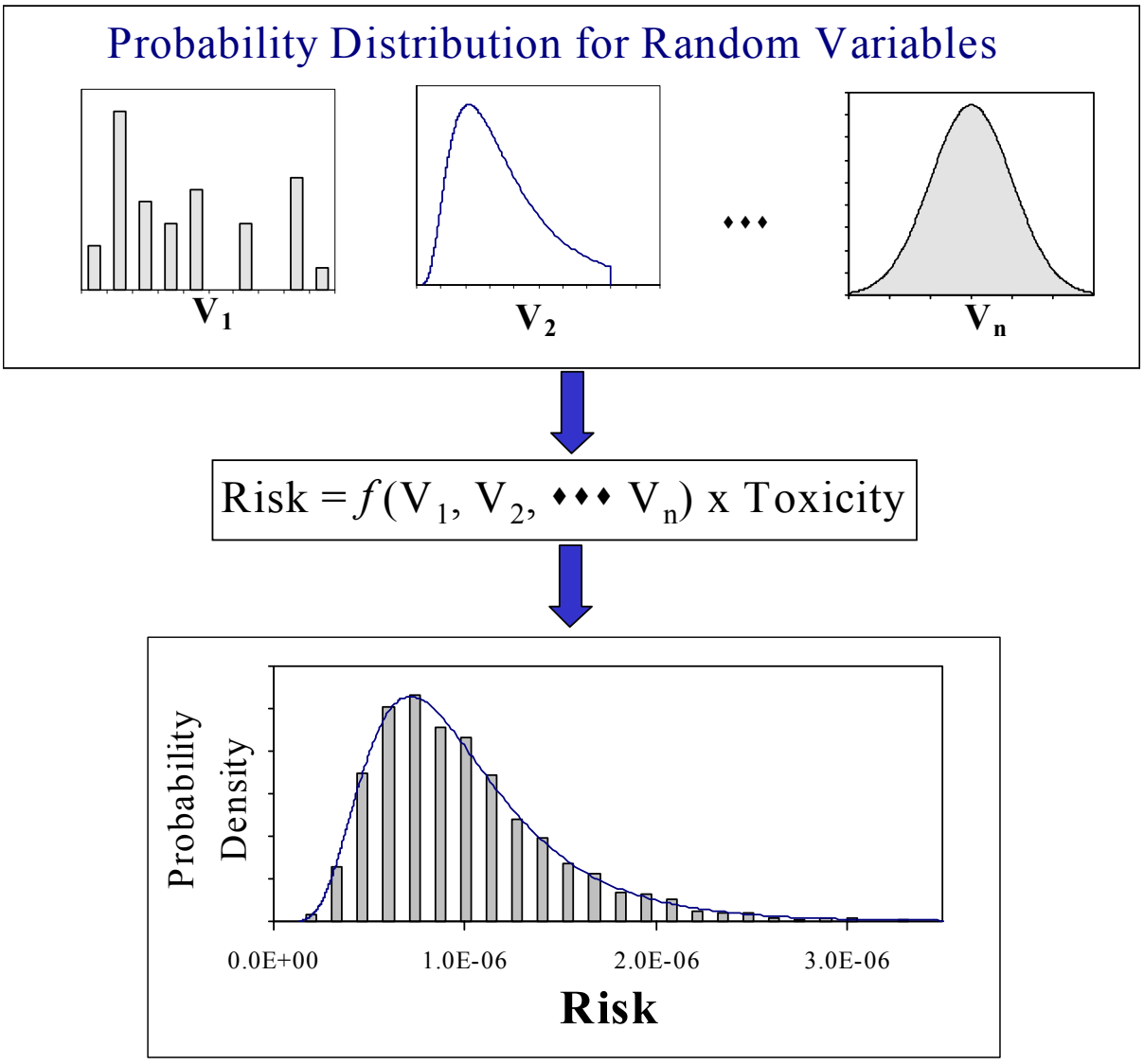


Figure 1-2. Conceptual model of Monte Carlo analysis. Random variables (V_1, V_2, ...V_n) refer to exposure variables (e.g., body weight, exposure frequency, ingestion rage) that are characterized by probability distributions. A unique risk estimate is calculated for each set of random values. Repeatedly sampling (V_i) results in a frequency distribution of risk, which can be described by a PDF. In human health risk assessments, the toxicity term should be expressed as a point estimate. In ecological risk assessment (see Sections 1.4 and 1.4.1) the toxicity term may be expressed as a point estimate or as a probability distribution.

The rapid evolution in computing power has greatly reduced concerns among regulators regarding the number iterations needed in MCA.

> ☞ *While this guidance does not prescribe specific criteria or set an arbitrary "minimum" number of iterations needed for PRA, a general rule of thumb is that a sufficient number of iterations should be run to obtain numerical stability in percentiles of the output (e.g., risk distribution) that are important for decision making.*

Numerical stability refers to the stochastic variability, or "wobble" associated with random sampling, and can be evaluated by running multiple simulations with the same set of input assumptions and calculating the average percent change in a specified percentile of the output (e.g., Maddalena et al., 2001). For example, it may be determined that 5,000 iterations are sufficient to achieve numerical stability in the 50th percentile, but insufficient for the 95th percentile risk estimate when a criteria of ± 1% is applied for multiple simulations. As discussed in Section 1.4, one of the eight conditions specified by EPA for the acceptance of PRA is that the numerical stability of the output be presented and discussed, since it will vary depending on what percentile of the risk distribution is evaluated. While some commercial software now have a feature to automatically stop simulations after a specified criterion for numerical stability is achieved (Burmaster and Udell, 1990), care should be taken to understand how this criterion is implemented across the entire range of the output distribution.

1.2.3 WHY IS VARIABILITY IMPORTANT IN RISK ASSESSMENT? HOW IS IT ADDRESSED BY THE POINT ESTIMATE AND PROBABILISTIC APPROACHES?

As noted previously, variability refers to true heterogeneity or diversity that occurs within a population or sample. Factors that lead to variability in exposure and risk include variability in contaminant concentrations in a medium (air, water, soil, etc.), differences in ingestion rates or exposure frequencies, or in the case of ecological assessments, inter- and intra-species variability in dose-response relationships. *Risk Assessment Guidance for Superfund Volume I* (Section 6.1.2 of U.S. EPA, 1989a) and the *NCP Preamble* (U.S. EPA, 1990) state that human health risk management decisions at Superfund sites will generally be based on an individual that has RME. Likewise, RME estimates of risk are the most appropriate basis for decision making using an ecological risk assessment. Use of the RME and CTE risk descriptors in ecological risk assessment are discussed in Chapter 4. The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures based on both quantitative information and professional judgment (Sections 6.1.2 and 6.4.1 of U.S. EPA, 1989a). In addition, the Agency released guidance in 1992 (U.S. EPA, 1992c) recommending the inclusion of a "central tendency" exposure estimate to an individual, as well as a high-end exposure estimate, in the risk assessment. Generally, the CTE is considered to be a measure of the mean or median exposure. The difference between the CTE and the RME gives an initial impression of the degree of variability in exposure or risk between individuals in an exposed population.

Depending on assessment needs at a site, a range of point estimates of risk can be developed to represent variability in exposures. To support the evaluation of RME risk estimates using the point estimate approach described in Section 1.3, the Superfund program developed guidance with recommended default values for exposure variables as inputs to the risk equations (U.S. EPA, 1992a, 1996, 1997a, 2001d). These standardized values are a combination of average (e.g., body weight, skin surface area) and high-end exposure assumptions (e.g., drinking water intake, exposure duration). A CTE risk estimate is based on central estimates (e.g., mean, 50th percentile) for each of the exposure variables.

Available site-specific data on plausible mean and upper range values for exposure variables should be used to support CTE and RME risk estimates. The point estimate approach to risk assessment does not determine where the CTE or RME risk estimates lie within the risk distribution. For example, the RME risk estimated with the point estimate approach could be the 90th percentile, the 99.9th percentile, or some other percentile of the risk distribution. Without knowing what percentile is represented by the RME risk estimate, the risk manager might be unsure about the likelihood of the RME risk occurring or being exceeded in the receptor population and about what level of remedial action is justified or necessary to achieve the protective objectives of CERCLA.

In a PRA, distributions used as inputs to the risk equations can characterize the inter-individual variability inherent in each of the exposure assumptions. By characterizing variability with one or more input distributions, the output from the Monte Carlo simulation is a distribution of risks that could occur in that population (Figure 1-3). The central tendency of the risk distribution (e.g., arithmetic mean, geometric mean, 50th percentile) may be characterized as the CTE risk estimate. Similarly, the high-end of the risk distribution (e.g., 90th to 99.9th percentiles) is representative of exposures to the RME individual. In addition to providing a better understanding of where the CTE and RME risks occur in the distribution, a PRA can also provide an estimate of the probability of occurrence associated with a particular risk level of concern (e.g., cancer risk of 1E-05). A PRA that quantifies variability can be used to address the question, "What is the likelihood (i.e., probability) that risks to an exposed individual will exceed 1E-05?" Based on the best available information regarding exposure and toxicity, a risk assessor might conclude, "The estimated distribution for variability in risk across the target population indicates that 10% of the individuals exposed under these circumstances have a risk exceeding 1E-05." This type of evaluation can be achieved using a technique known as one-dimensional Monte Carlo Analysis (1-D MCA). Guidelines for interpreting the high-end of the risk distribution in terms of the RME risk estimate are discussed further in Section 1.4.1 and Chapter 7.

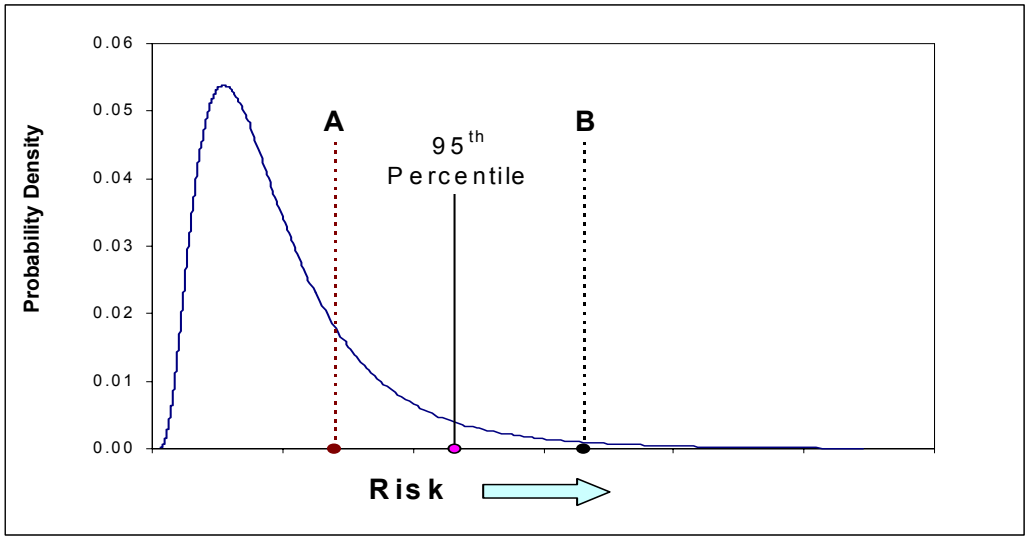


Figure 1-3. Example of a probability distribution for risk illustrating the 95th percentile and two different risk levels of concern (A and B). Assuming the 95th percentile corresponds to the RME, the need for remedial action depends on how the RME risk compares with the risk level of concern. For Case A (RME > level of concern), remedial action may be warranted. For Case B (RME < level of concern), remedial action may be unnecessary.

The agreement (or lack of agreement) between the results of the point estimate calculations and the PRA calculations is expected to vary as a function of the form of the exposure or risk model and the attributes of the input variables. In general, if the terms in the denominator of the exposure or risk equation have low variability and do not approach zero, then the CTE point estimate is likely to agree quite well with the arithmetic mean from the PRA simulation, and the RME point estimate is likely to correspond to the high-end of the risk distribution (see discussion of RME range in Section 1.2.5). However, if the exposure or risk model has terms in the denominator that are a significant source of variability, or if the terms approach zero, then the agreement between the point estimate values and the PRA values may be more substantial. In addition, since the RME point estimate of risk reflects a combination of central tendency and high-end input values, it is difficult to anticipate what percentile of a distribution of variability it represents.

> ☞ *If results of PRA calculations differ substantially from point estimate calculations, a risk manager may benefit from understanding the reasons for the differences and the relative strengths of the different approaches.*

Since point estimate and PRA approaches may yield different estimates of CTE and RME risks, the two approaches also may support different risk management decisions. This does not imply that either approach is invalid. Likewise, a correspondence between the point estimate and PRA results does not imply a greater accuracy or certainty in the modeling assumptions and inputs. Simply stated, PRA, based on the same risk equations and data as the point estimate approach, provides a different means of characterizing variability and uncertainty. Potential sources of variability and uncertainty in risk estimates should be identified, discussed, and to the extent practicable, quantified. Advantages and disadvantages of PRA and point estimate risk assessment are discussed in Section 1.2.4 and 1.3.

1.2.4 WHY IS UNCERTAINTY IMPORTANT IN RISK ASSESSMENT? HOW IS UNCERTAINTY ADDRESSED BY THE POINT ESTIMATE AND PROBABILISTIC APPROACHES?

Uncertainty derives from a lack of knowledge. Various taxonomies of uncertainty relevant to risk assessment have been presented (Finkel, 1990; Morgan and Henrion, 1990; Cullen and Frey, 1999). U.S. EPA guidance, including the *Final Guidelines Exposure Assessment Guidelines* (U.S. EPA, 1992a), *Exposure Factors Handbook* (U.S. EPA, 1997b,c,d), and *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997g) describe a variety of different types of uncertainty in risk assessment as well as modeling strategies for quantifying uncertainties. Potential sources of uncertainty in risk assessment can be divided into one of three broad categories:

(1) *Parameter uncertainty* - uncertainty in an estimate of an input variable in a model. In PRA, this may refer specifically to a statistical concept of uncertainty in estimates of population parameters (e.g., arithmetic mean, standard deviation) from random samples, due to the quality, quantity, and representativeness of available data as well as the statistical estimation method.

(2) *Model uncertainty* - uncertainty about a model structure (e.g., exposure equation) or intended use, including the relevance of simplifying assumptions to the endpoint of the risk assessment, the choice of probability distribution to characterize variability, and interpolation or extrapolation beyond the scale used to calibrate a model from empirical data.

(3) *Scenario uncertainty* - uncertainty regarding missing or incomplete information to fully define exposure. This may include descriptive errors regarding the magnitude and extent of chemical exposure or toxicity, temporal and spatial aggregation errors, incomplete analysis (i.e., missing exposure pathways), and potential mis-specification of the exposed population or exposure unit.

Sources of uncertainty described by these categories are important because they can influence risk management decisions in both point estimate and probabilistic risk assessment. As additional sources of uncertainty are quantified and included in the risk assessment, uncertainty in risk estimates may appear to increase, suggesting there may be little confidence in a risk management decision. This situation may appear to be counterintuitive for those managers who expect confidence to increase as uncertainty is quantified. However, as discussed below and in Chapter 6 (see Section 6.4.2), uncovering and quantifying these sources of uncertainty may help to provide perspective, and make the decisions using the tiered process more transparent. In PRA, there are a variety of methods that can be used to effectively quantify uncertainty as well as communicate confidence in risk estimates (see Chapter 3, Section 3.4; Chapter 6, Section 6.4, and Section 6.5).

Parameter uncertainty may be the most readily recognized source of uncertainty that is quantified in site-specific risk assessments at hazardous waste sites. Parameter uncertainty can occur in each step of the risk assessment process from data collection and evaluation, to the assessment of exposure and toxicity. Sources of parameter uncertainty may include systematic errors or bias in the data collection process, imprecision in the analytical measurements, inferences made from a limited database when that database may or may not be representative of the variable under study, and extrapolation or the use of surrogate measures to represent the parameter of interest.

In the point estimate approach, parameter uncertainty is addressed in a qualitative manner for most variables. For example, the uncertainty section of a point estimate risk assessment document might note that a soil sampling plan yielded a small sample size that may not be representative of overall contaminant concentrations and, as a result, the risk estimate may over- or under-estimate actual risk. Uncertainty in the concentration term is addressed quantitatively to a limited extent in a point estimate approach by using the 95% UCL for the arithmetic mean concentration in both CTE and RME risk estimates; this accounts for uncertainty associated with environmental sampling and site characterization (U.S. EPA, 1992d, 1997f). The 95% UCL is combined in the same risk calculation with various central tendency and high-end point estimates for other exposure factors.

Some examples of the models that EPA uses in the risk assessment process are the equations used to calculate exposure and risk, the linearized multistage model used to estimate cancer dose-response relationships, and media-specific models to estimate contaminant concentrations. All models are simplified, idealized representations of complicated physical or biological processes. Models can be very useful from a regulatory standpoint, as it is generally not possible to adequately monitor long term exposure for populations at contaminated sites. However, models that are too simplified may not adequately represent all aspects of the phenomena they were intended to approximate or may not capture important relationships among input variables. Other sources of model uncertainty can occur when important variables are excluded, interactions between inputs are ignored, or surrogate variables that are different from the variable under study are used.

In most probabilistic assessments, the first step of analysis is usually an analysis of variability in exposure or risk. However, PRA methods may also be used to characterize uncertainty around the best estimate of the exposure or risk distribution. This is done using "2-dimensional" MCA (2-D MCA) (see Appendix D). One convention that has been used to distinguish between probability distribution functions for variability and uncertainty is to use subscripts "v" and "u" to indicate PDFs that characterize variability (PDFv) or uncertainty (PDFu). Figure 1-4 shows an example of the results of this type of 2-D MCA. This analysis can provide a quantitative measure of the *confidence in the fraction of the population* with a risk exceeding a particular level; which is sometimes referred to as a *vertical confidence interval* (Figure 1-4). For example, a conclusion based on this type of output might be, "While the best estimate for the variability distribution for risk across the target population indicates that 10% of the individuals exposed under these circumstances have a risk exceeding 1E-06, the uncertainty is such that we can only be reasonably certain (e.g., 95% sure) that no more than 20% of the exposed population has a risk that exceeds 1E-06." Additionally, the output from a 2-D MCA can provide a quantitative measure of the *confidence in the risk estimate* for a particular fraction of the population; which is sometimes referred to as a *horizontal confidence interval*. This type of output might support the following type of conclusion, "While the best estimate for the variability distribution for risk across the target population indicates that 10% of the individuals exposed under these circumstances have a risk exceeding 1E-06, the uncertainty is such that we can only be reasonably certain (e.g., 95% sure) that the risk for this group of individuals does not exceed 2E-06." The term "confidence interval" is used loosely in this context to convey information about uncertainty; however, it is not the same as a statistical confidence interval that one might obtain by estimating a population parameter from a sample. The vertical and horizontal bars shown in Figure 1-4 represent a range of possible estimates for the percentile given one or more sources of uncertainty that were included in the simulation. If the target audience for this graphic has a greater understanding of statistics, it may be less confusing if alternative phrases are used to describe the results, such as "credible interval" or "probability band".

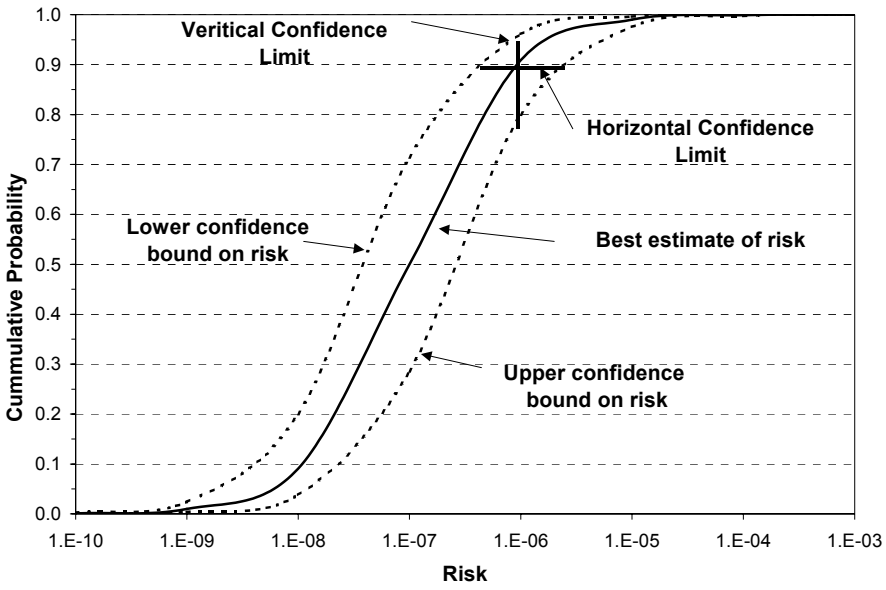


Figure 1-4. Illustration of "Vertical" and "Horizontal" Confidence Intervals (or limits) on a risk estimate. This type of output can be produced from a 2-D MCA in which probability distributions of uncertainty are introduced into the risk equation. See Chapter 3 and Appendix D for further discussion of 2-D MCA in quantitative uncertainty analysis.

In general, one should avoid developing input distributions to a PRA model that yield a single risk distribution that intermingles, or represents both variability and uncertainty. By separately characterizing variability and uncertainty, the output from a PRA will be easier to understand and communicate. A number of tools can aid in evaluating the uncertainty in estimated distributions for variability. Both simple and very complex approaches have been applied to this problem. Two basic

methods for quantifying variability and parameter uncertainty simultaneously are described in Exhibit 1-5. PRAs that use these approaches can provide quantitative estimates of uncertainty in percentiles of the risk distribution based on confidence intervals or credible intervals for one or more parameter estimates. Techniques for characterizing both variability and uncertainty in PRA are discussed in more detail in Chapters 3, 4, 5, and 7, and Appendices A, C, and D.

A common apprehension concerning the utility of PRA is that it may require more information and data than are available to generate credible PDFs. Risk assessors may feel that they can't specify a PDF because they don't have enough information to choose a distribution type, estimate parameters, or evaluate the representativeness to the site population of concern. However, if sufficient information exists to support a meaningful point estimate evaluation (i.e., if some sort of central tendency and upper bound values are available for each input variable), then it is usually possible to perform a screening level, or preliminary 1-D MCA that may provide additional useful information regarding variability. Likewise, an initial two-dimensional analysis may be performed that does not require collection of any new data, but simply characterizes uncertainty in the existing data. The results of such a 2-D MCA can help to identify the main sources of uncertainty in the risk results, and can support decisions to collect more data and/or proceed with additional tiers of analysis in order to improve the assessment. As with a preliminary 1-D MCA, the decision to conduct a more advanced probabilistic analysis does not always result in added data requirements.

EXHIBIT 1-5

QUANTIFYING VARIABILITY AND UNCERTAINTY

1. Single source of uncertainty

Run multiple one-dimensional Monte Carlo simulations (1-D MCA) in which each simulation uses a different point estimate for a parameter selected from an uncertainty distribution, combined with PDFv's for one or more variables. For example, separate simulations can be run in which the mean of the exposure concentration variability distribution is represented by either the 95% lower or upper confidence limit on the mean. A comparison of the output of these simulations would provide a partial characterization of the quantitative impact of uncertainty in the mean exposure concentration on the risk estimate (provided that certain conditions hold; i.e., risk increases with increasing exposure concentration) (see Chapter 3, Section 3.3.1).

2. Multiple sources of uncertainty

Run a single two-dimensional Monte Carlo simulation (2-D MCA), in which separate probability distributions are specified for variability and parameter uncertainty and values from these distributions are randomly selected and used in each iteration of the Monte Carlo simulation (see Appendix D).

Use of probabilistic methods (e.g., MCA) to propagate variability and uncertainty through risk models offers five key advantages over point estimate approaches in addressing uncertainty in risk estimates:

(1) Probabilistic methods may often provide a more complete and informative characterization of variability in exposure or risk than is usually achievable using point estimate techniques.

(2) Probabilistic methods can provide a more quantitative expression of the confidence in risk estimates than the point estimate approach.
(3) Sensitivity analysis methods using PRA may help risk assessors to better identify influential exposure factors.

(4) Probabilistic methods can account for dependencies between input variables (e.g., body weight and skin surface area).

(5) Probabilistic methods provide quantitative estimates of the expected value of additional information that might be obtained from data collection efforts (Morgan and Henrion, 1990). The importance of quantifying uncertainty in an *expected value of information* (EVOI) framework is discussed in Appendix D.

Since both point estimate and probabilistic approaches in risk assessment are applied to the same conceptual models (i.e., the same exposure and risk models), uncertainties in the conceptual model are generally addressed in the same manner. If other models are available to explain or characterize a given phenomenon, the risk estimates associated with each of those conceptual models could be compared to determine the sensitivity of the risk to the uncertainty in the choice of a model (see Chapter 2 and Appendix A). For example, when deciding on a contaminant concentration term for tetrachloroethylene in groundwater for a residential exposure assessment 10 years in the future, it would be appropriate to compare and contrast several fate and transport models and their results before deciding on a concentration term.

1.2.5 REASONABLE MAXIMUM EXPOSURE AT THE HIGH-END

Risk management decisions at Superfund sites should be based on an estimate of the risk to a reasonably maximum exposed receptor, considering both current and future land-use conditions. The RME is defined as the highest exposure that is reasonably expected to occur at a site. In general, risks corresponding to the 90^{th} to 99.9^{th} percentiles of the risk distribution estimated from a PRA are considered plausible high-end risks, and the RME risk should be selected within this range (see Section 1.2.4, Section 1.4.1, and Chapter 7 for further discussion). In comparison with point estimate risk assessments, PRA can provide the entire range of estimated risks as well as the likelihood of values within the range (i.e., the frequency distribution)

As noted in Chapter 7, estimates of risk become more uncertain at very high percentiles (e.g., the 99.9^{th}), so results of PRA calculations at these extreme values should be used with caution. Risk frequency distributions toward the 99.9^{th} percentile may be numerically unstable due to the uncertainties embedded in the input exposure assumptions. This guidance recommends that a risk manager select the RME in consultation with a risk assessor. One item for discussion should be the numerical stability of the high-end RME risk value (i.e., a stable value on the frequency distribution within the high-end range that could be reproduced in successive Monte Carlo simulations.)

1.3 ADVANTAGES AND DISADVANTAGES OF POINT ESTIMATE AND PROBABILISTIC APPROACHES

As discussed in Chapter 2, a PRA should not be conducted until adequate point estimate calculations have been completed. Once this has been done, the potential benefits of proceeding to a PRA evaluation should be based on an understanding of the potential advantages and limitations in each approach. Potential advantages and disadvantages of point estimate calculations are summarized in Exhibit 1-6 and potential advantages and disadvantages of PRA are listed in Exhibit 1-7.

In general, compared to a point estimate risk assessment, a PRA based on the same state of knowledge may offer a more complete characterization of variability in risk, can provide a quantitative evaluation of uncertainty, and may provide a number of advantages in assessing if and how to proceed to higher levels of analysis. However, there are also some real and perceived disadvantages regarding additional effort on the part of both the risk assessor and the risk manager, and the potential to cause confusion if the effort is not clearly presented.

In general, the key question to consider in deciding whether a PRA should be performed is whether or not the PRA analysis is likely to provide information that will help in the risk management decision making. For some sites, the additional information provided by a PRA will not affect the decision that would have been made with a point estimate approach alone, and a PRA will not be useful. However, when the decision whether or not to take action is not completely clear, PRA may be a valuable tool. The tiered process for PRA (Chapter 2) introduces the concept of scientific management decision points (SMDPs) to guide the complexity of analysis that may be needed for decision making. An SMDP marks a point in the process in which the potential that another analysis may influence the risk management decision is evaluated based on the problem formulation, the information available to define input variables, the results of previous analyses, and the feasibility of a subsequent analysis.

☞ *A point estimate approach is conducted for every risk assessment; a probabilistic analysis may not always be needed.*

EXHIBIT 1-6

ADVANTAGES AND DISADVANTAGES OF POINT ESTIMATE APPROACH

Advantages

- Calculations are simple and do not require any advanced software.
- EPA has established default inputs and methods to help standardize point estimate calculations between sites.
- Useful as a screening method—may allow risk management decisions with no additional work.
- Central tendency and RME estimates of risk provide a semi-quantitative measure of variability.
- Method is easily described and communicated.
- Requires less time to complete; not as resource intensive.

Disadvantages

- Computational simplifications may result in deviations from target values.
- Results are often viewed as "the answer"; importance of uncertainty is sometimes lost.
- Information from sensitivity analysis is generally limited to dominant exposure pathways and chemicals of concern; may not highlight the key exposure variables and uncertain parameters.
- Does not provide a measure of the probability that risk exceeds a regulatory level of concern, or the level of confidence in a risk estimate.
- Provides fewer incentives for collecting better or more complete information.
- May not utilize all available data for characterizing variability and uncertainty in risk estimates.

EXHIBIT 1-7

ADVANTAGES AND DISADVANTAGES OF PROBABILISTIC RISK ASSESSMENT

Advantages

- Can make more complete use of available data when defining inputs to the risk equation.
- Can provide a more comprehensive characterization of variability in risk estimates.
- Can provide a more comprehensive characterization of uncertainty in inputs, which may support statements regarding confidence in risk estimates. Communication of uncertainty in the risk assessment can help to build trust among stakeholders.
- Sensitivity analysis can identify the exposure variables, probability models, and model parameters that influence the estimates of risk.
- Puts the risk assessment in a *Value-of-Information* framework (see Appendix D). Can identify data gaps for further evaluation/data collection and can use wider variety of site-specific information.
- Allows available site-specific information to inform the choice of high-end percentile from the risk distribution that corresponds with RME risk.

Disadvantages

- Concepts and approaches may be unfamiliar; there is often apprehension regarding added costs and potential for inadvertent error and/or intentional misrepresentation.
- Places more burden on risk assessors to ensure the PRA is done correctly and on managers to understand and make decisions within a range of alternatives.
- May require more time and resources to select and fit probability distributions, and may require greater effort to communicate methodology and results.
- May convey false sense of accuracy when data are sparse.
- Complexities of the PRA approaches may obscure important assumptions or errors in basic exposure or risk models.
- If communication of the more complex PRA is unsuccessful, then it may generate mistrust of the assessment and risk management decisions.

1.4 CONDUCTING AN ACCEPTABLE PRA

In 1997, EPA issued a memorandum which contained its policy statement on PRA (U.S. EPA, 1997g). The 1997 EPA Policy Statement is as follows:

> It is the policy of the U.S. Environmental Protection Agency that such probabilistic analysis techniques as Monte Carlo analysis, given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments. As such, and provided that the conditions described below are met, risk assessments using Monte Carlo analysis or other probabilistic techniques will be evaluated and utilized in a manner that is consistent with other risk assessments submitted to the Agency for review or consideration. It is not the intent of this policy to recommend that probabilistic analysis be conducted for all risk assessments supporting risk management decisions. Such analysis should be a part of a tiered approach to risk assessment that progresses from simpler (e.g., deterministic) to more complex (e.g., probabilistic) analyses as the risk management situation requires. Use of Monte Carlo or other such techniques in risk assessments shall not be cause, *per se*, for rejection of the risk assessment by the Agency. 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, Conditions for Acceptance 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. In the case of ecological risk assessment, however, this policy applies to all aspects including stressor and dose-response assessment.

In support of this policy statement, EPA has outlined eight *conditions for acceptance* (in italics below), and good scientific practice of PRA. A PRA that is submitted to the Agency for review and evaluation should generally comply with each condition in order to ensure that adequate supporting data and credible assumptions are used in the assessment. These conditions are as follows:

(1) The purpose and scope of the assessment should be clearly articulated in a "problem formulation" section that includes a full discussion of any highly exposed or highly susceptible subpopulations evaluated (e.g., children, the elderly). The questions the assessment attempts to answer are to be discussed and the assessment endpoints are to be well defined.

(2) The methods used for the analysis (including all models used, all data upon which the assessment is based, and all assumptions that have a significant impact upon the results) are to be documented and easily located in the report. This documentation is to include a discussion of the degree to which the data used are representative of the population under study. Also, this documentation is to include the names of the models and software used to generate the analysis. Sufficient information is to be provided to allow the results of the analysis to be independently reproduced.

Possible sources of bias inherent in the input distributions should be discussed along with the expected impacts on the resulting risk estimates. For example, if a site-specific study of fish consumption indicated consumption rates are five to ten times higher than other studies from similar populations, this possible bias or inaccuracy should be discussed in the document. Computer programs should generally

be described in sufficient detail to allow the reviewer to understand all aspects of the analysis. Computer code/spreadsheets should provide adequate documentation and annotation.

(3) The results of sensitivity analyses are to be presented and discussed in the report. Probabilistic techniques should be applied to the compounds, pathways, and factors of importance to the assessment, as determined by sensitivity analyses or other basic requirements of the assessment.

Sensitivity analysis is a valuable tool in any tier of a PRA.

(4) The presence or absence of moderate to strong correlations or dependencies between the input variables is to be discussed and accounted for in the analysis, along with the effects these have on the output distribution.

(5) Information for each input and output distribution is to be provided in the report. This includes tabular and graphical representations of the distributions (e.g., probability density function and cumulative distribution function plots) that indicate the location of any point estimates of interest (e.g., mean, median, 95th percentile). The selection of distributions is to be explained and justified. For both the input and output distributions, variability and uncertainty are to be differentiated where possible.

(6) The numerical stability of the central tendency and the higher end (i.e., tail) of the output distributions are to be presented and discussed.

As discussed in Section 1.2.5, numerical stability refers to the observed numerical changes in parameters of the output distribution (e.g., median, 95th percentile) from a Monte Carlo simulation as the number of iterations increases. Because most risk equations are linear and multiplicative, distributions of risk will generally be right-skewed, and approximate a lognormal distribution. Values in the tails of the distribution typically are less stable than the central tendency, and the rate of convergence for the tails will depend on the form of the risk model, the skewness of the probability distributions selected for input variables and the numerical methods used to simulate probability distributions. Provided that appropriate numerical methods are employed, numerical stability is generally not a concern for most 1-D MCA models, which can be run with a sufficient number iterations in minutes with modern high speed computers; however, it can be an important consideration for more complex simulations, such as with 2-D MCA models.

(7) Calculations of exposures and risks using deterministic (e.g., point estimate) methods are to be reported if possible. Providing these values will allow comparisons between the probabilistic analysis and past or screening level risk assessments. Further, deterministic estimates may be used to answer scenario specific questions and to facilitate risk communication. When comparisons are made, it is important to explain the similarities and differences in the underlying data, assumptions, and models.

If results of PRA calculations differ substantially from point estimate calculations, a risk manager may benefit from understanding the reasons for the differences and the relative strengths of the different approaches. Sometimes, a closer look at uncertainties in the underlying data, assumptions, and models will lead a risk assessor to revisit parts of the assessment in order to provide a more consistent basis for comparison.

(8) Since fixed exposure assumptions (e.g., exposure duration, body weight) are sometimes embedded in the toxicity metrics (e.g., Reference Doses, Reference Concentrations, Cancer risk factors), the exposure estimates from the probabilistic output distribution are to be aligned with the toxicity metric.

1.4.1 KEY POLICIES FOR APPLYING PRA AT SUPERFUND SITES

EPA's recommended process for conducting an acceptable PRA generally follows the policy and guiding principles presented above. In addition, this section highlights four key policies for conducting acceptable PRAs at hazardous waste sites.

(1) Follow the Tiered Approach to PRA

In accordance with the *1997 EPA Policy Statement* (U.S. EPA, 1997g), this guidance recommends using a tiered approach when considering PRA to help with risk management decisions. A tiered approach begins with a relatively simple analysis and progresses stepwise to more complex analyses. The level of complexity should match the site-specific risk assessment objectives and the risk management goals. The tiered approach, with helpful suggestions on risk communication, is presented in Chapter 2. A brief introduction is given below.

The premise for recommending a tiered approach is that there is a balance between the benefits of conducting a more complex analysis, and the cost in terms of additional time, resources, and challenges for risk communication. PRA may require additional resources compared with the point estimate approach, and may not be used routinely for screening level assessment. At more complex hazardous waste sites, PRA may not be warranted if the investment of time and resources is unlikely to provide information on variability and uncertainty in risk that will affect the risk management decision.

This guidance recommends that a point estimate risk assessment be conducted in the first tier after completing the remedial investigation (RI) planning, site scoping, problem formulation, data collection, and the development of a site conceptual model. In general, when site decision making would benefit from additional analysis beyond the point estimate risk assessment, and when the risk manager needs more information to complete the RI/FS process, the risk manager would proceed to higher tiers. Sensitivity analysis should be conducted in each tier to guide decisions regarding data collection and the complexity of the analysis needed to characterize variability and/or uncertainty in risk. Sensitivity analysis can also play an important role in risk communication by supporting decisions to continue characterizing less influential variables with point estimates in higher tiers.

(2) Select the RME Risk from the RME Risk Range (90^{th} to 99.9^{th} percentile)

The RME is defined as the highest exposure that is reasonably expected to occur at a site. *Final Guidelines for Exposure Assessment* (EPA, 1992a) states that the "high-end" of exposure for a population occurs between the 90^{th} and 99.9^{th} percentiles, with the 99.9^{th} percentile considered a bounding estimate. Using a point estimate approach, the calculation of the RME risk would be based on high-end input values in combination with average input values. For example, for estimation of risks from the ingestion of groundwater, default exposure is based on a high-end water intake rate (2 L/day), a high-end exposure frequency and duration (350 days/year for 30 years), and an average body weight (70 kg).

With the probabilistic approach, the calculation of the RME risk would be based on a range of input values, or frequency distributions, including low, average, and high-end values for each of the input exposure factors. For example, for estimation of risks from ingestion of groundwater, exposure would be based on the combination of lognormal distributions for water intake rate, body weight, and exposure duration (each using a specified mean and standard deviation) and a triangular distribution for exposure frequency (using a specified minimum, most likely value, and maximum). As a result, the RME risk would become a probability distribution ranging from low- to high-end values based on varying a combination of input values. In PRA, a recommended starting point for risk management decisions regarding the RME is the 95th percentile of the risk distribution (see Chapter 7).

(3) Use PRA for Dose-Response in Ecological Assessment, not in Human Health Assessment

Approaches to characterizing variability and uncertainty in toxicological information should reflect both the latest developments in the science of hazard and dose-response evaluation and consistent application of EPA science policy. This statement is consistent with the *1997 EPA Policy Statement* presented in Section 1.4 above (U.S. EPA, 1997g). Probabilistic approaches to ecological dose-response assessment may be explored, as discussed and demonstrated in Chapter 4. This guidance does not develop or evaluate probabilistic approaches for dose-response in human health assessment and, further, *discourages undertaking such activities on a site-by-site basis*. Such activities require contaminant-specific national consensus development and national policy development. Parties wishing to undertake such activities should contact the OERR to explore ways in which they might contribute to a national process for the contaminant of interest to them.

(4) Prepare a Workplan for EPA Review and Approval

A workplan should be developed and submitted for review before commencement of a PRA. The workplan should document the combined decisions of the RPM and risk assessor involved in the risk assessment, and positions of the stakeholders. The workplan should address conditions and policies presented in this section of *RAGS Volume 3: Part A*, the software to be used, the exposure routes and models, and the input probability distributions and their basis, including appropriate literature references. The workplan is discussed in more detail in Chapter 2.

A checklist of some of the key considerations to assist in the review of a PRA is provided in Appendix F.

1.5 ORGANIZATION OF THE GUIDANCE

Subsequent chapters of *RAGS Volume 3: Part A* focus on the following topics:

Chapter 2 - The Tiered Approach to PRA

Chapter 2 includes information regarding organizational issues that may need to be considered by the RPM in developing a PRA. Examples, include: workplans, involvement of the Community Involvement Coordinator (CIC), additional meetings with communities, and review of PRA documents.

Chapter 2 also presents the tiered approach in full detail. The approach begins with RI planning, scoping, problem formulation, and data collection. Tier 1 entails a point estimate risk assessment and sensitivity analysis. Tier 2 proceeds with additional data collection, a MCA to characterize variability

and/or uncertainty, and a more in-depth sensitivity analysis. More advanced techniques are used in Tier 3 to simultaneously characterize variability and uncertainty. The endpoint of the tiered approach is to provide information that helps risk managers complete the RI/FS process.

Chapter 3 - Probabilistic Human Health Risk Assessment

Chapter 3 provides a discussion of how PRA approaches may be utilized in human health risk assessment. Probabilistic approaches focus on the exposure assessment, and an example is included to illustrate the application of the tiered approach to a human health risk assessment.

Chapter 4 - Probabilistic Ecological Risk Assessment

Chapter 4 provides a discussion of how PRA approaches may be utilized in ecological risk assessment. This includes a discussion of basic tactics, such as how to decide if, and when, a PRA is needed, along with technical discussions and examples of how to model variability and/or uncertainty in exposure, toxicity, and risk (characterized both as hazard quotients and responses) for different types of ecological receptors, both within and between species. The chapter also provides a discussion of how the results of an ecological PRA can be used in risk management decision making, and provides guidelines for planning and performing an ecological PRA.

Chapter 5 - PRA and Preliminary Remediation Goals (PRGs)

This chapter provides a discussion about issues associated with deriving PRGs from both point estimate risk assessment and PRA. Issues and limitations associated with back calculation are highlighted, along with an explanation and recommendation regarding the iterative forward calculations.

Chapter 6 - Communicating Risks and Uncertainties in PRA

Chapter 6 provides a basic overview of the current Superfund guidance on communicating with the public. With this as a basis, the chapter provides specific information regarding continuous involvement of stakeholders in the PRA process, various tools that may be useful in communicating the principles of PRA, organizational issues regarding planning of communication strategies, and examples of procedures that may be helpful at individual sites. This chapter also provides references to various documents on current approaches for communicating risk to the public.

Chapter 7 - Role of PRA in Decision Making

This chapter provides guidance on how to interpret the results of a PRA to determine if an unacceptable risk is present, and criteria to consider when moving from a risk-based PRG to a remedial goal.

Appendix A - Sensitivity Analysis

Important information from PRA includes the results of sensitivity analysis. This appendix outlines the methodology and interpretation of statistical methods used to conduct sensitivity analysis with point estimate and probabilistic models.

Appendix B - Selecting and Fitting Distributions

One of the more challenging aspects of PRA is choosing appropriate probability distributions to represent variability and uncertainty in the input variables. This appendix presents a process for selecting and fitting distributions to data, including hypothesizing families of distributions, parameter estimation techniques, and goodness-of-fit tests.

Appendix C - Exposure Point Concentration (EPC)

An important variable in most risk assessments is the concentration term. This appendix presents the basic principles of the EPC, and different methods for quantifying both variability and parameter uncertainty in the EPC.

Appendix D - Advanced PRA Models

Sometimes a more complex modeling approach can be used to improve the representativeness of the probabilistic risk estimates. These approaches are generally anticipated to be applied in Tier 3 of the tiered approach. Examples include the use of Microexposure Event modeling, geostatisics, and Bayesian Monte Carlo analysis.

Appendix E - Definitions

A list of definitions is provided at the beginning of each chapter. This appendix provides a compilation of all definitions presented in the guidance.

Appendix F - Generic Checklist

After a PRA has been submitted to the Agency, an efficient process is needed to evaluate the accuracy and clarity of the results. This appendix suggests a series of elements of the review process that can be adopted to structure the review of PRAs for both human health and ecological risk assessment.

Appendix G - Frequently Asked Questions (FAQ) about PRA

Risk assessors and risk managers who read *RAGS Volume 3: Part A* will find that probabilistic risk assessment covers a wide variety of topics ranging from statistical theory to practical applications and policy decisions. U.S. EPA OERR plans to maintain and periodically update a list of frequently asked questions and responses on an EPA Superfund web page at http://www.epa.gov/superfund/index.htm. This appendix provides a preliminary list of anticipated questions.

Appendix H - Index

This index includes keywords and concepts used throughout this guidance document. They are listed alphabetically with numbers indicating the appropriate chapter and page number(s) within each chapter. Commas separate page numbers within a chapter or appendix, while semi-colons separate chapters and appendices. For example: probability density function, 1-5, 6-8; 4-3, 10-12; C-1, 8-10. This would indicate Chapter 1, page 5, and pages 6-8; Chapter 4, page 3, and pages 10-12; Appendix C, page 1 and pages 8-10.

1.6 NEXT STEPS FOR PRA IMPLEMENTATION

This guidance has presented the current principles, including the tiered approach, and examples to aid in conducting acceptable PRAs at Superfund sites. Policies and practices will change over time as scientific advances continue in the future. The PRA Workgroup intends to keep current and provide new information on EPA Superfund web page at http://www.epa.gov/superfund/index.htm. EPA expects to make the following PRA support items available on-line in the near future:

- *RAGS Volume 3: Part B*: A workbook that serves as a companion to *RAGS Volume 3: Part A*; it will include case studies and examples in PRA.
- *Guidance on Probability Distributions*: Documents and/or spreadsheets to aid in selecting and fitting probability distributions for input variables.
- *Guidance on Data Representativeness*: A ranking methodology to evaluate data representativeness for various exposure scenarios.
- *Hands-On Training*: Basic MCA training materials, and limited computer hands-on training sessions available to Regional EPA and State staff.
- *Access to PRA Workgroup*: A workgroup to provide support on PRA to EPA regional risk assessors.
- *FAQs*: A list of Frequently Asked Questions (FAQs) about PRA and responses from the PRA Workgroup, maintained and periodically updated on-line.

REFERENCES FOR CHAPTER 1

Burmaster, D.E. and E.C. Udell. 1990. A Review of Crystal Ball®. *Software Review* 10: 343–345.

Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Plenum Press, NY.

Finley, B.L. and D.J. Paustenbach. 1994. The Benefits of Probabilistic Exposure Assessment: Three Case Studies Involving Contaminated Air, Water and Soil. *Risk Anal*. 14(1):53–73.

Finkel, A.M. 1990. Confronting Uncertainty in Risk Management: A Guide for Decision Makers. Center for Risk Management, Resources for the Future. Washington, DC.

Maddalena, R.L., T.E. McKone, D.P.H. Hsieh, and S. Geng. 2001. Influential Input Classification in Probabilistic Multimedia Models. *Stochastic Environmental Research and Risk Assessment* 15(1):1–17.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.

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

National Research Council (NRC). 1989. *Improving Risk Communication.* National Academy Press. Washington, DC.

National Research Council (NRC). 1994. *Science and Judgement in Risk Assessment.* National Academy Press. Washington, DC.

Presidential/Congressional Commission on Risk Assessment and Risk Management. 1997. *Risk Assessment and Risk Management in Regulatory Decision Making*. Final Report, Volume 2.

Rugen, P. and B. Callahan. 1996. An Overview of Monte Carlo, A Fifty Year Perspective. *Hum Ecol Risk Assess.* 2(4):671–680.

U.S. EPA. 1989a. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1–89/002. NTIS PB90-155581.

U.S. EPA. 1989b. *Risk Assessment Guidance for Superfund.* (RAGS)*: Volume II. Environmental Evaluation Manual*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/001.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991a. *Risk Assessment Guidance for Superfund (RAGS): Volume I–Human Health Evaluation Manual Supplemental Guidance: "Standard Default Exposure Factors."* Interim Final. Office of Solid and Emergency Response, Washington, DC. OSWER Directive No. 9285.6-03.

U.S. EPA. 1991b. *RAGS Volume I, Human Health Evaluation Manual (Part B: Development of Risk-based Preliminary Remediation Goals).* Office of Emergency and Remedial Response. Washington, DC. EPA/540/R-92/003. December.

U.S. EPA. 1991c. *RAGS Volume I, Human Health Evaluation Manual (Part C: Risk Evaluation of Remedial Alternatives).* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-01C. October.

U.S. EPA. 1991d. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9355.0-30.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. *57 Federal Register*, 22888-22938, May 29.

U.S. EPA. 1992b. *Framework for Ecological Risk Assessment.* EPA 630/R-92/001. February.

U.S. EPA. 1992c. *Guidance on Risk Characterization for Risk Managers and Risk Assessors.* Memorandum from F. Henry Habicht II, Deputy Administrator. Office of Solid Waste and Emergency Response, Washington, DC.

U.S. EPA. 1992d. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-081.

U.S. EPA. 1993. *Use of IRIS (*Integrated Risk Information System) *Values in Superfund Risk Assessment*. Memorandum from William H. Farland and Henry L. Longest II. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7.16, December 21.

U.S. EPA. 1994a. *Role of Ecological Risk Assessment in the Baseline Risk Assessment.* Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-17.

U.S. EPA. 1994b. *Use of Monte Carlo Simulation in Risk Assessments*. Region 3, Hazardous Waste Management Division. Office of Superfund programs, Philadelphia, PA. EPA/903/F-94/001.

U.S. EPA. 1995a. *Guidance for Risk Characterization*. Office of Research and Development. Washington, DC. http://www.epa.gov/ORD/spc/rcpolicy.htm.

U.S. EPA. 1995b. *Memorandum from Carol Browner on Risk Characterization*. Office of the Administrator, Washington, DC. February 22.

U.S. EPA. 1995c. *Policy for Risk Characterization.* Office of Research and Development. Washington, DC. http://www.epa.gov/ORD/spc/rcpolicy.htm.

U.S. EPA. 1995d. *Policy on Evaluating Health Risks to Children*. Office of Children's Health Protection. Washington, DC. http://www.epa.gov/children/whatwe/rrguide.pdf.

U.S. EPA. 1995e. *Use of Monte Carlo Simulation in Performing Risk Assessments* (Technical Section). Region 8, Hazardous Waste Management Division, Superfund Management Branch Technical Guidance, Denver, CO, RA-10.

U.S. EPA. 1996. *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 540/R-96/018.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments*. Interim Final. Environmental Response Team, Edison, NJ. EPA/540/R-97/006, OSWER Directive No. 9285.7-25, June.

U.S. EPA. 1997b. *Exposure Factors Handbook, Volume 1.* Office of Research and Development,Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 1997c. *Exposure Factors Handbook, Volume 2.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fb.

U.S. EPA. 1997d. *Exposure Factors Handbook, Volume 3.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fc.

U.S. EPA. 1997e. *Guidance on Cumulative Risk Assessment. Phase 1. Planning and Scoping*. Washington, DC.

U.S. EPA. 1997f. *Lognormal Distribution in Environmental Applications*. Office of Research and Development, and Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006. December.

U.S. EPA. 1997g. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment*. Final. National Center for Environmental Assessment, Washington, DC. EPA/630/R-95/002F.

U.S. EPA. 1999. *Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-28P.

U.S. EPA. 2000. *Risk Characterization Handbook*. Office of Science Policy. Office of Research and Development. EPA 100-B-00-002. December.

U.S. EPA. 2001a. *Risk Assessment Guidance for Superfund*: *Volume I. Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments).* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-47. December.

U.S. EPA. 2001b. *Risk Assessment Guidance for Superfund: Volume 1, Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment)*. Interim. Review Draft–For Public Comment. Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-02E-P. September.

U.S. EPA. 2001c. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9355.4-24. December.

U.S. EPA. 2001d. *The Role of Screening-Level Risk Assessments and Refining Contaminants of Concern Baseline Risk Assessments.* Office of Solid Waste and Emergency Response. 12th Intermittent Bulletin, ECO Update Series. EPA 540/F-01/014. June.

CHAPTER 2

WORKPLAN AND THE TIERED APPROACH

2.0 INTRODUCTION

While probabilistic risk assessment (PRA) can provide useful information for risk management, not all site decisions will benefit from probabilistic approaches. Similarly, not all PRAs need involve complex models and quantitative uncertainty analysis methods; often, very useful information can be obtained by taking the point estimate approach one step further to explore variability in selected input variables. The level of effort and complexity of the risk assessment should match site-specific needs. The use of a tiered approach for moving from a point estimate risk assessment to PRAs of varying levels of complexity is recommended (Figure 2-1 and 2-2). This chapter outlines the basic steps of a tiered approach for including PRA in a site risk assessment. The major feature of the tiered approach is an iterative evaluation of the risk estimates developed at each tier to determine if they are sufficient for risk management decisions. Built into the tiered approach are opportunities for communication with stakeholders with a view to saving time and costs, and facilitating a successful remedial process.

2.1 WORKPLAN

In practice, the potential value of PRA may be considered at various planning stages of a risk assessment. For some sites, PRA and point estimate risk assessment approaches may be discussed in the initial scoping of the risk assessment. For other sites, PRA may become a viable option only after the point estimate risk assessment results are available. Ideally, PRA should be considered as early as possible in the planning of risk assessment activities at a site so that sampling plans and data collection efforts may be appropriately directed. Initial PRA discussions should be included as part of the risk assessment workplan. If a PRA is being considered following completion of a point estimate risk assessment, the original workplan for the point estimate assessment should be expanded to include needs that are unique to PRA.

The methods and procedures used to prepare a workplan to gather additional information for a baseline point estimate risk assessment are documented in RAGS Volume I: Part A (U.S. EPA, 1989). This chapter of RAGS Volume 3: Part A describes the procedures that would be used to prepare a workplan to gather additional information to conduct a PRA. Separate workplans may be warranted for human health and ecological risk assessments.

Like the quality assurance project plan (QAPP), the workplan for a PRA should document the combined decisions of the remedial project manager (RPM) and the risk assessor. Meaningful involvement of stakeholders early in the decision-making process also will save time and effort.

EXHIBIT 2-1

DEFINITIONS FOR CHAPTER 2

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in a population, usually considered to be the mean or median of the distribution.

Countably Infinite - Used to describe some discrete random variables, this term refers to a set of numbers that can be counted with integers (e.g., one, two, three) and that has no upper limit. Examples include the number of tosses required for a coin to show a head—we can count each toss, but it is possible that at least one more toss is needed. The number of dust particles in a volume of air is another example. Countably finite implies there is an upper limit (e.g., days of work per year).

CTE Risk - The estimated risk corresponding to the central tendency exposure.

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - A technique for characterizing the uncertainty and variability in risk estimates by repeatedly sampling the probability distributions of the risk equation inputs and using these inputs to calculate a range of risk values.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Point Estimate - In statistical theory, a quantity calculated from values in a sample to estimate a fixed but unknown population parameter. Point estimates typically represent a central tendency or upper bound estimate of variability.

Point Estimate Risk Assessment - A risk assessment in which a point estimate of risk is calculated from a set of point estimates for exposure and toxicity. Such point estimates of risk can reflect the CTE, RME, or bounding risk estimate depending on the choice of inputs.

Potentially Responsible Party (PRP) - PRPs are individuals, companies, or any other party that are potentially liable for payment of Superfund cleanup costs.

Preliminary Remediation Goal (PRG) - Initially developed chemical concentration for an environmental medium that is expected to be protective of human health and ecosystems. PRGs may be developed based on applicable or relevant and appropriate requirements (ARARs), or exposure scenarios evaluated prior to or as a result of the baseline risk assessment. (U.S. EPA, 1991a, 1991b).

Probabilistic Risk Assessment (PRA) - A risk assessment that yields a probability distribution for risk, generally by assigning a probability distribution to represent variability or uncertainty in one or more inputs to the risk equation

Probability Density Function (PDF) - A graph that shows the probability of occurrence of an unknown or variable quantity. A PDF is used to characterize a continuous random variable, X. PDFs can be used to display the shape of the distribution for an input variable or output variable of a Monte Carlo simulation. The term *density* comes from the concept that a probability at a point, x, for a continuous distribution is equal to the area under the curve of the PDF associated with a narrow range of values around x.

Probability Distribution - A mathematical representation of the function that relates probabilities with specified intervals of values for a random variable. Also called a *probability model*.

Probability Mass Function (PMF) - A function representing the probability distribution for a discrete random variable. The mass at a point refers to the probability that the variable will have a value at that point.

Random Variable - A variable that may assume any value from a set of values according to chance. Discrete random variables can assume only a finite or countably infinite number of values (e.g., number of rainfall events per year). A random value is continuous if its set of possible values is an entire interval of numbers (e.g., quantity of rain in a year).

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Remedial Investigation/Feasibility Study (RI/FS) - Studies undertaken by EPA to delineate the nature and extent of contamination, to evaluate potential risk, and to develop alternatives for cleanup.

EXHIBIT 2-1

DEFINITIONS FOR CHAPTER *2*—*Continued*

RME Risk - The estimated risk corresponding to the reasonable maximum exposure.

Scientific/Management Decision Point (SMDP) - A point during the tiered process in PRA when the risk assessor communicates results of the assessment to the risk manager. At this point, the risk manager determines whether the information is sufficient to arrive at a decision or if additional data collection or analysis is needed. SMDPs provide a tool for transitioning to a subsequent tier or for exiting the tiered process.

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic r that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Uncertainty - Lack of knowledge about specific variables, parameters, models, or other factors. Examples include limited data regarding the concentration of a contaminant in an environmental medium and lack of information on local fish consumption practices. Uncertainty may be reduced through further study.

A PRA workplan should be developed early in the risk assessment planning process for the site, regardless of who will actually develop the PRA (e.g., Environmental Protection Agency (EPA), EPA contractor, or potentially responsible party (PRP)). If a PRP performs the PRA, the workplan should be submitted to EPA for review and approval prior to commencing the PRA. It should describe the intended PRA in sufficient detail so that EPA can determine if the work products will adequately address risk assessment and management needs (see Exhibit 2-2 for contents of a typical workplan). It is important that the risk assessor and RPM discuss the scope of the probabilistic analysis and the potential impact it may have on the remedial investigation/feasibility study (RI/FS).

> ☞ *Given the time and effort that can be expected to be invested in conducting a PRA, it is important that a workplan undergo review and approval by EPA, prior to proceeding with the assessment.*

In general, regions should not accept probabilistic analysis when a workplan for the analysis has not been submitted to the Agency, and approved by the regional risk assessor and RPM.

The tiered process for PRA, described in Section 2.3, is an iterative process. As new information becomes available, it should be used to evaluate the need to move to a higher tier. The decision to move an assessment to a higher tier of complexity should result in a revised workplan reflecting the greater complexity and demands of the higher tier. The proposed probabilistic sensitivity analysis developed at the lower tier should be included in the revised workplan, along with a point estimate risk assessment based on any data collected as part of a lower tier. The probabilistic methods used in a PRA can often be restricted to the chemicals and pathways of concern that contribute the greatest risk. The less sensitive chemicals and exposure pathways should still remain in the PRA using point estimates, unless there is a compelling reason to exclude them from the assessment altogether. As stated in Appendix A (Section A.1, *Risk Communication*), the decision to represent an input variable with a point estimate, rather than a probability distribution, will generally be made on a case-by-case basis. The decision will reflect an

attempt to balance the benefits of simplifying the analysis (e.g., easier to communicate; focuses discussion on more critical areas) with the potential for arbitrarily reducing the variance in the output distribution (e.g., discounting variability in multiple variables with negligible contributions to risk may end up having a non-negligible effect on the RME percentile).

Throughout the process of developing the PRA, EPA risk assessor and other contributors to the assessment should have a continuing dialogue to discuss the elements of the workplan and their potential impacts on the assessment. This dialogue, along with interim deliverables, will help to ensure that the risk assessment report will meet the needs of the Agency and that any problems are identified and corrected early in the process.

EXHIBIT 2-2

EXAMPLES OF IMPORTANT CONTENTS OF A PRA WORKPLAN

1. Statement of the ecological assessment endpoints and/or human risk
2. Summary of the point estimate risk assessment
3. Potential value added by conducting a PRA and proceeding to the subsequent tiers
4. Discussion of adequacy of environmental sampling for PRA or moving to a successive tier (e.g., data quality issues)
5. Description of the methods and models to be used (e.g., model and parameter selection criteria)
6. Proposal for obtaining and basis for using exposure factor distributions or ecological toxicity distributions
7. Methods for deriving the concentration term
8. Proposal for probabilistic sensitivity analysis
9. Software (e.g., date and version of product, random number generator)
10. Bibliography of relevant literature
11. Proposed schedule, discussion points, and expertise needed

2.2 SPECIAL ADMINISTRATIVE CONSIDERATIONS IN PRA

Inclusion of a PRA in the RI/FS will generate certain administrative activities for the RPM. The scope of these activities will depend on whether the PRA is conducted by EPA and its contractors or by the PRP. The following sections provide practical advice for the RPM who is considering applications of PRA at a site.

2.2.1 SCOPING OF PRA

The RPM will generally be involved in the discussions among EPA project team, as well as PRPs and other stakeholders, regarding the level of PRA that is appropriate for the site. As outlined in the tiered approach (see Section 2.3), the scope and complexity of the PRA should satisfy the risk assessment and management decision making needs of the site. Team members should meet to discuss the scope of the PRA, the anticipated community outreach, and the required level of review. These discussions can be useful for ascertaining the level of contractor involvement, specific requirements for deliverables from PRPs, and the anticipated number of responses to comments. These meetings should include consideration of funding, resources, and availability of personnel to work on the PRA.

2.2.1.1 PRA SCOPE OF WORK FOR FUND-LEAD SITES

A Statement of Work (SOW) should be developed before any work is started on a PRA, regardless of whether the PRA is to be submitted to the Agency or developed by the Agency. The SOW should outline the general approach that EPA and its contractor will use in developing the PRA. The SOW should include the general approaches for the following PRA items: selection of input probability distributions, documentation of methods and results, selection of computer programs, submission of

computer codes and outputs, comparison of the results from the point estimate and probabilistic assessments, and the format for presenting the final PRA in the RI/FS document. The SOW should be sufficiently detailed to support a milestone schedule, which should be submitted as part of the SOW. Based on the complexity of the PRA, and consistent with the RAGS Volume I: Part D principles of involving the risk assessor early and often in the risk assessment process (U.S. EPA, 2001), it may be appropriate to obtain submission of interim deliverables to allow the risk assessor the opportunity to identify potential problems early in the process.

Within the RI/FS workplan, additional resources may be required to hold additional meetings, to respond to comments specific to the PRA, and to develop handouts describing PRA in terms accessible to a wider audience than risk assessors. Where appropriate, these additional resource requirements should be included in the SOW along with interim and final deliverable dates. Chapter 6 provides guidance on communicating concepts and results of PRA to various audiences.

2.2.1.2 PRP SCOPE OF WORK FOR PRP-LEAD SITES

The SOW for PRP-lead sites should follow the same general outline as the SOW for fund-lead sites (Section 2.2.1.1). Legal documents such as Unilateral Orders, Administrative Orders of Consent, and Consent Decrees should contain language requiring the PRP to submit a workplan before any work on the PRA is started. It is also important that interim deliverables, including computer code or spreadsheet models, be submitted so that EPA can review and verify the results of the PRA. A comparison of the results of the PRA and the point estimate assessment should be included in the final RI/FS.

Depending on the complexity of the site and the anticipated PRA, the RPM may be involved in more extensive negotiations with the PRPs. These negotiations may involve both EPA staff and contractor support. These activities may need to be included in the appropriate SOWs.

If warranted by the complexity of the PRA, the RPM may consider the need to expand oversight contracts to include additional resources for the contractor to review and comment on the interim deliverables and finalize the PRA. This may require a specialized level of expertise that will need to be discussed with the contractor. Further, the contract section regarding community involvement may also need to be expanded to include additional resources for developing handouts describing PRA in terms accessible to a wider audience than risk assessors and for holding additional community meetings.

2.2.2 DEVELOPMENT OF PROBABILITY DISTRIBUTIONS

A key component of any PRA is the selection of representative probability distributions. The information available to support the characterization of variability or uncertainty with probability distributions may be an important factor in the decision to conduct a PRA. In some cases, this may require resources to conduct exploratory data analysis or to collect site-specific information. As part of this process, a PRA using preliminary distributions based on the available information may be considered to identify the variables and exposure pathways that may have the strongest effect on the risk estimates. Appendix B (Section B.2.0) provides a more detailed description of preliminary distributions and their potential role in the tiered process. All of these activities may require extensive discussions with the PRPs and the community. In addition, for PRP-lead sites, they may require additional resources to critically review the proposed distributions. The RPM should consider these potential activities in developing the SOW and legal documents to assure adequate resources are available to address them.

2.2.3 EPA REVIEW OF PRA DOCUMENTS

The review of PRA documents may require more time than is usually allocated for point estimate risk assessments. In part, the additional time is needed for reviewing and discussing input distributions, for developing and running computer simulations, and for discussing outcomes of the assessment with the PRP or EPA contractor. The early involvement of an EPA risk assessor may reduce the time needed for review of the final risk assessment documents, although additional review time may still be required, depending on the complexity of the PRA conducted.

In addition to EPA's review, it may also be important to include external reviewers with specialized expertise in PRA to aid in the review. This additional support may involve resources and time to review documents and verify simulation results, as well as additional contractual arrangements. As stated in Chapter 1, Section 1.4 (Conducting an Acceptable PRA), it is important that negotiations with the PRP address the assurance that adequate details will be included in the submission so that the methods can be evaluated, and the results independently reproduced.

2.2.4 PEER-REVIEW

Depending on the level of complexity of the PRA, and whether new science is being used, it may be necessary to conduct a peer review of the document. The Agency's guidance on peer review (U.S. EPA, 2000b) should be consulted for information regarding the criteria for determining whether or not a peer review is appropriate and, if it is, the process that should be followed.

2.2.5 RESPONSE TO COMMENTS ON PRA

The time and resources needed to respond to comments on a PRA may vary depending on the complexity of the PRA. In developing the SOW, workplan, and schedule for the RI/FS, it is important that the RPM include adequate resources and time for the thorough evaluation of the PRA. In developing the response to comments, it may be necessary to consider alternative PRAs submitted by reviewers. The RPM should plan for sufficient time and resources needed for such activities.

2.2.6 ADMINISTRATIVE RECORD

Criteria should be established for documentation to be included in the administrative record. Examples may include documentation regarding the basis for selection of input distributions, a description of the design of the PRA conducted, the computer codes used in simulations, how tiering decisions are made, and the results of the PRA. The RPM should consider using technologies such as a CD-ROM to document the appropriate information for the record.

2.2.7 COMMUNICATION WITH STAKEHOLDERS

Chapter 6 provides details regarding the goal of early involvement of the public in the PRA process. For example, Section 6.1 of Chapter 6 provides additional topics for consideration in development of community involvement plans (CIPs) where PRA is considered. In general, early involvement of the community in the RI/FS process is important, but such involvement should meet the site-specific needs. Important considerations include resources, funding, and the level of effort appropriate for the site.

2.2.8 COMMUNICATION WITH EPA MANAGEMENT

Communication with EPA managers regarding PRA is discussed in Chapter 6. The RPM may need to consider allocating additional resources for prebriefings of appropriate management levels, development of handouts, and follow-up to the management meetings. Coordination with appropriate EPA staff and contractors may be necessary to assure the communication is effective.

2.3 OVERVIEW OF THE TIERED APPROACH

The tiered approach presented in this guidance is a process for a systematic, informed progression to increasingly more complex risk assessment methods including PRA. A schematic presentation of the tiered approach is shown in Figure 2-1 and Figure 2-2. Higher tiers reflect increasing complexity and, in many cases, will require more time and resources. Higher tiers also reflect increasing characterization of variability and/or uncertainty in the risk estimate, which may be important for making risk management decisions. Central to the concept of a systematic, informed progression is an iterative process of evaluation, deliberation, data collection, work planning, and communication (see Figure 2-2). All of these steps should focus on deciding (1) whether or not the risk assessment, in its current state, is sufficient to support risk management decisions (a clear path to exiting the tiered process is available at each tier); and (2) if the assessment is determined to be insufficient, whether or not progression to a higher tier of complexity (or refinement of the current tier) would provide a sufficient benefit to warrant the additional effort.

The deliberation cycle provides an opportunity to evaluate the direction and goals of the assessment as new information becomes available. It may include evaluations of both scientific and policy information. The risk manager, in the decision-making process, is encouraged to seek input on a regular basis from EPA staff and other stakeholders. Exhibit 2-3 lists some of the potential stakeholders that may contribute to the deliberation process.

Although PRA may involve technical dialogue between EPA and outside "experts", input from members of the general public who may have an interest in the outcome of the remedial process should also be sought at appropriate stages of the process. Frequent and productive communication between EPA and stakeholders throughout the risk assessment process is important for enhancing the success of a PRA.

EXHIBIT 2-3

STAKEHOLDERS POTENTIALLY INVOLVED IN EPA'S DECISION-MAKING PROCESS FOR PRA

- EPA risk assessors and managers
- Members of the public
- Representatives from state or county environmental or health agencies
- Other federal agencies (e.g., health agencies, Natural Resource Damage Assessment trustees, etc.)
- Tribal government representatives
- Potentially responsible parties and their representatives
- Representatives from federal facilities (e.g., Department of Defense, Department of Energy, etc.)

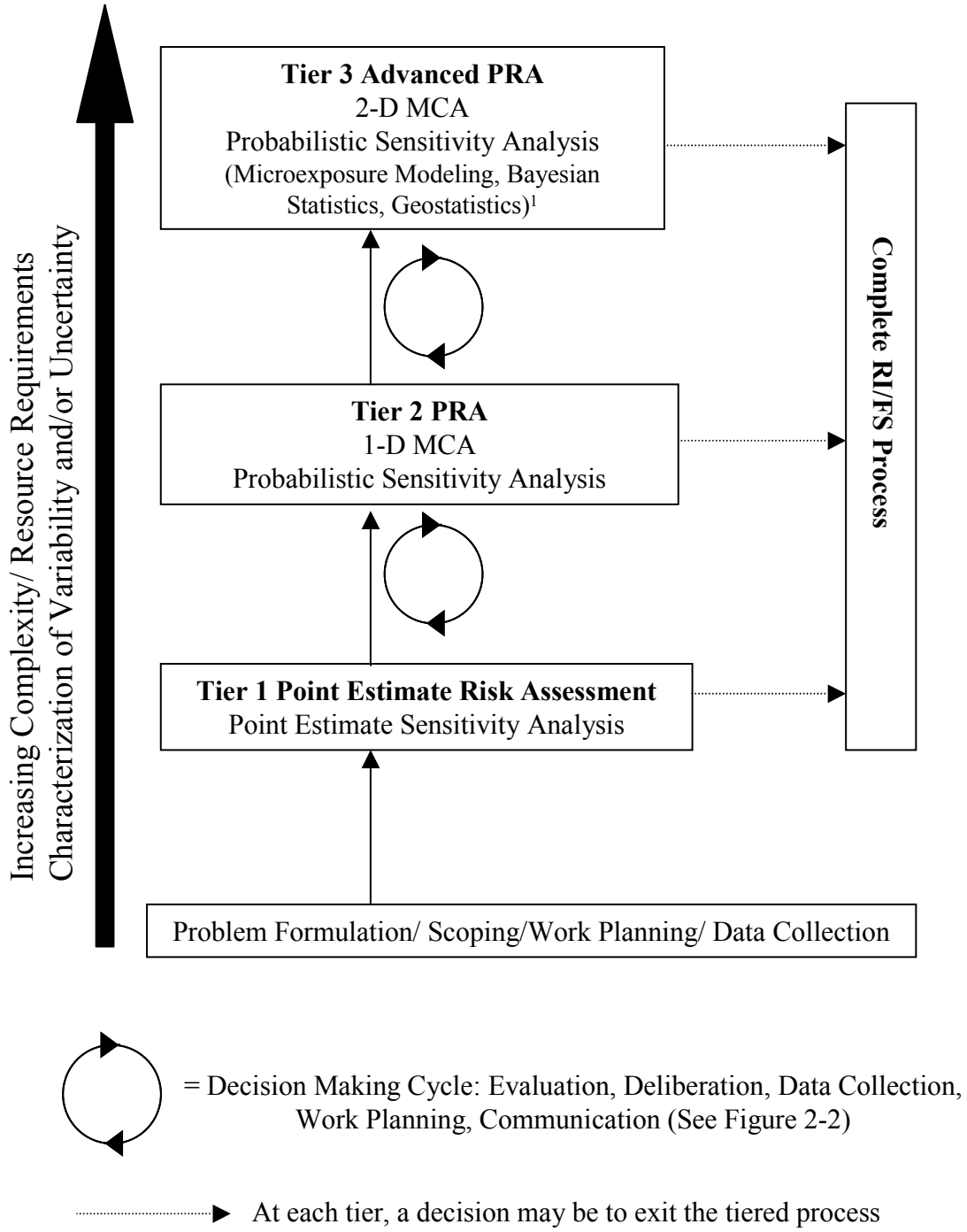


Figure 2-1. Schematic Diagram of Tiered Approach.

[1] Examples of advanced methods for quantifying temporal variability, spatial variability, and uncertainty (see Appendix D)

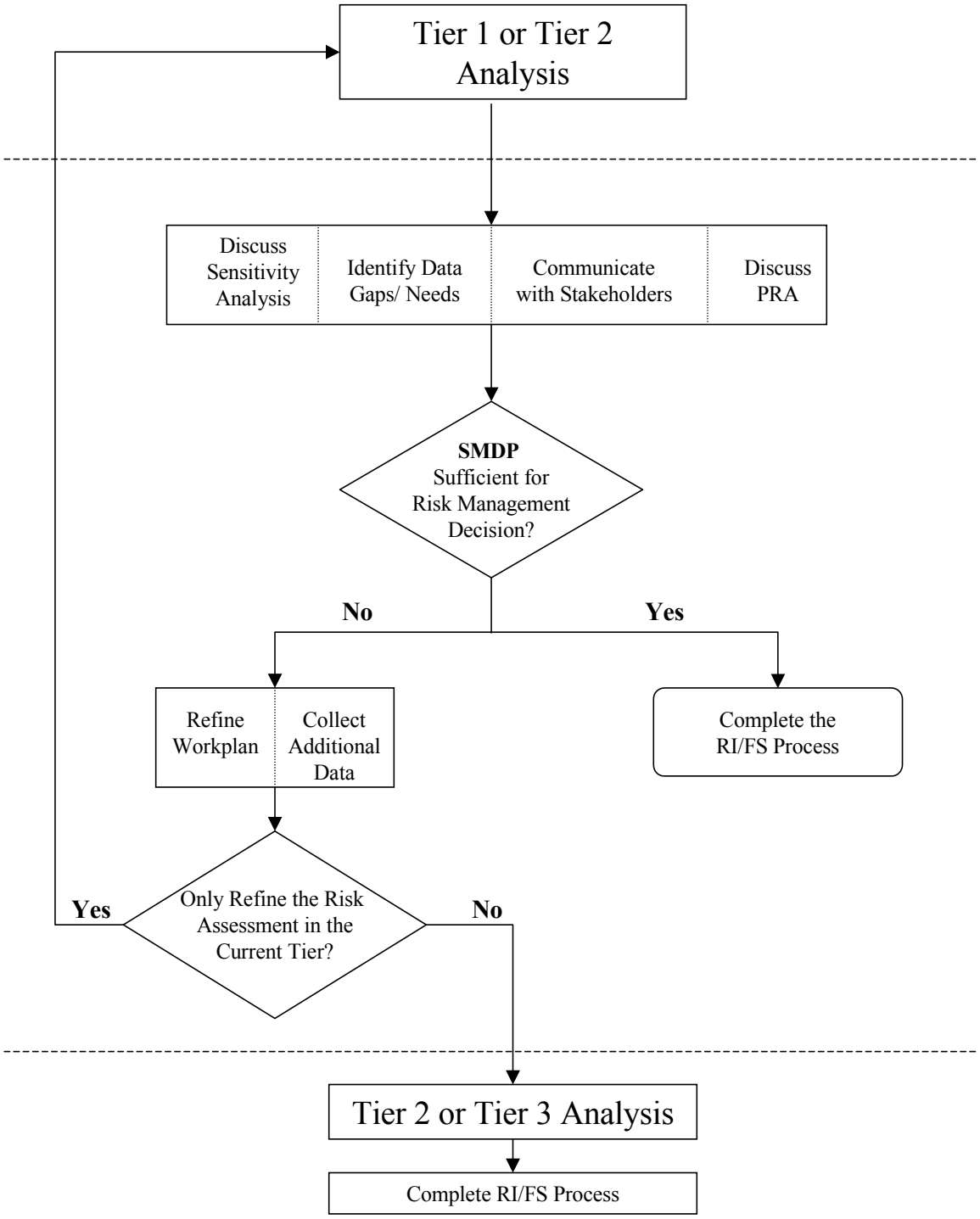


Figure 2-2. Schematic diagram of deliberation/decision cycle in the tiered process for PRA. SMDP refers to a scientific/management decision point, which implies that the decision involves consideration of not only the risk assessment, but also Agency policy, stakeholder concerns, cost, schedule, feasibility and other factors.

2.3.1 GETTING STARTED

All risk assessments should begin with problem formulation, scoping, preparation of a workplan (Section 2.1), and data collection. Problem formulation generally is an iterative process where reevaluation may occur as new information and data become available. The RPM should convene a scoping meeting prior to any risk assessment activities. Depending on the site-specific factors, discussion of performing a PRA may be appropriate at this initial scoping meeting. Alternatively, this discussion may be more productive at a later stage of the tiered process.

The risk manager should initiate discussions with EPA staff and other stakeholders early in the process, well before planning a risk assessment. Early communication with risk assessors or other EPA staff can help the risk manager evaluate the adequacy of the current information and plan additional data-gathering activities. Early communication with communities and other stakeholders should establish trust and facilitate a successful remedial process (see Chapter 6 on risk communication).

Generally, once the appropriate steps have been taken to adequately formulate and identify the problem and complete a workplan (Section 2.1), data collection efforts towards the point estimate risk assessment may begin. The process for conducting a point estimate risk assessment (Tier 1) is documented elsewhere in various RAGS volumes and related Superfund risk assessment guidance documents (e.g., U.S. EPA, 1989, 2001).

2.3.2 TIER 1

Tier 1 consists of the well-established process for planning and conducting human health and ecological point estimate risk assessments. Typical elements of a Tier 1 risk assessment, as they relate to higher tiers, are presented in Exhibit 2-4. A more detailed discussion of these elements can be found in Chapters 3 and 4 and Appendix A (Sensitivity Analysis).

EXHIBIT 2-4

TYPICAL ELEMENTS OF TIER 1 RISK ASSESSMENT

Analysis Tool - point estimate risk assessment

Variability Modeling - semi-quantitative, using central tendency exposure (CTE) and reasonable maximum exposure (RME) estimates as input variables

Uncertainty Modeling - semi-quantitative using confidence limits on certain point estimates (e.g., concentration term)

Sensitivity Analysis - point estimate calculation of percentage contribution of exposure pathways, for both CTE and RME risk. Systematically vary one input variable at a time across a plausible range and rank inputs based on sensitivity ratios or sensitivity scores.

Risk-Based Decision-Making Output - point estimate of risk—*Does the point estimate exceed the risk level of concern?*

A more detailed schematic presentation of the tiered process, showing the various elements of the deliberation/decision cycle and their linkage to higher tiers is shown in Figure 2-2. The two main factors to consider when determining whether the results of a risk assessment are sufficient for decision making are: (1) the results of a comparison of the risk estimate with the risk level of concern; and (2) the level of confidence in the risk estimate.

In Tier 1, comparison of risk estimates with risk levels of concern is relatively straightforward, since the outcome of a point estimate risk assessment is a single estimate of risk that either will exceed or not exceed the risk level of concern. Evaluating confidence in the Tier 1 risk estimates is more difficult because quantitative measures of uncertainty often are not easily obtained from a point estimate analysis. Uncertainty arises from two main

sources: (1) uncertainty in the inputs to the risk equations that stems from lack of knowledge (data gaps), and (2) uncertainty in the accuracy of the point estimate that stems from the mathematical simplifications that are inherent in point estimate computations.

There are usually many sources of uncertainty in the values used to calculate risk. One of the most familiar (but not always the most significant) is uncertainty in environmental concentration values of contaminants. This source of uncertainty is usually accounted for by calculating a 95% upper confidence limit (95% UCL) for the mean concentration in the exposure equation (U.S. EPA, 1992b). Chapter 5, Appendix C, and Appendix D provide more complete discussions of policies and methods for quantifying uncertainty in the exposure point concentration. Uncertainties in other variables in the risk equations (intake rates, exposure frequency and duration, toxicity factors, etc.) may also be significant, and are often addressed by choosing inputs that are more likely to yield an overestimate than an underestimate of risk. These sources of uncertainty are usually addressed qualitatively, by providing a discussion of the likely direction and magnitude of the error that may be associated with the use of the specific inputs (U.S. EPA, 1989). Stakeholders can provide useful information about uncertain variables and sources for site-specific data. This is an important reason to ensure that stakeholders are given the opportunity to review the risk assessment and be involved in the process.

Decision Alternatives

The evaluation of the point estimate risk assessment will yield one of two outcomes: (1) sufficient for risk management decisions; or (2) insufficient for risk management decisions. If the risk manager views the results of the point estimate risk assessment as sufficient for risk management decision making, the risk manager can exit the tiered approach and complete the RI/FS process (Figure 2-2). Depending on site-specific information, the results may support a decision for "no further action" or for a "remedial action." A "no further action" decision may result when the risk estimate is clearly below the level of concern (e.g., the National Oil and Hazardous Substances Pollution Contingency Plan (NCP) risk range of 1E-04 to 1E-06) and confidence in the risk estimate is high. A decision for remedial action may result when a national standard (e.g., maximum contaminant levels (MCLs) applied to groundwater) may be exceeded, or when the risk is clearly above the level of concern (e.g., the NCP risk range of 1E-04 to 1E-06) and confidence in the risk estimate is high. The decision for a specific remedial action involves consideration of the NCP's nine criteria for remedial decisions (U.S. EPA, 1990) and other site-specific factors.

An alternative conclusion would be that the results of the point estimate risk assessment are not sufficient for risk management decision making. For example, results may not be sufficient when the risk estimate is within the NCP risk range of 1E-04 to 1E-06 and confidence in the risk estimate is low. In this case, the risk manager should not exit the tiered approach. Instead, appropriate steps should be taken to increase the confidence that a management decision is protective. These steps may include discussing the point estimate sensitivity analysis, identifying data gaps, communicating with stakeholders (e.g., to obtain site-specific information), discussing the potential value of conducting a PRA (or a more advanced probabilistic analysis), work planning, and additional data collection (see Figure 2-2).

A sensitivity analysis can be a valuable component of the evaluation of a risk assessment. Sensitivity analysis can identify important variables and pathways that may be targets for further analysis and data collection. The type of information provided by a sensitivity analysis will vary with each tier of a PRA. Several methods are available at each tier, and the results of the analysis can vary greatly depending on the methods used. A comprehensive discussion of these methods is presented in Appendix A and briefly summarized here. Sensitivity analysis in Tier 1 will usually involve relatively

simple methods and will not involve Monte Carlo simulation. A typical approach would be to calculate the relative contributions of individual exposure pathways to the point estimate of risk. A more complex approach involves selecting values from a plausible range for a specific input variable to the exposure or risk equation and to use these values (i.e., low-end estimate and high-end estimate) to calculate corresponding point estimates of risk. The sensitivity of the risk estimate to each variable is then evaluated by calculating a sensitivity ratio, which is simply the percentage change in the risk estimate divided by the percentage change in the input variable value (see Appendix A, Section A.2.1.3, Sensitivity Ratios).

The sensitivity ratio (SR) approach is typically applied to one variable at a time because jointly varying point estimates for multiple variables can be cumbersome (see Chapter 3, Table 3-2 for an example of two jointly varied inputs). Information provided by the SR approach is generally limited to bounding estimates of risk based on small deviations and/or plausible ranges of point estimates for inputs. Because the point estimate approach does not generate a distribution of risk, SRs cannot provide quantitative information about the relative contributions of input variables to the variance in risk or the uncertainty in selected percentile of the risk distribution. This limitation of the SR approach may be particularly important if the ranking of input variables may change depending on the percentile range that is evaluated. For example, in a probabilistic analysis, the soil ingestion rate variable may contribute most to the variability in risk across the entire risk distribution, but the exposure duration may be the driver in the high-end ($> 90^{th}$ percentile) of the risk distribution, where the RME risk is defined. In addition, for standard product-quotient risk equations, the SR approach also has difficulty distinguishing the relative importance of exposure variables in the risk equation. Appendix A presents a hypothetical example to illustrate why this happens for the common risk equations. An improvement over the SR approach, called Sensitivity Score, involves weighting each ratio by the variance or coefficient of variation of the input variable when this information is available. In general, the most informative sensitivity analysis will involve Monte Carlo techniques (see Appendix A, Table A-1). Potential strengths and weaknesses of sensitivity analysis methods may be an important factor in deciding whether or not to conduct a probabilistic analysis in Tier 2.

Once data gaps have been identified, steps may be taken to gather additional data and revise the point estimates of risk based on these data. As with any data collection effort, the data quality objectives (DQO) process should be followed to obtain samples appropriate for the risk assessment and sufficient to support the remedial decision (U.S. EPA, 1992a, 1993, 1994, 2000a). The deliberation and decision cycle (Figure 2-2) should then be reiterated to determine if the refined risk assessment is sufficient to support risk management decisions. The collection of additional data may also provide a compelling reason to consider moving to Tier 2 and conducting a PRA. If, during the PRA discussions, it is determined that information from a PRA may influence the risk management decisions, PRA may be warranted. This iterative process of collecting data, recalculating point estimates, and reconsidering the potential value of PRA may continue until sufficient data are available to support risk management decisions, or data collection efforts are not possible due to resource constraints. For example, soil ingestion rate data may be limited to a few studies with small sample sizes, but a new soil ingestion study may be prohibitively expensive, time consuming, or difficult to conduct in a manner that will reduce the uncertainty in the risk estimate. Uncertainty due to data quantity is not necessarily a reason to exit the tiered process at Tier 1.

In cases where there is uncertainty in selecting a probability distribution because of small sample sizes, it may be informative to develop a preliminary probability distribution such as a triangular or uniform (see Appendix B, Section B.2.0). These preliminary distributions will contribute to the variability in the risk estimate, and can therefore be included in the probabilistic sensitivity analysis. Results of Monte Carlo simulations that include one or more preliminary distributions may lead to several

alternative decisions. If the sensitivity analysis suggests that the risk estimate is relatively insensitive to the variable described with the distribution, then the uncertainty associated with the choice of a distribution should not affect the risk management decision process using the tiered approach (e.g., choice of RME percentile, derivation of a PRG). In other words, the choice would be to continue with the tiered process. If, however, the variables described by preliminary distribution are important sources of variability or uncertainty in the risk estimate, then this information should be presented in the scientific management decision point (see Figure 2-2). The uncertainty may be sufficiently important in the risk management decision to warrant additional data collection efforts. Conversely, it may be necessary to exit the tiered process if the uncertainty cannot be reduced. Although the tiered process may be stopped at this point, it can still be informative to present the results from the PRA. For example, information about uncertainty may affect the choice of the percentile used to characterize the RME risk. In addition, it may be appropriate to weight the results of the point estimate analysis more heavily in the risk management decision when uncertainty in the PRA is high. Further guidance on appropriate choices for distributions based on the information available to characterize variability is given in Appendix B.

PRA also may be warranted if it would be beneficial to know where on the risk distribution the point estimate lies. An example of this would be a risk estimate that is within the NCP risk range of 1E-04 to 1E-06. The assessment may be sufficient to support risk management decisions if it could be shown that the point estimate of risk lies sufficiently high in the risk distribution. For example, a "no further action" decision may be strengthened if the point estimate is at the 99th percentile of the risk distribution, if risks in lower percentiles of the RME risk range are below the NCP risk range, and if there is high confidence in the risk result. This type of evaluation can be conducted using PRA techniques.

Even if the RME point estimate of risk exceeds the risk level of concern, and PRA is not needed to confirm this result, information from a PRA can be helpful in determining a strategy for achieving a protective preliminary remediation goal (PRG). A detailed discussion of the use of PRA in setting remediation action levels is given in Chapter 5. The advantages and disadvantages of the point estimate approach and PRA are presented in Chapter 1 (Exhibits 1-5 and 1-6).

2.3.3 TIER 2

Tier 2 of the tiered approach to risk assessment will generally consist of a simple probabilistic approach such as one-dimensional Monte Carlo analysis (1-D MCA). A 1-D MCA is a statistical technique that may combine point estimates and probability distributions to yield a probability distribution that characterizes variability or uncertainty in risks within a population (see Chapter 1). Guidance for selecting and fitting distributions is presented in Appendix B. Typical elements of a Tier 2 risk assessment, as they relate to higher and lower tiers are presented in Exhibit 2-5. A more detailed discussion of these elements can be found in Chapters 3 and 4, and Appendix A (Sensitivity Analysis).

While most of the Tier 2 assessments are expected to use 1-D MCA to characterize variability in risk, sometimes a 1-D MCA of uncertainty may be of interest. For example, as suggested in Exhibit 2-5, a probability distribution for uncertainty in the arithmetic mean or median (i.e., 50th percentile) for selected input variables may be specified in a 1-D MCA to yield a probability distribution for uncertainty for the central tendency risk estimate. However, as most Tier 2 assessments are expected to combine input distributions for variability, this guidance focuses on 1-D MCA for characterizing variability in the risk estimate.

EXHIBIT 2-5

TYPICAL ELEMENTS OF TIER 2 RISK ASSESSMENT

Analysis Tool - 1-D MCA

Variability Modeling - full characterization of variability in risk using PDFs or PMFs for input variables

Uncertainty Modeling - semi-quantitative estimate of uncertainty using iterative 1-D MCA simulations of variability, or a single 1-D MCA of uncertainty in the CTE risk

Sensitivity Analysis - varying multiple variables with probability distributions gives a quantitative ranking (e.g., correlation coefficient) of the relative contributions of exposure pathways and variables to CTE or RME risk

Risk-Based Decision-Making Output - risk distribution for variability: *Does the risk level of concern fall within an acceptable range on the risk distribution (i.e., RME range)?* Also, risk distribution for uncertainty: *What is the 90% confidence interval for the CTE risk?*

Decision Alternatives

Generally, the three main questions to consider when determining whether the results of a 1-D MCA are sufficient for risk management decisions are: (1) What is the RME risk range and how does it compare to the level of concern?; (2) Where does the point estimate risk lie on the risk distribution?; and (3) What is the level of confidence in the risk estimate? In Tier 2, similar to the point estimate approach, the level of confidence in a single 1-D MCA risk distribution is generally addressed in a qualitative or semi-quantitative way. As discussed in Chapter 1 (Section 1.2.4) and Chapter 3 (Section 3.4.1), one should avoid developing input distributions to a PRA model that yield a single risk distribution that intermingles, or represents both variability and uncertainty. In Tier 2, the preferred approach for characterizing uncertainty in the risk estimate is to perform multiple 1-D MCA simulations (of variability), which uses a different point estimate for uncertainty for one or more parameters, combined with probability distributions for variability for one or more variables. Chapter 3 (see Table 3-2 and Figures 3-3 and 3-4) presents an example of iterative 1-D MCA simulations using combinations of point estimates characterizing uncertainty for two variables. More advanced PRA techniques such as two-dimensional Monte Carlo analysis (2-D MCA), in which distributions for variability and uncertainty are propagated separately through an exposure model, can be undertaken in Tier 3 (Appendix D).

In order to use a PRA to determine if risks are unacceptable and to develop preliminary remediation goals (PRGs) that are protective of the RME individual (see Chapter 5), a single point from the RME risk range should be selected (e.g., 95th percentile). In general, this can be accomplished by selecting an estimate within the RME risk range based on the level of confidence in the output of the 1-D MCA. Uncertainty in risk estimates may be quantified or reduced by considering site-specific factors, biological data, and toxicity data. Stakeholders can provide useful information about uncertain variables and sources for site-specific data. More detailed guidance for choosing a percentile value within the RME range is provided in Chapter 7.

The evaluation of the risk assessment in a 1-D MCA in Tier 2 will yield one of two outcomes: (1) sufficient for risk management decisions; or (2) insufficient for risk management decisions. If determined to be sufficient, the risk manager can exit the tiered approach and complete the RI/FS process. The results of a 1-D MCA may support a decision for "no further action" or for a "remedial action." A "no further action" decision may result when the RME risk range (or a specified point in the RME risk range) is clearly below the level of concern (e.g., Hazard Index=1) and confidence in the risk distribution is high. A decision for remedial action may result when a national standard (e.g., MCLs applied to groundwater) may be exceeded, or when the RME risk range (or a specified point in the RME risk range) is clearly above the level of concern and confidence in the risk distribution is high. The decision for a specific remedial action involves consideration of the NCP's nine evaluation criteria for remedial decisions (U.S. EPA, 1990; see Chapter 1) and other site-specific factors.

An alternative conclusion at the end of a Tier 2 analysis would be that the results of the 1-D MCA are not sufficient for risk management decisions. There are several factors that might support this conclusion:

(1) The RME risk range is close to the NCP risk range and confidence in the risk distribution is low. In this case, the risk manager might decide to not exit the tiered approach, and instead continue taking appropriate steps to increase the confidence in the risk estimate.

(2) Uncertainty is high and it is believed that more than one variable is a major contributor to the uncertainty in the risk estimate. It can be difficult to explore uncertainty in more than one variable using 1-D MCA simulations of variability, even using iterative approaches discussed in Chapter 3 (Section 3.4.1).

(3) Results of the point estimate risk assessment differ significantly from the results of the 1-D MCA. While the RME risk estimates are not expected to be identical, typically the RME point estimate will correspond with a percentile value within the RME range (i.e, 90^{th} to 99.9^{th} percentile) of the risk distribution. If the RME point estimates fall outside this range, further steps may be warranted to evaluate the choices for input variables—both the RME point estimates, and the probability distributions and parameters (including truncation limits) for the 1-D MCA.

The deliberation/decision cycle (Figure 2-2) between Tier 2 and Tier 3 is similar to the cycle between Tier 1 and 2 and includes discussing the Tier 2 probabilistic sensitivity analysis, identifying data gaps, communicating with stakeholders (e.g., to obtain site-specific information), discussing the potential value of further analysis with probabilistic methods, work planning, and additional data collection. As with the Tier 1 assessment, additional data collection should follow the DQO process (U.S. EPA, 1992a, 1993, 1994, 2000a) and point estimates of risk should be revisited with the new data. The deliberation/decision cycle is an iterative process in which the level and complexity of the analysis increases until the scope of the analysis satisfies decision-making needs. This iterative process should continue until sufficient data are available to support risk management decisions. As in all tiers, stakeholder involvement should be encouraged. Once a 1-D MCA for variability or uncertainty is completed and is available for review and interpretation, a stakeholder meeting should be convened. Interested stakeholders should be given the opportunity to review the 1-D MCA and provide comments. Communication issues specific to PRA are discussed in Chapter 6 (Risk Communication).

In addition to identifying data gaps, consideration for a refined 1-D MCA or more advanced PRA techniques may begin as a means of determining what benefits they may confer to the decision-making

process. If, during further discussions of PRA, it is determined that information from a more advanced PRA may influence the risk management decision, the use of an advanced PRA may be warranted. If additional data have been collected, the point estimate and 1-D MCA should be refined. Specifically, an advanced PRA may be warranted if it would be beneficial to characterize uncertainty in more than one variable at a time. A 2-D MCA can simultaneously characterize variability and uncertainty in multiple variables and parameter estimates. The decision to employ such advanced methods should be balanced with considerations of resource constraints and the feasibility of reducing uncertainty in a given variable. A detailed discussion of advanced PRA methods, including 2-D MCA, is provided in Appendix D.

2.3.4 TIER 3

Tier 3 of the tiered approach to risk assessment consists of advanced PRA methods, such as 2-D MCA, Microexposure Event Analysis (MEE), geostatistical analysis of concentration data, and Bayesian statistics. Typical elements of a Tier 3 risk assessment are presented in Exhibit 2-6. A more detailed discussion of these elements is given in Appendix D. As in other tiers, Tier 3 includes an iterative process of deliberation and decision making in which the level and complexity of the analysis increases until the scope of the analysis satisfies decision-making needs. As in all tiers, stakeholder involvement is encouraged.

EXHIBIT 2-6

TYPICAL ELEMENTS OF TIER 3 RISK ASSESSMENT

Analysis Tool - 2-D MCA, MEE, geostatistics, and Bayesian statistics

Variability Modeling - full characterization using PDFs or PMFs for input variables

Uncertainty Modeling - quantitative, segregating uncertainty from variability, and associated with multiple variables simultaneously

Sensitivity Analysis - varying parameters of probability distributions to identify and rank order parameter uncertainty with the same sensitivity analysis methods used for Tier 2 (see Appendix A). Also, explore alternative choices of probability distributions and sources of model uncertainty.

Risk-based Decision-Making Criteria - risk distribution for variability with confidence limits—*Does the risk level of concern fall within an acceptable range on the risk distribution (i.e., RME range), and with an acceptable level of uncertainty?*

Generally, the various elements of the deliberation/decision cycle for Tier 3 are the same as those for Tier 1 and 2 (Figure 2-2). An advanced PRA would be conducted and made available for review to the risk manager and stakeholders. The risk manager must determine if the results of the advanced PRA are sufficient for risk management decision making. Issues to consider when making this determination are similar to those identified for evaluating point estimate risk results and 1-D MCA results, and focus on evaluating the sources and magnitude of uncertainty in relation to the established risk level of concern. If the results are sufficient for risk management decisions, the risk manager may exit the tiered approach and complete the RI/FS process. If the results are not found to be sufficient for risk management decisions, data gaps should be identified and if additional data are collected, all stages of the risk assessment, including the advanced PRA, the 1-D MCA, and the point estimate risk assessment, should be refined. Alternatively, additional advanced PRA methods may be explored. Refer to Appendix D for a discussion of more advanced PRA techniques. Overall, analysis should continue within Tier 3 until sufficiently informed risk management decisions can be made.

2.3.5 FLEXIBILITY IN DEFINING TIERS

The assignment of specific analytical tools to Tiers 1, 2, and 3 (Figure 2-1 and Exhibits 2-4 through 2-6) results in generalizations that may not be applicable to all site assessments. Upon completion of the deliberation phase between Tier 1 and Tier 2, the conclusion may be that analytical tools in Tier 3 would be applicable and beneficial for addressing decision making issues. For example, geospatial modeling may be beneficial for improving estimates of uncertainty in the exposure point concentration or in designing field sampling plans to further reduce uncertainty. An improved estimate of the 95% UCL from geospatial analysis (shown in Exhibit 2-6 as a Tier 3 analytical tool) would then be integrated into a Tier 2 assessment, or the complete distribution for uncertainty in the mean concentration could be incorporated into a 2-D MCA in Tier 3. Flexibility in defining the level of complexity of the analysis used in a given tier is essential to accommodating the wide range of risk assessment issues likely to be encountered. An important benefit gained from use of the tiered approach is to ensure a deliberative process in the advancement of the assessment to higher levels of complexity. It is far more important that a deliberative process take place and be documented, than it is to constrain a set of analytical tools to a specific tier.

REFERENCES FOR CHAPTER 2

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment)*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991a. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume 1: Human Health Evaluation Manual* (HHEM), Part B, Development of Risk-Based Preliminary Remediation Goals. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333

U.S. EPA. 1991b. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9355.0-30.

U.S. EPA. 1992a. Guidance on Data Usability in Risk Assessment. Part A. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7.09A. NTIS PB92-96336.

U.S. EPA. 1992b. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive 9285.7-081.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund: Interim Final Guidance*. Office of Research and Development, Washington, DC. EPA/540/R-93/071.

U.S. EPA. 1994. *Guidance for the Data Quality Objectives Process (EPA QA/G-4).* Office of Research and Development, Washington, DC. EPA/600/R-96/055. September.

U.S. EPA. 2000a. *Data Quality Objectives Process for Hazardous Waste Site Investigations*. Office of Environmental Information, Washington, DC. EPA/600/R-00/007. January.

U.S. EPA. 2000b. *Peer Review Handbook: 2nd Edition.* Science Policy Council. Washington, DC. EPA/100/B-00/001. December.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund*: *Volume I. Human Health Evaluation Manual, Part D: Standardized Planning, Reporting, and Review of Superfund Risk Assessments.* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285-47. December.

CHAPTER 3

USING PROBABILISTIC ANALYSIS IN HUMAN HEALTH ASSESSMENT

3.0 INTRODUCTION

This chapter outlines how probabilistic analysis may be applied to human health risk assessments in the Environmental Protection Agency's (EPA) Superfund program. The paradigm for human health risk assessment as described in EPA's *Risk Assessment Guidance for Superfund* (U.S. EPA, 1989), includes data collection/evaluation in addition to exposure and toxicity assessment and risk characterization. Although the strategies and methods used in collecting and analyzing data can significantly impact the uncertainty in a risk estimate, they are issues relevant to risk assessment in general, and are addressed in other guidance documents, such as EPA's *Guidance for Data Useability in Risk Assessment* (U.S. EPA, 1992b). RAGS Volume 3: Part A focuses on a tiered approach for incorporating quantitative information on variability and uncertainty into risk management decisions.

3.1 CHARACTERIZING VARIABILITY IN EXPOSURE VARIABLES

Exhibit 3-1 gives the general equation used for calculating exposure, often expressed as an average daily intake. In a point estimate approach, single values (typically a mixture of average and high-end values) are input into the equation. In probabilistic risk assessment (PRA), the only difference is that a probability distribution, rather than single value, is specified for one or more variables. A Monte Carlo simulation is executed by repeatedly selecting random values from each of these distributions and calculating the corresponding exposure and risk. For the majority of PRAs, it is expected that probability distributions will be used to characterize inter-individual variability, which refers to true heterogeneity or diversity in a population. Thus, variability in daily intake, for example, can be characterized by combining multiple sources of variability in exposure, such as ingestion rate, exposure frequency, exposure duration, and body weight. Variability in chemical concentrations (Chapter 5 and Appendix C) and the toxicity term in ecological risk assessment (Chapter 4) may also be considered in risk calculations.

EXHIBIT 3-1
GENERAL EQUATION FOR EXPOSURE

$$I = \frac{C \times CR \times EF \times ED}{BW \times AT} \qquad \text{Eq. 3-1}$$

where,

I	=	daily intake
C	=	contaminant concentration
CR	=	contact rate (ingestion, inhalation, dermal contact)
EF	=	exposure frequency
ED	=	exposure duration
BW	=	body weight
AT	=	averaging time

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3

95% UCL for mean - The one-sided 95% upper confidence limit for a population mean; if a sample of size (*n*) was repeatedly drawn from the population, the 95% UCL will equal or exceed the true population mean 95% of the time. It is a measure of uncertainty in the mean, not to be confused with the 95th percentile (see below), which is a measure of variability. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged.

95th percentile - The number in a distribution that is greater than 95% of the other values of the distribution, and less than 5% of the values. When estimated from a sample, this quantity may be equal to an observed value, or interpolated from among two values.

Arithmetic Mean (AM) - A number equal to the average value of a population or sample. Usually obtained by summing all the values in the sample and dividing by the number of values (i.e., sample size).

Assessment Endpoint - The specific expression of the population or ecosystem that is to be protected. It can be characterized both qualitatively and quantitatively in the risk assessment.

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in the population, usually considered to be the arithmetic mean or median of the risk distribution.

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

CTE Risk - The estimated risk corresponding to the central tendency exposure.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF or PMF, gives the cumulative probability of occurrence for a random independent variable. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*.

Exposure Point Concentration (EPC) - The average chemical concentration to which receptors are exposed within an exposure unit. Estimates of the EPC represent the concentration term used in exposure assessment.

Frequency Distribution/Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

High-end Risk - A risk descriptor representing the high-end, or upper tail of the risk distribution, usually considered to be equal to or greater than the 90th percentile.

Low-end Risk - A risk descriptor representing the low-end, or lower tail of the risk distribution, such as the 5th or 25th percentile.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Probability Density Function (PDF) - A function representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Probability Mass Function (PMF) - A function representing the probability distribution for a discrete random variable. The mass at a point refers to the probability that the variable will have a value at that point.

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3—*Continued*

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic *r* that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic *r* that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Target Population - The set of all receptors that are potentially at risk. Sometimes referred to as the "population of concern". A sample population is selected for statistical sampling in order to make inferences regarding the target population (see Appendix B, Section B.3.1, Concepts of Populations and Sampling).

Figure 3-1 shows a hypothetical example of an input distribution for drinking water ingestion rate. Assume that survey data for drinking water ingestion rates were compiled in order to select and fit a probability distribution. One of the first steps in exploring the data set may be to plot a frequency distribution. In the graph, the height of the bars (the y-axis) represents the relative frequency of ingestion rates in the population and the spread of the bars (the x-axis) is the varying amounts of water ingested (L/day). Since ingestion rate is a continuous random variable, the probability distribution can also be represented graphically with a probability density function (PDF). Assume that the following parameters are estimated from the sample: arithmetic mean=1.36, standard deviation=0.36, geometric mean=1.31, and geometric standard deviation=1.30. These parameter estimates may be used to define a variety of probability distributions, including a 2-parameter lognormal distribution. The fit of the lognormal distribution can be evaluated by visual inspection using the PDF given by Figure 3-1, or by a lognormal probability plot (see Appendix B).

The y-axis for a PDF is referred to as the *probability density*, where the density at a point on the x-axis represents the probability that a variable will have a value within a narrow range about the point. This type of graph shows, for example, that there is a greater area under the curve (greater probability density) in the 1-2 L/day range than 0-1 L/day or 2-3 L/day. That is, most people reported consuming 1-2 L/day of drinking water. By selecting a lognormal distribution to characterize inter-individual variability, we can state more precisely that 1 L/day corresponds to the 15th percentile and 2 L/day corresponds to the 95th percentile, so approximately 80% (i.e., 0.95–0.15=0.80) of the population is likely to consume between 1 and 2 L/day of drinking water.

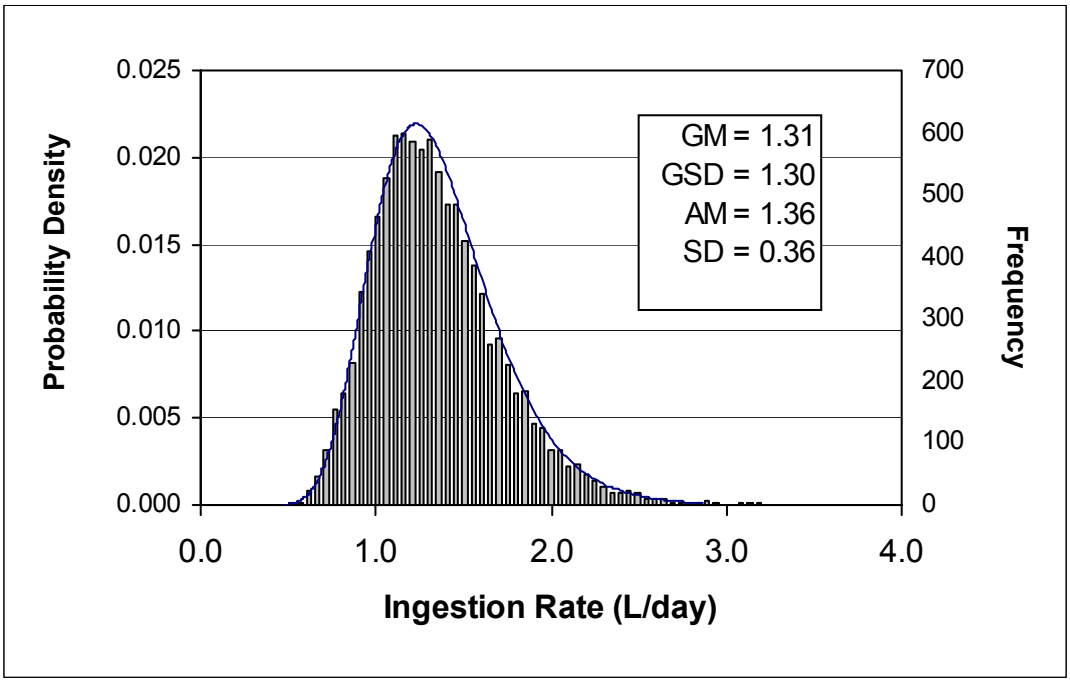


Figure 3-1. Example of a frequency distribution for adult drinking water ingestion rates, overlaid by a graph of the probability density function (PDF) for a lognormal distribution defined by the sample statistics. The distribution represents inter-individual variability in water intakes and is characterized by two parameters. Typically, the geometric mean (GM) and geometric standard deviation (GSD), or the arithmetic mean (AM) and arithmetic standard deviation (SD) are presented to characterize a lognormal distribution.

3.1.1 DEVELOPING DISTRIBUTIONS FOR EXPOSURE VARIABLES

When site-specific data or representative surrogate data are available, a probability distribution can be fit to that data to characterize variability. Appendix B describes how to fit distributions to data, how to assess the quality of the fit and discusses topics such as the sensitivity of the tails of the distribution to various PDFs, and correlations among variables. Many of the issues discussed below regarding the use of site-specific data or surrogate data are relevant to both point estimate risk assessment and PRA.

For the majority of the exposure variables, such as exposure duration, water intake rates, and body weight, site-specific data will not be available. The risk assessor will have to either select a distribution from existing sources, or develop a distribution from published data sets and data summaries. Examples of sources for these distributions and data sets are EPA's *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), Oregon Department of Environmental Quality's *Guidance for Use of Probabilistic Analysis in Human Health Risk Assessment* (Oregon DEQ, 1998), and the scientific literature. An appropriate PDF should be determined in collaboration with the regional risk assessor. The process by which PDFs are to be selected and evaluated should be described in the workplan. EPA's Superfund program is in the process of developing a ranking methodology to evaluate data representativeness relevant to various exposures scenarios. Following peer review and project completion, the results will be posted on EPA Superfund web page.

> ☞ *At this time, EPA does not recommend generic or default probability distributions for exposure variables.*

Regardless of whether a PDF is derived from site-specific measurements or obtained from the open literature, the risk assessor should carefully evaluate the applicability of the distribution to the target population at the site. The distribution selected should be derived from the target population or from a surrogate population that is representative of the target population at the site. For example, a distribution based on homegrown vegetable consumption in an urban population would not be representative for a farming population in the Midwest. If such a distribution were to be used, (and no other data were available), the uncertainty and bias that this PDF would impart to the risk estimate should be communicated to the risk decision makers.

For purposes of risk management decision making, the significance of not having site-specific data should be evaluated in the context of representativeness and sensitivity analysis. If published data are representative of the potentially exposed population, then site-specific data may be unnecessary. For example, body weights of children and adults have been well studied from national surveys and can generally be considered reasonable surrogates for use in site risk assessments. Furthermore, even if a variable is likely to vary among different exposed populations, it may not contribute greatly to the variance or uncertainty in risk estimates. In this case, surrogate data may also be used with confidence in the risk estimate. In addition, the PRA may be simplified by using point estimates instead of probability distributions for the "less sensitive" exposure variables. In part, the decision to use a point estimate in lieu of a probability distribution must balance the benefit of simplifying the analysis and the communication process (see Chapter 6), against the reduction (however small) in the variance of the risk distribution. The utility of sensitivity analysis in identifying the important factors in a risk estimate is discussed further below and in Appendix A.

It is also important to evaluate the sample design and sample size when deciding to apply a distribution to a specific site. Depending on the situation, a very large data set derived from a national

population may be more useful than a site-specific data set derived from a small, incomplete, or poorly designed study. Appendix B provides additional discussion on how to evaluate the data and studies that form the basis for a distribution. Often, the question arises regarding the appropriateness of combining data sets to derive a PDF. Before combining data sets, a careful evaluation should be made of the representativeness of the study populations, and the similarity in study designs and quality. In addition, statistical tests may be used to determine whether or not data sets are compatible with a common probability distribution (Hedges and Olkin, 1985; Stiteler et al., 1993). In general, risk assessors should be reluctant to combine data sets for the purpose of developing a PDF that characterizes variability. Due to the number of potential differences inherent in the study design, alternative data sets may provide a better measure of uncertainty in the probability distribution and parameter estimates, rather than a means of increasing the overall sample size for defining a single probability distribution. For example, if multiple data sets are available, a more informative approach may be to incorporate each data set into the PRA in a separate analysis, as a form of sensitivity analysis on the choice of alternative data sets.

Each probability distribution used in a Monte Carlo Analysis (MCA) should be presented with sufficient detail that the analysis can be reproduced (see Chapter 1, Section 1.4, Condition #2). This information may be presented in tabular and/or graphical summaries. Important information for a summary table would include a description of the distribution type (e.g., lognormal, gamma, etc.), the parameters that define the distribution (e.g., mean and standard deviation, and possibly upper and lower truncation limits for a normal distribution), units, and appropriate references (see Table 3-6, for example). The table should also indicate whether the distribution describes variability or uncertainty. The report should discuss the representativeness of the data and why a particular data set was selected if alternatives were available. Graphical summaries of the distributions may include both PDFs and cumulative distribution functions (CDFs), and should generally be used to document distributions that characterize site-specific data.

3.1.2 CHARACTERIZING RISK USING PRA

Quantitative risk characterization involves evaluating exposure (or intake) estimates against a benchmark of toxicity, such as a cancer slope factor or a noncancer hazard quotient. The general equation used for quantifying cancer risk from ingestion of contaminated soil is shown in Exhibit 3-3, and the equation for noncarcinogenic hazard is shown in Exhibit 3-4. A Hazard Index is equal to the sum of chemical-specific Hazard Quotients.

At this time, this guidance does not propose probabilistic approaches for dose-response in human health assessment and, further, *discourages undertaking such activities on a site-by-site basis*. Such activities require contaminant-specific national consensus development and national policy development (see Chapter 1, Section 1.4.1). Chapter 4 discusses methods for applying probabilistic approaches to ecological dose-response assessment.

The probabilistic calculation of risk involves random sampling from each of the exposure variable distributions. The output of this process is a distribution of risk estimates. When the calculation of risk (or any other model endpoint) is repeated many times using Monte Carlo techniques to sample the variables at random, the resulting distribution of risk estimates can be displayed in a similar fashion. The type of summary graph used to convey the results of a MCA depends on the risk management needs. For example, Chapter 1, Figure 1-3 shows how a PDF for risk might be used to compare the probabilistic estimate of the RME risk (e.g., 95th percentile) with a risk level of concern. This type of summary can also be used to effectively illustrate the relationship between the RME risk determined from point estimate and probabilistic approaches.

EXHIBIT 3-3

EQUATION FOR CANCER RISK

$$Risk = Dose \times CSF$$

Example for Soil Ingestion

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

where,

C	=	concentration in soil (mg/kg)	ED	=	exposure duration (years)
IR	=	soil ingestion rate (mg/day)	BW	=	body weight (kg)
CF	=	conversion factor (1E-06 kg/mg)	AT	=	averaging time (days)
EF	=	exposure frequency (days/year)	CSF	=	oral cancer slope factor $(mg/kg\text{-}day)^{-1}$

EXHIBIT 3-4

EQUATION FOR NONCANCER HAZARD QUOTIENT

$$Hazard\ Quotient = \frac{Dose}{RfD}\ or\ \frac{Concentration}{RfC}$$

where,

RfD	=	reference dose, oral or dermally adjusted (mg/kg-day)
RfC	=	reference concentration, inhalation ($\mu g/m^3$)

In addition, the CDF can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., cancer risk of 1E-04 or Hazard Index of 1). Figure 3-2 illustrates both the PDF and CDF for risk for a hypothetical scenario. Factors to consider when applying the PDF or CDF are discussed in Chapter 1, Exhibit 1-3. When in doubt about the appropriate type of summary to use, both the PDF and CDF should be provided for all risk distributions. At a minimum, each summary output for risk should highlight the risk descriptors of concern (e.g., 50^{th}, 90^{th}, 95^{th}, and 99.9^{th} percentiles). It can also be informative to include the results of the point estimate analysis—the risks corresponding to the central tendency exposure (CTE) and the reasonable maximum exposure (RME).

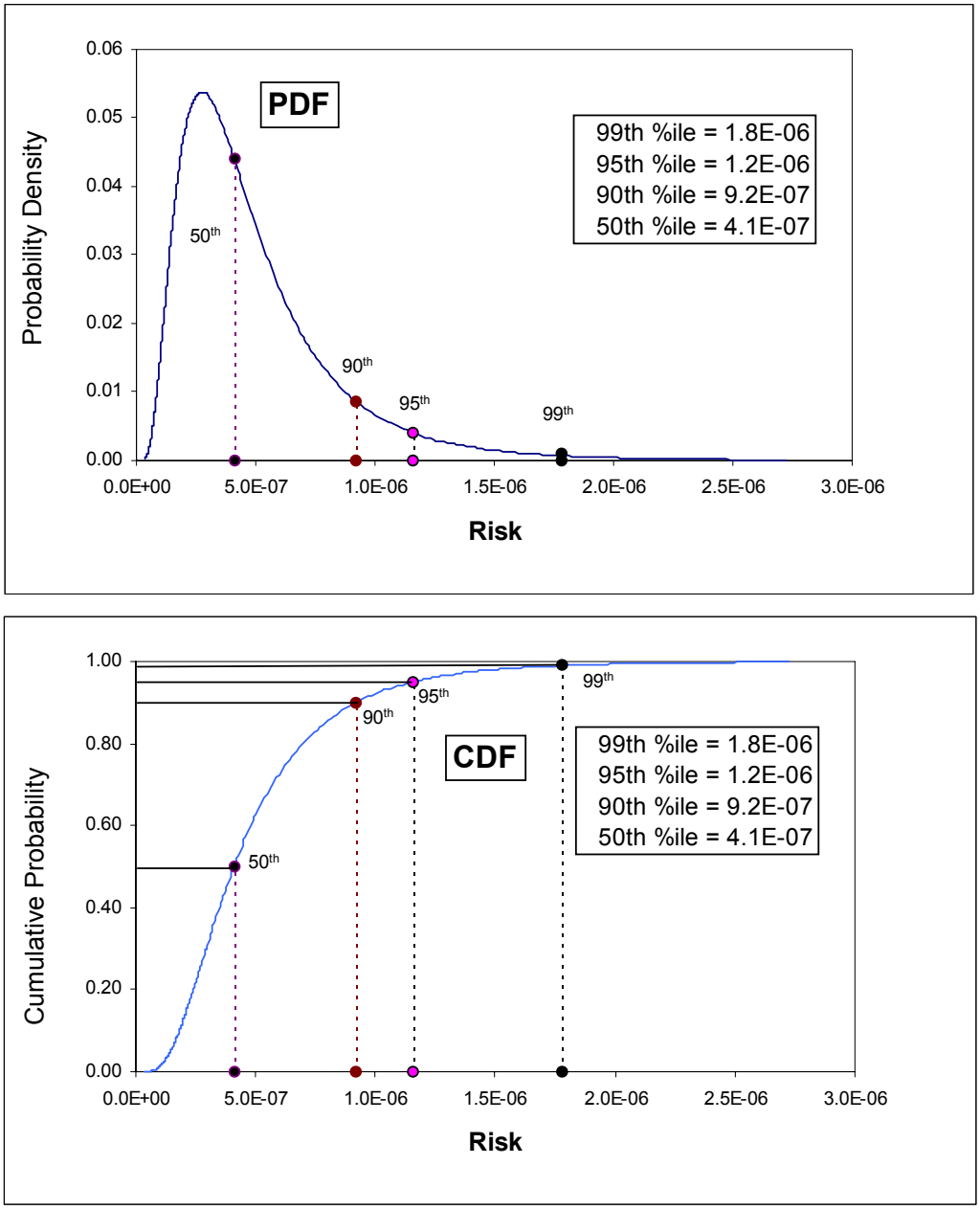


Figure 3-2. Hypothetical PRA results showing a PDF (top panel) and CDF (bottom panel) for cancer risk with selected summary statistics. The CDF rises to a maximum cumulative probability of 1.0. The CDF clearly shows that the level of regulatory concern chosen for this example (1E-06) falls between the 90th and 95th percentiles of the risk distribution.

3.2 ROLE OF THE SENSITIVITY ANALYSIS

Prior to conducting a PRA, it is worthwhile to review several points pertaining to the sensitivity analysis. As shown in Chapter 2 (Figures 2-1 and 2-2), sensitivity analysis can play an important role in decision making at each tier of the tiered process. Beginning with Tier 1, a point estimate for risk should be calculated prior to conducting a PRA. Based on the results of the point estimate, the risk assessor and risk decision makers should determine whether a probabilistic analysis will offer additional benefit. One factor in this decision may be the results of a sensitivity analysis. A primary objective of the sensitivity analysis is to determine which variables and pathways most strongly influence the risk estimate. At many Superfund sites, an estimate of cumulative risk considers contamination in multiple media, moving through multiple pathways and interacting with a number of receptors. Depending on the complexity of the site, and the modeling approaches, a risk assessment may involve one exposure pathway and few variables, or multiple pathways with many variables (e.g., multimedia fate and transport models). However, resources and time are often limited. The sensitivity analysis is invaluable in focusing these limited resources on the most influential variables and pathways.

Several methods for conducting sensitivity analysis are described in Appendix A. It is important to note that when a sensitivity analysis is performed and the major variables are identified, this does not mean that the less influential pathways and variables should be eliminated from the risk assessment. It means that because they are not major contributors to the variability or uncertainty in risk, they can be described with point estimates without affecting the risk management decision. If distributions are readily available for these less influential variables, one may use distributions. The key goal is to provide a comprehensive risk characterization that is scientifically credible and sufficient for risk decision making. The time and effort required to achieve various levels of complexity should be weighed against the value of the information provided to the risk managers.

Additionally, if a variable is specified as influential in the sensitivity analysis, this does not automatically mean that a distribution has to be developed for this variable. If the risk assessor feels that data are simply not sufficient from which to develop a distribution, then a plausible point estimate can be used. The risk assessor should be aware of a possible problem arising from using point estimates in the absence of data adequate to support a distribution. If a variable has the potential to significantly impact the risk outcome, and a very high-end or low-end point estimate is used in the PRA, this has the potential to right-shift or left-shift the final distribution of risk. Even though there might not be enough data to develop a distribution of variability for an influential variable, it would be prudent to communicate the importance of this data gap to the risk decision makers, and perhaps run multiple simulations with several plausible input distributions for that variable. Communication of this uncertainty may persuade the risk decision makers to collect additional data to better define the variable.

3.3 EXPOSURE POINT CONCENTRATION TERM

A brief discussion of the concentration term is provided below. A more complete discussion of the concentration term in PRA is provided in Appendix C. The reader is also referred to Chapter 5 on development of PRGs.

The major source of uncertainty in Superfund risk assessments is often incomplete knowledge of the concentration of one or more chemicals in various exposure media. In any risk assessment, the derivation of the concentration term will reflect assumptions about: (1) properties of the contaminant, (2) the spatial and temporal variability in contamination, (3) the behavior of the receptor, and (4) the time scale of the toxicity of the chemical(s).

Contaminant concentrations contacted by a receptor are likely to vary depending on the spatial variability of contamination and the movements of the receptor. Different individuals may be exposed to different concentrations based on inter-individual variability in activity patterns. If information regarding activity patterns is unavailable, receptors are typically assumed to exhibit random movement such that there is an equal probability of contacting any area within the exposure unit (EU). An EU is defined as the geographical area in which a receptor moves and contacts contaminated medium during the period of the exposure duration. In addition, in Superfund risk assessments, the toxicity criteria are often based on health effects associated with chronic exposure (e.g., lifetime risk of cancer following chronic daily intake over a period of 30 years). Hence, the most appropriate expression for the concentration term, for the majority of risk assessments, is one that characterizes the long-term average exposure point concentration within the EU.

> ☞ *The most appropriate expression of the exposure point concentration term for chronic exposure will characterize the long-term average concentration experienced by a receptor within the exposure unit.*

In point estimate risk assessments, the exposure point concentration term is usually calculated as the 95% upper confidence limit (95% UCL) of the arithmetic mean because of the uncertainty associated with estimating the true (i.e., population) mean concentration at a site. If the sampling density is sparse relative to the size of the EU, the uncertainty may be high due to the relatively small number of measurements available to estimate the mean concentration within the EU. The decision to use the upper confidence limit to define the concentration term introduces a measure of protectiveness by reducing the chance of underestimating the mean. Although there will be situations in which modeling variability in concentration will be the appropriate choice (e.g., non-random movement within an EU, acute exposure events, migration of groundwater contaminant plume, migration of fish, etc.), in most cases, characterization of the concentration term will focus on uncertainty. Appendix C provides a more complete discussion on characterizing both variability and uncertainty in the concentration term. Table 3-1 summarizes a number of appropriate methods for characterizing uncertainty in the parameter of an exposure variable, such as the arithmetic mean of the concentration term.

3.4 CHARACTERIZING UNCERTAINTY IN EXPOSURE VARIABLES

Uncertainty is described as a lack of knowledge about factors affecting exposure or risk. To evaluate regulatory options, risk assessors are expected to translate the available evidence, however tentative, into a probability of occurrence of an adverse health effect. Data from a sample or surrogate population are used to develop estimates of exposure and risk in a specific target population (see Section 3.1.4 and Appendix B, Section B.3.1). This extrapolation requires assumptions and inferences that have inherent strengths and limitations, and may bias the outcome of the risk estimate. For example, a common assumption in risk assessments for carcinogens is that a contaminant concentration within the boundaries of a hazardous waste site represents the concentration that a receptor is exposed to throughout the period of exposure, with the corresponding dose averaged over the course of a lifetime. This assumption may be conservative (i.e., result in overestimation of exposure) if it is unlikely that receptors will be exposed at the hazardous waste site for the entire exposure duration. It is incumbent on the risk assessor to clearly present the rationale for the assumptions used in a risk assessment, as well as their implications and limitations.

U.S. EPA guidance, including the *Exposure Assessment Guidelines* (U.S. EPA, 1992a), *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), and *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997d) have classified uncertainty in exposure assessment into three broad categories:

(1) *Parameter uncertainty* - uncertainty in values used to estimate variables of a model;
(2) *Model uncertainty* - uncertainty about a model structure (e.g., exposure equation) or intended use; and
(3) *Scenario uncertainty* - uncertainty regarding missing or incomplete information to fully define exposure.

Each source of uncertainty is described in detail below, along with strategies for addressing them in PRA.

3.4.1 PARAMETER UNCERTAINTY

Parameter uncertainty may be the most readily recognized source of uncertainty that is quantified in site-specific risk assessments at hazardous waste sites. Parameter uncertainty can occur in each step of the risk assessment process from data collection and evaluation, to the assessment of exposure and toxicity. Sources of parameter uncertainty may include systematic errors or bias in the data collection process, imprecision in the analytical measurements, and extrapolation from surrogate measures to represent the parameter of interest. For example, soil data collected only from the areas of highest contamination, rather than the entire area that a receptor is expected to come into contact, will result in a biased estimate of exposure.

In general, parameter uncertainty can be quantified at any stage of the tiered process, including point estimate analysis (Tier 1), one-dimensional Monte Carlo analysis (1-D MCA) (Tier 2), and two-dimensional Monte Carlo analysis (2-D MCA) (Tier 3). In the point estimate approach, parameter uncertainty may be addressed in a qualitative manner for most variables. For example, the uncertainty section of a point estimate risk assessment document might state that an absorption fraction of 100% was used to represent the amount of contaminant in soil absorbed from the gastrointestinal (GI) tract, and as a result, the risk estimate may overestimate actual risk. In addition, a sensitivity analysis may be performed, wherein one input variable at a time is changed, while leaving the others constant, to examine the effect on the outcome. In the case of absorption from the GI tract, different plausible estimates of the

high-end, or RME absorption fraction might be used as inputs to the risk equation. The differences in the risk estimates would reflect uncertainty in the RME absorption fraction.

Quantitative approaches for characterizing parameter uncertainty in exposure variables in a Monte Carlo simulation are summarized in Table 3-1. If uncertainty in only a few parameter values is of interest, multiple 1-D MCA simulations can yield the same results as a 2-D MCA simulation, but without the time and effort of a 2-D MCA. An example illustrating this concept is given in Table 3-2. With multiple 1-D MCA simulations, variability is characterized in one or more variables using probability distributions for variability (PDFv's), and uncertainty in a parameter is characterized with a series of different point estimates from a probability distribution for uncertainty (PDFu) (e.g., 95% lower confidence limit LCL [95% LCL], sample mean, and 95% UCL). In a 2-D MCA simulation, variability is characterized in one or more variables using PDFv's, and uncertainty in one or more parameters is characterized with PDFu's. With both approaches, the influence of the parameter uncertainty can be presented as a credible interval or confidence interval (CI) around the risk distribution, depending on how the PDFu's are defined. When only a few sources of parameter uncertainty are quantified, multiple 1-D MCA simulations are preferred over a 2-D MCA because the approach is easier to use and communicate. However, if the goal is to explore the effect that many sources of parameter uncertainty may have on the risk estimates simultaneously, a 2-D MCA is preferred. Iterative 1-D MCA simulations with different combinations of confidence limits may be impractical.

Table 3-1. Methods for Characterizing Parameter Uncertainty with Monte Carlo Simulations.

Approach	Example of Model Input	Method	Example of Model Output
Single Point Estimate	• 95% UCL	1-D MCA	PDFv[1] for risk, calculated using the 95% UCL for one parameter.
Multiple Point Estimates	• 95% LCL • sample mean • 95% UCL	1-D MCA	Three PDFv's for risk, representing the 90% CI for each percentile of the risk distribution.[2] The 90% CI only accounts for uncertainty in a single parameter (not multiple parameters).
Parametric PDFu[1]	PDFu for the mean based on the sampling distribution, derived from a Student's *t*-distribution.	2-D MCA	One PDFv for risk with confidence intervals at each percentile of the risk distribution. The CI reflects uncertainty in one or more parameters.
Non-parametric PDFu	PDFu for the mean based on bootstrap resampling methods.	2-D MCA	Same as parametric probability distribution for uncertainty.

[1]Probability distribution for uncertainty (PDFu) and probability distribution for variability (PDFv).
[2]The 95% UCL for the concentration term represents a 1-sided confidence interval (CI), meaning there is a 95% probability that the value is *greater* than or equal to the mean. Similarly, the 95% LCL would represent the 1-sided CI in which there is a 95% probability that the value is *less* than or equal to the mean. Both values are percentiles on the probability distribution for uncertainty (PDFu), also called the sampling distribution for the mean. Together, the 95% LCL and 95% UCL are equal to the 2-sided 90% confidence interval only for cases in which the PDFu is symmetric. For example, the sampling distribution for the arithmetic mean of a sample from a normal distribution with an unknown variance is described with the symmetric Student's *t*-distribution, whereas the PDFu for the mean of a lognormal distribution is asymmetric. In order to compare the results of multiple 1-D MCA simulations and a 2-D MCA simulation, the same methodology should be employed to define the PDFu and the corresponding confidence limits.

It is generally incorrect to combine a PDFu for one parameter (e.g., mean of the concentration term) with one or more PDFv's in other exposure factors when conducting a 1-D MCA for variability.

However, distributions for uncertainty and variability may be appropriately combined in a 2-D MCA. As discussed in Appendix D, with 2-D MCA, a clear distinction should be made between probability distributions that characterize variability (PDFv) and parameter uncertainty (PDFu). A 2-D MCA propagates the uncertainty and variability distributions separately through an exposure model, thereby making it possible to evaluate the effect of each on the risk estimates.

Example: Comparison of Multiple Point Estimates of Uncertainty in 1-D MCA, and Distributions of Uncertainty in 2-D MCA

Table 3-2 illustrates an application of the approaches presented in Table 3-1 for quantifying variability and parameter uncertainty. This is a hypothetical example, and no attempt was made to use standard default assumptions for exposure variables. Two sources of variability are quantified: (1) inter-individual variability in exposure frequency (EF), characterized by a triangular distribution, and (2) inter-individual variability in exposure duration (ED), characterized by a truncated lognormal distribution. In addition, two sources of uncertainty are presented: (1) a point estimate for soil and dust ingestion rate, intended to characterize the RME; and (2) an upper truncation limit of the lognormal distribution for ED, intended to represent a plausible upper bound for the exposed population. Methods for quantifying these sources of uncertainty are discussed below. Additional sources of uncertainty may also have been explored. For example, the choice of a triangular distribution for a PDFv may be provocative for some risk assessors, since there are few cases in which empirical data suggest a random sample is from a triangular distribution. Nevertheless, triangular distributions may be considered rough, or "preliminary" distributions (see Chapter 2 and Appendix B, Section B.2) for cases when the available information supports a plausible range and central tendency.

The choice of distributions is a potential source of uncertainty that can be explored by rerunning simulations with each alternative, plausible choice, and examining the effect on the RME risk. Simulations with preliminary simulations may yield at least three different outcomes. First, this type of sensitivity analysis can help guide efforts to improve characterizations of variability for selected variables that have the greatest affect on the risk estimates. Second, results may provide justification to exit the tiered process without continuing with additional Monte Carlo simulations since further effort would be unlikely to change the risk management decision. Finally, if the major sources of uncertainty can be clearly identified, a subset of the less sensitive variables may be defined by point estimates without significantly reducing the uncertainty in the risk estimates.

Parameter uncertainty can be quantified for both point estimates and PDFv's. In this example, both types of inputs (i.e., point estimates and PDFv's) are presented as sources of parameter uncertainty: the RME point estimate for soil and dust ingestion rate (IRsd), and the upper truncation limit on a PDFv for ED. For IRsd, assume that three different studies provide equally plausible values for the RME: 50, 100, and 200 mg/day. A uniform PDFu is specified to characterize this range of plausible values. For ED, assume that the maximum value reported from a site-specific survey was 26 years, but surrogate data for other populations suggest the maximum may be as long as 40 years. A uniform PDFu is specified to characterize this range of plausible values as well.

In Cases 1-3, the impact of uncertainty in IRsd and ED was evaluated using a series 1-D MCA simulations. Inputs for uncertain parameters associated with IRsd and ED in Case 1, 2, and 3 represent the minimum, central tendency, and maximum values, respectively. Each simulation yields a different risk distribution based on different combinations of point estimates for parameters. Although a PDFu was specified for IRsd, it would have been incorrect to combine the PDFu with the PDFv's for EF and ED in a

1-D MCA because the result would have been a single distribution of risk that co-mingled uncertainty and variability.

In Case 4, a single 2-D MCA simulation was run using the PDFu's for uncertainty and the PDFv's for variability. By propagating variability and uncertainty separately, the 2-D MCA yields a series of distributions of risk, from which credible intervals can be calculated for each percentile of the CDF.

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

Table 3-2. Example of 1-D MCA and 2-D MCA.

Variable	Type of Input	1-D MCA			2-D MCA
		Case 1	Case 2	Case 3	Case 4
C (mg/kg)	pt estimate	500	500	500	500
IRsd (mg/day)	pt estimate	50	100	200	see below
	PDFu for pt estimate	--	--	--	uniform (50, 200)[a]
CF (kg/mg)	pt estimate	1E-06	1E-06	1E-06	1E-06
EF (days/year)	PDFv	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350
ED (years)	PDFv	T-lognormal mean = 9 stdv = 10 max = 26	T-lognormal mean = 9 stdv = 10 max = 33	T-lognormal mean = 9 stdv = 10 max = 40	T-lognormal mean = 9 stdev = 10 max = PDFu (see below)
	PDFu for parameter of PDFv	--	--	--	max ~ uniform (26, 40)[b]
BW (kg)	pt estimate	70	70	70	70
AT (days)	pt estimate	25550	25550	25550	25550
CSF $(mg/kg\text{-}day)^{-1}$	pt estimate	1E-01	1E-01	1E-01	1E-01

[a]Uncertainty in the RME point estimate, defined by a uniform distribution with parameters (minimum, maximum).
[b]Uncertainty in the upper truncation limit of the lognormal distribution, defined by a PDFv with parameters (mean, standard deviation, maximum) and a PDFu for the maximum defined by a uniform distribution with parameters (minimum, maximum).

Monte Carlo Simulation Results

Figures 3-3 and 3-4 illustrate CDFs for risk produced from Monte Carlo simulations using *Crystal Ball*® 2000. The 1-D MCA simulations (Figure 3-3) were run with 10,000 iterations and Latin Hypercube sampling. The 2-D MCA simulation (Figure 3-4) was run with 250 iterations of the outer loop (uncertainty) and 2,000 iterations of the inner loop (variability). Details regarding 2-D MCA simulation are given in Appendix D.

Figure 3-3 shows CDFs for risk based on three simulations of a 1-D MCA simulation. Each simulation used a different combination of plausible estimates of the RME value for IRsd and the upper truncation limit for ED, as discussed above. The results provide a bounding estimate on the risk distribution given these two sources of uncertainty. The 95th percentile risk, highlighted as an example of the RME risk estimate, may range from approximately 7E-06 to 3.5E-05.

Figure 3-4 shows a single CDF for risk, representing the central tendency risk distribution. This CDF was derived by simulating uncertainty in the risk distribution using 2-D MCA. For this example, the 2-D MCA yields 250 simulations of the risk distributions for variability, so that there are 250 plausible estimates of each percentile of the risk distribution. In practice, more than 250 simulations may be needed to adequately quantify uncertainty in the risk distribution. Results of a 2-D MCA can be presented as probability distributions of uncertainty, or box-and-whisker plots of uncertainty at selected percentiles of the risk distributions. Figure 3-4 shows the central tendency (50th percentile) estimate of uncertainty for the entire CDF of risk. In addition, a box-and-whisker plot is shown at the 95th percentile of the CDF. Selected statistics for the box-and-whisker plot are included in a text box on the graphic (i.e., minimum; 5th, 50th, and 95th percentiles, and maximum). The 90% credible interval is given by the 5th and 95th percentiles. For this example, the 90% credible interval for the 95th percentile of the risk distribution is: [9.1E-06, 3.1E-05].

Figures 3-3 and 3-4 demonstrate that the two approaches (i.e., multiple 1-D MCA and 2-D MCA) can yield the same results. However, when there are numerous sources of uncertainty, 2-D MCA offers at least two advantages over multiple 1-D MCA simulations: (1) 2-D MCA allows the multiple sources of uncertainty to be included simultaneously so the approach is more efficient than a series of 1-D MCA simulations; and (2) multiple 1-D MCA simulations yield multiple estimates of the RME risk, but it is not possible to characterize the uncertainty in the RME risk in quantitative terms; a 2-D MCA yields a PDFu for RME risk, which allows for statements regarding the level of certainty that the RME risk is above or below a risk level of concern.

The 95th percentile is a focus of this example because it is a recommended starting point for determining the risk corresponding to the RME. Chapter 7 provides guidance to the risk decision makers on choosing an appropriate percentile (on a distribution of variability) within the RME risk range (90th to 99.9th percentiles). The chapter also includes a qualitative consideration of the uncertainty or confidence surrounding a risk estimate in the decision-making process.

Figure 3-3

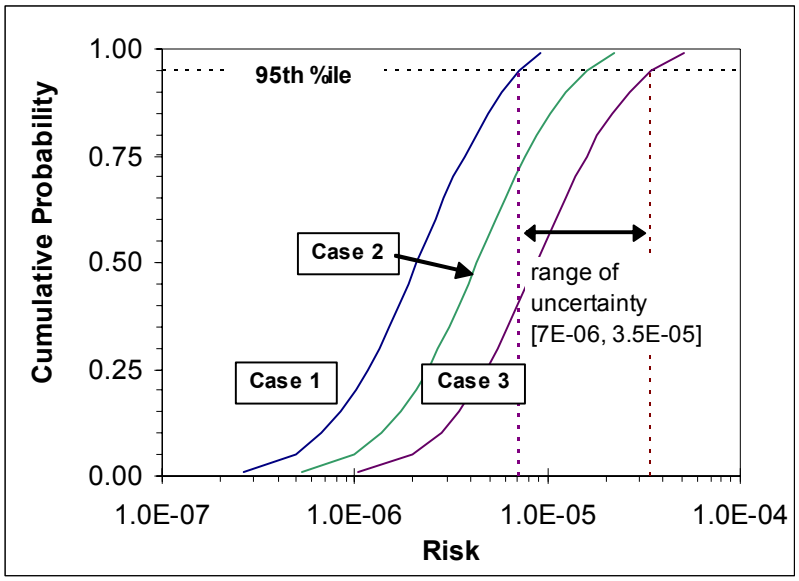
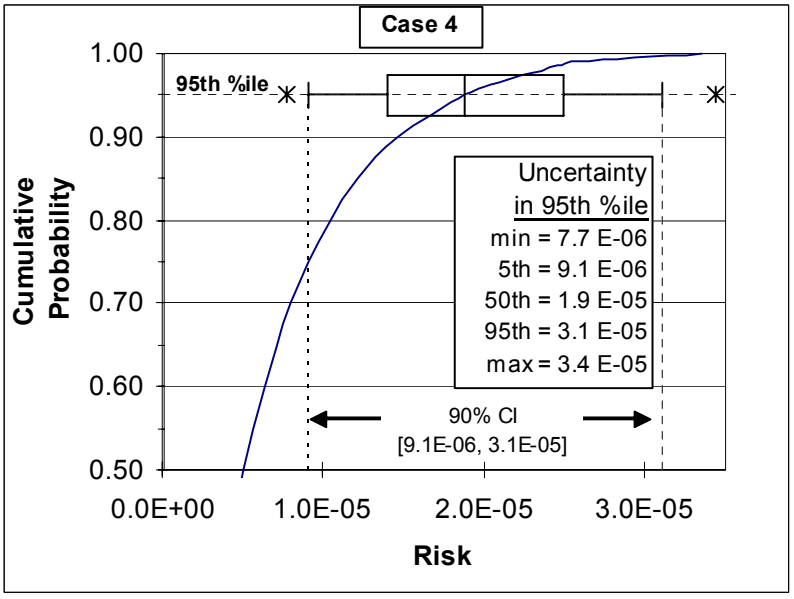


Figure 3-4



3.4.2 SCENARIO AND MODEL UNCERTAINTY

All models are simplified representations of complex biological and physical processes. As such, they, and the scenarios to which they are applied, may introduce a significant source of uncertainty into an exposure and risk estimate. Models may exclude important variables or important pathways of exposure, ignore interactions between inputs, use surrogate variables that are different from the target variables, or they may be designed for specific scenarios and not others. As a result, a model may not adequately represent all aspects of the phenomena it was intended to approximate or it may not be appropriate to predict outcomes for a different type of scenario. For example, a model intended to estimate risk from continuous, steady state exposures to a contaminant may not be appropriate or applicable for estimating risk from acute or subchronic exposure events. In any risk assessment, it is important to understand the original intent of a model, the assumptions being made in a model, what the parameters represent, and how they interact. Based on this knowledge, one can begin to understand how representative and applicable (or inapplicable) a model may be to a given scenario. If multiple models exist that can be applied to a given scenario, it may be useful to compare and contrast results in order to understand the potential implications of the differences. The use of multiple models, or models with varying levels of sophistication, may provide valuable information on the uncertainty introduced into a risk estimate as the result of model or scenario uncertainty. The collection of measured data as a reality check against a given parameter or the predicted model outcome (such as the collection of vegetable and fruit contaminant data to compare against modeled uptake into plants) is also useful in attempting to reduce or at least gain a better understanding of model and scenario uncertainty.

3.5 EXAMPLE OF PRA FOR HUMAN HEALTH

The following hypothetical example provides a conceptual walk-through of the tiered approach for PRA in Superfund risk assessment. The example begins with a baseline human health point estimate risk assessment (Tier 1) and moves to Tier 2, in which multiple iterations of a 1-D MCA are run using default and site-specific assumptions for input distributions. The general concepts associated with the tiered approach are discussed in Chapter 2, and a similar example for ecological risk assessment is given in Chapter 4. The 1-D MCA results are based on simulations with *Crystal Ball*® 2000 using 10,000 iterations and Latin Hypercube sampling. These settings were sufficient to obtain stability (i.e., <1% difference) in the 95% percentile risk estimate. The example is presented in Exhibit 3-5. Tables and figures supporting the example are given immediately following the exhibit.

EXHIBIT 3-5
USING THE TIERED PROCESS FOR PRA
HYPOTHETICAL CASE STUDY FOR HUMAN HEALTH RISK ASSESSMENT

RI Planning/Scoping/Problem Formulation/Data Collection

- Site Description: Former federal facility
- Site Size: 100 acres (5 acres within spill area (ISA); 95 acres outside spill area (OSA))
- Stakeholders: Refuge employees, environmental activists, etc.
- Land Use: Future wildlife refuge
- Receptors: Future wildlife refuge workers (i.e., ornithologists and fishery biologists)
- Sampling Data: n=35 surface soil samples (see Figure 3-5 for sample locations)
- Chemical of Concern: ChemX
- Chemical Properties: Nonvolatile
- Toxicological Properties: Carcinogen: CSF_{oral} and CSF_{dermal}= 5.5E-02, CSF_{inh}= 2.73E-02; Noncarcinogenic health data are lacking
- Risk Level of Concern: 1E-04 for cancer

↓

Tier 1 Point Estimate - Baseline Risk Assessment

- Exposure Unit: (see Figure 3-5) ornithologist (exposed in OSA) and fishery biologist (exposed in ISA)
- Exposure Pathways: Ingestion of soil/dust; inhalation of fugitive dust, dermal absorption
- Concentration Term: 95% UCL for arithmetic mean (Table 3-3)
- Risk Equations: Exhibit 3-6
- Exposure Parameters: Table 3-4
- Results: Table 3-5

↓

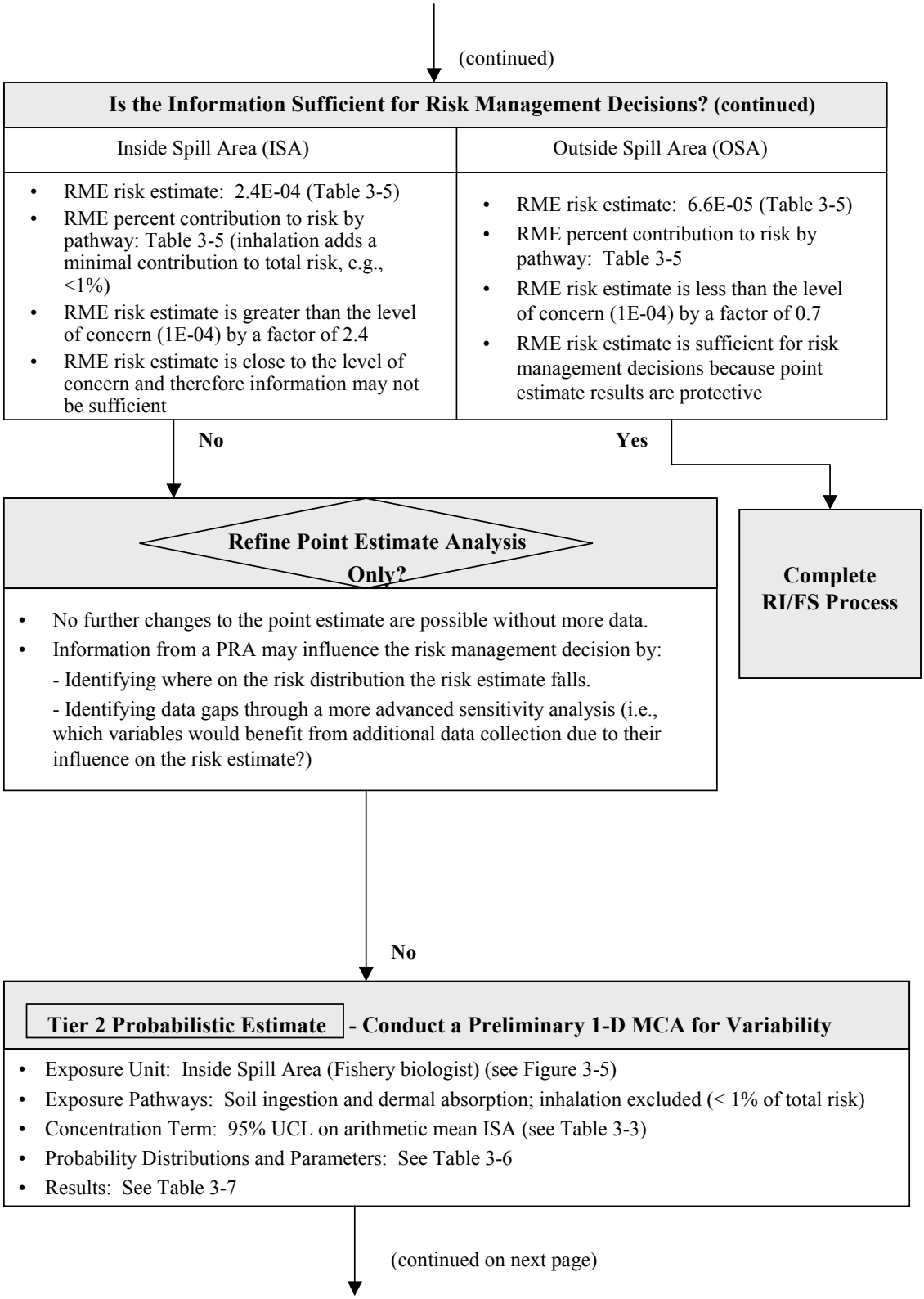
SMDP Is the Information Sufficient for Risk Management Decisions?

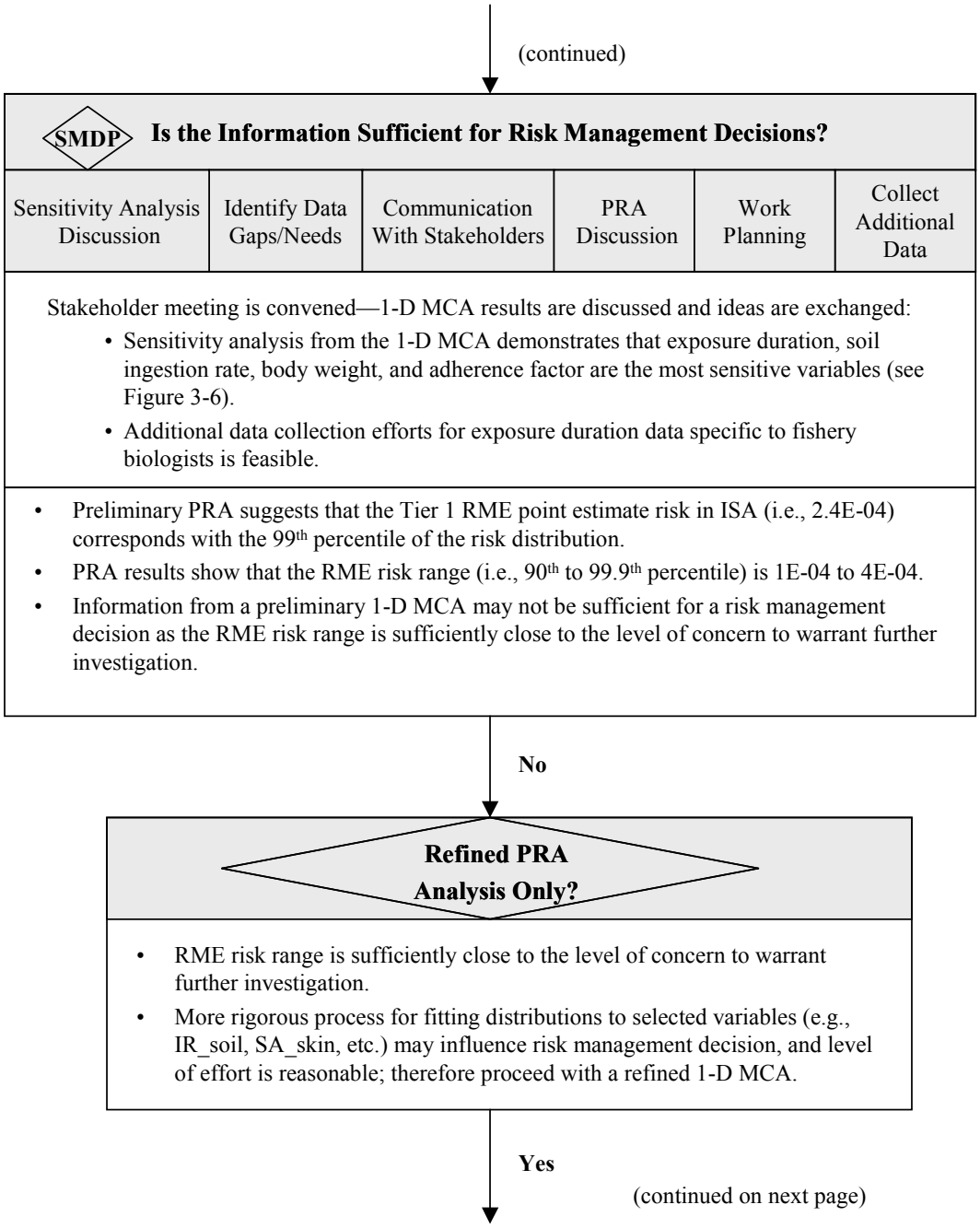
Sensitivity Analysis Discussion	Identify Data Gaps/Needs	Communication With Stakeholders	PRA Discussion	Work Planning	Collect Additional Data

Stakeholder meeting is convened—point estimate results are discussed and ideas are exchanged as follows:

- Risk estimates are expected to be conservative due to the use of standard default exposure parameters, but are the defaults representative?
- Stakeholders are concerned about risk to workers and about the consequences of remediation (e.g., negative impacts on habitat and potential job losses).
- Stakeholders are concerned about the relevance of some nonsite-specific exposure variables (e.g., exposure duration), but are not sure which variables to investigate further (i.e., which is the most influential?).
- Results of the sensitivity analysis from point estimate risk assessment cannot identify where the high end risk estimate falls on the risk distribution.
- There is sufficient information (e.g., arithmetic mean, standard deviation, percentiles) for some of the exposure variables to develop initial probability distributions to characterize variability.

(continued on next page)


(continued)
Is the Information Sufficient for Risk Management Decisions? (continued)
Inside Spill Area (ISA)
Outside Spill Area (OSA)
• RME risk estimate: 2.4E-04 (Table 3-5)
• RME percent contribution to risk by pathway: Table 3-5 (inhalation adds a minimal contribution to total risk, e.g., <1%)
• RME risk estimate is greater than the level of concern (1E-04) by a factor of 2.4
• RME risk estimate is close to the level of concern and therefore information may not be sufficient
• RME risk estimate: 6.6E-05 (Table 3-5)
• RME percent contribution to risk by pathway: Table 3-5
• RME risk estimate is less than the level of concern (1E-04) by a factor of 0.7
• RME risk estimate is sufficient for risk management decisions because point estimate results are protective
No
Yes
Refine Point Estimate Analysis Only?
Complete RI/FS Process
• No further changes to the point estimate are possible without more data.
• Information from a PRA may influence the risk management decision by:
- Identifying where on the risk distribution the risk estimate falls.
- Identifying data gaps through a more advanced sensitivity analysis (i.e., which variables would benefit from additional data collection due to their influence on the risk estimate?)
No
Tier 2 Probabilistic Estimate - Conduct a Preliminary 1-D MCA for Variability
• Exposure Unit: Inside Spill Area (Fishery biologist) (see Figure 3-5)
• Exposure Pathways: Soil ingestion and dermal absorption; inhalation excluded (< 1% of total risk)
• Concentration Term: 95% UCL on arithmetic mean ISA (see Table 3-3)
• Probability Distributions and Parameters: See Table 3-6
• Results: See Table 3-7
(continued on next page)


(continued)
SMDP
Is the Information Sufficient for Risk Management Decisions?
Sensitivity Analysis Discussion
Identify Data Gaps/Needs
Communication With Stakeholders
PRA Discussion
Work Planning
Collect Additional Data
Stakeholder meeting is convened—1-D MCA results are discussed and ideas are exchanged:
• Sensitivity analysis from the 1-D MCA demonstrates that exposure duration, soil ingestion rate, body weight, and adherence factor are the most sensitive variables (see Figure 3-6).
• Additional data collection efforts for exposure duration data specific to fishery biologists is feasible.
• Preliminary PRA suggests that the Tier 1 RME point estimate risk in ISA (i.e., 2.4E-04) corresponds with the 99th percentile of the risk distribution.
• PRA results show that the RME risk range (i.e., 90th to 99.9th percentile) is 1E-04 to 4E-04.
• Information from a preliminary 1-D MCA may not be sufficient for a risk management decision as the RME risk range is sufficiently close to the level of concern to warrant further investigation.
No
Refined PRA Analysis Only?
• RME risk range is sufficiently close to the level of concern to warrant further investigation.
• More rigorous process for fitting distributions to selected variables (e.g., IR_soil, SA_skin, etc.) may influence risk management decision, and level of effort is reasonable; therefore proceed with a refined 1-D MCA.
Yes
(continued on next page)

Yes

(continued)

Tier 2 Refined PRA - Conduct Refined 1-D MCA and Refined Point Estimate

- Exposure Unit: Fishery biologist-inside spill area (ISA) (see Figure 3-5)
- Exposure Pathways: Ingestion of soil and dust, and dermal absorption
- Concentration Term: 95% UCL on arithmetic mean
- Probability Distributions/Parameters: see Table 3-8 for sample data and summary statistics; exposure duration defined by lognormal PDF (arithmetic mean=14, SD=9.4, upper truncation of 44 years)
- Results: see Table 3-9

SMDP Is the Information Sufficient for Risk Management Decisions?

Sensitivity Analysis Discussion	Identify Data Gaps/Needs	Communication With Stakeholders	PRA Discussion	Work Planning	Collect Additional Data

Stakeholders meeting is convened. Refined 1-D MCA results are discussed and ideas are exchanged as follows:

- Sensitivity analysis from refined 1-D MCA indicates that the use of site-specific data did not significantly alter the relative ranking or magnitude of rank correlations for input variables (similar graphic as Figure 3-6).
- Refined 1-D MCA results suggest that the refined RME point estimate risk corresponds with the 99th percentile of the risk distribution (Table 3-9).
- Refined 1-D MCA results show that the RME range (i.e., 90th to 99.9th percentile) is 1.6E-04 to 5E-04, with 95th percentile of 2.1E-04.
- Information from refined 1-D MCA is sufficient for risk management decision because the RME risk (95th percentile) is above the level of concern of 1E-04 using site specific exposure duration data, and additional data collection on IR_soil term is not warranted. Complete RI/FS process.

Yes

Complete RI/FS Process

- Stakeholders and RPM decide that the best remedial alternative is to remove surface soil in the 5 acre spill area and cover the refuge area with clean fill before beginning refuge construction.

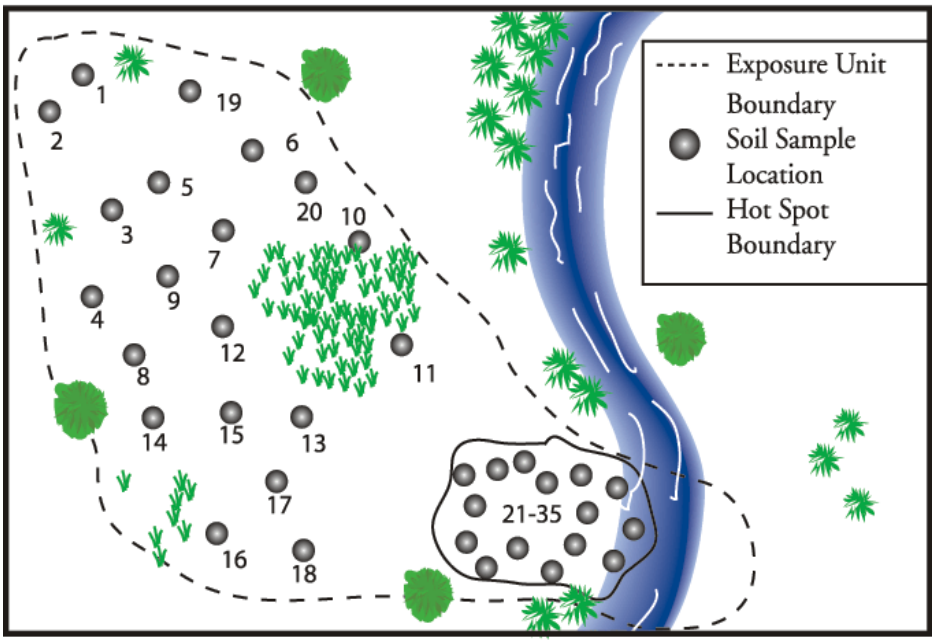


Figure 3-5. Site map for future wildlife refuge showing boundaries for the exposure unit and potential hotspot, as well as sampling locations (n=35). Sample numbers correspond with concentration data given in Table 3-3.

[1]The 95% UCL was estimated using the Land method (see Appendix C).

Table 3-3. Concentrations in Surface Soil (mg/kg).

Outside Spill Area (n=20)		Inside Spill Area (n=15)	
1088	305	1934	970
646	2787	402	985
3943	760	4215	743
149	149	1121	158
3704	1088	629	21296
845	837	2293	
488	1295	257	
387	1239	288	
1438	1006	57	
2502	283	228	

Summary Statistics	Outside Spill Area	Inside Spill Area
Mean	1247	2372
Standard Deviation	1121	5348
95% UCL[1]	2303	8444

EXHIBIT 3-6

RISK EQUATIONS

Soil Ingestion

$$\text{Risk} = \frac{Cs \times CF \times IRs \times FI \times EF \times ED}{BW \times AT} \times \text{Oral CSF}$$

Dermal Absorption

$$\text{Risk} = \frac{Cs \times CF \times SA \times AF \times ABS \times EF \times ED}{BW \times AT} \times \text{Dermal-Adjusted CSF}$$

Inhalation of Fugitive Dust

$$\text{Risk} = \frac{Cs \times 1/PEF \times IRa \times ET \times EF \times ED}{BW \times AT} \times \text{Inhalation CSF}$$

Total Risk = Sum of risks from each exposure pathway (soil + dermal + inhalation)

Where:

Cs = Concentration of ChemX in soil (mg/kg)
IRs = Soil ingestion rate for receptor (mg/day)
FI = Fraction ingested from contaminated source (unitless)
CF = Conversion factor (1E-06 kg/mg)
SA = Skin surface area available for exposure (cm^2/event)
AF = Soil to skin adherence factor for ChemX (mg/cm^2)
ABS = Absorption factor for ChemX (unitless)
IRa = Inhalation rate for receptor (m^3/hr)
PEF = Soil-to-air particulate emission factor (kg/m^3)
ET = Exposure time for receptor (hours/day)
EF = Exposure frequency for receptor (days/year)
ED = Exposure duration for receptor (years)
BW = Body weight of receptor (kg)
AT = Averaging time (years)
CSF = Cancer slope factor (oral, dermal, inhalation) $(mg/kg\text{-}day)^{-1}$

Table 3-4. Exposure Parameters used in Point Estimate Analysis.

Exposure Variable	CTE Value	RME Value	Units	Reference
IRs	50	100	mg/day	CTE: U.S. EPA, 1997a, p. 4–25 RME: U.S. EPA, 2001
FI	0.5	1	unitless	Site-specific
CF	1E-06	1E-06	kg/mg	Constant
SA	3300	3300	cm^2/event	U.S. EPA, 2001, 50th percentile value for all adult workers—exposure to face, forearms, and hands
AF	0.1	0.2	mg/cm^2	CTE: U.S. EPA, 1998; Table 3.3, value for gardeners RME: U.S. EPA, 2001
ABS	0.1	0.1	unitless	U.S. EPA, 1998, default for semi-volatile organic compounds (SVOCs)
IRa	1.3	3.3	m^3/hr	U.S. EPA, 1997a, p. 5–24, outdoor worker hourly average: mean and upper percentile
PEF	1.36E+09	1.36E+09	kg/m^3	U.S. EPA, 2001
ET	8	8	hours/day	Site-specific
EF	200	225	days/year	CTE: Site-specific assumption RME: U.S. EPA, 2001
ED	5	25	years	CTE: U.S. EPA, 1993, p. 6 RME: U.S. EPA, 2001
BW	70	70	kg	U.S. EPA, 1993, p. 7
AT	25550	25550	days	constant

CTE = central tendency exposure; RME = reasonable maximum exposure.

Table 3-5. Point Estimate Risks and Exposure Pathway Contributions**.**

Risk Estimate by Exposure Pathway	Inside Spill Area (n = 15)		Outside Spill Area (n = 20)	
	CTE	RME	CTE	RME
Soil Ingestion	6.5E-06 (43 %)	1.5E-04 (60 %)	1.7E-06 (43 %)	4.0E-05 (60 %)
Dermal Absorption	8.6E-06 (57 %)	9.6E-05 (40 %)	2.3E-06 (57 %)	2.6E-05 (40 %)
Inhalation	9.9E-10 (< 1 %)	1.4E-08 (< 1 %)	2.7E-10 (< 1 %)	3.8E-09 (< 1 %)
Total Risk	**1.5E-05**	**2.4E-04**	**4.1E-06**	**6.6E-05**

Example of % contribution: % Soil for RME risk inside spill area = (Soil risk / Total risk) x 100%
= (1.46E-04 / 2.42E-04) x 100% = 60%

Table 3-6. Input Distributions for Exposure Variables used in 1-D MCA for Variability.

Exposure Variable[1]	Distribution Type	Parameters[2]	Units	Reference
IR_soil	Triangular	0, 50, 100	mg/day	U.S. EPA, 1993, 2001
SA_skin[3]	Lognormal	18150, 37.4	cm^2	U.S. EPA, 1997a, Table 6-4 (Total male/female body surface area)
Absorption Fraction	Uniform	0.1, 0.2	mg/cm^2	U.S. EPA, 2001; minimum truncation limit is professional judgment
IR_air	Lognormal	1.68, 0.72	m^3/hour	U.S. EPA, 1996, p.5–10
EF	Triangular	200, 225, 250	days	U.S. EPA, 2001; truncation limits are professional judgment
ED	Lognormal[4]	11.7, 7.0	years	U.S. EPA, 1997b, Table 15-161 and U.S. EPA, 2001 (Mean value is based on average of total median tenure for professional specialty and farming, forestry, and fishing)
	Truncated Lognormal[5]	14.0, 9.4, 44.0	years	Site-specific survey data, used in refined 1-D MCA
BW	Lognormal	71.75, 14.2	kg	U.S. EPA, 1997a, Tables 7-4 and 7-5; (Combined male/female body weight distributions)

[1]All other exposure parameters are inputted as point estimates (see Table 3-4).
[2]Parameters for lognormal PDF are X ~ Lognormal (arithmetic mean, arithmetic standard deviation) unless otherwise stated. Parameters for triangular PDF are X ~ Triangular (minimum, mode, maximum). Parameters for uniform PDF are X ~ Uniform (minimum, maximum).
[3]A point estimate of 0.189 was used to adjust the surface area skin (SA_skin) distribution, which is based on total body surface area, to account for skin exposures limited to face, forearms, and hands (U.S. EPA, 1997a, Vol. I).
[4]Parameters for preliminary lognormal PDF for ED were converted from a geometric mean of 10 and a 95th percentile of 25.
[5]Parameters for site-specific lognormal PDF for ED are arithmetic mean, standard deviation, and upper truncation limit.

Table 3-7. 1-D MCA Risk Estimates using Preliminary Inputs.

Cumulative Percentile	Spill Area Risk
50th	5.7E-05
90th	1.3E-04
95th	1.6E-04
99th	2.4E-04
99.9th	3.9E-04

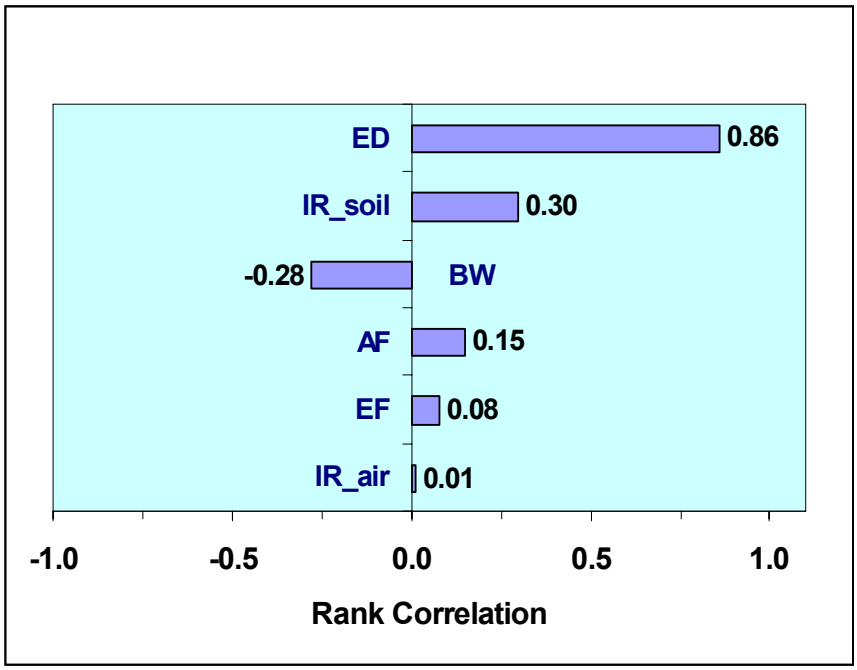


Figure 3-6. Results of sensitivity analysis for preliminary 1-D MCA (Tier 2) showing the Spearman Rank correlations (see Appendix A and B) between input variables and risk estimates.

Table 3-8. Exposure Duration Survey Results.

Survey Results (years)			Summary Statistics	
24.9	20.3	17.2	n	20
8.4	11.7	6.5	min	3.0
3.0	4.7	16.5	max	44.2
6.8	20.9	6.0	arithmetic mean	14.0
18.5	10.6	18.8	standard dev	9.4
9.1	12.7	11.7	median/GM	11.7
7.2	44.2		GSD	1.8

Table 3-9. Refined Point Estimate and 1-D MCA Risk Estimates.

Cumulative Percentile	Spill Area Risk
Refined RME Point Estimate	3.1E-04
50th	6.7E-05
90th	1.6E-04
95th	2.1E-04
99th	3.2E-04
99.9th	5.3E-04

REFERENCES FOR CHAPTER 3

Hedges, L.V. and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. Academic Press, Inc. Orlando.

Oregon DEQ. 1998. *Guidance for the Use of Probabilistic Analysis in Human Health Exposure Assessments*. Waste Management and Cleanup Division. Interim Final. November.

Stiteler, W.M., L.A. Knauf, R.C. Hertzberg, and R.S. Schoeny. 1993. A Statistical Test of Compatibility of Data Sets to a Common Dose-Response Model. *Regulatory Tox. Pharm*. 18: 392–402.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. 57 *Federal Register*, 22888-22938. May 29.

U.S. EPA. 1992b. *Guidance on Data Usability in Risk Assessment*. Part A. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7.09A. NTIS PB92-96336.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund*. Office of Solid Waste and Emergency Response. Washington, DC.

U.S. EPA. 1996. *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 540/R-96/018.

U.S. EPA. 1997a. *Exposure Factors Handbook, Volume 1.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 1997b. *Exposure Factors Handbook, Volume 2.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fb.

U.S. EPA. 1997c. *Exposure Factors Handbook, Volume 3.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fc.

U.S. EPA. 1997d. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment*. Final. National Center for Environmental Assessment, Washington, DC. EPA/630/R-95/002F.

U.S. EPA. 2001. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Office of Solid Waste and Emergency Response. Washington, DC. OSWER Directive No. 9355.4-24. December.

CHAPTER 4

PROBABILISTIC ANALYSIS IN ECOLOGICAL RISK ASSESSMENT

4.1 INTRODUCTION

4.1.1 BASIC APPROACH FOR PERFORMING ECOLOGICAL RISK ASSESSMENTS

Ecological risk assessment (ERA) is defined by the 1997 Environmental Protection Agency's (EPA) *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments (ERAGS)* (U.S. EPA, 1997a) as an evaluation of the "likelihood that adverse ecological effects are occurring or may occur as a result of exposure to one or more stressors". The *ERAGS* document is generally similar to, and consistent with the earlier framework guidance and approach (U.S. EPA, 1992a) which was expanded upon and superceded by the *Guidelines for Ecological Risk Assessment* (U.S. EPA, 1998). The EPA has developed extensive technical and policy guidance on how ERAs should be planned and performed (see Exhibit 4-2). In general, this process has three main elements, as shown in Figure 4-1:

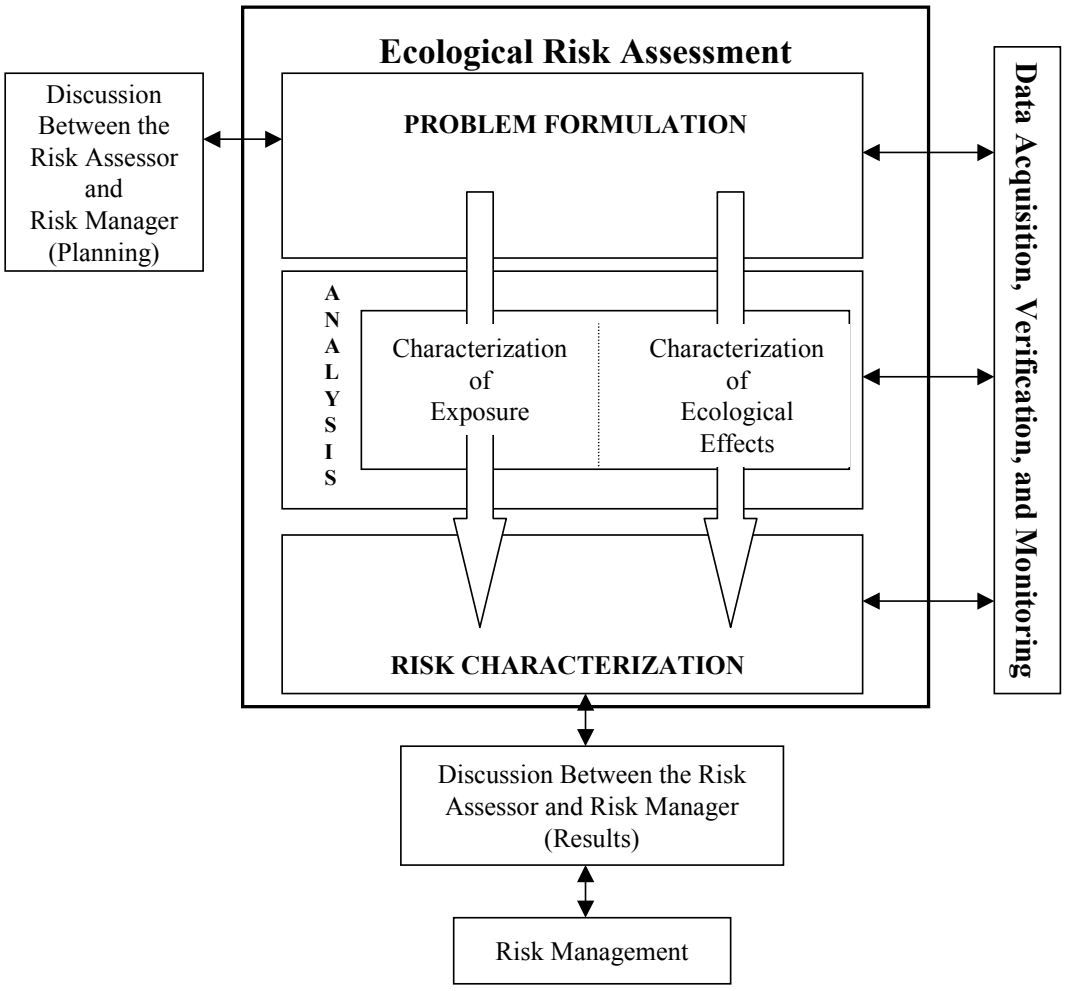


Figure 4-1. Ecological Risk Assessment Framework (U.S. EPA, 1992a)

Problem Formulation provides a foundation for the entire risk assessment. This element includes the specification of risk management goals and assessment endpoints, the development of a site conceptual model with exposure pathways and receptors, and the development of a sampling and analysis plan to collect data on exposures and measures of effects that are needed to support the ERA. In general, problem formulation serves as the foundation of an ERA and often is an iterative process, whereby substantial re-evaluation may occur as new information and data are collected during the site investigations. Collection of data in subsequent iterations is often triggered by identification of major data gaps and uncertainties in the risk characterization that prevent confident decision making by risk managers.

Analysis includes two principal measurement steps that are based upon the problem formulation: Assessment of exposures and assessment of ecological effects. Assessment of exposures includes the identification of stressors at the site that may affect ecological receptors, a characterization of the spatial and/or temporal pattern of the stressors in the environment at the site, and an analysis of the level of contact or co-occurrence between the stressors and the ecological receptors. Assessment of ecological effects includes identification of the types of effects which different stressors may have on ecological receptors, along with a characterization of the relationship between the level of exposure to the stressor and the expected biological or ecological response. This is referred to as the stressor-response relationship.

Risk Characterization combines the exposure characterization and the effects characterization in order to provide a quantitative likelihood or qualitative description of the nature, frequency, and severity of ecological risks attributable to exposure to stressors at a site, as well as an evaluation of the ecological relevance of the effects. Good risk characterizations express results clearly, articulate major assumptions and uncertainties, identify reasonable alternative interpretations, and separate scientific conclusions from policy judgments (U.S. EPA, 1995, 1998).

EXHIBIT 4-1

DEFINITIONS FOR CHAPTER 4

Assessment Endpoint - An explicit expression of an environmental value (ecological resource) that is to be protected, operationally defined by risk managers and risk assessors as valuable attributes of an ecological entity.

Benchmark Dose (BMD) - The dose which causes a specified level of response. The lower confidence limit on the BMD is usually referred to as the BMDL.

Community - An assemblage of populations of different species specified by locales in space and time.

Conceptual Model - A site conceptual model (SCM) in the problem formulation for an ecological risk assessment is a written description and visual representation of predicted relationships between ecological entities and the stressors to which they may be exposed, including sources and pathways of stressors.

Ecological Risk Assessment (ERA) - The process that evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors.

Lines of Evidence - Information derived from different sources or techniques that can be used to characterize the level of risk posed to exposed receptors; weight-of-evidence generally refers to the quantity of science, while strength of evidence generally refers to the quality of science.

Lowest-Observed-Adverse-Effect Level (LOAEL) - The lowest level of a stressor evaluated in a test that caused a statistically significant effect on one or more measurement endpoints linked to undesirable (adverse) biological changes.

Measurement Endpoint (Measure of Effect) - A measurable ecological property that is related to the valued characteristic chosen as the assessment endpoint. Measurement endpoints (also called measures of effect) often are expressed as the statistical or numeric summaries of the observations that make up the measurement.

No-Observed-Adverse-Effect Level (NOAEL) - The highest level of a stressor administered in a test that did not cause a statistically significant effect in any measurement endpoint linked to an undesirable (adverse) biological change.

Population - An aggregate of individuals of a species within a specified location in space and time.

Receptor - The ecological entity (with various levels of organization) exposed to the stressor.

Risk Characterization (ecological) - The third and last phase of ERA that integrates the analyses of exposure to stressors with associated ecological effects to evaluate likelihoods of adverse ecological effects. The ecological relevance of the adverse effects is discussed, including consideration of the types, severity, and magnitudes of the effects, their spatial and temporal patterns, and the likelihood of recovery.

Scientific/Management Decision Point (SMDP) - A time during the ERA when a risk assessor communicates results or plans of the assessment at that stage to a risk manager. The risk manager decides if information is sufficient to proceed with risk management strategies or whether more information is needed to characterize risk.

Species - A group of organisms that actually or potentially interbreed and are reproductively isolated from similar groups; also, a taxonomic grouping of morphologically similar individuals.

Stressor - Any chemical, physical or biological entity that can induce an adverse response in an ecological receptor; Superfund considers all stressors, but focuses on chemical (toxicant) stressors.

Toxicity Reference Value (TRV) - A dose or concentration used to approximate the exposure threshold for a specified effect in a specified receptor. A TRV is often based on a NOAEL or LOAEL from a laboratory-based test in a relevant receptor species.

EXHIBIT 4-2

ECOLOGICAL RISK ASSESSMENT GUIDANCE AND POLICY DIRECTIVES

EPA has developed extensive guidance and policies on methods and approaches for performing ERAs, including the following:

(1) *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments ("ERAGS"), Interim Final (U.S. EPA, 1997a)*. This document includes processes and steps specifically selected for use in ERAs at Superfund sites. This document supersedes the 1989 *EPA RAGS, Volume II, Environmental Evaluation Manual,* Interim Final (U.S. EPA, 1989). Supplements to ERAGS include the *EcoUpdates* (U.S. EPA, 1991-present, Intermittent Bulletin Series, 1991 to present), which provide brief recommendations on common issues for Superfund ERAs.

(2) *Guidelines for Ecological Risk Assessment ("Guidelines") (U.S. EPA, 1998).* This document updates general (nonprogram specific) guidance that expands upon and replaces the earlier *Framework for Ecological Risk Assessment* (U.S. EPA, 1992a). The approaches and methods outlined in the *Guidelines* and in *ERAGS* are generally consistent with each other.

(3) *Risk Assessment Guidance for Superfund (RAGS): Volume 1–Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments)*, (U.S. EPA, 2001). This guidance specifies formats that are required to present data and results in baseline risk assessments (both human and ecological) at Superfund sites.

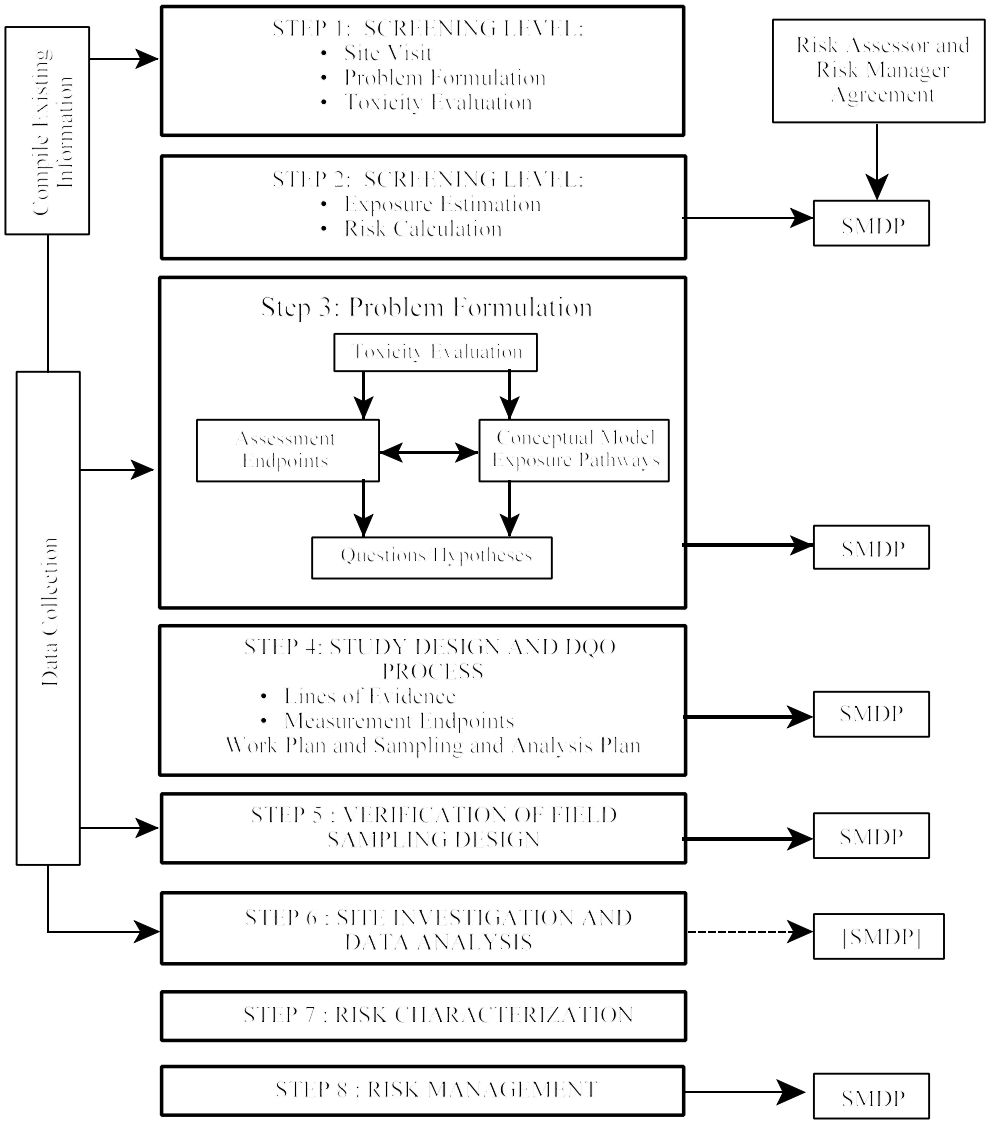
(4) Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors*, F. Henry Habicht, Deputy Administrator, Feb. 26, 1992 (U.S. EPA, 1992b). This policy requires baseline risk assessments to present ranges of risks based on "central tendency" and "reasonable maximum" (RME) or "high-end" exposures with corresponding risk estimates.

(5) Policy Memorandum: *Role of the Ecological Risk Assessment in the Baseline Risk Assessment*, Elliott Laws, Assistant Administrator, August 12, 1994 (U.S. EPA, 1994). This policy requires the same high level of effort and quality for ERAs as commonly performed for human health risk assessments at Superfund sites.

(6) Policy Memorandum: *EPA Risk Characterization Program*, Carol Browner, Administrator, March 21, 1995 (U.S. EPA, 1995). This policy clarifies the presentation of hazards and uncertainty in human and ERAs, calling for clarity, transparency, reasonableness, and consistency.

(7) Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites. Stephen D. Luftig for Larry D. Reed, October 7, 1999 (U.S. EPA, 1999). This document presents six key principles in ecological risk management and decision making at Superfund sites.

ERA is a key component of the remedial investigation process that EPA uses at Superfund sites. *ERAGS* is a program-specific guidance for Superfund that focuses on chemical stressors released into the environment from hazardous waste sites. This guidance refers to ERA as a "qualitative and/or quantitative appraisal of the actual or potential impacts of contaminants from a hazardous waste site on plants and animals other than humans and domesticated species. An excess risk does not exist unless: (1) the stressor has the ability to cause one or more adverse effects, and (2) the stressor co-occurs with or contacts an ecological component long enough and at a sufficient intensity to elicit the identified adverse effect." The *ERAGS* document provides guidance on using an eight-step process for completing an ERA for the Superfund Program, as shown in Figure 4-2.



SMDP= Scientific/Management Decision Point

Figure 4-2. Eight-step Ecological Risk Assessment Process for Superfund (U.S. EPA, 1997a).

4.1.2 PREDICTIVE *VS* OBSERVATIONAL APPROACHES

In general, conclusions about ecological hazards from environmental contamination may be based on information derived from two different techniques: the predictive approach (a comparison of calculated exposures with a set of toxicity reference values), and the observational approach (direct evaluation of the range of potential exposures, coupled with site-specific toxicity testing and population demographic estimates).

Predictive Approach: The core of all Superfund ERAs is the predictive approach, including exposure assessment, toxicity assessment, and risk characterization. The predictive approach is based on a comparison of calculated estimates of chemical exposure of a receptor to one or more Toxicity Reference Values (TRVs) appropriate for that chemical and that receptor. The ratio of exposure at the site to the TRV is referred to as the Hazard Quotient (HQ). The predictive approach has always been used at Superfund sites because it is relatively easy to implement, and because it can be used to evaluate not only current risks, but also risks that might exist in the future if any important changes were to occur in the level of contamination (e.g., due to on-going fate and transport processes), or to changes in land use (a change in land use might alter a number of habitat factors that influence the number and identify of ecological receptors). The predictive approach, however, has the inherent uncertainties of the assumptions in the exposure and toxicity models which are seldom site-specific and thus can lead to either over-protective or under-protective estimates of risk.

Direct Observation: If there is a need to reduce uncertainties in the predictive approach, direct observations of exposure and effects can be collected at Superfund hazardous waste sites. The predictive approach used in ERA does not negate the use of descriptive toxicological approaches or the use of site-specific exposure data, such as toxicity testing or bioaccumulation measurements. Site-specific observations, such as toxicity testing of invertebrates over a gradient of site contaminant exposure levels, may be used to develop site-specific and chemical-specific toxicological relationships. Site-specific measures of exposure or ecosystem characteristics can be used to reduce uncertainty in the exposure assessment and aid in the development of cleanup goals in the Baseline ERA. The direct observation of the exposure and effects on ecological receptors does not however constitute a complete risk assessment. If field or laboratory studies are NOT designed appropriately to elicit stressor-response relationships, direct impacts should not be used as the sole measure of risk because of the difficulty in interpreting and using these results to develop cleanup goals in the ERA. Furthermore, poorly designed toxicological evaluations of environmental media from the site may not allow a definitive identification of the cause of adverse response. For example, receptor abundance and diversity as demographic data reflect many factors (habitat suitability, availability of food, predator-prey relationships among others). If these factors are not properly controlled in the experimental design of the study collecting the observational data, conclusions regarding chemical stressors can be confounded. In addition, direct observation provides information about current risks only and not potential risks should land use or exposure change in the future. Hence, direct observations may be used as a line of evidence in an ERA, but should not be the sole evidence used to characterize the presence or absence of the risks of an adverse effect in the future.

4.1.3 POTENTIAL ADVANTAGES AND LIMITATIONS OF PROBABILISTIC METHODS IN ERA

Probabilistic risk assessment (PRA) is a computational tool that may help increase the strength of the *predictive* evaluation of ecological risks, as well as sometimes helping to better evaluate distributions of observational data for an ERA. The potential advantages of PRA compared to, or possible benefits in augmentation of, the conventional point estimate approach for characterizing variability in exposure or risk are discussed in Chapter 1 and Exhibits 1-6 and 1-7. In brief, point estimate calculations utilize simplifications and assumptions in order to deal with the complex mathematics of combining inputs that are inherently variable. Probabilistic models, in contrast, are designed to combine sets of information on inputs that are expressed as probability distributions. Therefore, PRA generally can yield risk estimates that allow for a more complete characterization of variability and uncertainty, and a potentially more useful sensitivity analysis as compared to estimating sensitivities of inputs from point estimates (see Appendix A). For example, sensitivity analysis can help determine major contributors to exposure factors and sources of uncertainty that could help to design better sampling and analysis plans in later iterations to help fill data gaps and reduce uncertainties for risk characterization.

Because of the inherent differences in the computational approach, as in the case with any additional risk assessment information, PRA may sometimes lead to a different risk assessment outcome and risk management decision than would be derived from the use of point estimate calculations alone. The differences in the decisions stemming from the two approaches will vary from case to case, depending mainly on the form of the exposure or risk model, the attributes of the distributions of the input values, and the quality, quantity, and representativeness of the data on which the input distributions are derived. Sometimes the differences between the two approaches will be quite large, and the information gained from a PRA can play an important role as weight-of-evidence in communicating risks to stakeholders and risk managers.

Even though PRA may have some advantages, it also has limitations and potential for misuse. PRA can not fill basic data gaps and can not eliminate all of the potential concerns associated with those data gaps. That is, if one or more of the input distributions are not well characterized and the distribution(s) must be estimated or assumed, then the results of the PRA approach will share the same uncertainty as the point estimate values. However, given equal states of knowledge, the PRA approach may yield a more complete characterization of the exposure or risk distribution than the point estimate approach.

Of course, any prediction of exposure or risk is based on the use of mathematical models to represent very complex environmental, biological, and ecological systems. No matter how sophisticated the computations, questions will always exist as to whether the calculated values are a good approximation of the truth. Therefore, even when PRA is used as a supplemental tool to point estimations (deterministic) of risks in the ERA process, a weight-of-evidence approach that combines the predictive approach with direct observations will still provide the most appropriate basis for decision making.

A second application of PRA in ERA, besides the characterization and incorporation of distributions of data for ERA, is the characterization of uncertainty in calculated estimates of exposure or risk. In this application, whatever uncertainty may exist in one or more of the input distributions is characterized, and quantitative estimates of the confidence limits around the mean, upper bound, or any other percentile of the output distribution are calculated. This use of PRA is often especially important in risk management decision making, since the range of uncertainty around central tendency exposure (CTE) and reasonable maximum exposure (RME) or other upper bound estimates of exposure or risk can

sometimes be quite large. As stated before, the point estimate approach can also provide estimates of uncertainty, but the PRA approach often provides a more complete characterization of the uncertainty.

4.1.4 FOCUS OF THIS CHAPTER

This chapter focuses on the application of PRA as a tool for predicting ecological risks at Superfund sites. Some of the methods and approaches described in this chapter are similar to those that have been developed by U.S. EPA's Office of Pesticide Programs Committee on Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) Risk Assessment Methods (ECOFRAM, 1999a, 1999b) for use in assessing environmental hazards of pesticide products. However, the methods described in this chapter are specifically designed to be applicable at Superfund sites and to be consistent with other Superfund guidance.

This chapter does not seek to provide guidance on the many basic issues that must be faced in planning and performing any ERA. Prior to considering the use of PRA in an ERA, fundamental concepts will already have been developed, such as a problem formulation with a conceptual site model, selection of representative receptors, definition of exposed populations, definition of risk management objectives and goals, selection of assessment endpoints, calculation of TRVs and development of site sampling plans, etc. Likewise, this chapter does not repeat the presentation of basic statistical and mathematical methods used in PRA, since these are described in other chapters and appendices of this document. In summary:

- ☞ *This chapter focuses on application of PRA techniques to ERA at Superfund sites.*
- ☞ *The reader is assumed to be familiar with the basic methods used in ERA at Superfund sites, and this chapter does not address basic tactical and technical issues in ERA.*
- ☞ *The reader is assumed to be familiar with the basic mathematical principles and techniques of PRA as described in other chapters and appendices of this document.*

4.2 DECIDING IF AND WHEN TO USE PRA IN ECOLOGICAL RISK ASSESSMENT

As shown in Figure 4-2, the ERA process for Superfund includes a number of scientific/ management decision points (SMDPs) (U.S. EPA, 1997a). The SMDP is a point of consultation between the risk manager, EPA Regional Biological Technical Assistance Group (BTAG) coordinator, EPA regional ecotoxicologist, and other stakeholders, and is intended to provide an opportunity for re-evaluation of direction and goals of the assessment at critical points in the process. It is during the SMDP discussions that it is important to decide whether or not a PRA is likely to be useful in decision making. If so, the pursuit of distributed data is justified. Within the 8-step process of developing the ERA, PRA could provide insight at several steps. A decision to move forward with distributional analyses should be considered within the BTAG context during the documentation of the outcome of the SMDPs after Step 3 within the process. As a reminder, PRA is NOT intended to be a replacement for point estimate analyses; rather PRA supplements the required presentation of point estimates of risk. It is also emphasized that the use of PRA should never be viewed as or used in an attempt to simply generate an alternative risk estimate or PRG, compared to that which was derived by a point estimate ERA; instead, PRA should be

used to provide insightful information on distributions of various factors (exposure, toxicity, and hazards) which can provide weight-of-evidence evaluations of potential risks in conjunction with a point estimate ERA. There are a number of factors to consider in making these decisions, as discussed below.

4.2.1 TECHNICAL CONSIDERATIONS

The fundamental reason for performing any predictive risk assessment (point estimate or probabilistic) is to provide information to risk managers in order to help support the risk management decision-making process. As noted above, a properly performed PRA may help to yield more description of variability in exposure and risk than can be achieved using the point estimate approach. Therefore, if any of a site's data may be better described and evaluated by distributions, then a PRA can be applied to any part of an ERA or even to the entire ERA for expressing risk characterization in probabilistic terms; again, always in conjunction with the required point estimate ERA. However, when risk estimates derived from the point estimate approach are either far below or far above a level of risk management concern, any such potential improvements in risk characterization are not likely to influence risk management decision making. In these cases, PRA is not likely to be as useful in decision making. Even so, PRA may help in these situations by providing information that may be useful in better deciding where the gradient of excess risks are reduced to acceptable levels. Rather, it is more common for a PRA to be useful when point estimates of risks are close to the decision threshold (such that PRA-based refinements in the risk estimates might be important in making risk management decisions). It is for this reason that PRA may be useful to apply either during the development of the ERA after the screen (Steps 3 to 6, U.S. EPA, 1997a), or after point estimate results from the baseline ERA have been completed (Steps 1 to 7, U.S. EPA, 1997a).

The results of a point estimate risk assessment will normally present the range of risks based on central tendency exposure and reasonable maximum exposure input assumptions and on the no-observed-adverse-effect-level (NOAEL)- and lowest-observed-adverse-effect-level (LOAEL)-based TRVs (U.S. EPA, 1992b, 1997b). The bounds for the highest HQ are derived from the ratio of the RME compared to the NOAEL-based TRV, and the bounds for the lowest HQ are based on the ratio of the CTE compared to the LOAEL-based TRV. These two bounded extreme estimates of risk can be used to screen out cases where PRA is not likely to be as useful. That is, if the risk to the RME receptor is clearly below a level of concern using the NOAEL-based TRV, then risks to the exposed population are likely to be low and PRA analysis is likely not needed. Likewise, if risks to the CTE receptor are clearly above a level of concern using the LOAEL-based TRV, then risks to the exposed population are likely to be of definite concern, and a PRA may not provide as much additional useful information to the risk manager, except in the case where uncertainties remain high and the derivation of an appropriate and realistic clean-up goal may be difficult. If the risks are intermediate between these two bounds (e.g., risks to the CTE receptor are below a level of concern based on the LOAEL-based TRV but are above a level of concern based on the NOAEL-based TRV), then PRA might be helpful in further characterizing the site risks in balance with the point estimates of risks and in supporting decision making or in deciding if additional iterations of analyses would be needed. This concept is illustrated graphically in Figure 4-3.

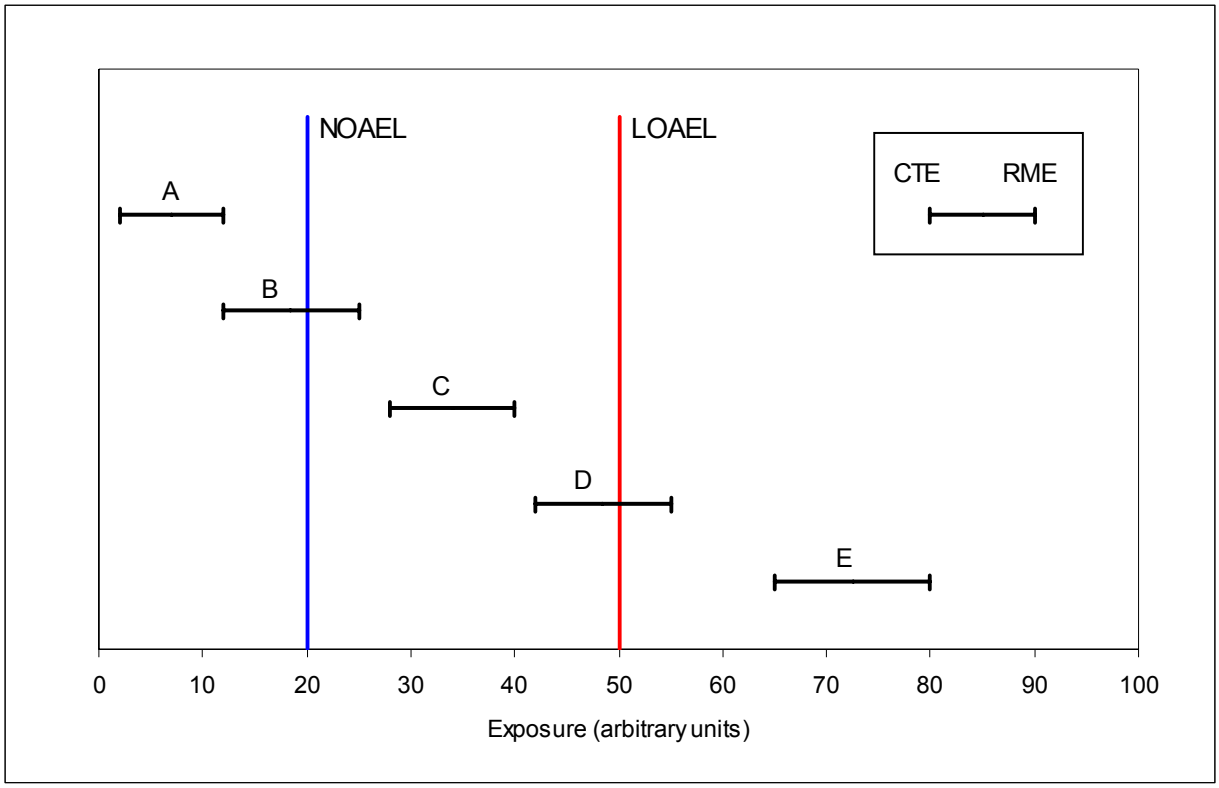


Figure 4-3. Example of cases where use of PRA may be helpful. In cases A and E, the range of risks (CTE to RME) estimated by the point estimate method are either well below (Case A) or well above (Case E) the likely level of concern based on the NOAEL-LOAEL range, and PRA is not likely to alter risk management decisions regarding the potential need for remediation. In cases B, C, and D, the point estimates of risk overlap or fall within the range of potential concern, suggesting that PRA-based risk estimates might be helpful in supporting risk management decisions.

The second main technical reason to consider conducting PRA is that the PRA methodology can help characterize and quantify the degree of variability and uncertainty around any particular estimate of exposure or risk (e.g., the CTE or RME). The purpose of the analysis would be to estimate the uncertainty around an exposure or toxicity or risk estimate, generally with little or no additional data acquisition. The only additional information needed to perform the analysis is an estimate of the uncertainty in the true parameter values of the key variables in the variability model. In some cases, these estimates of uncertainty around parameter values may be developed from statistical analysis of the available data. Alternatively, professional judgment may be used to establish credible bounds on the parameters, especially when relevant data are sparse.

☞ *Even in the presence of data gaps, uncertainty analysis using PRA can provide useful information. Indeed, it is when data are limiting or absent that a quantitative probabilistic analysis of uncertainty may be most helpful.*

4.2.2 COST AND SCHEDULE CONSIDERATIONS

Performing a PRA can sometimes add time and cost to an ERA. As discussed in Chapter 2, in part, the decision to progress from a point estimate assessment to a PRA reflects a belief that the potential value of the PRA for risk management decision making outweighs the additional time and costs. The tiered process encourages a systematic approach for both the point estimate and probabilistic assessments, whereby the least complex methods are applied first. For example, the initial Tier 2 assessment may be conducted with a set of preliminary probability distributions for variability (PDFv), developed with much the same information and assumptions that were applied to develop point estimates in Tier 1. Parameter values can be estimated by setting the arithmetic mean equal to the CTE point estimate, and the 95th percentile equal to the RME point estimate. The choice of distributions may differ depending on the state of knowledge for a particular variable (see Appendix B). For example, unbounded variables might be characterized with lognormal distributions while bounded distributions are characterized by beta or Johnson Sb distributions. Certain variables may continue to be characterized by point estimates, especially if the sensitivity analysis suggests that the chemical, pathway, and/or exposure variables are relatively minor contributors to total exposure and risk. The decision to collect additional data or explore alternative methods for developing probability distributions can be reexamined in an iterative fashion by evaluating the expected benefits of the added information to the risk management decision-making process. These concepts are presented in greater detail in Chapter 2 (see Figures 2-1 and 2-2).

4.3 PROBLEM FORMULATION

Once a decision has been made to include PRA in an ERA, the first step should be to re-visit the problem formulation step and carefully determine the scope and objectives of the PRA. Typically, a considerable amount of knowledge will have been gained during the screening level and baseline point estimate evaluations, and this knowledge should be used to help focus and narrow the scope of the PRA. That is, the PRA will generally utilize the same basic exposure and risk models used in the point estimate approach, but the PRA will typically evaluate only a sub-set of the scenarios considered. For example, chemicals, pathways, and/or receptors that are found to contribute a negligible level of exposure or risk may usually be omitted from the PRA, while those factors that contribute significantly to an excess level of risk concern in the point estimate approach should generally be retained. As noted previously, when a chemical or pathway is omitted from a PRA analysis, this does not mean that it is eliminated from the overall risk assessment; rather, it may be kept in the assessment as a point estimate.

The next step in problem formulation for a PRA should be to define whether the goal of the analysis is to characterize variability alone, or to characterize both variability and uncertainty. In either case, sensitivity analysis (as summarized in the preceding paragraph, or for more details see Appendix A) should be used to help identify which of the input variables contribute the most to the variability in the outputs (exposure, toxic effects, or risk), and the initial PRA should focus on defining the probability density functions (PDFs) for those input variables. An analysis of uncertainty, if thought to provide additional useful information, may also be included at the initial level, or may be delayed until the initial analysis of variability is completed.

As always, problem formulation should be viewed as an iterative process, and it is reasonable and appropriate that decisions regarding the scope and direction of the PRA should be reassessed (at SMDPs) as information becomes available from the initial evaluations. As stressed above, the fundamental criterion which should be used is whether or not further PRA evaluations are likely to provide additional information to a point estimate ERA that will help strengthen and support the risk management decision-making process.

4.4 MODELING VARIABILITY IN EXPOSURE

There are two main types of descriptors of exposure that may be used in ERA: dose and concentration. For terrestrial receptors such as mammals or birds, exposure is most often described in terms of ingested dose (mg/kg-day). In most cases, this will be based on chemical ingested from drinking water and/or the diet, including incidental soil ingestion, but could also include amounts of chemical taken up across the skin or through inhalation as additional routes of exposure. The exposure levels are most often expressed as doses, since that term tends to normalize the confounding factors of variable daily intake rates and body weights that occur if/when one only evaluates concentrations. For aquatic receptors, the main route of exposure is usually by direct contact and less often by ingestion, so exposure is usually characterized in terms of concentration of contaminants in surface water, pore water and/or sediment. Likewise, exposure of terrestrial plants and terrestrial invertebrates, such as earthworms, is usually described in terms of concentration of contaminants in soil. In some cases, exposure of terrestrial receptors is characterized in terms of specific tissue or whole-body concentrations of contaminants. Examples of calculating and presenting dose-based and concentration-based distributions of exposure are presented below.

4.4.1 CHARACTERIZING VARIABILITY IN DOSE

The general equation used for calculating the **dose** of a contaminant of concern in a specified environmental medium (e.g., water, soil, air, diet, etc.) by a particular member of a population of exposed receptors is:

$DI_{i,j} = C_i \times IR_{i,j} / BW_j$

where:

$DI_{i,j}$ = Average daily intake of chemical due to ingestion of medium "I" by a population member "j" of the exposed population (mg/kg-day)

C_i = Concentration of chemical in environmental medium "I" (mg/unit medium)

$IR_{i,j}$ = Intake rate of medium "I" at the site by population member "j" (units of medium per day)

BW_j = Body weight of population member "j" (kg)

Total exposure of a population member "j" is then the sum of the exposures across the different media:

$DI_{total,j} = \sum DI_{i,j}$

In this basic equation, $IR_{i,j}$ and BW_j are random variables (i.e., they have different measurable values for different members of the exposed population) that are often correlated. For example, a receptor with a relatively low intake rate can also be expected to have a low body weight. Some studies utilize paired measurements of IR and BW by individual, and present a distribution of the ratio ($IR_{i,j}$ /BW_j), referred to as a body weight-normalized intake rate (mg/kg-day). This expression provides an alternative to using a correlation coefficient to relate two input variables (see Appendix B), and can be entered into the dose equation as follows:

$$DI_{i,j} = C_i \times \left(\frac{IR_{i,j}}{BW_j} \right)$$

where the ratio is characterized by a single probability distribution. Because the variability in this ratio is likely to be different than the variability in the ratio of the IR and BW variables treated independently,

accounting for the correlation can affect the distribution of dose and risk. If empirical data for quantifying the ratio are limited but a relationship is expected, plausible ranges of correlations may be explored as a source of uncertainty in the risk estimates.

The concentration term (C_i) may be characterized by a point estimate or a probability distribution, depending on the relationship between the geographic scales of the measurement data and receptor home range (see Appendix C, Section C.3.1). If the home range of the receptor is small compared to the spatial distribution of sampling locations, C_i may be characterized by the probability distribution for variability in measured concentrations. Alternatively, if the home range is large compared with the exposure area evaluated, then a point estimate (e.g., mean or uncertainty in the mean) may be more appropriate.

In the PRA approach, PDFs should be defined for as many of the input variables as reasonable, especially for those variables that are judged (via sensitivity analysis) to contribute the most to the variability in total exposure. The basic principles for selecting the key variables to model as PDFs are presented in Appendix A, and the basic methods used for selecting and fitting distributions are described in detail in Appendix B.

Figure 4-4 shows several examples of graphical formats which may be used to present the estimated distribution of ingested doses in an exposed population. If a single distribution is plotted (top panel), the PDF format is usually the most familiar and useful for risk assessors and managers, but the cumulative distribution function (CDF) format tends to be less cluttered when multiple distributions are shown (e.g., compare the middle graph to the bottom graph). In addition, percentiles can be read directly from a CDF format, but not from a PDF format graph. In all cases, it is very useful to superimpose the CTE and RME point estimate ranges of exposure directly on the same graph as is used to show the distribution of exposures estimated by PRA. This provides a convenient way to compare the results of the two alternative computational methods, and interpret additional information that the PRA can add to the point estimate ERA.

☞ *A conventional point estimate, range of exposure (CTE to RME) or toxicity (NOAEL to LOAEL) and corresponding risk ranges should be calculated and presented for comparison with the PRA results.*

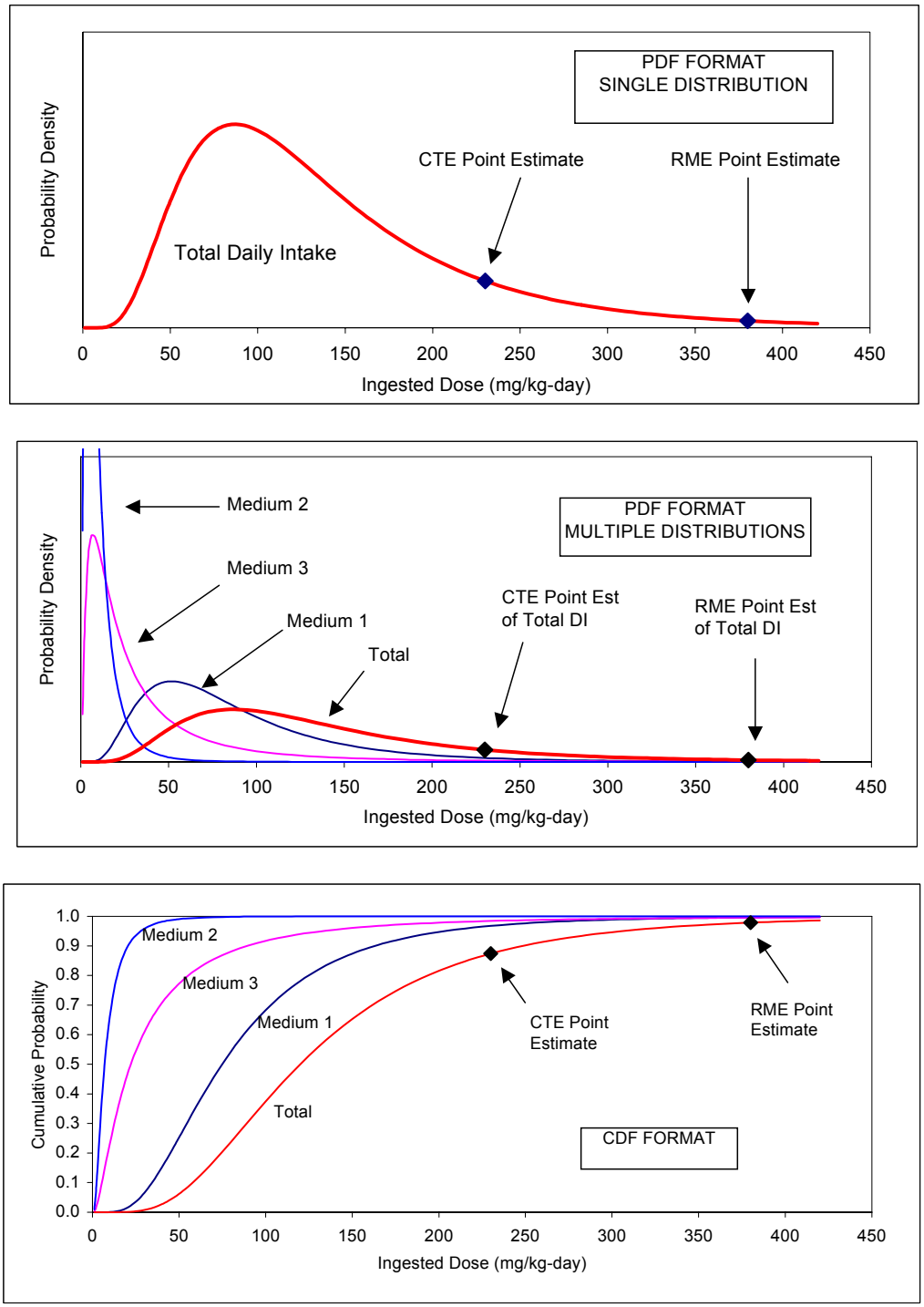


Figure 4-4. Example Graphical Presentations of Dose Distributions.

4.4.2 CHARACTERIZING VARIABILITY IN EXPOSURE CONCENTRATION

As noted above, in some cases the most appropriate descriptor of exposure is concentration (either in an abiotic medium such as water, soil, or sediment, or in the tissues of the receptor), rather than ingested dose. Assuming that the concentration values in the medium of concern are measured rather than modeled, PRA is not required to generate the distribution of concentrations. Rather, the available data may be used to define an appropriate theoretical or empirical distribution function (EDF), as described in Appendix B. If concentrations in the medium are modeled (calculated by PRA) rather than measured, then the exposure distribution may be estimated by using distribution functions (PDFs or CDFs, rather than using point estimates as inputs to the fate and transport model(s) and/or uptake models that predict the concentration levels in the medium of concern. The resulting distribution(s) of concentration may be displayed graphically using the same formats as illustrated in Figure 4-4, except that the x-axis has units of concentration rather than dose. As above, the point estimate ranges of concentration used in the CTE and RME calculations should be plotted on the same graphs to provide a convenient basis for comparing the results of the two approaches and to help interpret the additional information that the PRA can add to the point estimate outputs.

4.5 MODELING VARIABILITY IN TOXICITY

4.5.1 VARIABILITY IN RESPONSE AMONG MEMBERS OF A POPULATION

Data on the toxicity of a chemical usually comes from laboratory studies whereby groups of organisms (laboratory mammals, fish, benthic organisms, plants, earthworms, etc.) are exposed to differing levels of chemical, and one or more responses (endpoints) are measured (survival, growth, reproduction, etc.). These toxicological observations define the exposure-based stressor-response curve that is characteristic for that specific receptor, chemical, and response.

In the point estimate approach, information from the dose/stressor-response curve is generally converted to one or more TRVs, each representing a specific point on the dose-based or concentration-based stressor-response curve. For example, the highest dose or concentration that did not cause a statistically significant change in a toxicologically significant endpoint is defined as either the NOAEL dose or the no-observed-effect concentration (NOEC), while the lowest dose or concentration that did cause a statistically significant effect on a relevant endpoint is the LOAEL dose or the lowest-observed-effect concentration (LOEC). Generally, exposures below NOAEL- or NOEC-based TRVs are interpreted to pose acceptable risk, while exposures above LOAEL- or LOEC-based exposures are judged to pose potentially unacceptable risk. It is essential to note the need for high quality toxicity data to derive reliable and confident TRVs. Strong sampling and study designs, that generate data for site exposure factors and toxicological stressor-response relationships, are of critical importance for producing high quality ERAs by either point estimate or PRA approaches. Shortcomings in either area could be major data gaps or uncertainties that detract from the confidence in the risk characterization of the ERA, and may be a basis for pursuing additional iterations of sampling or studies that are more strongly designed to fill those critical data gaps and reduce uncertainty.

Use of the TRV approach, however, does have some potential limitations. Most important is that the ability of a study to detect an adverse effect depends on both the range of doses tested and the statistical power of the study (i.e., the ability to detect an effect if it occurs). Thus, studies with low power (e.g., those with only a few test animals per dose group) tend to yield NOAEL or NOEC values that are higher than studies with good power (those with many animals per dose group). In addition, the choice of the TRV is restricted to doses or concentrations that were tested, which may or may not be close

to the true threshold for adverse effects, and this uncertainty increases as the interval between doses increases. Finally, it is not always easy to interpret the significance of an exposure that exceeds some particular TRV, since the severity and incidence of response depends on the shape and slope of the exposure response curve (information that is not captured in a point estimate TRV).

One way to resolve some of these stressor-response limitations is to apply uncertainty factors to the NOAEL or NOEC and LOAEL or LOEC, which calculates an adjusted TRV that reduces the study's exposure level of concern to account for those uncertainties, so that there is a lesser chance of overlooking possible adverse exposures (i.e., avoiding a false negative conclusion). Another way to resolve some of the stressor-response limitations is to fit a mathematical equation to the available exposure-response data and describe the entire exposure-response curve. This may be done using any convenient data fitting software, but EPA has developed a software package specifically designed for this type of effort. This software is referred to as the Benchmark Dose Software (BMDS), and is available along with detailed documentation and explanation of the methodology at **www.epa.gov/ncea/bmds.htm**.

The most appropriate mathematical form of the exposure-response model depends on whether the endpoint measured is discrete and dichotomous (e.g., survival) or continuous (e.g., growth rate). For a dichotomous endpoint, the result of the fitting exercise is a mathematical exposure-response model P that yields the probability of a response in an individual exposed at any specified level of exposure (expressed either as dose or concentration). Exhibit 4-3 shows an example of this process using hypothetical data. Thus, for an individual with an exposure level of "x", the probability of a response in that individual is simply P(x). In a population of individuals with exposures x1, x2, x3, ...xi, the expected number of responses (e.g., deaths) in the exposed population is the sum of the probabilities across all individuals in the population. Stated another way, the average fraction of the population that will experience the response is given by the expected value of P (i.e., the average value of P(x)).

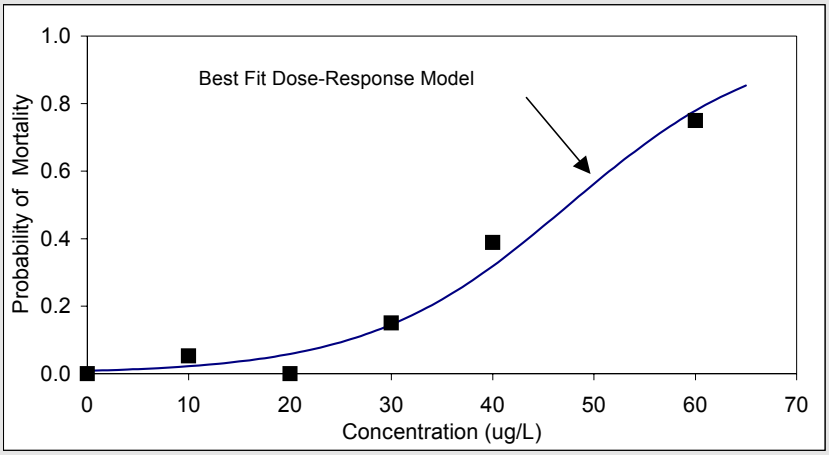
EXHIBIT 4-3

MODELING VARIABILITY IN RESPONSE FOR A DICHOTOMOUS ENDPOINT

The following data are from a hypothetical study of the acute lethality (24 hour) of a chemical using fathead minnows as the test organism:

Concentration ug/L	Number Tested	Survival	
		Dead	Alive
0	20	0	20
10	19	1	18
20	20	0	20
30	20	3	17
40	18	7	11
60	20	15	5

These data were fit to each of the dichotomous models available in BMDS. The best-fit model was the logistic equation. A graph of the best fit curve is shown below.



Basic Equation

Probability of mortality (conc) = 1 / (1 + exp(-a - b*conc))

Best fit parameters

a	-4.80
b	0.101

Goodness of Fit

P	0.604	P=Chi Square Goodness of Fit test statistic
AIC	79.12	AIC=Akaike's Information Criterion

For a continuous endpoint, the BMDS software yields equations that give the expected mean response m(x) at a specified exposure level, along with the standard deviation s(x) that characterizes how variable the response is among different individuals exposed at that same exposure level. The standard deviation may be modeled either as a constant (homogeneous variance) or a function of the exposure level (heterogeneous variance), with the choice depending on which approach yields the best agreement with the observed variances. In most cases there will not be sufficient data to allow a meaningful analysis of the true shape of the underlying distribution of responses at a given exposure, so the choice of the distributional form of the variability in response will require an assumption. In the absence of any clear evidence to the contrary, it is considered likely that the distribution of responses will not be strongly skewed, and that the distribution may be reasonably well modeled using a normal PDF (truncated as necessary to prohibit selection of biologically impossible or implausible values). Thus, variability in response at dose "x" may generally be modeled as:

Response(x) ~ NORMAL[m(x), s(x), min, max]

However, if available data suggest some other distributional form is more appropriate, that form should be used and justified.

Exhibit 4-4 shows an example of this process using hypothetical data. In this case, the mean response was found to be well modeled by the Hill equation, and the standard deviation was found to be best characterized as a constant (rho=0). Thus, given an exposure level "x", the mean response m(x) may be calculated from the model, and this value along with the standard deviation may then be used as parameters for an appropriate type of PDF (e.g., normal) to describe the expected distribution of responses in a population of different individuals exposed at level "x". Section 4.7.2 describes methods that may be used to characterize and quantify the uncertainty associated with this approach.

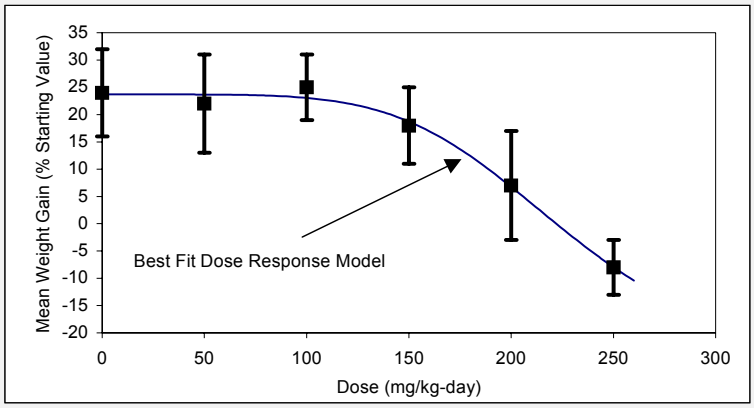
EXHIBIT 4-4

MODELING VARIABILITY IN RESPONSE FOR A CONTINUOUS ENDPOINT

The following data are from a hypothetical study of the effects of a chemical on the growth of laboratory mice. Animals were exposed to the chemical via drinking water for 21 days. The measurement endpoint was weight gain, expressed as a percentage of the starting weight of each animal.

Ingested dose mg/kg-day	Number Tested	Weight Gain (% Starting Value)	
		Mean	Stdev
0	5	24	8
50	5	22	9
100	5	25	6
150	5	18	7
200	5	7	10
250	5	-8	5

These data were fit to each of the continuous models available in BMDS. The best-fit model was the Hill equation with constant variance. A graph of the best fit curve is shown below.



Basic Equations

Mean Response(d) = int + v*d^n / (k^n + d^n)
Variance(d) = alpha*mean response(d)^rho

Best fit parameters

int	23.70
v	-51.41
n	5.295
k	228.7
alpha	48.5
rho	0 (constant variance)

Goodness of Fit

P	0.685	P=Chi Square Goodness of Fit test statistic
AIC	154.5	AIC=Akaike's Information Criterion

4.5.2 VARIABILITY IN RESPONSE AMONG SPECIES

In some cases, risk management decisions may also consider community-level effects as well as population-level or sub-populations effects. That is, a stressor might be considered to be below a level of concern for the sustainability of a community if only a small fraction of the total number of exposed species are affected. In this case, toxicological responses may be best characterized by the distribution of toxicity values across species. This is referred to as a Species Sensitivity Distribution (SSD). This type of approach is generally used for communities of aquatic receptors, since all of the different species that make up the community (e.g., all fish, benthic invertebrates, aquatic plants, and amphibians that reside in a stream) will be exposed to approximately the same concentration of contaminant in the water. The process for generating an SSD consists of the following steps:

(1) Select an appropriate type of endpoint (lethality, growth, reproduction, etc.), and select an appropriate type of point estimate from the exposure-response curve for each species. For example, the TRV might be the LC50 for lethality or the EC20 for growth. The key requirement is that the SSD be composed of TRV endpoints that are all of the same type, not a mixture.

(2) Collect all reliable values for that type of TRV from the literature for as many relevant species as possible. When more than one value is available for a particular species, either select the value that is judged to be of highest quality and/or highest relevance, or combine the values across studies to derive a single composite TRV for each species. It is important to have only one value per species to maintain equal weighting across species.

(3) Characterize the distribution of TRVs across species with an appropriate CDF. Note that there is no *a priori* reason to expect that an SSD will be well characterized by a parametric distribution, so both parametric and empirical distributions should be considered.

Once an SSD has been developed, the fraction of species in the exposed community that may be affected at some specified concentration may be determined either from the empirical distribution or from the fitted distribution. Exhibit 4-5 shows examples of this approach. In this hypothetical case, the TRV selected for use was the LC_{low} (in this case, the LC_{low} is defined as all LC values <=LC10). A total of 13 such values were located. The first graphical presentation is the empirical distribution function, where the Rank Order Statistic (ROS) of each value is plotted as a function of the log of the corresponding value. This may be used directly to estimate the fraction of the species in a community that will be affected by any particular environmental concentration. For example, in this case, it may be seen that a concentration of 10 ug/L would be expected to exceed the LC_{low} for about 33% of the aquatic species for which toxicity data are available. The second graph shows how the data may be characterized by fitting to a continuous distribution. In this case, a lognormal distribution was selected as a matter of convenience, but other distributions may also yield acceptable fits. Based on the best fit lognormal distribution for the SSD data, it is calculated that a concentration of 10 ug/L would be expected to impact about 31% of the exposed species. However, as noted above, there is no special reason to expect that an SSD will be well characterized by a continuous parametric distribution, so some caution should be used in the use of a continuous distribution to fit an SSD, especially when the SSD is based on a limited number of species and when the purpose of the SSD is to estimate percentiles and exposures outside the observed range. The risk assessor should always present an evaluation of the robustness of an SSD to aid in the decision process.

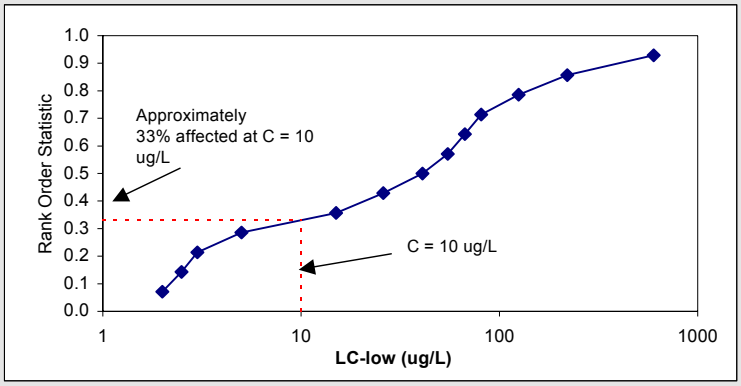
EXHIBIT 4-5

HYPOTHETICAL SPECIES SENSITIVITY DISTRIBUTION

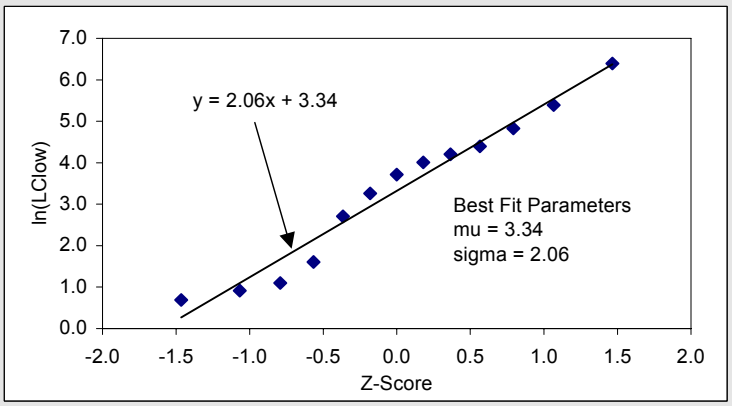
Hypothetical Data

Species	LC_{low}	$ln(LC_{low})$	Rank	ROS	z-score
a	2	0.693	1	0.07	-1.465
b	2.5	0.916	2	0.14	-1.068
c	3	1.099	3	0.21	-0.792
d	5	1.609	4	0.29	-0.566
e	15	2.708	5	0.36	-0.366
f	26	3.258	6	0.43	-0.180
g	41	3.714	7	0.50	0.000
h	55	4.007	8	0.57	0.180
i	67	4.205	9	0.64	0.366
j	81	4.394	10	0.71	0.566
k	125	4.828	11	0.79	0.792
l	220	5.394	12	0.86	1.068
m	600	6.397	13	0.93	1.465

Example EDF: ROS vs LC_{low} (log-scale)



Example Parametric Fit: (Lognormal)



4.6 MODELING VARIABILITY IN RISK

4.6.1 VARIABILITY IN HAZARD QUOTIENT

As noted above, the most common descriptor of risk used in predictive risk assessments is the Hazard Quotient (HQ). The HQ is the ratio of the exposure for some generalized or typical hypothetical member of the receptor population at a site, compared to an appropriate TRV value that equates to an acceptable level of risk for that receptor and chemical. Usually the HQ approach is not based on a single value, but on a range of values in which different levels of exposure (CTE and RME) are compared to both the NOAEL to LOAEL benchmarks. In general, HQ values below 1 are interpreted as indicating acceptable risk, while HQ values above 1 are interpreted as indicating the potential for adverse effects.

Because exposure varies among different members of an exposed population of receptors, HQ values also vary among members of the exposed population. Several alternative approaches for characterizing this variability by PRA methods are presented below.

Variability Within a Population

Figure 4-5 illustrates the simplest approach for summarizing variability in HQ values among the members of an exposed population. In this format, the TRV values appropriate for a particular exposure are simply superimposed on the graph illustrating the distribution of exposures. This may be done either for a dose-based (as shown in the figure) or for a concentration-based exposure parameter. This format allows an easy evaluation of the fraction of the population above (HQ > 1) and below (HQ < 1) each TRV, especially when presented in CDF format. However, this format does not allow for a quantitative estimate of the fraction of the population with HQ values above any value other than 1, although a similar calculation and presentation could be made for any multiple of the TRVs, which would directly equate to that multiple of the HQ (e.g., depicting the

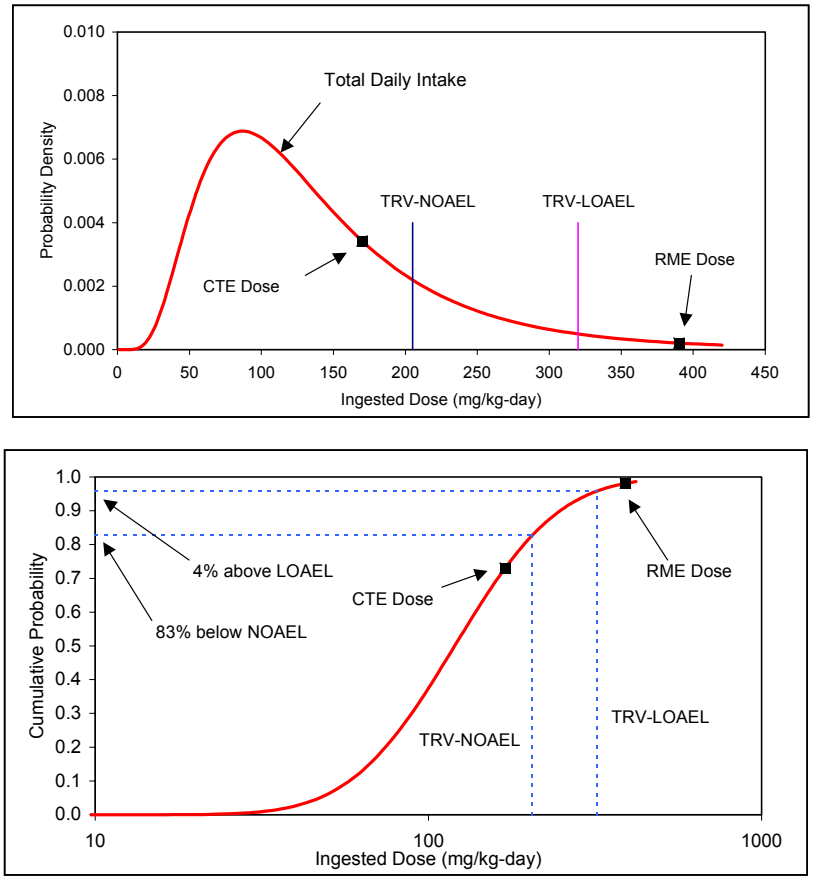


Figure 4-5. Example Comparison of Exposure Distribution to TRV.

results for a value equal to 10-times the TRV would show the fraction of the population with an HQ greater than 10).

More directly, the distribution of HQ values may be calculated by dividing each exposure value by one or all of the TRVs based on the NOAEL, LOAEL, BMDL, etc., as shown in Figure 4-6. Note that dividing a distribution by a constant does not change the shape of the distribution (only its scale), so the shape of the HQ distribution will appear identical to that of the exposure distribution. Figure 4-6 illustrates two HQ distributions; one calculated using the NOAEL-based TRV, the other using the LOAEL-based TRV. In a case such as this where there are two or more HQ distributions, a CDF format is generally easier to evaluate than a PDF format, since overlap between the curves is minimized. The CDF format allows an easy quantitative evaluation of the fraction of the population above and below any particular HQ level. For example, in the case shown in Figure 4-6, it may be seen that 83% of the population is expected to have HQ values below 1 based on the NOAEL-based TRV, while 4% are expected to have HQ values above 1 based on the LOAEL-based TRV. This type of description (percentage of the population with HQ values within a specified range) is very helpful in predicting proportions of a population exposed to specified doses of concern.

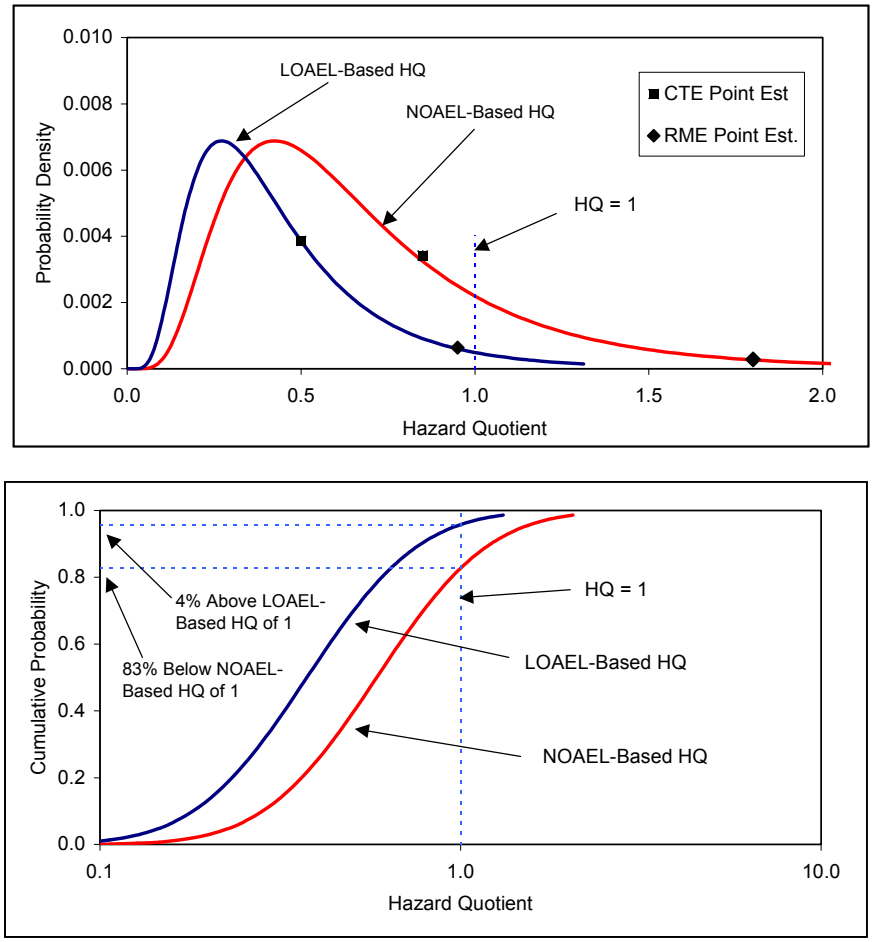
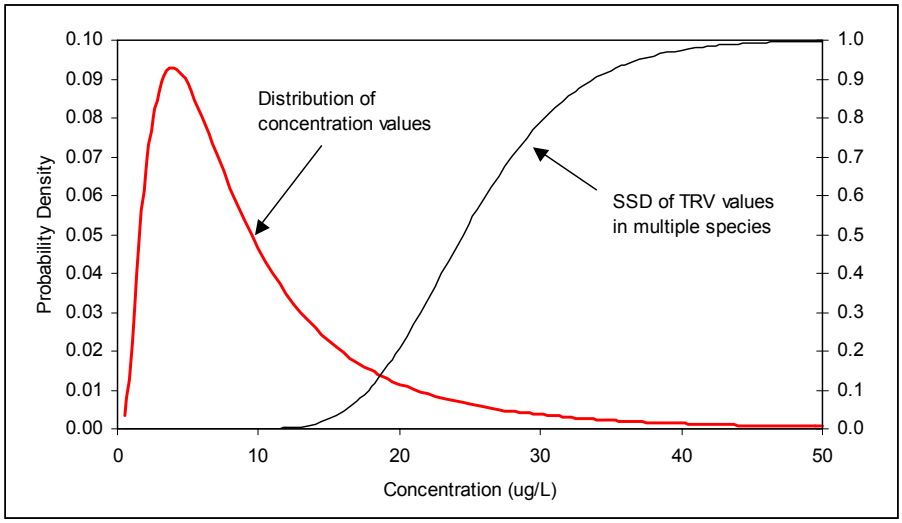
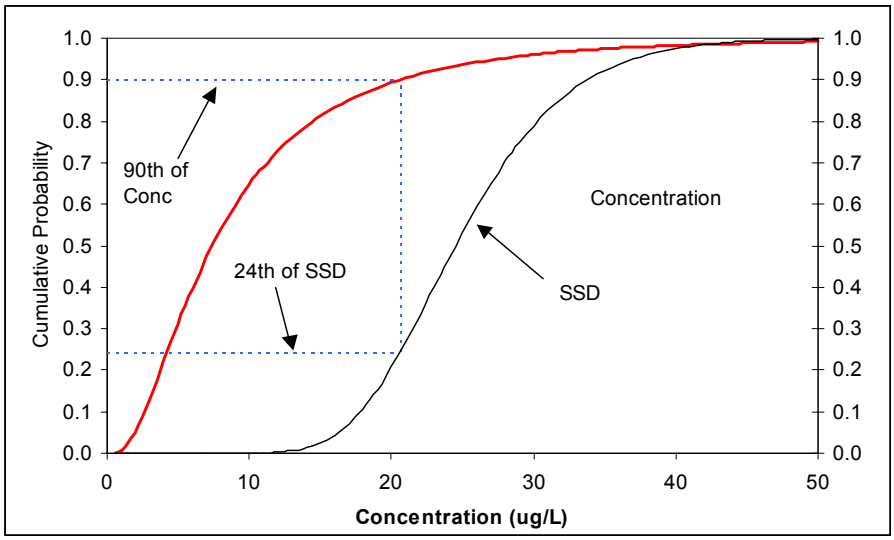


Figure 4-6. Example Distribution of HQ Values.

Variability Between Species

A similar approach may be used for characterizing the variability in risks among different species in a community. Figure 4-7 is an example that compares the distribution of concentration values in a water body (the variability might represent either time or space) to an appropriate SSD of TRVs for different species of aquatic receptors that might reside in that water body. Three different graphical formats are illustrated. In the upper panel, the PDF of concentration is compared to the CDF of the SSD. This format is easy to understand and may be interpreted visually, but is difficult to interpret quantitatively. The middle panel shows that same information, but with both distributions presented in CDF format. This allows for a quantitative evaluation of the fraction of the species that will be above their respective TRVs at any specified part of the exposure distribution. For example, using a simple graphical interpolation process (shown by the dashed lines), it may be seen that the 90th percentile of concentration (21 ug/L) will impact approximately 24% of the exposed species. The bottom panel shows the results when this same process is repeated (mathematically) for each of the concentration percentiles. As seen, hazards to the community of receptor species is quite low until concentration values reach the 80th to 85th percentile, but then rise rapidly. For example, a concentration value equal to the 95th percentile (about 28 ug/L, which will occur approximately 5% of the time) is expected to impact approximately 68% of the exposed species, and the 99th percentile (which will occur about 1% of the time) is expected to impact nearly all of the exposed species.





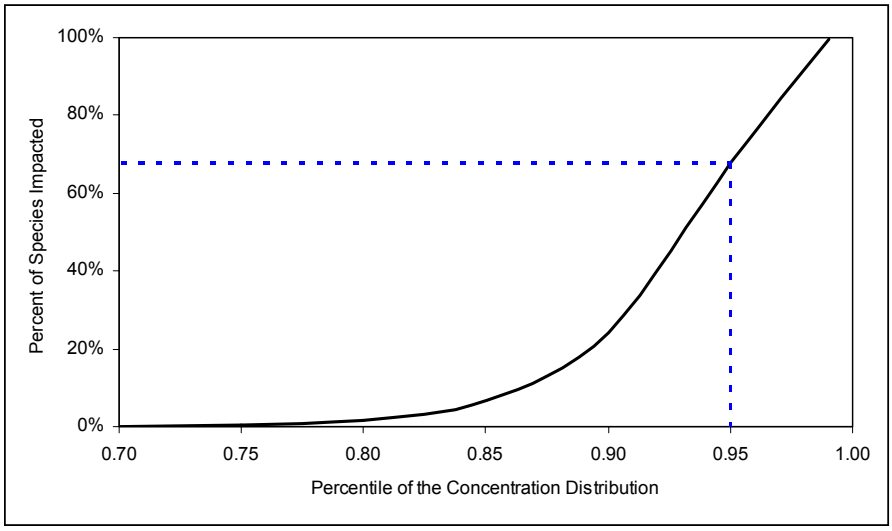


Figure 4-7. Example Presentation of Species Sensitivity Distribution.

4.6.2 VARIABILITY IN RESPONSE

As noted above, HQ and Hazard Index (HI) (where appropriate) values are a convenient way to characterize risk to ecological receptors, but interpreting the biological significance of the ranges of HQ values greater than 1 is not always easy. One of the main advantages to the PRA approach is that distributions of exposure may be combined with exposure-response distributions in order to generate distributions that characterize the frequency and magnitude (severity) of responses in an exposed population. Two examples of this approach are presented below.

Example 1: Dichotomous Response

In this hypothetical example, a toxic chemical is being transported by surface water run-off from a Superfund site into a nearby stream. Because of short-term and seasonal variability in rainfall levels (which influences both run-off rate and stream flow), the concentration of the chemical in the stream has been observed to vary as a function of time. The risk manager at the site wants to know two things: (1) How often will the concentration enter a range that can cause acute lethality in fish?; and (2) When that happens, what percent of the fish population is likely to die? Exhibit 4-6 summarizes the hypothetical concentration data and illustrates the basic approach. In this case, the concentration data are most conveniently modeled as an empirical PDF. Next, assume that the acute concentration-lethality curve is available for the chemical of interest in a relevant indicator species of fish. For convenience, assume the response function is the same as that shown in Exhibit 4-3. Then, the PDF for acute mortality may be generated by repeated sampling from the concentration distribution and calculating the probability of response (acute mortality) for each concentration value selected. Because this is a case where the entire population of fish at the exposure location may be assumed to be exposed to the same concentration in water, the probability of mortality in a single fish is equivalent to the average fraction of the population that is expected to die as a result of the exposure. The resulting PDF is shown in the graph in Exhibit 4-6. As seen, lethality is expected to be low or absent about 95% of the time, but about 5% of the time the concentration may enter a range where acute lethality may occur. The extent of mortality within the exposed population is expected to range from about 20% at the 97^{th} percentile of exposure (i.e., this is expected to occur about 3% of the time), up to about 70% at the 99^{th} percentile of exposure (i.e., this is expected to occur about 1% of the time).

EXHIBIT 4-6

MODELING VARIABILITY IN A DICHOTOMOUS RESPONSE

Scenario

Exposure of a population of fish to concentration values in a stream that vary over time

Hypothetical Concentration Data in Water

Value		Percentile
0.5	(1/2 DL)	0.00
1.1		0.10
2.5		0.25
5.1		0.50
9.2		0.75
15.8		0.90
24.7		0.95
52.6		0.99
83.1	(max)	1.00

Response

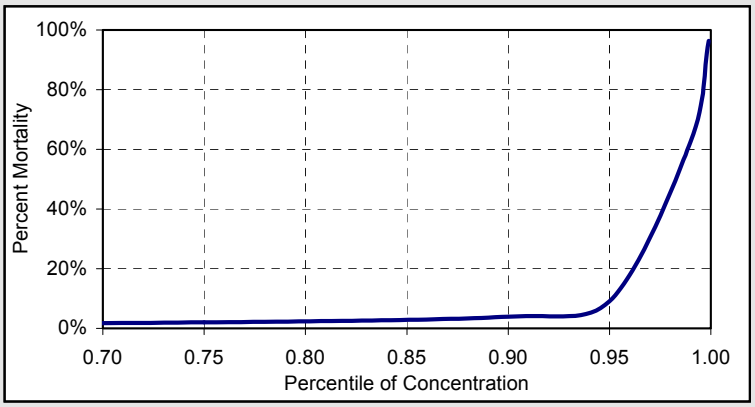
Endpoint = acute mortality
Stressor-response model fit (see Exhibit 4-2)
$P(c) = 1/(1+\exp(4.8 - 0.1*c))$

PRA Simulation

Step 1 Draw a concentration at random from the empiric distribution
Step 2 Calculate the probability of mortality at that dose
Track this as the forecast cell

Example Output

Percentile	% Lethality
0.050	0.9%
0.250	1.0%
0.500	1.4%
0.750	2.0%
0.900	3.9%
0.950	9.1%
0.990	63%
0.999	96%



Example 2: Continuous Response

Exhibit 4-7 provides a hypothetical example of modeling variability in response for a continuous endpoint. In this example, assume that a toxic chemical has been released by a Superfund site and has been transported in low levels by air to a nearby meadow. Among the receptors of potential concern in the meadow are a number of different types of small mammal, and the field mouse has been selected to serve as an indicator species for this group. The goal of the PRA is to characterize the effects of the chemical on the growth of field mice in the meadow. Exposure occurs mainly by ingestion of seeds that have been contaminated by uptake of the chemical from soil, and it has been determined that the variability in average daily intake (DI) of chemical from the diet can be modeled as a lognormal distribution with mean of 104 mg/kg-day, and a standard deviation of 127 mg/kg-day. Assume for convenience that the exposure-response curve for growth inhibition in mice by the chemical is the same as that presented previously in Exhibit 4-4. Given these inputs, the expected distribution of responses is derived as follows:

Step 1: Draw a random value for the DI of a random member of the population
Step 2: Calculate the mean response m(d) and the standard deviation of the response s(d) for a group of individuals exposed at that dose (d)
Step 3: Define the distribution of responses at that dose as NORMAL[m(d), s(d)]
Step 4: Draw a response from that distribution, and track this as the output variable

An example of the output for this example is shown in the two graphs at the bottom of Exhibit 4-7. As seen, mice that are not exposed to the chemical display a range of growth rates ranging from about +10% to +40%. Many of the mice (about 90%) residing in the contaminated field are experiencing a range of growth rates that are only slightly decreased from rates expected for unexposed animals. However, about 10% of the animals have weight gains that are markedly less than for unexposed animals, ranging from about +5% to -30% (i.e., a net weight loss of 30% compared to the starting weight).

It should be noted that the response distribution calculated in this way is what would be expected for a large population of exposed receptors. If the actual exposed population is small, then the actual response distribution may vary somewhat compared to the typical response shown in Exhibit 4-7. In cases where it is important to evaluate this variability about the expected average pattern of response, this may be done by running repeated Monte Carlo simulations using a number of trials (iterations) within each simulation that is equal to the expected size of the exposed population. Each simulation will thus represent a possible response distribution in the exposed population, and the range of responses across different populations may be evaluated by comparing the multiple simulations. As noted above, the magnitude of the variability between populations is expected to be small if the population size (number of trials) is large, although this depends on the characteristics of the exposure and response functions.

EXHIBIT 4-7

MODELING VARIABILITY IN A CONTINUOUS RESPONSE

Scenario

Exposure of a population of field mice to a chemical ingested via the food chain

Example Inputs

Exposure

Distribution of Average DI LN(104,127)

Response (see Exhibit 4-3)

Endpoint = Growth (% increase in 21 days)

Stressor-response model fit

Mean response(dose) = 23.7 - 51.4*dose^n / (228.7^5.29 + dose^5.29)

Stdev (dose) = 7.0 (constant)

PRA Simulation

Step 1 Draw a dose at random from the lognormal distribution of dose

Step 2 Calculate the mean response [m(d)] and standard deviation of the response (s(d) at that dose

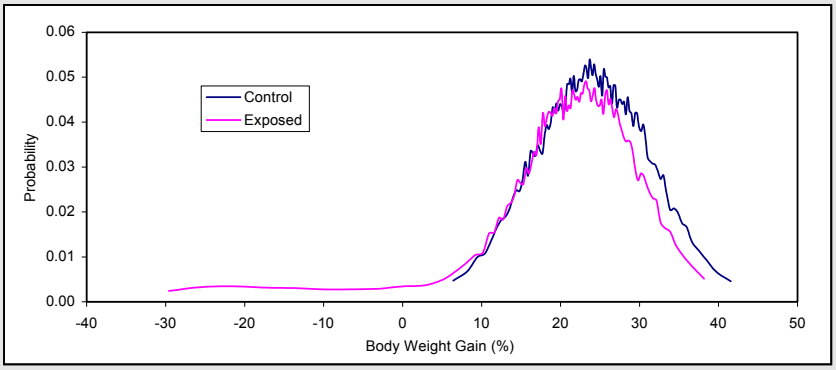
Step 3 Define the PDF for response at dose d: NORMAL(m(d), s(d))

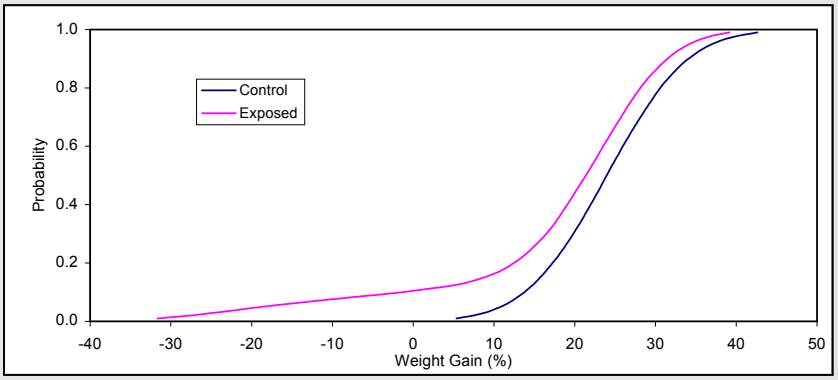
Step 4 Draw a response at random from this PDF

Track this as the forecast cell

Example Output

Percentile	Control	Exposed
0.05	10.9	-18.6
0.25	18.6	14.7
0.50	24.0	21.4
0.75	29.3	26.9
0.90	34.1	31.5
0.95	37.0	34.2
0.99	42.6	39.1





4.6.3 JOINT PROBABILITY CURVES

In this approach, if data are available to characterize the probability of a particular exposure occurring, and an exposure-response curve is available, these may be combined to yield a curve (referred to as a Joint Probability Curve) that shows the probability that a response greater than some specified magnitude will occur. An example is shown in Figure 4-8. The upper panel shows a hypothetical cumulative exposure probability distribution (plotted on the primary y-axis) along with the exposure-response curve (plotted on the secondary y-axis). The steps needed to generate the Joint Probability Curve are as follows:

Step 1: Select an exposure level "x" and record the probability (P_x) of exceeding that exposure. For example, in Figure 4-8, at an exposure of 12 units, the cumulative probability of exposure is 84%. Thus, the probability of exceeding that exposure is 16%.

Step 2: Find the expected response at that same exposure (R_x). In this case, the response at an exposure of 12 is 2.2.

Step 3: Plot a data point at R_x on the x-axis and P_x on the y-axis.

Step 4: Repeat this process for many different exposure levels, being sure to draw samples that adequately cover all parts of the probability scale.

The lower panel of Figure 4-8 shows the results obtained using the hypothetical data in the upper panel. The advantage of this format is that it gives a clear visual display of both the probability and magnitude (severity, extent) of response. Further, the area to the left of the curve is a relative index of the population-level or community-level risk, and comparison of this area across different scenarios is helpful in comparing different risk scenarios (both in risk characterization and risk management). However, this approach is based on the mean response at a dose, and does not account for variability in response between multiple individuals all exposed at that dose. Employing a two-dimensional Monte Carlo analysis (2-D MCA) procedure could help to display this variability in response between the individuals at a given dose.

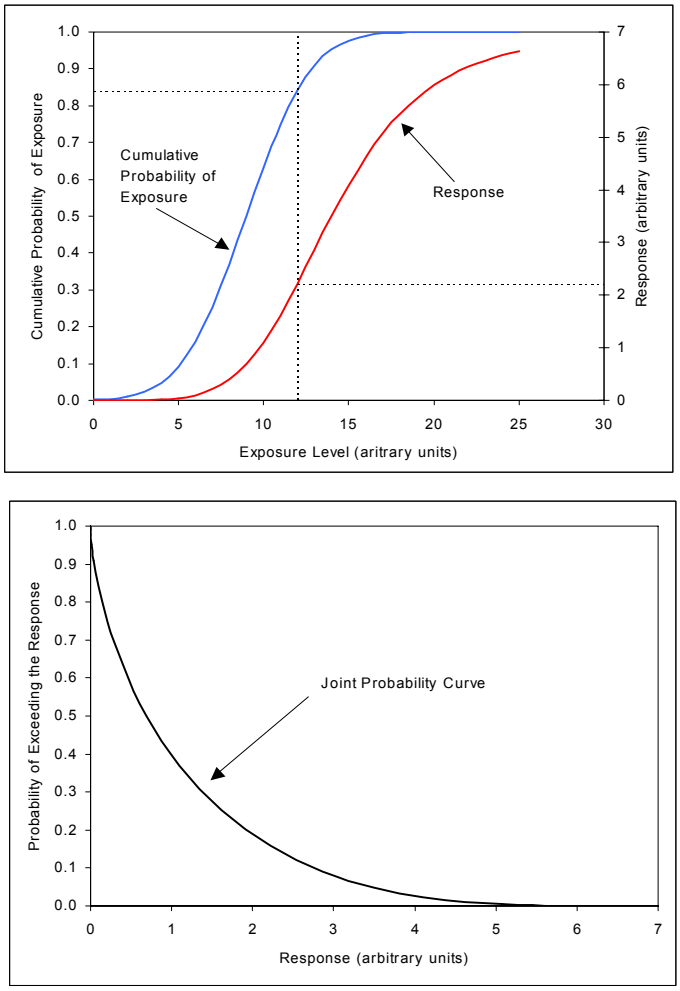


Figure 4-8. Example Joint Probability Curve.

Note that unless 2-D MCA is used, this approach does not require Monte Carlo modeling. Rather, the calculations can usually be performed in a spreadsheet format using built-in spreadsheet functions.

4.7 MODELING UNCERTAINTY IN ECOLOGICAL RISK ASSESSMENTS

As emphasized above, one of the greatest potential benefits of the PRA approach is the ability to combine estimates of uncertainty associated with different components of the exposure and risk models in order to describe the overall uncertainty in final exposure or risk estimates. Some basic options for characterizing and presenting uncertainty in exposure, toxicity, HQ, and response are presented below.

4.7.1 UNCERTAINTY IN EXPOSURE

Most estimates of dose-based exposure for terrestrial receptors (birds, mammals) are based on calculated estimates of chemical intake using simple or complex food web models, sometimes coupled with environmental fate and transport models that can link risk to a receptor with a source of contamination. In cases where receptors are exposed mainly by direct contact rather than ingestion (e.g., fish, soil invertebrates, etc.), concentration-based (as opposed to dose-based) descriptors of exposures may be derived using mathematical fate and transport models. The basic principles for modeling uncertainty in ecological exposure models (either dose-based or concentration-based) are the same as discussed in Appendix D. In brief, probability distribution functions of uncertainty (PDFu's) are used to characterize the uncertainty in the parameters of the probability distribution functions of variability (PDFv's) for some or all variables in the exposure model. Then, a 2-D MCA is used to derive quantitative estimates of the uncertainty around each percentile of the variability distribution of exposure. Figure 4-9 illustrates the type of tabular and graphic outputs that this approach generates.

Variability Percentile	Uncertainty Percentiles		
	5th	Mean	95th
0.05	0.4	1.1	2.0
0.10	0.7	1.6	2.8
0.15	0.9	2.1	3.5
0.20	1.2	2.6	4.2
0.25	1.5	3.1	5.0
0.30	1.8	3.7	5.9
0.35	2.1	4.3	6.7
0.40	2.6	5.0	7.6
0.45	3.0	5.8	8.7
0.50	3.6	6.6	9.9
0.55	4.2	7.7	11.3
0.60	5.0	8.8	12.9
0.65	5.9	10.3	14.8
0.70	7.2	12.1	17.2
0.75	8.8	14.4	20.3
0.80	10.9	17.5	24.1
0.85	14.5	22.0	30.1
0.90	20.1	29.6	39.4
0.95	32.9	46.5	60.0

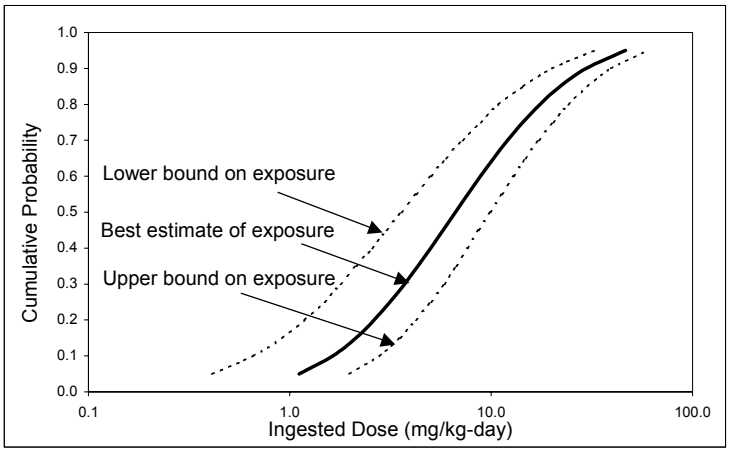


Figure 4-9. Example Presentation of Uncertainty in Exposure.

If exposure is based on measured rather than calculated values by PRA (e.g., measured concentrations in an abiotic medium,

measured concentrations in receptor tissues), uncertainty in the empirical or best-fit continuous distribution through the data can be quantified using the statistical methods detailed in Appendix B.

As discussed in Chapter 1, it is important to understand that there are many sources of uncertainty and that this approach to uncertainty analysis focuses mainly on parameter uncertainty and uncertainty in the true shape of input variable distributions. It does not capture other sources of uncertainty relating to the fundamental adequacy of the exposure and risk models used to describe the behavior of complex biological systems or of sampling and analytical errors and uncertainties, so the uncertainty estimates should always be interpreted in this light as being somewhat incomplete.

4.7.2 UNCERTAINTY IN TOXICITY

Toxicity information used for ERAs is often a source of uncertainty in the risk assessment process. This uncertainty may arise from multiple areas and may include both quantitative uncertainty in the dose-response data (involving toxicokinetics and study designs) and qualitative uncertainty in the relevance of the data (involving toxicodynamics). Methods for characterizing the quantitative uncertainty in both point estimates of toxicity (TRVs) and in full exposure-response curves are outlined below.

Uncertainty in TRVs

TRVs for a chemical are point estimates of exposure levels that do not cause an unacceptable effect in an exposed receptor population. Ideally, all TRVs would be based on NOAEL and LOAEL values derived from studies in which the receptor, endpoint, exposure route and duration were all matched to the assessment endpoints defined for the site. However, such exact matches are seldom available. Therefore, it is often necessary to extrapolate available toxicity data across route, duration, endpoint and/or species, leading to uncertainty in the most appropriate value to use as the NOAEL or LOAEL. There are no default methods for developing TRVs on a site. However, some options include the use of allometric dose scaling models, physiologically-based biokinetic models, benchmark dose estimates or other approaches based mainly on policy and/or professional judgment. Guidelines for dealing with the uncertainty in components of the TRV derivation by uses of PRA are provided below.

Uncertainty in NOAELs and LOAELs

Uncertainty in the NOAEL or LOAEL for a chemical has two components: (1) uncertainty within a study; and (2) uncertainty between studies, under exact specified conditions of exposure.

Assuming that a single study has been selected to provide the NOAEL and/or LOAEL values to be used in deriving a TRV for a chemical, it is customary to define the NOAEL as the highest exposure that did not cause a statistically significant effect, and the LOAEL is the lowest exposure that did cause a statistically significant effect. As noted earlier (see Section 4.5.1), this approach has a number of limitations, and there may be substantial uncertainty as to whether the observed NOAEL and LOAEL values actually bracket the true threshold effect level. One way to quantify uncertainty in the exposure levels that cause some specified level of adverse effect is through the use of exposure-response curve-fitting software such as EPA's BMDS package. In this approach, the risk assessor selects some level of effect that is judged to be below a level of concern, and another level of effect that would be of concern. The choice of these response levels is a matter of judgment, and depends on the nature and severity of the endpoint being evaluated. A specified level of effect is referred to as a Benchmark Response (BMR), and the exposure that causes that response is referred to as the Benchmark Dose (BMD). Given information on the number of test organisms in each test group and on the variability of the response in those

organisms, the BMD software uses maximum likelihood methods to derive the 5% lower confidence bound on the exposure that causes the BMR. This is referred to as the BMDL. This uncertainty bound may be used to quantify the uncertainty in the BMD, and hence to characterize this source of uncertainty in the TRV. The simplest method for approximating the uncertainty distribution around the BMD is to assume the distribution is approximately normal, with mean equal to the BMD and standard deviation (standard error) given by:

Stdev=(BMD - BMDL) / 1.645

For advanced analyses, a more accurate characterization of the uncertainty distribution around the BMD may be derived by Monte Carlo simulation. In this approach, each model parameter is assumed to be normally distributed, with mean and standard error values provided by the BMDS output. Monte Carlo simulation is then used to select alternative model parameter sets, being sure to account for the covariance between parameters (the covariance matrix is also provided by the BMDS output). For each parameter data set, the BMD is calculated, and the distribution of BMD values across many iterations is a better approximation of the uncertainty in the BMD.

Uncertainty in the effect level (NOAEL or LOAEL) for a chemical may also arise because there is more than one study available for the chemical, and the studies do not yield equal estimates of the effect level. It is important to note that the process of reviewing available toxicity studies, choosing the most relevant endpoint for use in deriving a TRV, and identifying the most relevant study is a process requiring basic toxicological expertise (not probability or statistics), and this process must be completed both for point estimate and probabilistic risk assessments. In general, studies based on different receptors, endpoints, exposure routes and/or durations are not equally relevant for evaluating a particular assessment endpoint in a particular indicator species. However, in some cases, multiple studies of the same endpoint in the same species will be available. In such a case, assuming that all the studies are judged to be equally reliable, the best estimate of the LC50 may be derived by calculating the geometric mean of the available alternative values (after adjustment to constant hardness). Uncertainty around the best estimate may then be based on the observed inter-study variability, using the basic principles for choosing PDFu's as described in Appendix B.

Uncertainty in Extrapolation of TRVs

In general, extrapolation of TRVs across species or endpoints is not desirable, since the magnitude and direction of any potential error is generally not known. Sometimes, extrapolations between species are attempted based on allometric scaling models that seek to adjust toxicity values accounting for differences in body weight. Alternatively, physiologically-based pharmacokinetic (PBPK) models that seek to account for differences in a number of other physiological variables (metabolism rate, organ size, blood flow, etc.) can be used. However, the validity of these models is often not well established. In those cases where these models are used, and where the uncertainty in the model is judged to warrant quantitative evaluation, the primary source of the model should be consulted in order to derive an estimate of the uncertainty in the quality of the extrapolation and in the parameters of the model. As noted earlier, PRA may capture uncertainty associated with model input parameters, but does not usually capture all sources of uncertainty in the model. In particular, most models of this sort are designed to extrapolate only the average response as a function of dose, and are not intended to extrapolate variability between individuals at a specified dose. When no mathematical model is available to support quantitative extrapolation across species, exposure duration or endpoint, professional judgment and/or policy may be used to select extrapolation factors to account for the uncertainty.

The risk assessor should ensure that the risk manager understands the uncertainty associated with any model selected and applied, and that the results of the calculations (point estimate or PRA) are conditional upon the model selected.

Uncertainty in Parameters of the Dose-Response Models

When toxicological exposure-response data are fit to mathematical equations, the fitting software will usually provide quantitative information on the uncertainty in the best estimates for each of the model parameters. For example, in the dichotomous model illustrated in Exhibit 4-3, the output from the BMDS software included the following information on the uncertainty in the parameters of the best-fit logistic equation:

Parameter	Best Est	Std Error (SE)
a	-4.80	0.83
b	0.101	0.019

Because the uncertainty in the best estimate of each model parameter is asymptotically normally, uncertainty in the parameters may be modeled as:

PDFu (parameter i)=NORMAL(best estimate of parameter i, SE of parameter i)

Note that the parameters of the model are generally not independent, and generally should not be treated as such. Thus, when modeling the uncertainty in the parameters of the best-fit exposure-response model, the PDFv's for the parameters should be correlated according to the correlation matrix or the variance-covariance matrix, as provided by the modeling software.

4.7.4 UNCERTAINTY IN RESPONSE

If the risk characterization phase of the risk assessment focuses on an estimation of the distribution of responses rather than the distribution of HQ values, the uncertainty in the distribution of responses can be evaluated using two-dimensional Monte Carlo techniques using PDFu's for the parameters of the exposure and exposure-response models derived as described above. The same graphical output may be used for this presentation as was illustrated in Figure 4-9, except that the x-axis is response rather than HQ. This format is illustrated in Figure 4-10 for a dichotomous endpoint (e.g., acute lethality). In this example, the average probability of response among the members of the exposed population (shown in the graph by the black diamond symbols) is 8.2%, with a confidence bound around the mean of 4.9 to 12.8%. This is equivalent to concluding that about 8.2% of the population is expected to suffer acute lethality, but the true fraction dying could range from 4.9 to 12.8%.

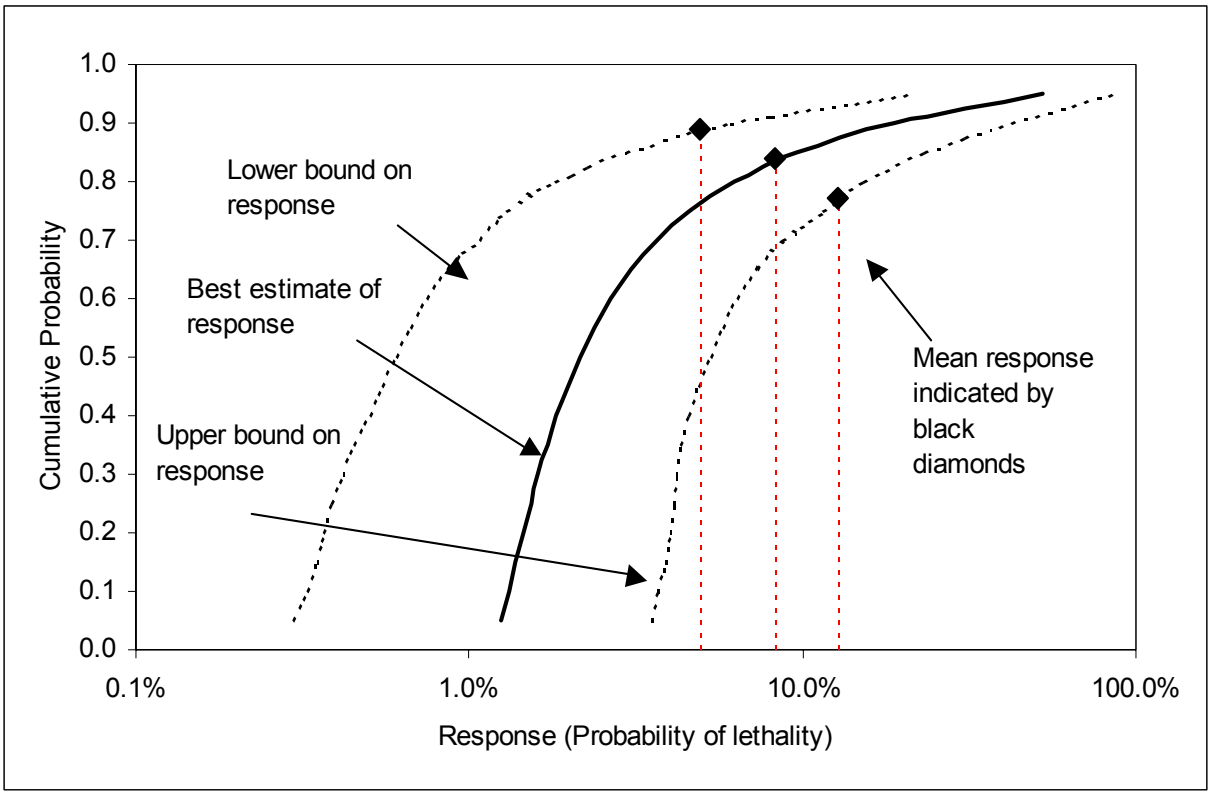
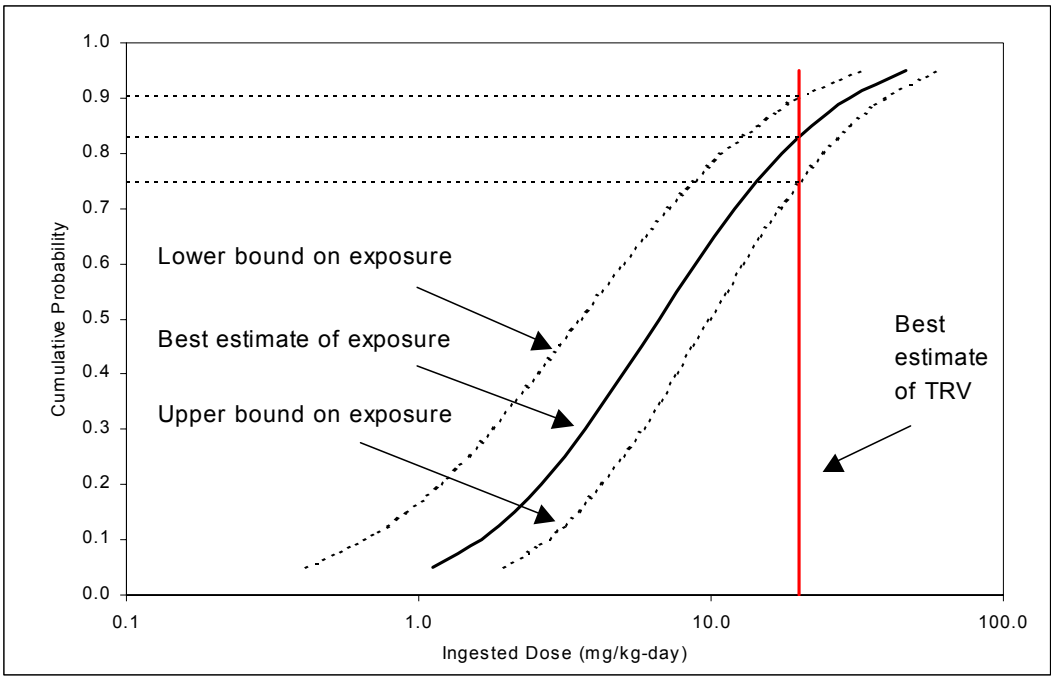


Figure 4-10. Example Presentation of Uncertainty in Response.

4.7.3 UNCERTAINTY IN HAZARD QUOTIENT

Once the uncertainty in exposure and/or toxicity distributions has been characterized as described above, there are a number of options for presenting the resultant uncertainty in the HQ (or HI, if appropriate and applicable for summing HQs) distributions. Figure 4-11 shows one simple graphical format, where the point estimate of the TRV is superimposed on the uncertainty bounds of the exposure distribution (upper panel), or the uncertainty bounds of the TRV are superimposed on the best estimate of exposure (lower panel). One could also superimpose the range of TRVs over the range of exposures, to capture most of the uncertainty in the HQ. Furthermore, such distributional outputs should always show the point estimate ranges of CTE and RME exposures in respect to the ranges of TRVs, for use in weight-of-evidence to help interpret the PRA and point estimate results. The advantage of this format is that no additional Monte Carlo modeling is needed to derive initial descriptors of uncertainty in risk. For example, in the upper panel it may be seen that the best estimate of the fraction of the population exposed at a level below the TRV is about 83%, but that this is uncertain due to uncertainty in the exposure estimates, and the true percent below the TRV might range from 74 to 90%. Similarly, in the bottom panel, the best estimate of the fraction of the population below the TRV is also about 83%, but due to uncertainty in the TRV the actual value could range from 64 to 91%. Uncertainty could also be presented by showing a combined graph with both ranges of exposure and TRVs, such as described below.



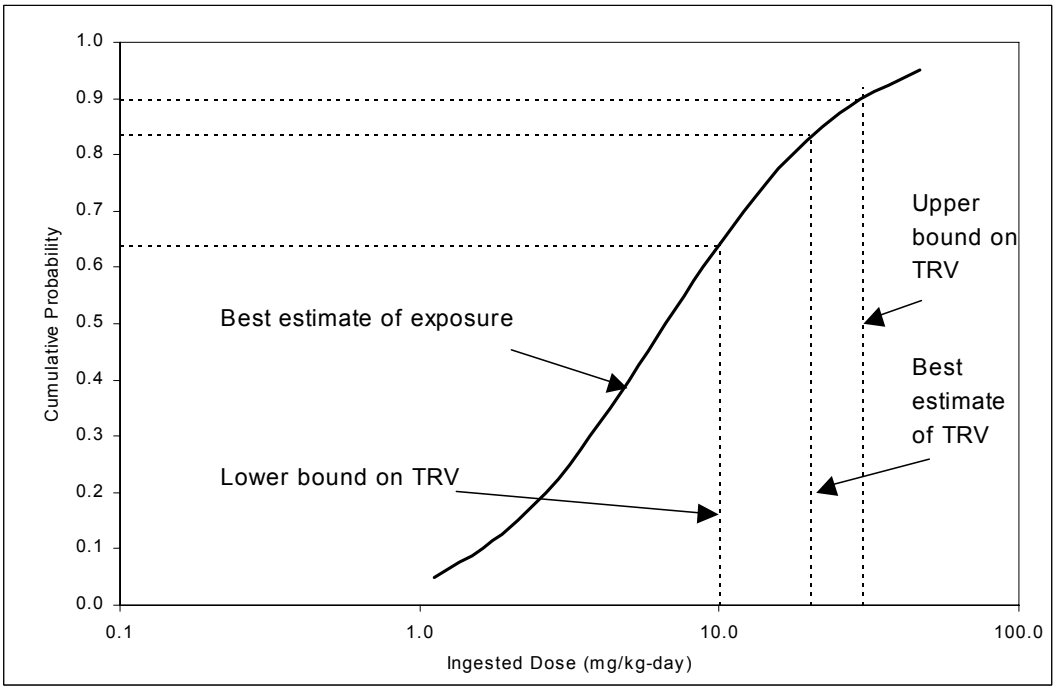


Figure 4-11. Example Presentation of Uncertainty in Exposure and TRV.

A more complete characterization of uncertainty in HQ may be achieved by using PRA to combine the uncertainty in both the exposure and the TRV terms, resulting in the uncertainty bounds on the HQ distribution itself (see Figure 4-12). In this example, it may be seen that 63% of the exposed population is estimated to have an HQ below 1.0, but that this is uncertain due to uncertainty in both the exposure distribution and the TRV, and that the true fraction of the population below a level of concern (HQ < 1) could range from 45 to 81%.

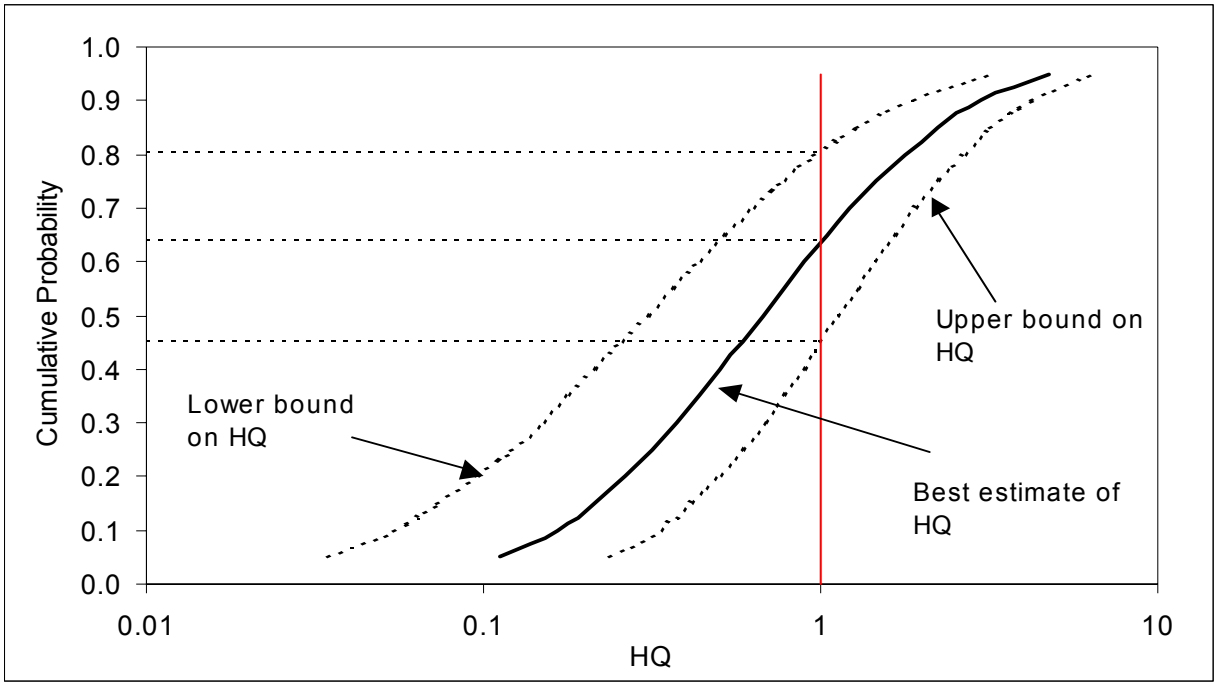


Figure 4-12. Example Presentation of Uncertainty in HQ Estimates.

4.8 INTERPRETING RESULTS OF AN ECOLOGICAL PRA

In some cases, the information contributed by a PRA may provide a more complete characterization of risks to a population of receptors than can be obtained by using point estimate methods. However, whether by PRA or by point estimate or a combination, the results of the risk assessment must be interpreted to reach a risk management decision.

In contrast to the case for human health risk assessments (where default risk-based decision rules are well established), there are no established default decision rules for identifying when risks to ecological receptors are and are not of concern. In the point estimate approach, EPA guidance (U.S. EPA 1992b, 1995) recommends consideration of both the RME and CTE exposure/dose estimates along with TRVs based on both LOAELs and NOAELs (U.S. EPA 1997a) to reach a risk management decision. The same principle applies to probabilistic ERAs.

In some cases, interpretation of an ecological PRA is relatively simple. For example, if the distribution of HQ values calculated using an appropriate NOAEL-based TRV are less than 1.0 for nearly all members of the population, then it is likely that risks are within an acceptable range for the population. Conversely, if the distribution of HQ values calculated using a LOAEL-based TRV are significantly greater than 1.0 for most members of an exposed population, then it is likely that risks are not acceptable

for the population. However, for cases which fall between these bounding conditions (and for cases where one needs to clearly define the boundaries of potential excess risks for a gradient of contamination and exposures), the level of risk or response that is considered acceptable must be defined by the risk assessor and the risk manager on a site-specific and receptor-specific basis. This evaluation should take the following factors into account:

(1) The Risk Management Goal

The risk management objective for most Superfund ERAs is defined as population sustainability (U.S. EPA, 1999). In this case, harm to some members of the exposed population may be acceptable, if that harm does not lead to an overall reduction in population viability. This situation (protection of a population rather than protection of individuals) is sometimes equated with use of the CTE (average) receptor as the basis for risk management decision making. That is, if the HQ for the CTE receptor is below a level of concern, it is sometimes assumed that population risks are acceptable.

However, the choice of the CTE receptor as the basis for risk management decision making may not be sufficiently protective in all cases. For the vast majority of wild populations, the proportion of the population that must be protected to ensure population stability will be unknown. At a small number of sites, a population biologist may be able to provide some information. Moreover, the percentile of the CTE receptor in the exposure or risk distribution may vary depending on the shape of the distribution. The proportion of the population experiencing exposure greater than that of the CTE receptor could range from less than 10% up to 50% or even higher. Also, the ecological significance of an adverse effect on some members of a population depends on the nature of the stressors and on the life history and population biology of the receptor species. Because of these complexities, use of the CTE as a decision threshold for nonthreatened or endangered species may be appropriate in a small number of cases, but risk assessors and risk managers should realize that the choice of the CTE receptor requires a species- and endpoint-specific justification and the CTE should not be used as the default basis for a risk management decision. Rather, for the majority of ERAs, the risk management decision should be based on the RME receptor or an upper percentile of the distribution of variability in risk/exposure.

(2) The Toxicological Basis of the TRV

The biological significance of a distribution of variability in HQ cannot be interpreted without a proper understanding of the nature of the TRV being used to evaluate the distribution. This includes the nature of the toxicological endpoint underlying the TRV, its relevance to the assessment endpoint, and the shape (steepness) of the dose-response curve. For example, an HQ of 2 based on an EC20 for reduction in reproductive success would likely be interpreted as more significant toxicologically than an HQ of 2 based on the EC20 for an increase in liver weight. Likewise, an HQ of 2 based on an LC_{low} for acute lethality would be more significant if the dose-response curve for lethality were steep than if it were shallow, since it would be easier to cause greater response with smaller increases in exposure to contaminants.

(3) The Characteristics of the Receptor

Ultimately, the question which must be assessed is whether an effect of degree "x" occurring in "y" percent of the population is biologically and ecologically significant. This, in turn depends on the attributes of the receptor being evaluated. For example, a reduction of 10% in the reproductive success of a fecund and common species (e.g., the field mouse) might not lead to a significant reduction in population number, while the same effect could be of concern in a species with lower fecundity and/or lower population density (e.g., the moose). Thus, the interpretation of an analysis of variability in exposure and/or effect often requires the input of a trained population biologist with expertise in the receptor of concern.

Because of these issues, there is no default rule for what level of effect is and is not acceptable for an exposed ecological population; except for the case of no potential excess risks where the RME exposures do not exceed the TRV based on a NOAEL, assuming there is reasonable confidence in those exposure and toxicity values. In some cases, mathematical models may be available for predicting the population-level consequences of a given pattern of effects (e.g., see ECOFRAM 1999a for some aquatic population models), but in general the extrapolation from a distribution of individual responses to an estimation of population-level effects is difficult. For this reason, close consultation between the risk manager and the ecological risk assessor is necessary for translating results of an ERA into an appropriate and successful risk management decision.

4.9 GUIDELINES FOR PLANNING AND PERFORMING A PROBABILISTIC ERA

4.9.1 PLANNING AN ECOLOGICAL PRA

Chapter 2 provides a general discussion of the key steps that should be followed when planning a PRA. These guidelines are equally applicable to ecological PRA as to human health PRA. Of the key steps in the process, most important are the following:

Dialogue Among Stakeholders

As discussed in Section 4.2, the decision if and when to perform an ecological PRA is an SMDP shared by risk assessors, risk managers, and stakeholders, including members of the public, representatives from state or county environmental agencies, tribal government representatives, natural resource trustees, private contractors, and potentially responsible parties (PRPs) and their representatives. A scoping meeting should be held after the completion of the baseline risk assessment in order to discuss the potential purpose and objectives of a PRA, and to identify the potential value of the analysis to the risk management process. If it is decided to perform at least an initial PRA evaluation, subsequent meetings of a similar type should occur iteratively in order to assess whether any further effort is warranted.

Preparation of a Workplan

Any PRA beyond the simplest screening level evaluation should always be accompanied by a workplan. The purpose of the workplan is to ensure that all parties agree on the purpose and scope of the effort, and on the specific methods, data, and procedures that will be used in the PRA. Workplans should be developed according to available guidance for workplans for nonprobabilistic ERA (U.S. EPA, 1992b,

1997a) and should consider three elements: (1) the 16 guiding principles of MCA (U.S. EPA, 1997b); (2) the eight guidelines for PRA report submission (U.S. EPA, 1997b); and (3) the tiered approach to ERA (U.S. EPA, 1997a). Development of a workplan for PRA is discussed in greater detail in Chapter 2, and Exhibit 4-8 summarizes the key elements of a proper workplan. The workplan must be submitted to the BTAG coordinator and/or regional ecotoxicologist for review and for approval by the risk manager. The EPA strongly recommends that PRPs who wish to perform PRAs of ecological risk involve the Agency in the development of a workplan in order to minimize chances of significant disagreement, as is required by EPA policy.

EXHIBIT 4-8

EXAMPLE ELEMENTS OF A WORKPLAN FOR ECOLOGICAL PRA

1. Introduction/Overview
   - Conceptual site model
   - Assessment endpoints
   - Indicator species
   - Measures of exposure and effect

2. Description of Exposure and Risk Models
   - Basic exposure models (fate and transport, uptake, food web, intake, etc.)
   - Basic risk models (HQ, dichotomous response, continuous response)

3. Results from a Point Estimate Assessment
   - CTE and RME risk estimates from baseline evaluation

4. Rationale why a PRA will be helpful
   - Goals of the assessment (variability, uncertainty, both)
   - Expected benefit to risk manager

5. Description of the Proposed PRA
   - Exposure scenarios to be evaluated
   - Output variables to be modeled in variability and/or uncertainty space

6. Proposed PDFs, and their basis
   - Method for performing sensitivity analysis and for selecting key variables
   - Data source for characterizing key variables
   - Approach for selecting and parameterizing key variables
   - Proposed list of PDFs for exposure variables (optional but desirable)
   - Method for dealing with the concentration term
   - Method for dealing with correlations

7. Proposed Software and Simulation Approach
   - Commercial or custom
   - Monte Carlo or Latin Hypercube
   - Number of Iterations
   - Method(s) for sensitivity analysis

8. Preliminary Results (optional, but helpful)
   - Results of a screening level evaluation
   - Identification of variables where more effort is needed to improve the distribution function

4.9.2 EVALUATING AN ECOLOGICAL PRA

When an ecological PRA is submitted to EPA for consideration, it will be reviewed in order to determine if it has been performed in accord with sound principles of ERA (U.S. EPA, 1997a, 1998), and with sound principles of PRA (U.S. EPA, 1997b). A general checklist that may be helpful to reviewers is provided in Appendix F, and key features of this checklist are summarized in Exhibit 4-9. Eight specific conditions for acceptance of a PRA submitted to EPA are provided in U.S. EPA (1997b).

At the discretion of EPA risk assessor or risk manager, the PRA report may be submitted for additional EPA internal review and/or an external review process in accord with Agency guidelines for conducting peer reviews (U.S. EPA, 2001). The external peer review may be used in cases where the issues are complex or contentious and the opinions of outside expert peer reviewers can improve the PRA.

EXHIBIT 4-9

CHECKLIST FOR INCLUDING A PRA AS PART OF THE ERA (SEE APPENDIX F)

- All risk assessments should include point estimates prepared according to current Superfund national and regional guidance.
- A workplan must be submitted for review and approval by the appropriate EPA regional project manager (RPM) and/or BTAG coordinator prior to submission of the PRA.
- A tiered approach should be used to determine the level of complexity appropriate for the ERA. The decision to ascend to a higher level of complexity should be made with the risk manager, regional risk assessor and other stakeholders.
- The eight conditions for acceptance presented in the EPA policy on PRA (U.S. EPA, 1997b) should be clearly addressed by each PRA submitted to the Agency.
- Information in the PRA should possess sufficient detail that a reviewer can recreate both the input distributions and all facets of the analysis. This includes copies of published papers, electronic versions of necessary data and other materials deemed appropriate by EPA.

4.10 EXAMPLE OF THE TIERED PROCESS IN ERA

As discussed in detail in Chapter 2, one of the key elements in the risk assessment process is deciding if and when further analysis is warranted. This includes decisions regarding whether to employ PRA calculations to supplement point estimate calculation, and if so, what level of effort to invest in those PRA calculations. The following section presents a relatively simple hypothetical example illustrating how the tiered approach might operate at a site where ecological risk is an important concern.

Problem Formulation

PestCorp is a former chemical manufacturing facility that produced mainly chlorinated pesticides 10 to 20 years ago. Data collected on the PestCorp property indicate that a number of spills or releases of chlorinated pesticides took place when the facility was in operation, and that site soils are broadly contaminated, especially with pesticide X. This contaminated soil has lead to impacts on a nearby lake of about 300 acres that receives surface water runoff from the PestCorp site. Samples from the lake reveal low but detectable levels of pesticide X in water, with relatively high values in sediment and in the tissues of a variety of aquatic organisms (crayfish, snails, benthic macroinvertebrates and fish). The concentration values in all media (water, sediment, aquatic organisms) tend to be highest in the part of the lake receiving runoff from the PestCorp property, with a gradient of diminishing values at locations further away from the area where runoff enters the lake.

A BTAG committee formed by EPA to identify potential ecological concerns at the site recognized that many different species could be exposed to the contaminants in the lake, including aquatic receptors residing in the lake (fish, invertebrates, aquatic plants), as well as mammals and birds that frequent the lake for food or water. Because pesticide X is lipophilic and tends to biomagnify in the food web, the BTAG decided that the highest risks would likely occur in higher-level predators such as mammalian omnivores, and selected the racoon as a good indicator species to represent this trophic group. Pathways of exposure that were identified as warranting quantitative evaluation included (a) ingestion of water, (b) ingestion of aquatic food items, and (c) incidental ingestion of sediment while feeding or drinking at the lake. The BTAG determined that the assessment endpoint was protection of mammalian omnivore populations.

Point Estimate Risk Evaluation

A series of iterative screening-level point estimate calculations (Steps 1 to 2 of the 8-step ERAGS process) were performed to investigate whether or not there was a basis for concern at the site. Initial calculations using simplified and conservative inputs (i.e., exposure based on the maximum measured concentration in each medium, an area use factor of 1, and the most conservative available TRVs) indicated that the HQ value for pesticide X could be quite large. Therefore, a refined screening level evaluation was performed in which point estimates of CTE and RME risk were derived using the best information currently available. Key elements of the approach are summarized below:

- The CTE receptor was assumed to be exposed at a location where concentration values were the average for the whole lake, and the RME receptor was assumed to be exposed at a location where concentrations were equal to the 95^{th} percentile of values from the lake.

- Because only limited data were available for measured concentrations of pesticide X in aquatic prey items, the concentration values in aquatic prey were estimated using a linear bioaccumulation model: C(prey)=C(sed) x BAF. The BAF was estimated from the existing data by finding the best fit correlation between the concentration values in sediment and crayfish at 7 locations in the lake: C(crayfish)=5.04 x C(sed) (R^2=0.792).

- The TRV values were based on a study in mink in which the toxicity endpoint was the percent inhibition of reproductive success.

These inputs and the resulting HQ values are shown in Exhibit 4-10. As seen, estimated risks to the CTE receptor approach or slightly exceed a level of concern (HQ=4.7E-01 to 1.4E+00), and risks to an RME receptor are well above a level of concern (9.1E+00 to 2.7E+01). The chief pathway contributing to the dose and risk is ingestion of contaminant in aquatic food web items (crayfish, fish, amphibians, etc.).

EXHIBIT 4-10

REFINED SCREENING POINT ESTIMATE INPUTS AND RESULTS

Basic model

HQ = DI(total) / TRV
DI(total) = DI(water) + DI(food) + DI(sed)
DI(i) = C(i) * IR(i)*AUF(i)

Other Assumptions

C(diet) = C(sed) * BAF
IR(sed) = IR(diet) * F(sed)
IR(diet) = IR(total) * F(diet)

Category	Variable	Variable	Units	Point Est. Values	
				CTE	RME
Inputs	Concentration	Concentration in water	mg/L	0.12	0.38
		Concentration in sediment	mg/kg	24	77
		BAF (sediment to aquatic prey)	--	5	5
		Concentration in aquatic prey	mg/kg	120	385
	Intake Rates	Total water intake rate	L/kg-day	0.082	0.12
		Total food intake rate	kg/kg-day	0.06	0.09
		Fraction of diet that is sed	--	0.03	0.06
		Fraction of diet that is aquatic prey	--	0.15	0.25
	Area Use Factors	Fraction of total water ingested at the lake	--	0.3	0.6
		Fraction of total diet from the lake	--	0.25	0.6
	TRVs	LOAEL-based TRV	mg/kg-day	0.6	0.6
		NOAEL-based TRV	mg/kg-day	0.2	0.2
Results	Daily Intake	Water ingestion	mg/kg-day	3.0E-03	2.7E-02
		Sediment ingestion	mg/kg-day	1.1E-02	2.5E-01
		Aquatic prey ingestion	mg/kg-day	2.7E-01	5.2E+00
		Total	mg/kg-day	2.8E-01	5.5E+00
	HQ (LOAEL-Based)	Water ingestion		4.9E-03	4.6E-02
		Sediment ingestion		1.8E-02	4.2E-01
		Aquatic prey ingestion		4.5E-01	8.7E+00
		Total		**4.7E-01**	**9.1E+00**
	HQ (NOAEL-Based)	Water ingestion		1.5E-02	1.4E-01
		Sediment ingestion		5.4E-02	1.2E+00
		Aquatic prey ingestion		1.4E+00	2.6E+01
		Total		**1.4E+00**	**2.7E+01**

SMDP 1 at Step 2 of ERAGS

The BTAG considered these results to indicate that inhibition of reproduction was possible in at least some members of the exposed population, but that the fraction of the population that was affected and the degree of impact on the population was difficult to judge from the point estimate calculations. Based on this, a decision was made to conduct a screening level PRA in order to provide some additional information on the magnitude and probability of risk.

Workplan 1

The contractor performing the risk assessment developed a brief workplan that proposed an approach for a screening level PRA. The plan called for a Monte Carlo-based evaluation of variability in exposure and risk among different members of the exposed mammalian omnivore (racoon) population. In brief, all exposure inputs that were treated as constants in the point estimate approach (i.e., were the same for CTE and RME exposure) were also treated as constants in the PRA evaluation. Because water contributed so little to dose or HQ, this pathway was not evaluated in the PRA, but was accounted for by adding in the point estimate values to the PRA results. All variables that are fractions (i.e, may only assume values between zero and one) were modeled as beta distributions, and all other variables were modeled as lognormal. For screening purposes, the parameters for all distributions were selected so that the mean and 95th percentile values of the PDF's matched the corresponding CTE and RME point estimates. The BTAG reviewed this proposed approach and authorized PRA work to begin.

Screening Level PRA Results

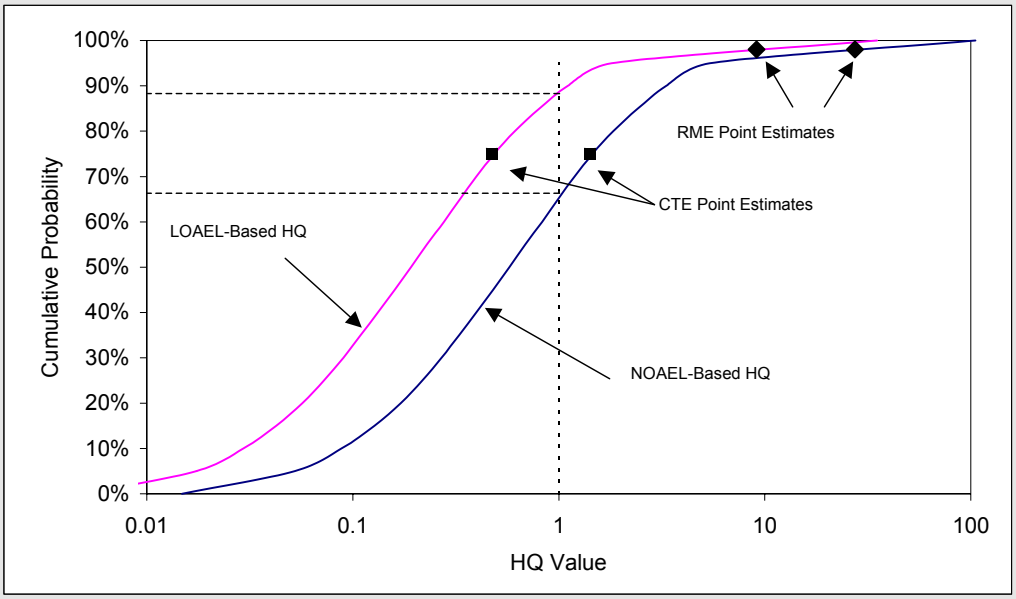
The screening level PRA inputs and the resulting estimates of the variability in HQ are shown in Exhibit 4-11. The CTE and RME point estimates are also shown for comparison. As seen, the PRA distribution of HQ values indicates that about 68% of the individuals in the population are likely to have HQ values below 1E+00, while 32% have HQ values above 1E+00.

Comparison of the CTE point estimates of HQ to the mean HQ values derived by PRA reveals the values are very close. This is expected because both depend on the mean values of the input variables, and the same mean values were used in both sets of calculations. With regard to upper-bound estimates, the RME point estimate values are at the 98th percentile of the PRA HQ distribution, within the target range (90th to 99th) usually considered appropriate. Note, however, that the 98th percentile is about 5-fold higher than the 95th percentile, emphasizing the high sensitivity of the RME HQ values to the precise percentile of the RME.

EXHIBIT 4-11

SCREENING LEVEL PRA CALCULATIONS OF HQ DISTRIBUTION

Data Category	Variable	Units	Screening Level Distribution		
			Type	param 1	param 2
Concentrations	Concentration in water	mg/L	Not evaluated in PRA		
	Concentration in sediment	mg/kg	LN	24	33
	BAF	--	Const	5	
	Concentration in aquatic prey	mg/kg	Calculated		
Intake Rates	Total water intake rate	L/kg-day	Not evaluated in PRA		
	Total food intake rate	kg/kg-day	LN	0.060	0.060
	Fraction of diet that is sed	--	Beta	3.42	110.7
	Fraction of diet that is aquatic prey	--	Beta	6.10	34.6
Area Use Factors	Fraction of total water ingested from lake	--	Not evaluated in PRA		
	Fraction of total diet from the lake	--	Beta	1.20	3.59
TRVs	LOAEL-based TRV	mg/kg-day	Const	0.6	
	NOAEL-based TRV	mg/kg-day	Const	0.2	



TRV Basis	Central Tendency			Upper Bound		
	Mean of PRA	Point Est CTE	Ratio	95th of PRA	Point Est. RME	Ratio
NOAEL	1.44	1.42	0.99	5.4	27.4	5.06
LOAEL	0.48	0.47	0.99	1.80	9.12	5.06

SMDP 2

The BTAG considered these results, and decided that it was very probable that pesticide X was causing an effect in some members of the exposed population, but decided that a final risk management decision would be facilitated by characterizing the distribution of responses (rather than the distribution of HQ values). The BTAG asked the contractor performing the work to develop a proposed approach for characterizing the distribution of responses.

Workplan 2

The contractor obtained a copy of the toxicity report upon which the TRVs were based, and determined that the study did include sufficient dose-response data to support reliable dose-response modeling. The contractor recommended that this be done using EPA's BMDS. The BTAG approved this proposed approach and authorized work to proceed.

PRA Refinement 1

The contractor fit the raw dose-response data (inhibition of reproduction in mink) to a number of alternative models available in BMDS, and found that the dose-response curve could be well characterized by the Hill Equation with nonconstant variance, as follows:

Mean Response at dose d (% decrease in reproduction)=$(100 \times d^{2.5})/(0.9^{2.5} + d^{2.5})$
Std. Dev. in Response at dose d (%)=SQRT$[1.6 \cdot (\text{mean response at dose d})^{1.3}]$

Based on this model, the point estimate LOAEL value (0.6 mg/kg-day) corresponds to an effect level of about 27%, and the NOAEL of 0.2 mg/kg-day corresponds to an effect level of about 2%.

Using this exposure-response model in place of the point-estimate TRV values, the refined PRA predicted a distribution of responses in the exposed population as shown in Exhibit 4-12. As seen, approximately 81% of the population was predicted to experience an effect on reproduction smaller than 10%, while 9% were expected to have a reduction of 10 to 30%, 4% a reduction of 30 to 50%, and 6% a reduction of more than 50%. On average across all members of the exposed population, the predicted reduction in reproductive success was about 9%.

EXHIBIT 4-12

SIMULATED DISTRIBUTION OF RESPONSES

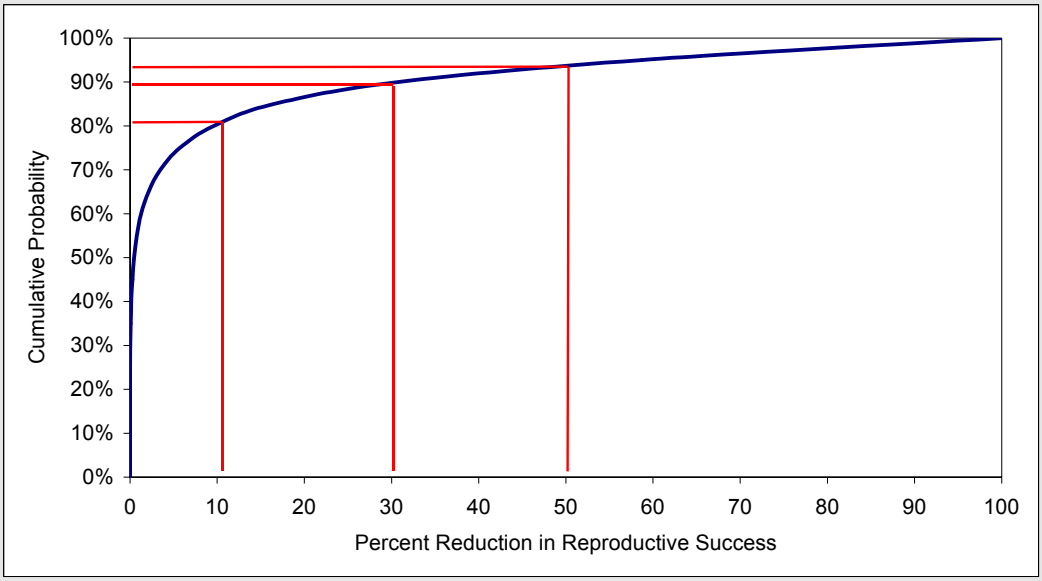
Exposure-Response Model

Resp = Normal(Mean,Stdev)

Mean = a + b*x^n / (x^n + k^n)

Stdev = alpha*mean^rho

x	Total daily intake
a	0
b	100
k	0.9
n	2.5
alpha	1.6
rho	1.3



Percent Reduction	Percent of Population
0-10%	81%
10-30%	9%
30-50%	4%
>50%	6%

SMDP3

The BTAG debated the likely population-level consequences of this predicted distribution of responses in members of the exposed population. After consulting with a field biologist with experience in the population dynamics of mammals such as racoons, the BTAG decided that the distribution of responses in the exposed population would cause a continued stress on the mammalian omnivore community and that reductions in population number were likely over time. Based on this, the risk manager and the BTAG agreed that remedial action was desirable and that a range of alternative clean-up strategies should be investigated. This was performed using the methods described in Chapter 5 (see Exhibit 5-5).

REFERENCES FOR CHAPTER 4

ECOFRAM. 1999a. ECOFRAM Aquatic Report (Draft). Ecological Committee on FIFRA Risk Assessment Methods. Draft report available online at http://www.epa.gov/oppefed1/ecorisk. Report dated May 4.

ECOFRAM. 1999b. ECOFRAM Terrestrial Draft Report. Ecological Committee on FIFRA Risk Assessment Methods. Draft report available online at http://www.epa.gov/oppefed1/ecorisk. Report dated May 10.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund, Volume II: Environmental Evaluation Manual.* Interim Final. Office of Emergency and Remedial Response. Washington, D.C. EPA/540/1-89/001. March.

U.S. EPA. 1991-present. *Eco Update*. Intermittent Bulletin Series. Office of Emergency and Remedial Response. 1991 to present.

U.S. EPA. 1992a. *Framework for Ecological Risk Assessment.* EPA Risk Assessment Forum. EPA/630/R-92/001. February.

U.S. EPA. 1992b. Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors* from F. Henry Habicht, Deputy Administrator, February 26.

U.S. EPA. 1994. Memorandum*: Role of the Ecological Risk Assessment in the Baseline Risk Assessment*. Elliott Laws, Assistant Administrator, Office of Solid Waste and Emergency Response. OSWER Directive No. 9285.7-17. August 12.

U.S. EPA. 1995. *EPA Risk Characterization Program*. Memorandum from the Administrator. March 21.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments.* Interim Final. Solid Waste and Emergency Response. OSWER Directive No. 9285.7-25. June 5.

U.S. EPA. 1997b. *Guiding Principles for Monte Carlo Analyses*. Risk Assessment Forum. EPA/630/R-97-001.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment.* Risk Assessment Forum. U.S. Environmental Protection Agency, Washington DC. EPA/630/R-95/002F. April. Published May 14. *Federal Register* 63(93):26846-26924.

U.S. EPA. 1999. Memorandum: *Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. P. Stephen D. Luftig for Larry D. Reed, Office of Emergency and Remedial Response. OSWER Directive No. 9285.7-28. October 7.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund*: *Volume I. Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments).* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-47. December.

CHAPTER 5

PROBABILISTIC RISK ASSESSMENT AND PRELIMINARY REMEDIATION GOALS

5.0 INTRODUCTION

According to the National Contingency Plan (NCP) (U.S. EPA, 1990a, 40CFR §300.430(d)(4)), risk assessment and risk management decision making go hand-in-hand: data from the remedial investigation are used to characterize risk, and results of the baseline risk assessment help to establish acceptable exposure levels for use in developing remedial alternatives. In practice, risk managers may identify two major objectives of risk assessment: (1) to determine if remediation is necessary (i.e., *Is there unacceptable risk at the site?*); and (2) if remediation is necessary, to determine a preliminary remediation goal (PRG) (i.e., *What chemical concentrations would result in a risk estimate that will be adequately protective of human health and the environment?*). The answer to the first question *(is there unacceptable risk?*) depends upon a number of factors, including the measured or estimated concentration levels of contaminants in site media, and takes uncertainty in the measurements into account. In contrast, the answer to the second question (*what is the PRG needed to achieve a specified level of protection?*) does not necessarily depend on any knowledge of the actual level or pattern of site-specific concentration data, and does not necessarily depend on the uncertainty in site concentration data. Thus, while exposure point concentrations (EPCs) and PRGs are closely related to each other, they have important differences (see Section 5.1 for further elaboration on EPCs and PRGs).

Once a risk manager has selected a PRG at a site, determining whether a particular area meets or will meet the PRG requires careful comparison of site data with the PRG, including a consideration of the uncertainty in the site data. For a further discussion on variability and uncertainty in the concentration term, readers are urged to consult Appendix C in this guidance.

EXHIBIT 5-1

SUMMARIES OF SOME KEY TERMS

Preliminary Remediation Goal (PRG) - initially developed chemical concentration for an environmental medium that is expected to be protective of human health and ecosystems. PRGs may be developed based on applicable or relevant and appropriate requirements, or exposure scenarios evaluated prior to or as a result of the baseline risk assessment. (U.S. EPA, 1991a).

Generic PRG - a chemical concentration protective of human health developed prior to the baseline risk assessment that uses default exposure assumptions representing common exposure scenarios, e.g., Region 3 risk-based concentrations (RBCs) or Region 9 PRGs.

Site-specific PRG - site-specific chemical concentration, protective of human health and ecosystems, based on exposure scenarios in the baseline risk assessment. Generally calculated for the various exposure scenarios considered in the baseline risk assessment.

Remediation Goals (RG) - site-specific chemical concentration, protective of human health and ecosystems, chosen by the risk manager as appropriate for a likely land use scenario.

Remediation Action Level (RAL) - the "not-to-exceed" level; a concentration such that remediation of all concentrations above this level in an exposure unit lowers the EPC sufficiently to achieve a target risk level. The RAL will depend on the mean, variance, and sample size of the concentrations within an exposure unit as well as considerations of short-term effects of the chemicals of concern.

Cleanup Level (Final Remediation Level) - chemical concentration chosen by the risk manager after considering both RGs and the nine remedy selection criteria of the NCP (U.S. EPA, 1990a). Also referred to as Final Remediation Levels (U.S. EPA, 1991a), chemical-specific cleanup levels are documented in the Record of Decision (ROD). A cleanup level may differ from a PRG because risk managers may consider details of the site-specific exposure, various uncertainties in the risk estimate, and implementation issues (e.g., the technical feasibility of achieving the PRG).

EXHIBIT 5-2

DEFINITIONS FOR CHAPTER 5

95% UCL for mean - The one-sided 95% upper confidence limit for a population mean; if a sample of size (*n*) was repeatedly drawn from the population, the 95% UCL will equal or exceed the true population mean 95% of the time. It is a measure of uncertainty in the mean, not to be confused with the 95th percentile (see below), which is a measure of variability. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged.

95th Percentile -The number in a distribution that is greater than 95% of the other values of the distribution, and less than 5%of the values. When estimated from a sample, this quantity may be equal to an observed value, or interpolated from among two values.

Applicable or Relevant and Appropriate Requirements (ARARs) - Federal or state environmental standards; the NCP states that ARARs should be considered in determining remediation goals. ARARs may be selected as site-specific cleanup levels.

Backcalculation - A method of calculating a PRG that involves algebraic rearrangement of the risk equation to solve for concentration as a function of risk, exposure, and toxicity.

Bootstrap Methods - Parametric and non-parametric methods for estimating confidence intervals for a statistic by resampling directly from the data set with replacement.

Coverage - Confidence intervals are expected to enclose a true but unknown parameter according to a specified probability, such as 90% or 95%. This is the expected coverage of the confidence interval, given a specified significance level (alpha). The difference between the expected coverage and the actual coverage is one metric for evaluating statistical methods that yield different confidence intervals.

Exposure Point Concentration (EPC) - The average chemical concentration to which receptors are exposed within an exposure unit. Estimates of the EPC represent the concentration term used in exposure assessment.

Exposure Unit (EU) - For Superfund risk assessment, the geographical area about which a receptor moves and contacts a contaminated medium during the period of the exposure duration.

Forward Calculation - A method of calculating a risk estimate that involves the standard arrangement of the risk equation to solve for risk as a function of concentration, exposure, and toxicity.

Iterative Reduction (IR) - A method of calculating a PRG that involves successively lowering the concentration term until the calculated risk is acceptable. This method can be applied to any medium.

Iterative Truncation (IT) - A method of calculating a PRG that involves developing an expression for the concentration term in which higher values of concentration are removed or "truncated" to reduce the maximum concentration, and re-calculating risks associated with the reduced concentration. The method may be repeated with consecutively lower truncation limits until risk is acceptable.

Land Method - The conventional method for calculating uncertainty in the mean concentration (e.g., 95% UCL) when the sample data are obtained from a lognormal distribution (U.S. EPA, 1992).

Maximum Detected Concentration (MDC) - The maximum concentration detected in a sample.

True Mean Concentration - The actual average concentration in an exposure unit. Even with extensive sampling, the true mean cannot be known. Only an estimate of the true mean is possible. A greater number of representative samples increases confidence that the estimate of the mean more closely represents the true mean.

Two Office of Solid Waste and Emergency Response (OSWER) guidance documents in preparation: (1) *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites* (U.S. EPA, 2001a), and (2) *Draft Guidance on Surface Soil Cleanup at Superfund Sites: Applying Cleanup Levels* (U.S. EPA, 2001b), also address topics related to the calculation of EPCs and comparison of those EPCs to a PRG.

In practice, calculations of risks, given concentration data, are commonly referred to as "forward calculations", while calculations of PRGs, based on chosen target risk levels, are referred to as "back-calculations". This terminology reflects the algebraic rearrangement of the standard risk equation needed to solve for the concentration term when point estimates are used to characterize exposure and toxicity input variables. For probabilistic risk assessment (PRA), the process for developing a PRG can be more complex. This chapter presents methods and recommendations for developing site-specific PRGs within the framework of PRA.

Are there different types of PRGs?

Generic PRGs have been developed for some chemicals and exposure media using point estimates based on standard default exposure assumptions (e.g., U.S. EPA, 1991b) and toxicity criteria available in the Integrated Risk Information System (IRIS) or Health Effects Assessment Summary Table(s) (HEAST) or from Environmental Protection Agency's (EPA's) National Center for Environmental Assessment. Soil Screening Guidance levels, Region 9's PRG table and Region 3's Risk Based Concentrations (RBCs) table are examples of generic point estimate PRGs. Generic PRGs are often used for screening chemicals of potential concern in Data Evaluation and Hazard Identification steps of the risk assessment process.

- ☞ *There is a clear distinction between generic PRGs, site-specific PRGs, remediation goals (RGs), and cleanup levels. The focus of this chapter is on site-specific PRGs.*

At this time, EPA does not recommend the use of PRA to develop generic PRGs. Until the science and policy decisions associated with the use of default assumptions in PRA have evolved, generic PRGs should only be developed from point estimate methods, as was done in the examples listed above.

As indicated in Exhibit 5-1, site-specific PRGs generally are developed after the baseline risk assessment. However, during the feasibility study or even later in the Superfund process, the methods described in this chapter may be used to modify cleanup levels at the discretion of the risk manager. However, it is generally not appropriate to use PRA for modifying cleanup levels during the feasibility study if PRA was not used in the baseline risk assessment.

- ☞ *Risk-based PRGs are initial guidelines and do not represent final cleanup levels.*

Only after appropriate analysis in the remedial investigation/feasibility study (RI/FS), consideration of public comments, and issuance of the record of decision (ROD) does a RG become a final cleanup level. A cleanup level may differ from a RG because risk managers may consider various uncertainties in the risk estimate. While the two main criteria for determining a cleanup level are: (1) protection of human health and the environment, and (2) compliance with applicable or relevant and appropriate requirements (ARARs), a cleanup level may differ from the RG because of modifying criteria, such as feasibility, permanence, state and community acceptance, and cost effectiveness. These and other factors are reflected in the nine evaluation criteria outlined in the NCP (U.S. EPA, 1990a; 40CFR §300.430(e)(9)(iii)) (see Chapter 1, Exhibit 1-2).

This chapter and Appendix C provide a comprehensive description of the issues associated with developing site-specific PRGs with both point estimate and probabilistic approaches, including the use of geostatistics. Because methods for calculating a 95% upper confidence limit for the mean (95% UCL) are discussed fully in the *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites* (U.S. EPA, 2001a) and *Draft Guidance on Surface Soil Cleanup at Superfund Sites: Applying Cleanup Levels* (U.S. EPA, 2001b), they are covered only briefly in this guidance. In general, this chapter, Appendix C, and the Superfund guidance under development should be consulted by risk assessors when developing site-specific PRGs.

5.1 GENERAL CONCEPTS REGARDING EPCS AND PRGS

PRGs developed from point estimate risk assessments and PRAs will be discussed in this section to compare and contrast the two approaches. The PRG is a special case of the concentration term (or EPC) in the risk equation. The intent of the EPC is to represent the average chemical concentration in an environmental medium in an exposure unit (EU) (i.e., the area throughout which a receptor moves for the duration of exposure). The EPC should be determined for individual EUs within a site. Because an EPC is calculated from a sample, there is uncertainty that the sample mean equals the true mean concentration within the EU; therefore, to account for associated uncertainty, the 95% upper confidence limit for the mean (95% UCL) is generally used for Superfund risk assessments (U.S. EPA, 1992). For both point estimate and probabilistic approaches, the PRG is an assumed value of the EPC that yields a risk estimate that is at or below an acceptable risk level.

> ☞ *The EPC usually represents the average concentration within the EU estimated from a sample; the PRG usually represents the average concentration within the EU that corresponds to an acceptable level of risk.*

The PRG may be thought of as a goal for the post-remediation EPC (see Section 5.1.2). Specifically, after remediation is completed, the average concentration (or the 95% UCL used as a measure of uncertainty in the average) for the EU should be sufficiently low to be protective of human health and the ecosystem. While the methods used to calculate the pre- and post-remediation EPC may differ, the interpretation of the EPC remains constant. For example, if the 95% UCL is used to represent the EPC before remediation, then the EPC following remediation (e.g., the PRG) should also represent a 95% UCL (Bowers et al., 1996).

Risk assessors may consider both variability and uncertainty in the development of an EPC. The calculation of a 95% UCL generally requires knowledge of not only chemical concentration measurements within the EU but also the receptor's behavior. Relevant information may include the variability in concentrations in the given sample, the sampling locations, and variability in the movement and activity patterns of receptors within the EU. A discussion of spatial and temporal variability associated with characterizing contamination in different exposure media is presented in Appendix C, and important sources of uncertainty in the EPC are discussed in Section 5.1.1.

For all risk assessments, chemical concentration measurements should be collected in a manner that is consistent with an understanding of both the source of contamination and the definition of the exposure unit. An investment of time and resources should be made in planning, scoping, and problem formulation. Part of this investment is to follow the Data Quality Objectives (DQO) process to obtain samples appropriate for the risk assessment and sufficient to support the remedial decision (U.S. EPA, 1993, 1994, 2000). Using new methods of sample collection and analysis such as dynamic workplans and real-time analysis may enable risk managers to get the most "bang for the buck" from the resources available for site characterization. Information about these methods and the DQO process is available from EPA's Office of Emergency and Remedial Response (U.S. EPA,

2001c) and Technology Innovation Office (U.S. EPA, 2001d, 2001e). The world wide web address is http://clu-in.org/char1_edu.cfm#syst_plan.

5.1.1 SOURCES OF UNCERTAINTY IN THE EPC

The 95% UCL is generally used as the EPC to represent uncertainty in the mean concentration in both the central tendency exposure (CTE) and reasonable maximum exposure (RME) risk estimates for Superfund (U.S. EPA, 1992). Similarly, in PRA, a probability distribution for uncertainty may be used in a two-dimensional Monte Carlo analysis (2-D MCA) simulation (see Appendix D) to represent a source of uncertainty in the EPC. There are numerous potential sources of uncertainty in the estimate of the true mean concentration within the EU. The sources of uncertainty when the EPC is expressed as either a single number or a distribution are the same and can be grouped into the following four broad categories:

(1) ***Uncertainty in the sample data.*** A limited number of measurements in the sample are used to make inferences about the EPC and the spatial distribution of concentrations at a site. Uncertainties may arise from many factors, including both sampling variability and measurement error. As the number of samples increases, the uncertainty generally decreases (e.g., more information will be available to characterize the spatial distribution and variation in concentration). In point estimate risk assessments, the 95% UCL is generally used as the EPC to account for the uncertainty in estimating the average concentration within an EU.

(2) ***Uncertainty about the location of the EU.*** When the size of a receptor's EU is less than the size of the site, the placement of the EU may be a source of uncertainty, especially when the contamination is distributed unevenly across the site and the PRA includes exposure scenarios for future land uses.

(3) ***Uncertainty in the behavior of the receptor.*** Even in the case of extremely well characterized sites, it remains uncertain whether the receptor will contact the environmental medium in a temporal and/or spatial distribution that can be adequately represented by the environmental samples collected.

(4) ***Uncertainty in chemical concentrations over time.*** The concentration in a given medium may undergo temporal changes, which may introduce uncertainty in estimates of a long-term average. Examples include the movement or attenuation of a solvent plume in groundwater; aerobic or anaerobic degradation; the change in the average concentration in a fish population due to changes in population dynamics; and the mixing of surface and subsurface soil over time.

A lack of knowledge in all four categories may be considered when selecting approaches to quantify uncertainty in the concentration term. One of the first steps in quantifying uncertainty is to define the EU, or the geographical area in which individual receptors are randomly exposed for a relevant exposure duration. Depending on the receptor's movement and activities, an EU may be as small as a child's play area (e.g., sandbox) or as large as the foraging area of an upper trophic level animal predator (e.g., an entire military base). The relationship between the size of the EU, the movements of the target receptor, and health endpoint of concern (i.e., acute or chronic) may dictate the appropriate use of sample data in developing an EPC. One of the assumptions generally made for the concentration term in Superfund risk assessment is that receptors contact all parts of an EU at random, and that measurements are obtained from a simple (or stratified) random sample. If an individual is randomly exposed within the same EU over a long period of time, the most appropriate metric for the EPC would be the true (but unknown) population mean of the concentrations within the EU (e.g., 95% UCL).

Often, the scale of the EU will be different (smaller or larger) than the scale of the sample data. For example, an ecological receptor population may have a small home range relative to the size of the entire site, or the endpoint of concern may be acute toxicity, requiring an evaluation of a short-term exposure scenario. If the receptors are not expected to contact all parts of the site with equal probability, then the EU may be redefined so that only a subset of the data collected for site characterization are used to estimate the EPC. In addition, the location of the EU may be unspecified within the site because there may be multiple areas that provide suitable habitat for the receptor population. Departing from the assumption of random exposure within one unique geographic area presents an additional challenge to estimating an EPC. In some cases, it may be informative to develop multiple estimates of the EPC in a PRA. By treating the EPC as a random variable, risk assessors can explore the effect of uncertainty in the location of the EU. A variety of modeling approaches are available to calculate an EPC (e.g., arithmetic mean, or 95% UCL) based on the spatial variability in chemical concentrations measured over an area larger than the EU. Methods such as geostatistics (see Section 5.5.2 and Appendix D), Microexposure Event Modeling (MEE) (see Appendix D), and random walk scenarios (Hope, 2000, 2001) may be used to quantify both the spatial and temporal variability in exposure to varying concentrations. Using these methods, risk assessors may redefine the EU to be more representative of the random movement of the receptor during the period of exposure. Because these modeling approaches may be considered more advanced methods for quantifying the EPC, they are generally considered in Tier 3 of the PRA process (see Chapter 2).

5.1.2 PRE- AND POST-REMEDIATION EXPOSURE POINT CONCENTRATIONS

The differences between pre- and post-remediation EPCs are discussed below. In general, both estimates of the EPC are based on the same concepts regarding the exposed population and the definition of the EU. However, the post-remediation EPC will tend to yield lower estimates of (post-remediation) risk and can require more advanced methods for calculating uncertainty (e.g., 95% UCL).

The pre-remediation EPC is determined based on existing site sampling at the time of the remedial investigation, prior to remediation. By contrast, the post-remediation EPC generally is determined based on a prediction of site conditions after remediation. For example, in surface soil, the post-remediation EPC can be determined by substituting the nondetect level (generally, half the laboratory reporting limit) for some of the high concentrations in the sample and recalculating the EPC. The underlying assumption in calculating a post-remediation EPC is that remediation will have sufficiently reduced the chemical concentrations at the site, and the risk existing after remediation is complete will be equal to or less than the target risk level of concern.

The preceding discussion is most applicable to surface soil PRGs. In general, compared with other exposure media (e.g., groundwater, air), surface soil is stationary with relatively constant chemical concentrations within an EU. For other environmental media, more complex approaches may be needed to estimate the post-remediation EPC. Modeling of the remediation process may introduce additional uncertainty not encountered in risk estimates based on the pre-remediation EPC.

5.1.3 REMEDIATION ACTION LEVELS (RALS) AND 95% UCL CALCULATION METHODS

The EPC should incorporate knowledge about the spatial distribution of contamination, the behavior of the receptor, the location of the EU, land use, and other factors. These factors affect both the numerical value of an EPC and uncertainty associated with this estimate. In many cases, it is presumed factors associated with land use will not change after remediation.

The remediation action level (RAL) is the maximum concentration that may be left in place at any location within an EU such that the average concentration (or 95% UCL as a measure of the average) will not

present a risk above levels of concern. This RAL may be considered a "not-to-exceed" threshold or action level for the purposes of site remediation. Using surface soil as an example, areas within the EU that have concentrations greater than the RAL may be excavated and replaced with clean fill (e.g., nondetect surrogate values). To obtain a post-remediation EPC, the 95% UCL is calculated after substituting the surrogate nondetect value for all measurements located within the EU that are greater than the RAL.

When appropriate, the same statistical method of uncertainty should be used to estimate UCLs for both the pre- and post-remediation EPCs. However, in some instances, the method used for calculating the pre-remediation EPC will be inappropriate for calculating the post-remediation EPC, because the distribution of contaminant concentration will have changed. For example, pre-remediation site sampling may suggest that variability in concentrations can be reasonably characterized by a lognormal distribution, which would support the use of the Land method for estimating the 95% UCL. The post-remediation site conditions, however, may reflect a mixture of clean fill and contamination, resulting in a poor fit to a lognormal distribution (see Figure 5-3, Section 5.5.3). In this case, the Land method would not be appropriate. Because of the difference in the statistical distribution of concentration measurements used to estimate the pre-remediation EPC and post-remediation EPC, a non-parametric (i.e., distribution free) method should be considered for calculating uncertainty in the average concentrations in both pre- and post-remediation scenarios. In general, when the method used to calculate the 95% UCL for a post-remediation scenario is different than that of the pre-remediation scenario, the 95% UCL for the pre-remediation scenario should be recalculated with the post-remediation method. Results of this change in methodology can be presented as part of a quantitative uncertainty analysis. Specifically, this recalculation will allow for an evaluation of the effect that a RAL has on the confidence interval for the mean. The discordance between pre- and post-remediation distributions can be expected to increase as the degree of remediation needed to achieve a target risk level of concern increases.

In general, risk assessors should be aware of the practical and statistical issues associated with the various methods of calculating the 95% UCL, and the application of these methods to both the pre- and post-remediation concentration distribution. Different methods can yield very different confidence intervals, some of which are expected to yield more accurate coverage (i.e., likelihood that the confidence interval includes the parameter) depending on characteristics of the underlying distribution of concentrations, such as distribution shape, sample size, and variance (Gilbert, 1987; Hall, 1988). Information about a variety of parametric and non-parametric methods, such as bootstrap resampling, can be found in *The Lognormal Distribution in Environmental Applications* (U.S. EPA, 1997), *Estimating EPCs When the Distribution is Neither Normal nor Lognormal* (Schulz and Griffin, 1999) and a Superfund guidance document currently under development, *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites* (U.S. EPA, 2001a).

5.1.4 CONSIDERATION OF RISK FROM ACUTE TOXICITY

Sometimes a risk assessment will need to address more than one health endpoint of concern (e.g., cancer and noncancer). The RAL should be sufficiently low so that it is simultaneously protective of each endpoint of concern. Generally, when acute toxicity is a concern, the long-term average concentration across the entire EU may not be the appropriate metric for assessing risks. For example, a single episode of a child ingesting a handful of soil containing malathion may result in an acute toxic effect to that child. Therefore, the RAL must not only be low enough to reduce the post-remediation EPC to acceptable long-term average levels, but also low enough that acute toxicity will not be an issue. This consideration applies to both point estimate and probabilistic estimates of PRGs.

☞ *For consideration of acute toxicity, the risk assessor should consult, as appropriate, with a toxicologist in the development of RALs.*

For a small number of chemicals, toxicity values have been determined based on acute effects (e.g., nitrate in drinking water). However, at present, EPA does not have acute toxicity criteria or guidance on acute toxicity applied to the RAL. Hence, consultation with a toxicologist is vital.

5.1.5 CHARACTERIZATION OF UNCERTAINTY IN THE EPC: POINT ESTIMATES AND DISTRIBUTIONS

In point estimate risk assessments, the 95% UCL is typically used to characterize uncertainty in the EPC (U.S. EPA, 1992). In PRA, either a point estimate (e.g., 95% UCL) or a probability distribution may be used to characterize uncertainty in the concentration term. The probability distribution may characterize either variability or uncertainty. The terms probability distribution for variability (PDFv) and probability distribution for uncertainty (PDFu) can be used to distinguish between probability distributions for variability and uncertainty, respectively.

The decision to use a point estimate, PDFv, or PDFu, as the input for the concentration term in a Monte Carlo model will depend on the goals of the Monte Carlo simulation, as determined by the tiered process (see Chapter 2). If the goal is to characterize variability in risk, in general, a one-dimensional Monte Carlo analysis (1-D MCA) will be used and the appropriate input for the concentration term will be a point estimate that characterizes uncertainty in the mean concentration within the EU. As explained in Section 5.1.1, risk assessors will need to consider the relationship between the size of the EU, the movements of the target receptor, and health endpoint of concern (i.e., acute or chronic) to determine how to use the available sample data to define the EPC. A PDFu is typically not an appropriate choice for the concentration term in a 1-D MCA when the goal is to characterize variability in risk. Mixing of a PDFu for the concentration term with PDFv's for other exposure variables in 1-D MCA would yield a single risk distribution from which the relative contributions of variability and uncertainty could not be evaluated. Use of a PDFu for the concentration term may be considered in 2-D MCA simulations (see Appendix D), where the goal may be to characterize both variability and uncertainty in risk.

When the sample size is small and the variance is large, the 95% UCL may exceed the maximum detected concentration (MDC). In such a case, the MDC is generally used to estimate the EPC, although the true mean may still be higher than this maximum value (U.S. EPA, 1992). For poorly characterized sites, there may be considerable uncertainty that site remediation will be sufficient to reduce the 95% UCL to a health-protective level. Poor site characterization may provide an impetus for the risk manager to opt for a more health-protective remedial alternative or to collect additional data.

To ensure that actual cleanup based on a RAL is protective generally requires post-remediation confirmation sampling. This step in the risk management process is emphasized further in Section 5.8 on measurement of attainment.

5.1.6 MULTIPLE CHEMICALS

Developing PRGs for multiple chemicals in one or more environmental media is particularly challenging. When multiple chemicals are present, the total risk level should be considered for regulatory purposes with each chemical contributing a portion of the total risk. This issue is quite complex and usually will affect both the calculation of the risk and development of site-specific PRGs. Chemicals may exhibit different spatial and temporal variability within the EU. Fate and transport characteristics may vary between chemicals as well as between different areas of the site. Co-located sampling, or geostatistical techniques (e.g., co-kriging) may

provide insights regarding relationships in spatial patterns for different chemicals (see Appendices C and D) and the corresponding exposures for receptors.

5.2 WHEN TO USE PRA FOR DEVELOPING PRGS

Because point estimate risk assessments and PRA employ different approaches to characterize variability and uncertainty, the resulting RME risk estimates and calculations of PRGs are often different. The magnitude of the difference can depend on many factors, including the number of input variables described with probability distributions in the PRA, the choice of distributions used to characterize variability or uncertainty (especially for those variables that are highly ranked in a sensitivity analysis), the percentile of the probability distribution that corresponds with RME point estimate for each input variable, and the choice of percentile from the PRA used to represent the RME risk (e.g., 95th percentile). Since the results of a point estimate approach and PRA can be expected to differ, but the magnitude of the difference is not known *a priori*, this can present a challenge in deciding whether or not to conduct a PRA to develop a PRG. The potential advantages and disadvantages of both the point estimate approach and the PRA can be factored into the decision (see Chapter 1, Exhibits 1-6 and 1-7).

In general, PRA may be appropriate for developing site-specific PRGs in cases where PRA has also been used to estimate site-specific risks. As indicated by the tiered approach (see Chapter 2), if the risk manager determines that quantifying variability and uncertainty may enhance risk management decision making, PRA may be warranted. If a PRA is feasible, the risk manager should proceed to Tier 2 and employ PRA to complete the RI/FS process. Usually, embedded in a site-specific PRG are all of the exposure assumptions and toxicity metrics used in the risk assessment. Hence, introducing the use of PRA for PRGs in the feasibility study (or any time after the remedial investigation and baseline risk assessment are complete) would, in effect, undermine the tiered approach.

☞ *If only point estimates were used in the risk assessment, probabilistic methods should not be used for PRG development.*

If additional data have been collected to conduct PRA, the point estimate risk assessment should be revisited with the new data as well. As discussed in Chapter 2, a point estimate risk assessment (Tier 1) should always accompany a PRA. PRA is intended to enhance risk management decision making, and should not be viewed as a substitute for point estimate approaches. Using the tiered approach, a risk assessor can determine the appropriate level of complexity that is supported by the available information to conduct the risk assessment and to calculate a PRG.

5.3 METHODS FOR DEVELOPING PRGS

Risk assessors may use PRA to quantify sources of uncertainty and variability in the calculation of PRGs as well as risks. Two of the common methods for calculating PRGs in PRA include: (1) backcalculation (see Section 5.4), which is equivalent in concept to the point estimate calculation of a PRG; and (2) iterative forward calculation methods, including iterative reduction and iterative truncation (see Section 5.5). Backcalculation can be used in PRA when the target risk and concentration terms are expressed as point estimates. Iterative methods can be more involved, but unlike backcalculation, there are no constraints on their application to PRA. The two approaches yield the same result when the same assumptions are used in the risk assessment.

5.4 BACKCALCULATION

Traditionally, risk is calculated as a function of multiple exposure variables, including the concentration term, and toxicity value (Equation 5-1). If one or more of the exposure variables is described by a PDF, a Monte Carlo simulation will yield a distribution for risk (see Chapter 1).

Backcalculation methods can be envisioned as setting a target risk level (e.g., RME risk equal to 10^{-6} or Hazard Index equal to 1) and then algebraically reversing the risk equation to solve for the concentration term (Equation 5-2). A Monte Carlo simulation using Equation 5-2 will yield a distribution of concentrations that reflects the combination of distributions from all other exposure variables.

$$\frac{C \times IR \times EF \times ED}{BW \times AT} = Intake$$
$$Intake \times Toxicity = Risk$$
$$C \times V = Risk$$

Equation 5-1

$$C = Risk \times V^{-1}$$

Equation 5-2

where,

Toxicity	=	toxicity term representing either the cancer slope factor (CSF) or reference dose (1/RfD) for the chemical in the exposure medium
C	=	concentration term
V	=	algebraic combination of the toxicity term with all exposure variables except *C*
IR	=	ingestion or inhalation rate
AT	=	averaging time
BW	=	body weight
ED	=	exposure duration
EF	=	exposure frequency

This calculation produces a distribution of PRGs that represents the same sources of variability as a forward calculation of risk. Each percentile of the PRG distribution (i.e., the α percentile) corresponds to the 1-α percentile from the distribution of risk estimates. For example, if the 95th percentile of the distribution of risk estimates was chosen to represent the RME individual, the 5th percentile (1-0.95=0.05) would be the corresponding concentration value from the distribution of PRGs (Bowers, 1999). The correspondence between the risk distribution and the PRG distribution is intuitive—just as selecting a higher percentile on the risk distribution is more protective, a lower percentile on the PRG distribution is more protective. The RME range for the risk distribution 90th to 99.9th percentile is analogous to an RME range for the PRG distribution of 0.1st to 10th percentile.

Backcalculation has been a familiar method of developing PRGs and may be appropriate in some situations for the sake of clarity and transparency due to the general understanding of this method among risk assessment practitioners. Once a backcalculation has been performed to determine a PRG, the PRG should be used as the concentration term in a forward calculation to ensure that the risk at the PRG is acceptable.

5.4.1 DIFFICULTIES WITH BACKCALCULATION

There are limitations in the use of backcalculation in PRA (Ferson, 1996). Simple rearrangement of Equation 5-1 does not suffice when the variable (i.e., the concentration or risk term) that is backcalculated is represented by a probability distribution (Burmaster et al., 1995; Ferson, 1996). The difficulty for PRA arises because each risk estimate from an MCA that uses the familiar "forward-facing" risk equation represents a combination of random values selected from the input distributions. Therefore, the output can be considered conditional on all of the inputs. Rearranging the risk equation does not maintain the same conditional probabilities; therefore, the distribution for risk estimated as a function of the distribution for concentration in Equation 5-1 does not return the same distribution for concentration when applied in Equation 5-2. While there are techniques that can maintain the dependencies and correlations between exposure factors when the risk equation is rearranged (e.g., deconvolution), they are complex and beyond the scope of this guidance.

Backcalculation methods may also be difficult to implement in situations in which complex fate-and-transport considerations are present. Leaching of soil contamination to groundwater, bioconcentration of chemicals at higher trophic levels, and other multimedia processes that result in exposure via several environmental media are situations in which backcalculation may not be useful. Note that these difficulties are not unique to backcalculation. Uncertainty in fate-and-transport considerations makes any type of PRG determination challenging.

Further, the backcalculation approach only provides information on the EPC that corresponds to a risk level of concern; it does not specify an RAL that would achieve this EPC. For example, when a risk equation is algebraically solved for concentration (see Equation 5-2), a PRG is developed without a corresponding RAL. Thus, there is no information associated with the PRG value to indicate the highest concentration in the EU that must be removed so that the average concentration (or 95% UCL) within the EU is at or below the PRG. Hence, additional efforts are needed. In addition, post-remediation concentrations may need to satisfy more than one regulatory constraint. For example, the average (or 95% UCL) concentration within an EU may need to be less than a concentration associated with chronic toxicity or cancer and simultaneously, the RAL concentration may need to be less than a concentration that might cause acute toxicity.

In spite of these caveats, backcalculation methods may be appropriate for some sites. For example, when the target risk is specified by a single numerical value and the risk manager has chosen a percentile of variability to represent the RME individual, then a backcalculated PRG can be derived from a PRA.

Although backcalculation methods may be appropriate for some sites, risk assessors should be familiar with their limitations. Because of these limitations, this guidance recommends iterative forward calculations as the primary method for calculating PRGs when performing a PRA. Iterative methods avoid difficulties associated with applying MCA to a backcalculation, and can provide more information for the risk manager.

5.5 ITERATIVE METHODS

Iterative methods simply involve calculating risk with the "forward-facing" equation (see Equation 5-1) a number of times (iteratively) using progressively lower values for the concentration term until the risk is sufficiently protective. This iterative method has also been called the "repeated runs" method. Note that iterative methods for calculating a PRG are not uniquely applicable to PRA. Iterative methods also may be used to develop PRGs in point estimate risk assessments.

☞ *EPA recommends iterative simulations as a general approach for calculating PRGs from probabilistic risk assessments.*

Most often, iterative forward calculations are performed using a systematic trial-and-error method until the percentile of variability in risk chosen to represent the RME individual is at or below acceptable risk levels. Sometimes, a short cut can be used to reduce the number of simulations needed with the trial-and-error method. If successive "guesses" of the EPC are plotted with the corresponding risk estimate, the exact solution can be determined from the best-fit line, thereby significantly reducing the effort required to implement this method. An example is given in Figure 5-1. For many risk equations, the relationship between the EPC and the RME risk will be approximately linear. Nevertheless, the final estimate of the EPC should be checked by running another simulation for risk with this estimate.

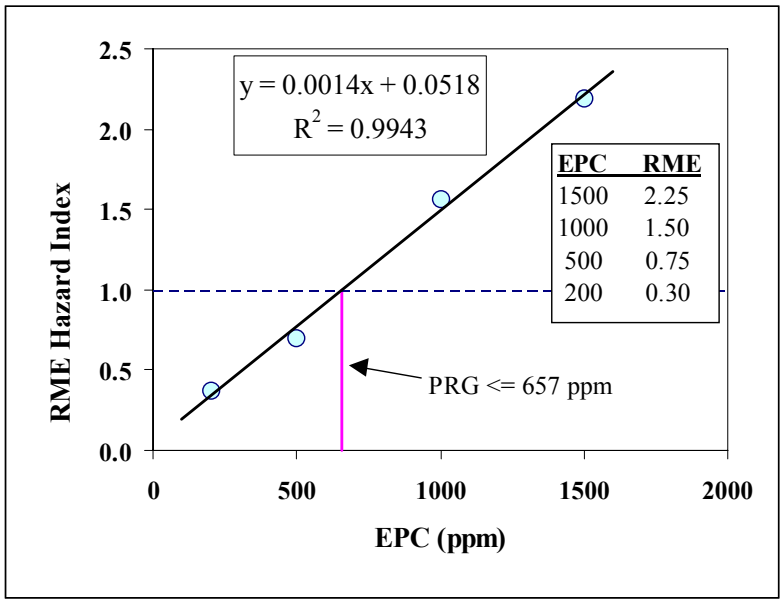


Figure 5-1. A hypothetical example of the use of iterative methods to determine the EPC that corresponds with a target RME Hazard Index (HI) of 1.0. Assume that the EPC is represented by the 95% UCL and the RME HI is the 95th percentile of the output distribution. In this case, four separate Monte Carlo simulations were run with iteratively decreasing values for the EPC. The least-squares, best-fit line to these four data points suggest that a reasonable PRG would be approximately 660 ppm.

A possible and significant advantage of iterative forward calculations over back-calculations is that the method is intuitive and yields a distribution of risks rather than a distribution of PRGs (as with a back-calculation method). The distribution of risks will be more familiar to the public and other stakeholders, and thus, both the method and the resulting output may be easier to communicate to senior level managers and stakeholders (see Chapter 6).

Two general types of iterative methods are described in more detail in Sections 5.5.1 and 5.5.2. The main difference between the methods is in the interpretation of the concentration term that is being reduced. With iterative reduction, the concentration is assumed to be the post-remediation EPC, whereas with iterative truncation, it represents the RAL needed to achieve a post-remediation EPC.

5.5.1 ITERATIVE REDUCTION

Iterative reduction can be applied to any medium. Generally, a point estimate representing the EPC (e.g., 95% UCL) is successively lowered, each time repeating the Monte Carlo simulation of variability in risk. When the EPC is reduced until the endpoint of concern (e.g., RME risk corresponding to the 95th percentile) is at or below an acceptable level of risk, the PRG is set at the corresponding EPC. The goal is to identify the point estimate that corresponds to a target risk level. Note that the PRG is not the same as the RAL. The RAL is the maximum concentration that may be left in place within an EU to achieve the PRG.

The concentration at which the risk is acceptable defines the PRG. Therefore, the PRG bears the same uncertainties as the EPC. For example, assume that a risk assessor examined the carcinogenic effects from

chronic consumption of a chemical in groundwater, then the exposure unit may be determined by the long-term average concentration at any well that potentially draws drinking water from the contaminated groundwater. Uncertainty in the long-term average concentration can reflect a number of factors that contribute to spatial and temporal variability, including the direction of groundwater flow, natural attenuation, and other fate and transport variables. Remediation by a pump-and-treat system for a prolonged period of time may be used to lower the concentrations at the wells. Even though the remediation strategy may be complicated by spatial and temporal variability, iterative reduction can be used to establish a PRG. A remediation strategy may be considered a potential candidate if it can achieve the PRG by reducing the average concentration at each of the well locations. The concept of "hot-spot" removal, or truncation of the highest concentrations first, would not be an option under this scenario (see Section 5.5.2).

5.5.2 ITERATIVE TRUNCATION

Iterative truncation is a method of calculating a PRG that involves developing an expression for the concentration term in which higher values of concentration are removed or "truncated" to reduce the maximum concentration. These higher values are replaced by the surrogate nondetect value. The risk is recalculated for each successive reduction in the highest value. The method is repeated with consecutively lower truncation limits until risk is acceptable.

Iterative truncation is most applicable to surface soil cleanup as the spatial variability over time is minimal compared to other media (e.g., surface water). With each iteration of the risk equation (e.g., Equation 5-1), the highest concentration value is truncated corresponding to a different RAL. In this way a"not-to-exceed" level is specified and the PRG is recalculated the same way in each iteration. The process continues until the risk distribution yields risk estimates at or below the level of concern.

Iterative truncation can be applied to either the empirical distribution function (EDF) for the concentration term, or a fitted distribution for variability in concentrations within the EU. Applied to the EDF, the maximum detected concentration within the EU is replaced with a surrogate value for a nondetect (e.g., half the reporting limit or the background value for some chemicals), and the EPC (e.g., 95% UCL) is recalculated for this altered data set. If this new EPC yields unacceptable risk, then the two highest detected concentrations are replaced by the nondetect value and the EPC is recalculated. In the third iteration, the three highest detections are replaced, and so on, until the target risk level is achieved. Alternatively, the sample data may be fit to a probability distribution for variability, and the process would be repeated with decreasing values in the high-end tail of the continuous distribution.

When the concentration term is a distribution representing uncertainty in the mean concentration, then, similar to the recalculation of the point estimate 95% UCL described above, this distribution of uncertainty in the mean concentration should be determined anew each time a datum is replaced with the nondetect value.

When a distribution of variability in concentration is used for the EPC, for example, in an ecological risk assessment where sampling may be sparse relative to the foraging area of a small home range receptor (see Appendix C), then the distribution developed in an identical way with the high values replaced by the surrogate nondetect value should be used in the iterative determination of a PRG.

The decision to apply iterative truncation should be made after considering a variety of characteristics of the sample data and post-remediation scenario (see Exhibit 5-3). For example, small sample size may result in high uncertainty in the 95% UCL, thereby limiting the use of iterative truncation. Quantitative criteria regarding these factors are not provided in this guidance given that the level of certainty required for decision making will

vary on a case-by-case basis. Use of geostatistical methods (Appendices C and D) may aid in interpreting site data or improving sampling design. Geostatistics is capable of describing the spatial distribution of a contaminant in a quantitative fashion. These methods establish a correspondence between the actual sampling locations and the locations a receptor would be expected to frequent. Additionally, it enables the estimation of concentrations in unsampled locations. Hence, for determination of concentrations at specific locations at a site or within EUs of various sizes and shapes, geostatistics may provide an invaluable tool. Geostatistics has applications both to developing the EPC and PRG and has been recommended and used at some sites for characterization of soil and groundwater contamination (U.S. EPA, 1990b, 1991c).

Although the consideration and use of geostatistics is encouraged, a full consideration of geostatistics is beyond the scope of this guidance. Those interested in greater detail than provided in Appendices C and D are urged to consult the Superfund guidance document currently under development, *Draft Guidance on Surface Soil Cleanup at Superfund Sites: Applying Cleanup Levels* (U.S. EPA, 2001b), for additional discussion of how geostatistics can be used to quantify the concentration term or the PRG.

Generally, iterative truncation methods fail to produce adequate cleanup strategies when site characterization is incomplete. This problem, however, is not specific to PRA. Both point estimate and probabilistic methods are sensitive to poor site characterization.

Risk assessors should realize that application of iterative truncation may result in areas on-site that have concentrations higher than the PRG. This is because the PRG will reflect an average concentration (or 95% UCL) from a distribution of concentrations in which the maximum is truncated at the RAL. For example, Figure 5-3 (see Section 5.5.3) shows how the concentration distribution can be truncated at an RAL, while still leaving behind concentrations greater than the PRG.

EXHIBIT 5-3

CRITERIA FOR ITERATIVE TRUNCATION

1. **Sample size (*n*) is sufficient.** Small sample sizes lead to large estimates of uncertainty in the concentration term. Small sample size may cause the risk assessor to overlook some sources of uncertainty.

2. **Concentration distribution is not highly skewed.** A highly skewed distribution may yield unreliable estimates of uncertainty, especially for small sample sizes.

3. **Sampling design yields a representative distribution of measurements within the exposure unit.** Simple random sampling may fail to represent a patchy spatial distribution of contaminants. Similarly, hotspot (e.g., cluster) sampling may fail to represent random movement of receptors. To evaluate potential biases in sampling, analyses with both standard statistical methods and geostatistical methods may be required.

4. **Assumptions about the post-remedial distribution of concentration are reasonable.** If these assumptions are shown to be incorrect by subsequent sampling events, the process for developing a PRG may need to be repeated and additional remedial activities may be required.

5.5.3 EXAMPLE OF ITERATIVE METHODS

The iterative truncation method is easiest to think about with regard to soil cleanup when contaminated soil is removed and replaced with clean fill soil. This replacement would reduce both the mean and 95% UCL. In most cases, risk assessors may assume that the concentrations of chemicals in clean fill soil can be represented by the surrogate nondetect value (e.g., half the detection limit). Alternatively, the fill may be sampled so that the measured concentrations in the fill dirt may be used to calculate the post-remediation

concentration term. Generally, metals and other inorganic chemicals will be present in clean fill, albeit at lower concentrations than on site.

A simple example using the 95% UCL as a point estimate for the EPC is given in Exhibit 5-4. In this example, background concentrations of chemical X were very low and hence, the fill was assumed to have a concentration of half the detection limit. The risk management objective is to identify a PRG in which the 95th percentile risk estimate is below 1E-04 and to determine the RAL necessary to achieve this PRG. This example illustrates how iterative truncation is applied to the empirical distribution function, rather than fitting the concentrations to a parametric distribution.

Assume that iterative reduction of the 95% UCL demonstrated that a post-remediation EPC of no greater than 33 mg/kg is needed to achieve a RME risk of 1E-04. What is the RAL that yields this EPC? The risk assessor recognizes that the post-remediation concentration distribution is very often a mixed distribution, consisting of a group of nondetect values and a truncated parametric distribution. Because of the complex nature of mixed distributions (Roeder, 1994), non-parametric methods for calculating the 95% UCL of the arithmetic mean (e.g., bootstrap resampling) were determined to be appropriate (U.S. EPA, 1997; Section 5.1.3).

EXHIBIT 5-4

EXAMPLE OF ITERATIVE METHODS

Scoping and Problem Formulation

Chromium contamination was present at a 12-acre industrial facility. In scoping and problem formulation, all stakeholders agreed that the facility would maintain itself and the current land use would continue into the foreseeable future. Most of the facility area was maintained as green space and as a buffer with the surrounding community. Surrounding the facility to the fence line were lawns and ornamental shrubs tended by landscape workers. These landscape workers were considered to be the high risk group as they would move freely and randomly over the entire area of the facility outside the buildings. Hence, the landscape workers would be exposed to an average concentration over the entire area of the facility outside the buildings. The management of the facility was very cooperative and concerned about their workers. Nonetheless, the facility management did not wish to bear more cost than necessary.

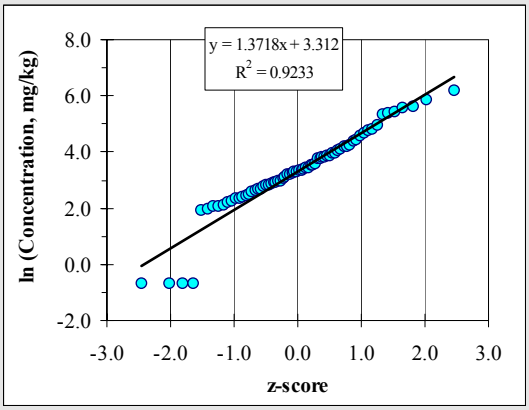
Site Characterization - Soil Sample (n=70)

Seventy surface soil samples were obtained using a sampling grid placed over all 12 acres. Five or six sampling locations were placed in each acre. None of the samples was composited. The grid-based sampling permits a rough estimation of the percentage of the site that would need active remediation. The detection limit for the chromium was 1 mg/kg. Four of the samples were nondetects. Sampling results are shown in Table 5-1. Although the samples from the site appeared to occur in a lognormal distribution (Figure 5-2), the presumed post-remediation distribution would be a mixed distribution, consisting of a truncated lognormal distribution and a group of data at the surrogate nondetect value.

Figure 5-2. Lognormal probability plot of soil concentrations, including 4 nondetects.

Table 5-1. Soil sample (n=70) (mg/kg).

0.5	9.7	16.2	25.1	34.0	54.1	120.6
0.5	10.6	17.1	25.4	34.0	57.8	122.2
0.5	10.8	17.4	26.4	36.5	60.2	140.7
0.5	11.0	17.9	26.9	43.3	65.7	211.9
6.8	11.8	18.4	27.1	43.3	66.1	224.1
7.2	12.0	18.6	28.2	45.3	71.8	235.6
7.8	13.7	19.7	28.3	46.4	82.7	266.8
8.0	13.9	19.8	30.3	48.2	84.7	284.0
8.2	14.7	22.0	30.9	49.3	98.1	361.2
9.3	15.0	22.8	31.1	52.6	107.7	486.6



In this example, a series of iterative truncations showed that removal of all sample results greater than 100 mg/kg (n=11) and replacement of these with the nondetect surrogate of 0.5 mg/kg yielded a 95% UCL of 33 mg/kg and RME risk below 1E-04. Table 5-2 summarizes the results of the calculations for the three conditions: (1) pre-remediation concentrations; (2) post-remediation concentrations using iterative truncation to achieve an RAL of 100 mg/kg; and (3) post-remediation concentrations assuming the 95% UCL calculated is used as the RAL. Note that if the PRG of 33 mg/kg was applied as a "not-to-exceed" level (i.e., RAL), the resulting remediation effort would increase from 15 to 40% of the site, yielding a 95% UCL of 14 mg/kg. While this would be a protective decision, other information was used to support the selection of the second scenario instead. A toxicologist was consulted, who indicated that acute exposure to the workers at levels of 100 mg/kg would not present a health risk. To build additional protectiveness into the remedy, the management also indicated scheduling for the landscape workers would be performed so the areas tended would be rotated among all the workers.

Table 5-2. Pre- and Post-Remediation EPCs (95% UCLs) for Chemical X in Surface Soil Samples.

Remediation Scenario	RAL (mg/kg)	EPC (mg/kg) 95% UCL	Percent of Site to be Remediated
1. Pre-remediation	NA	93	NA
2. Post-remediation using the PRG as the 95% UCL	100	33	15%
3. Post-remediation using the PRG as the RAL (i.e., "not-to-exceed")	33	14	40%

NA=not applicable for a pre-remediation scenario.

Figure 5-3 shows a conceptual framework for considering the post-remediation distribution as a mixture between a group of nondetects and a distribution of contamination truncated at the RAL. Prior to remediation, the EPC exceeds a level that would be protective of human health and ecosystems. If the high-end soil concentrations are removed and the soil is replaced with clean fill, the resulting distribution will be bimodal, with one peak occurring at the nondetect concentration, and the second occurring near the mean of the post-remediation distribution.

5.5.4 MULTIPLE EXPOSURE UNITS AND ITERATIVE METHODS

When multiple EUs are present at the site, there may be a small number of samples within a given EU and the uncertainty in the concentration term generally will be large. It may be possible to use knowledge of the mechanism of how the contamination occurred along with spatial patterns in the sampling results in other nearby EUs to quantify uncertainty. Geostatistical techniques for estimating the mean concentration may provide useful insights into the importance of accounting for spatial relationships among the sample data. Appendix C also provides a discussion of the situation of multiple EUs within a larger site.

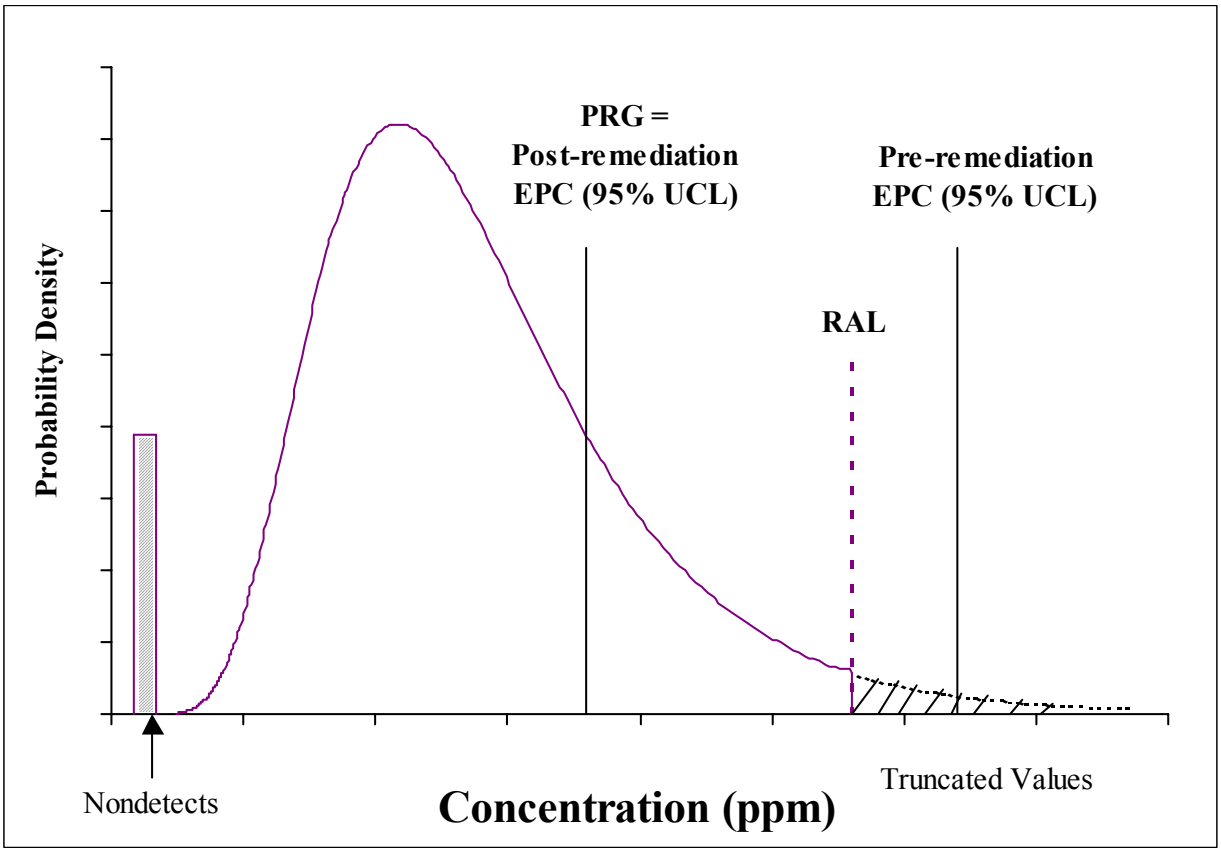


Figure 5-3. Hypothetical example of a mixed, bimodal distribution that represents a combination of the pre-remediation distribution truncated at the remediation action level (RAL) and a uniform distribution representing clean fill at the surrogate nondetect concentration. Shaded portions represent equal areas. In this example, the PRG is defined by the post-remediation EPC (95% UCL).

5.6 PRGs FOR GROUNDWATER

For some chemicals encountered at hazardous waste sites, chemical-specific ARARs may exist, and may be considered as PRGs. ARARs may be selected as site-specific cleanup levels. The maximum contaminant levels of the Safe Drinking Water Act are examples of ARARs.

☞ *For groundwater contamination, ARARs should be applied as RALs if they are protective.*

Of course, for cases in which an ARAR is less protective than a remediation goal determined from a risk assessment, then a risk-based PRG may be developed in accordance with the NCP (U.S. EPA, 1990a).

As an exposure medium, groundwater is the opposite of soil in that groundwater is not static, and receptors are usually exposed at one location (i.e., the well head). Often, a single well can be considered the EU when assessing risks associated with either the residential or industrial/occupational scenarios. The EPC may still reflect the concept of averaging over a long time period (e.g., years) due to potential changes in concentrations in

well water over time. For example, chemical fate and transport modeling may suggest that concentrations are decreasing over time. Similarly, there may be temporal and spatial variability depending on the seasonal fluctuations of the water table. Ideally, the risk assessment would focus on individuals who may be exposed at locations nearest to the center of the contaminant plume, where concentrations are likely to be highest (Freeze and Cherry, 1979; Sposito, et al., 1986).

Because of the uncertainty in the movement of groundwater and the necessity of sampling the medium at fixed locations, identifying a meaningful RAL needed to achieve a given PRG is difficult. In most cases, ARARs will be applicable as RALs or "not-to-exceed" levels.

5.7 PRGs FOR OTHER CONTAMINATED MEDIA

Iterative truncation techniques are generally applied to a static medium, such as soil, rather than dynamic or fluid media such as water and air. This is simply because it is difficult to design a method that will selectively remove high concentrations from a fluid medium. Iterative reduction may be more relevant than iterative truncation when an RAL cannot be developed. These issues are discussed below with respect to sediment, surface water, and fish.

Sediment

Sediment may be transported over time more readily than soils. If it can be assumed that the sediment remains in place, then iterative truncation techniques may be applied. However, at some sites, sediment may be considered a fluid medium. For example, sediment may be resuspended by the movement of water craft, waves, changing tides, or erosion. Similarly, the depth of the contaminated sediment may change over time as new layers of sediment are deposited above more contaminated sediment.

Exhibit 5-5 gives an example of the use of iterative truncation to evaluate alternative RALs for sediment of a lake contaminated by pesticide runoff. In this example, the RAL is related to both the ecological endpoint of concern (i.e., reduction in reproductive success of mammalian omnivores at the lake) and the fraction of areal extent of the lake that would require remediation at that RAL.

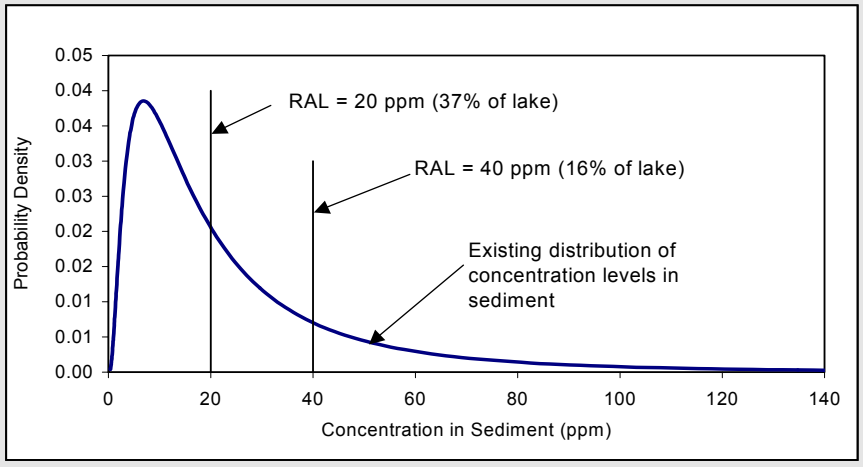
EXHIBIT 5-5

EVALUATION OF ALTERNATIVE RALS USING ITERATIVE TRUNCATION

Risks to a population of mammalian omnivores residing near a lake contaminated with pesticide "X" were judged to be sufficiently high that a reduction in population number over time was expected (see Chapter 4, Exhibit 4-12). The primary reservoir of pesticide X in the lake is sediment. The BTAG committee decided to use the iterative truncation method to estimate the beneficial effects of a series of different Remedial Action Levels (RALs). PRA was used to predict the distribution of responses (percent reduction in population success) and the areal extent of the lake requiring remediation as a function of RAL. The results are summarized below.

RAL in Sediment	Reduction in Reproductive Success			Fraction of Lake
	Mean	90th	95th	
None	8.9%	31%	59%	0%
100 ppm	7.6%	24%	48%	3%
80 ppm	7.0%	27%	45%	5%
60 ppm	5.9%	18%	36%	8%
40 ppm	4.4%	12%	26%	16%
20 ppm	1.9%	4.7%	10%	37%

The BTAG reviewed these results and concluded that while an RAL of 20 ppm would be needed to provide nearly complete protection of the exposed population, an RAL of 40 ppm would provide a good reduction in effect level while tending to minimize the areal extent of the lake that required remediation, which in turn would tend to minimize disturbance of the ecosystem during remediation. Based on this, the risk manager identified 40 ppm as the RAL and initiated a feasibility study to investigate ways of achieving this objective.



Biota (Fish, Aquatic Invertebrates, Plants)

Biota, such as fish, aquatic invertebrates, and plants can serve as bioindicators or indirect estimators of contamination in other exposure media that would be targets for remediation. The concentration of chemicals fish may reflect a combination of exposures via sediment, the water column, and food source (e.g., prey). Therefore, the use of bioindicators to develop PRGs in other media introduces a sources of uncertainty. If there is a high correlation between concentrations in fish and sediment, then sediment concentrations may be considered when developing PRGs to protect the receptor population. The EU, in this case, is the area where the angler population, or ecological predator population, harvests fish. However, in risk assessments that include a fish ingestion exposure pathway, there may be high uncertainty about the true concentration term. Concentrations may be affected by many factors, including changes in the fish population and changes in fish preferences, which may be difficult to address in risk assessments. The choice of fish species consumed by a given individual may also affect the concentration term.

Fish population studies and fate and transport considerations of the contaminants may indicate if and when a fish population will reach a calculated cleanup level. For many sites, it may be difficult to obtain this level of site-specific data due to resource and time constraints.

Although remediation may not immediately reduce contaminant concentrations in biota, the determination of a cleanup level can serve as a target for any future decline in concentrations. In general, iterative reduction methods are applicable for developing PRGs to protect aquatic ecosystems; however, under some conditions iterative truncation may also be used. For example, if contamination is correlated to relatively static sediment, and the home-range of the fish is relatively small (e.g., nonmigratory) then iterative truncation may be applicable.

Surface Water

The development of PRGs for surface water is also difficult with iterative truncation. For fluid media (e.g., groundwater or surface water), iterative reduction can be performed using a range of EPCs to determine a PRG with acceptable risk at the target RME percentile.

5.8 MEASUREMENT OF ATTAINMENT

The NCP (U.S. EPA, 1990a) provides for continued monitoring for groundwater cleanups to ensure attainment of the remedial action objectives. In addition, it is common practice among remedial project managers to conduct confirmation sampling after completing a remedy for soil contamination. However, completion of the remedial action according to this strategy does not necessarily mean that risks within EUs at the site have been reduced to levels specified in the ROD. The degree of uncertainty about whether the remedial action at the site has achieved the cleanup level should determine whether confirmation sampling is warranted. In general, confirmation sampling following cleanup activities is recommended. Sampling after the remedial investigation is complete may show additional areas needing remediation (i.e., where additional contamination exists).

If additional sampling is conducted after the remedial investigation, the concentration term and corresponding estimates of risk should be recalculated. The PRG developed in the remedial investigation may not be health-protective in light of the additional contamination. The same concepts that relate the concentration term to the PRG should be applied in this situation.

Confirmation sampling activities are included in remedial design/remedial action plans to ensure the remedy is successful. In addition, the five-year review presents a second opportunity to ensure that any contamination left on site does not pose an unacceptable risk.

☞ *If confirmation sampling indicates an insufficient reduction in risk, a more extensive remediation effort may be needed. Possible reasons for not achieving remedial action objectives can include inadequate site characterization or the discovery of unknown contamination.*

For post-remediation sampling, the DQO process should generally be followed. If the post-remediation risk associated with the confirmation sample indicates risk exceeds a level of concern, then additional remediation may be warranted.

5.9 SUMMARY OF RECOMMENDED METHODS

Table 5-3 summarizes the possible methods for developing PRGs for various environmental media. It should be noted that iterative reduction (IR) can be used in all cases, whereas iterative truncation (IT) is limited to situations where the highest concentrations can be identified and removed. Backcalculation may be applicable in all cases, but because of caveats noted in Section 5.4.1, iterative approaches are generally recommended in this document.

Table 5-3. Summary of Potential Methods for PRG Development by Environmental Medium.

Potential Exposure Medium	Back-calculation	Iterative Reduction (IR)	Iterative Truncation (IT)	Explanations for IT
Soil	X	X	X	Applicable if soil is relatively fixed.
Sediment	X	X	X	Applicable if sediment is relatively fixed. In some situations, sediment transport may be a better assumption due to current velocity, tides, resuspension, etc.
Biota (Fish, Aquatic Invertebrates, Plants) - bioindicators of contamination in sediment	X	X	SA	Depends on home-range of fish relative to the scale of the sampling design. If contamination is correlated to relatively static sediment, and the home-range of the fish is relatively small (e.g., non-migratory) then IT may be applicable.
Surface Water	X	X	NA	Not applicable as surface water is a fluid medium.
Groundwater (GW)	X	X	NA	Not applicable as GW is a fluid medium. Generally, ARARs must also be satisfied.
Home-grown produce, milk, livestock, other food items	X	X	SA	Depends on relative contributions of soil uptake (applicable) vs. foliar deposition (not applicable).

X=applicable
NA=not applicable
SA=sometimes applicable

REFERENCES FOR CHAPTER 5

Bowers, T.S., N.S. Shifrin, and B.L. Murphy. 1996. Statistical Approach to Meeting Soil Cleanup Goals. *Environ. Sci. Technol.* 30:1437–1444.

Bowers, T.S. 1999. The Concentration Term and Derivation of Cleanup Goals Using Probabilistic Risk Assessment. *Hum. Ecol. Risk Assess*. 5(4):809–821.

Burmaster, D.E., K.J. Lloyd, and K.M. Thompson. 1995. The Need for New Methods to Backcalculate Soil Cleanup Targets in Interval and Probabilistic Cancer Risk Assessments. *Hum. Ecol. Risk Assess.* 1(1):89–100.

Ferson, S. 1996. What Monte Carlo Methods Cannot Do. *Hum. Ecol. Risk Assess.* 2:990–1007.

Freeze, R.A. and J.A. Cherry. 1979. Groundwater. Prentice Hall, Inc., NJ.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, NY.

Hall, P. 1988. Theoretical Comparison of Bootstrap Confidence Intervals. *Ann. Statist.* 16:927–953.

Hope, B.K. 2000. Generating Probabilistic Spatially-Explicit Individual and Population Exposure Estimates for Ecological Risk Assessments. *Risk Anal.* 20(5):573–589.

Hope, B.K. 2001. A Case Study Comparing Static and Spatially Explicit Ecological Exposure Analysis Methods. *Risk Anal*. 21(6):1001–1010.

Roeder, Kathryn. 1994. A Graphical Technique for Determining the Number of Components in a Mixture of Normals. *J. Amer. Stat. Assoc.* 89(426):487–495.

Schulz, T.W. and S. Griffin. 1999. Estimating Risk Assessment Exposure Point Concentrations When the Data are not Normal or Lognormal. *Risk Anal.* 19: 577–584.

Sposito, G., W.A. Jury, and V.K. Gupta. 1986. Fundamental Problems in the Stochastic Convection-Dispersion Model of Solute Transport in Aquifers and Field Soils. *Water Res.* 22(1):77–88.

U.S. EPA. 1990a. *National Oil and Hazardous Substances Pollution Contingency Plan*. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, Thursday, March 8.

U.S. EPA. 1990b. *Geostatistics for Waste Management*. A Users Manual for the GEOPACK Geostatistical Software. EPA/600/8-90/004, January.

U.S. EPA. 1991a. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume I: Human Health Evaluation Manual* (HHEM), Part B, Development of Risk-Based Preliminary Remediation Goals. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333.

U.S. EPA. 1991b. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume I: Human Health Evaluation Manual* (HHEM), Supplemental Guidance: Standard Default Exposure Factors, Interim Final. Office of Emergency and Remedial Response, Washington, DC. OSWER Directive No. 9285.6-03. June.

U.S. EPA. 1991c. GEO-EAS 1.2.1 Users Guide. EPA/600/8-91/008. April.

U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-081.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund: Interim Final Guidance*. Office of Research and Development, Washington, DC. EPA/540/R-93/071.

U.S. EPA. 1994. *Guidance for the Data Quality Objectives Process (EPA QA/G-4).* Office of Research and Development, Washington, DC. EPA/600/R-96/055. September.

U.S. EPA. 1997. *The Lognormal Distribution in Environmental Applications*. Office of Research and Development, and Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006. December.

U.S. EPA. 2000. *Data Quality Objectives Process for Hazardous Waste Site Investigations*. Office of Environmental Information, Washington, DC. EPA/600/R-00/007. January.

U.S. EPA. 2001a. *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites*. Office of Emergency and Remedial Response, Washington, DC.

U.S. EPA. 2001b. *Draft Guidance on Surface Soil Cleanup at Superfund Sites: Applying Cleanup Levels.* Office of Emergency and Remedial Response, Washington, DC.

U.S. EPA. 2001c. *Integrating Dynamic Field Activities into the Superfund Response Process: A Guide For Project Managers.* Final Draft. Office of Emergency and Remedial Response, Washington, DC. OSWER Directive No. 9200.1-40. December.

U.S. EPA. 2001d. *Improving Sampling, Analysis, and Data Management for Site Investigation and Cleanup.* Technology Innovation Office. EPA/542/F-01/030a. April.

U.S. EPA. 2001e. *Resources for Strategic Site Investigation and Monitoring.* Technology Innovation Office, Washington, DC. EPA/542/F-01/030b. September.

CHAPTER 6

COMMUNICATING RISKS AND UNCERTAINTIES IN PROBABILISTIC RISK ASSESSMENTS

6.0 INTRODUCTION

The Environmental Protection Agency (EPA) has developed a guidance document, *Risk Assessment Guidance for Superfund: Volume I–Human Health Evaluation Manual, Supplement to Part A: Community Involvement in Superfund Risk Assessments* (U.S. EPA, 1999a) and two videotapes, "*Superfund Risk Assessment and How You Can Help, An Overview"* (10 minutes) (U.S. EPA, 1999b) and *"Superfund Risk Assessment and How You Can Help"* (40 minutes) (U.S. EPA, 2000b), to improve community involvement in the Superfund risk assessment process. The videotapes (available in both English and Spanish) show examples of how regions have involved communities in the risk assessment process at several Superfund sites. The guidance document and videotapes, along with the *Superfund Community Involvement Handbook and Toolkit* (U.S. EPA, 1998), should serve as a primary community involvement resource for risk assessors and remedial project managers (RPMs). The *Handbook and Toolkit* offers the following specific guidance:

- Provides suggestions for how Superfund staff and community members can work together during the early stages of Superfund remedial investigation and feasibility study (RI/FS) and later cleanup

- Identifies where, within the framework of the human health risk assessment methodology, community input can augment and improve EPA's estimates of exposure and risk.

- Recommends questions the site team (risk assessor, RPM, and community involvement coordinator [CIC]) should ask the community.

- Illustrates why community involvement is valuable during the human health risk assessment at Superfund sites.

This chapter provides guidance and suggestions on how to deal with risk communication issues that arise during a probabilistic risk assessment (PRA). Specifically, the concepts of uncertainty and variability may present additional communication challenges for PRA. For example, whereas discussions of uncertainty for point estimate risk assessments are often qualitative in nature, PRA opens the floor for discussion and presentation of quantitative uncertainty analysis. Concepts associated with quantitative characterizations of uncertainty may be more difficult to communicate and may not be well received due to stakeholder desires for certainty (Slovic et al., 1979). As such, this chapter highlights appropriate stakeholder involvement and principal risk communication skills that are effective for communicating PRA concepts and risk information. Key factors for successful communication of PRA include early and continuous involvement of stakeholders, a well-developed communication plan, good graphics, a working knowledge of the factors that may influence perceptions of risk and uncertainty, and a foundation of trust and credibility.

EXHIBIT 6-1

DEFINITIONS FOR CHAPTER 6

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in a population, usually considered to be the mean or median of the distribution.

Community Advisory Group (CAG) - A group formed to provide a public forum for community members to present and discuss their needs and concerns related to the Superfund decision-making process. A CAG serves as the focal point for the exchange of information among the local community, EPA, State regulatory agency, and other pertinent Federal agencies involved in the cleanup of a Superfund site.

Community Involvement Coordinator (CIC) - As a member of the CAG and site team, the CIC coordinates communication plans (i.e., the CIP) and addresses site-specific CAG organizational issues.

Community Involvement Plan (CIP) - A plan that identifies community concerns and the preferences of the community for the communication of site-related issues.

Confidence Interval - A range of values that are likely to include a population parameter. Confidence intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). When used to characterize uncertainty in a risk estimate, it is assumed that methods used to quantify uncertainty in the model inputs are based on statistical principles such as sampling distributions or Bayesian approaches. For example, given a randomly sampled data set, a 95% confidence interval for the mean can be estimated by deriving a sampling distribution from a Student's t distribution.

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF, gives the cumulative probability of occurrence for a random independent variable. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*.

Hazard Quotient (HQ) - The ratio of estimated site-specific exposure to a single chemical from a site over a specified period to the estimated daily exposure level, at which no adverse health effects are likely to occur.

Hazardous Substance Research Centers (HSRC) - Research centers providing free technical assistance to communities with environmental contamination programs through two distinct outreach programs: Technical Outreach Services for Communities (TOSC) and Technical Assistance to Brownfields Community (TAB).

Histogram - A graphing technique which groups the data into intervals and displays the count of the observations within each interval. It conveys the range of values and the relative frequency (or proportion of the sample) that was observed across that range.

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - A technique for characterizing the uncertainty and variability in risk estimates by repeatedly sampling the probability distributions of the risk equation inputs and using these inputs to calculate a distribution of risk values. A set of iterations or calculations from Monte Carlo sampling is a simulation. For example, a single iteration for risk from ingestion of water may represent a hypothetical individual who drinks 2 L/day and weighs 65 kg; another iteration may represent a hypothetical individual who drinks 1 L/day and weighs 72 kg.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Percentile - A number in a distribution such that X % of the values are less than the number and 1-X % are greater. For example, the 95th percentile is a number in a distribution such that 95% of the values are less than the number and 5% are greater.

EXHIBIT 6-1

DEFINITIONS FOR CHAPTER 6—*Continued*

Point Estimate Risk Assessment - A risk assessment in which a point estimate of risk is calculated from a set of point estimates for exposure and toxicity. Such point estimates of risk can reflect the CTE or RME, depending on the choice of inputs.

Potentially Responsible Party (PRP) - Individuals, companies, or any other party that is potentially liable for Superfund cleanup costs.

Preliminary Remediation Goal (PRG) - Initially developed chemical concentration for an environmental medium that is expected to be protective of human health and ecosystems. PRGs may be developed based on applicable or relevant and appropriate requirements (ARARs), or exposure scenarios evaluated prior to or as a result of the baseline risk assessment. (U.S. EPA, 1991a, 1991b).

Probabilistic Risk Assessment (PRA) - A risk assessment that yields a probability distribution for risk, generally by assigning a probability distribution to represent variability or uncertainty in one or more inputs to the risk equation.

Probability Density Function (PDF) - A function or graph representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Rank Correlation (Spearman Rank Order Correlation Coefficient) - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables.

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Remedial Investigation/Feasibility Study (RI/FS) - Studies undertaken by EPA to delineate the nature and extent of contamination, to evaluate potential risk, and to develop alternatives for cleanup.

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic r that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Stakeholder - Any individual or group who has an interest in or may be affected by EPA's site decision-making process.

Technical Assistance Grant (TAG) A federal grant that is intended to provide a community with the opportunity to hire independent experts to help evaluate and explain the results of a risk assessment.

Technical Outreach Services for Communities (TOSC) - A service of the HSRC with the aim to provide independent technical information and assistance to help communities with hazardous substance pollution problems.

Uncertainty - Lack of knowledge about specific variables, parameters, models, or other factors. Examples include limited data regarding the concentration of a contaminant in an environmental medium and lack of information on local fish consumption practices. Uncertainty may be reduced through further study.

Variable - A quantity that can assume many values.

Section 6.1 discusses the need for early and continuing stakeholder involvement. Section 6.2 recommends a seven-step process for communicating PRA results to stakeholders, and Sections 6.3 and 6.4 provide guidance on specific techniques for communicating information. The success of risk communication efforts will depend on the extent to which the communication strategy addresses the needs of a diverse audience, with different perceptions of risk and uncertainty (Section 6.5), and the degree of trust and credibility that is established from the outset of the process (Section 6.6). Section 6.7 provides a discussion of risk communication issues that are uniquely relevant to RPMs.

6.1 STAKEHOLDER INVOLVEMENT

Many stakeholders may be interested in a risk assessment (see Exhibit 6-2). It is generally important to *involve and engage interested stakeholders early and continuously* throughout the decision-making process (U.S. EPA, 2001).

EXHIBIT 6-2

STAKEHOLDERS POTENTIALLY INVOLVED IN THE DECISION-MAKING PROCESS FOR PRA

- EPA risk assessors and managers
- Members of the public
- Representatives from state or county environmental or health agencies
- Other federal agencies (e.g., health agencies, Natural Resources Damage Assessment (NRDA), trustees, etc.)
- Tribal government representatives
- Potentially responsible parties (PRPs) and their representatives
- Representatives from federal facilities (e.g., Department of Defense, Department of Energy, etc.)

Public involvement activities should be tailored to the needs of the community and described in the site communications strategy. The CIC should coordinate these first steps through the development of a Community Involvement Plan (CIP). Coordination between the RPM, risk assessor, and CIC is needed to determine the appropriate points in the RI/FS process to communicate with the community, and plan for the appropriate level of communication. The CIP should identify community concerns and the preferences of the community for the communication of site-related issues. The CIP may be updated during the RI/FS as needed.

Examples of outreach activities include giving oral presentations and poster sessions at public meetings, coordinating group meetings or focused workshops, conducting interviews with community members on specific issues, and distributing fact sheets.

Ideally, the public and other interested stakeholders would be involved early in the site-specific decision-making process. If the community has not been previously involved, efforts should be made, in coordination with the CIC, to identify and communicate with the appropriate individuals in the community prior to the Agency's receipt of the PRA workplan. The public and other stakeholders should be given the opportunity to provide input to the workplan for a PRA (see Chapter 2, Section 2.1).

The initial community meeting can serve to establish a rapport between EPA and the community and facilitate the exchange of information needed to support a PRA. This information may include policy decisions associated with both point estimate and probabilistic approaches, as well as technical details regarding the conceptual exposure model and the selection of distributions. A discussion of these topics may increase certainty about the assumptions made in the risk assessment. For example, the community may be able to offer insights regarding site-specific activities and sources of exposure data not readily

available to the risk assessor. This type of discussion should allow for the free exchange of information with the public and sets the stage for future discussions. It is important that an appropriate level of detail be presented at the first meeting. Instead of overloading the audience with information, it is generally better to coordinate several meetings so that complex policy and technical concepts can be broken down into smaller discussion topics.

Following the approval of the PRA workplan, the public and other interested stakeholders should be involved in various stages of the PRA development, including providing and/or reviewing data, reviewing the selected distributions (e.g., selected creel survey) and commenting on PRA documents as appropriate during public comment periods. On-going community involvement may require consideration of EPA's resources including the availability of personnel and contractor support. Other considerations include EPA's compliance with provision in the National Contingency Plan (NCP) for involving the community. The appropriate level of community involvement in the PRA should be based on a number of factors including the nature and extent of contamination at the site, the expressed interests of the community members, the complexity of the PRA, and the role of PRA in site-specific remediation or cleanup decisions.

6.2 COMMUNICATION AND PRESENTATION

Communication is a two-way process that should involve the transfer of information between the Agency and the stakeholders, as well as active listening by the Agency to the stakeholder's ideas and concerns. The goals of risk communication are to present risk information in an understandable manner through an open, honest, frank, and transparent presentation and discussion of risks, including uncertainties. In meeting these goals, it is important that the RPMs and risk assessors be sincere and direct in their presentation of the results of the PRA, accept the public and other interested stakeholders as valuable contributors to the process, and listen to the concerns and ideas that are raised.

One goal of communication should be to respect the stakeholder's concerns. The public and other interested stakeholders should have the opportunity to understand the PRA and its effects on the decision-making process. Technical Assistance Grants (TAGs) may be one way to advance this goal by providing the community the opportunity to hire independent experts to help evaluate and explain the results of the PRA. Alternatively, the RPM and risk assessor may use the tools outlined in Sections 6.3 to 6.6 to present PRA concepts and the results of the PRA to the community in a manner that is easily understood. This may require significant up-front planning, testing, and post-evaluation to identify the appropriate messages to communicate and to determine how well this information was communicated.

The site-specific PRA communication plan should be consistent with the NCP's provisions on community involvement. It is important to recognize that community involvement is part of a regulatory process and that EPA generally will consider all timely public input, but may not implement all of it. Ultimately, EPA must meet the legal requirements of the Superfund law in making decisions regarding remedial actions.

A vast body of literature exists regarding risk communication. Since the early 1980's, a number of researchers have developed models for communicating risk to the public. These models are available in the scientific literature, and a list of supplemental references is provided at the end of this chapter.

6.2.1 COMMUNICATION OF PRA WITH CONCERNED CITIZENS, OTHER STAKEHOLDERS, AND MANAGERS: AN OVERVIEW

Before the decision to conduct a PRA is made, a CIP should be in place. Generally, when a decision is made to conduct a PRA, an important step should be to work with citizens to develop a communication strategy for PRA and its application within the Superfund process (see Chapter 1). The initial introduction of the community to the RI/FS process should include a discussion of the principles of risk assessment. This discussion may be best presented in an informal setting such as a public availability session. Because of the potentially complex nature of PRA and quantitative uncertainty analysis, a small group meeting may be an appropriate forum in which to discuss issues and facilitate an exchange of ideas. If there is interest among a large group of stakeholders, multiple small group sessions may be scheduled. Such meetings may provide the foundation for building trust and credibility (see Section 6.6).

In general, it is important to identify whether a Community Advisory Group (CAG) should be formed. The purpose of a CAG is to provide a public forum for community members to present and discuss their needs and concerns related to the Superfund decision-making process. The CIC is an important member of the team and may coordinate communication plans, hand-out materials, and address site-specific organizational issues.

A number of resources may be available to the community to aid in understanding technical material in a PRA. In addition to the TAG program, which provides funds for qualified citizens' groups affected by a Superfund site to hire independent technical advisors, another program is the Technical Outreach Services for Communities (TOSC), which uses university educational and technical resources to help communities understand the technical issues involved in hazardous waste sites in their communities. This is a no-cost, non-advocate, technical assistance program supported by the Hazardous Substance Research Centers.

The tiered approach for PRA presented in Chapter 2 (Figures 2-1 and 2-2) encourages risk assessors and RPMs to participate in discussions with stakeholders early in the process of developing point estimate and probabilistic approaches. If a decision is made to perform a PRA, a continuing dialogue should be useful to evaluate interim results of the PRA and determine if additional activities are warranted (e.g., data collection, further modeling). These on-going discussions should help assure that RPMs are aware of the details of the PRA analysis and are comfortable with the material that will be shared with the community, other interested stakeholders, and senior managers.

6.2.2 STEPS FOR COMMUNICATION OF THE RESULTS OF THE PRA

The complexity of a PRA will vary depending on the site-specific nature of the assessment performed. For example, PRAs may include an analysis of variability, uncertainty, or both. Some analyses may involve simulations to evaluate temporal variability (e.g., Microexposure Event analysis) and spatial variability (e.g., geostatistics). The challenge for presenters is to *identify the critical information and level of detail to be presented to various audiences that may be involved in the Superfund decision-making process* (e.g., senior risk managers, concerned citizens, congressional staff, and PRPs).

The 7-step process, described below (and summarized in Exhibit 6-3), may be repeated many times during the performance of a PRA. For communication purposes, a PRA normally will involve more interaction with stakeholders than a point estimate risk assessment because PRA concepts and results are often more difficult to communicate.

EXHIBIT 6-3

IMPORTANT STEPS FOR COMMUNICATING PRA RESULTS

(1) Identify the audience
(2) Identify the needs of the audience
(3) Develop a communication plan
(4) Practice to assure clarity of presentation
(5) Present information
(6) Post-meeting review of presentation and community feedback
(7) Update information as needed for future assessments and presentations

(1) Identify the Audience

The first step should be to identify the audience of potentially interested stakeholders. Strategies for presenting PRA information normally will be tailored to the audience. Participants in the audience may change during the tiered process depending on the complexity of the PRA (see Chapter 2) and the specific site-management decisions being made.

(2) Identify the Needs of the Audience

The second step should be to identify the needs of the audience. The relevant information and the appropriate level of detail will vary depending on the audience. For example, some participants may be well informed about PRA concepts and will not need much introductory PRA information. For other audiences, PRA concepts may be new, so it may be beneficial to hold an informal meeting to discuss the general objectives and methods used to conduct a PRA. Once introductory PRA concepts have been discussed and are understood by the audience, more advanced discussions may be warranted on topics such as the sources of data used in the PRA, the most critical variables in the PRA (identified during the sensitivity analysis), the selection of distributions, and the level of characterization of uncertainty (see also Section 6.5). The risk assessor should select the key information for each topic and discuss the significance of this information based on the intended audience.

(3) Develop a Communication Plan

The third step should be to develop a plan to communicate significant information to the public in an easily understandable format (Exhibit 6-4). Adequate planning in the presentation of PRA information is essential. A thorough understanding of the design and results of the PRA will help to place the information in proper context and understandable format (U.S. EPA, 1994). Even more importantly, the risk assessors and RPMs should clearly identify the main messages to be presented.

EXHIBIT 6-4

KEY CONSIDERATIONS IN DEVELOPING UNDERSTANDABLE MATERIAL

- Identify main messages
- Place information in appropriate context
- Use clear formats
- Use examples and graphs
- Provide handouts and glossaries
- Present information with minimum jargon

Section 6.4 provides examples of graphics that may be useful in presentations of PRA. Handouts, glossaries, and other materials may complement a presentation and provide information for discussion following the meetings. In addition, examples designed to help demonstrate concepts unique to PRA (e.g., using one probability distribution to describe variability and a second distribution to describe parameter uncertainty) may help facilitate the flow of communication and increase the level of understanding. One useful technique in public meetings is to involve members of the audience to illustrate a concept. For example, the topic of discussion may be the method used to select and fit a probability distribution used to characterize variability in a PRA. To demonstrate this concept, a risk assessor can draw a bell-shaped curve on a flip chart and label the *x-axis*, "number of liters of water consumed per day", and the *y-axis*, "number of people who consume a specific amount of water in a day". Next, each meeting participant can be asked to identify their own consumption pattern, perhaps by holding up a 0.5 liter bottle and asking how many such bottles are consumed on an average day. This community-specific information can then be plotted on a new graph in the form of a histogram and the bars can be connected to form a curve or distribution similar to the one first drawn. The resulting distribution (for an example, see Figure 6-1) can then be used to discuss the following PRA concepts in more detail:

- Variability (between individuals)
- Shape of the distribution and plausible range of values
- Central tendency exposure (CTE) and reasonable maximum exposure (RME) estimation
- Uncertainty in the distribution (sample size, potential response bias, differences in activity patterns)
- Uncertainty in a parameter estimate (difference between the 95% upper confidence limit (UCL) for a mean and the 95th percentile)

Using this information as a basis, the risk assessor can compare the results from the community analysis with data from various geographic areas in the U.S. where water consumption patterns may differ. The risk assessor can then lead a discussion with the community regarding the various sources of uncertainty in selecting and fitting exposure distributions, including:

(a) **Extent of Representation** - Are the available data representative of the target population? For example, would the data on water consumption collected during the meeting be representative for various population groups?

(b) **Data Quantity** - What sample size is needed to develop a distribution? This discussion will introduce the concept that uncertainty in both point estimates and probability distributions may be reduced by increasing the sample size

(c) **Data Quality** - Are the data collected using acceptable study protocols? Is the information available from the peer-reviewed literature? An example can be made of the data collected during the meeting to highlight issues associated with survey design, and methods for controlling for potential bias or error. For example, if the survey data were to be used in a risk assessment for a drinking water scenario, the data quality may be improved by repeat sampling over time

Other exposure variables that can be used in this distribution example include: fish consumption rates, chemical concentrations in soil, and fraction of time spent indoors. In general, examples should focus on variables that may be of interest, are easily illustrated, and are unlikely to make participants uncomfortable divulging personal information such as age.

(4) Practice to Assure Clarity of Presentation

The fourth step should be to practice the presentation to assure that the information is presented clearly to the intended audience. Staff from communication groups or public information offices within EPA regional offices may help to determine whether or not the presentation addresses the needs of various audiences. Also, practicing the presentation with co-workers who are unfamiliar with the site can help assure that the appropriate messages are being conveyed, and will help the team prepare for potential questions that will arise during the meeting.

(5) Present Information

A number of factors should be considered when developing a plan to present the PRA in a meeting. Although the size of the public meeting can sometimes be unpredictable, typically individuals will feel more comfortable asking questions and expressing opinions in small, informal settings. For any audience, it is usually helpful to have general fact sheets on PRA available for distribution. The fact sheets may contain information that describes the PRA process, how information from the PRA will be used at the site, and how the community may comment on the PRA report. The meeting team should usually include the CIC, RPM, Risk Assessor, and additional support as necessary.

Audio-visual materials and equipment should be checked prior to the start of the meeting. For example, overheads should be viewed from the audience seating to assure that information is accessible and readable. Presentations using portable computers can be effective for showing how the results of the PRA may differ with changes in modeling assumptions.

(6) Post-meeting Review of Presentation and Community Feedback

At the end of a meeting, it can be helpful to encourage participants to provide feedback regarding effective and ineffective communication techniques. Not only can this information be used to improve presentations offered to similar audiences in the future, it also provides a sense for how well the main messages and specific technical issues were communicated.

(7) Update Information as Needed for Future Assessments and Presentations

Shortly after the meeting or briefing, modifications should be made to the materials for future presentations where appropriate. In addition, if information is obtained that is relevant to the risk assessment, this information may be included in a subsequent analysis, and the process would be repeated.

6.3 COMMUNICATING DIFFERENCES BETWEEN POINT ESTIMATE AND PRA

One method for effectively explaining the PRA approach to quantifying variability and uncertainty is to employ comparisons to the more easily understood point estimate methodology. These comparisons can focus on either the inputs or the outputs associated with the two approaches. The communicator may focus on a specific input variable, such as drinking water intake, and explain that with the point estimate methodology, a single average or high-end value (e.g., 2 liters per day for adults) normally is used to quantify exposure, whereas with PRA, a probability distribution (e.g., lognormal) is used to characterize variability in exposure among a population. In addition, the outcomes (e.g., cancer risk estimates) can be compared by showing where the point estimate(s) of risk fall within the distribution of risks generated with PRA.

When communicating results from point estimate and PRA models, an important concept to keep in mind is that both methods yield risk estimates with varying degrees of uncertainty. Continuing with the above example, concepts associated with uncertainty (e.g., representativeness, data quantity, and data quality) can be introduced by asking the audience if their estimate of water consumption on a specific day would be equal to their average daily consumption rate over a 1-year period. This example highlights a common source of uncertainty in exposure data (i.e., using short-term survey data to estimate long-term behavior). Section 6.5 discusses different perceptions of uncertainty.

It is common to accept output from quantitative models without fully understanding or appreciating the corresponding uncertainties and underlying assumptions. One challenge in presenting PRA results is to determine the most effective way to communicate sources of uncertainty without undermining the credibility of the assessment (see Section 6.6). For example, it may be counterintuitive that the more sources of uncertainty that are accounted for in a PRA, the wider the confidence intervals tend to be in the risk estimates (see Section 6.4.2). The audience may question the utility of a method that appears to introduce more complexity in a risk management decision. It may be useful to point out that many sources of uncertainty are present, and methods available to acknowledge and quantify them may differ in point estimate and probabilistic risk assessments.

The basic concepts of PRAs described in Chapter 1 may be used in developing presentations. Exhibits 1-5 and 1-6 in Chapter 1 summarize some of the advantages and disadvantages of point estimates and probabilistic approaches that should be considered when evaluating differences in the risk estimates of the two approaches. For example, point estimates of risk do not specify the proportion of the population that may experience unacceptable risks. In contrast, PRA methods allow statements to be made regarding both the probability of exceeding a target risk, and the level of confidence in the risk estimate.

When summarizing results of PRA, graphs and tables generally should include the results of the point estimates of risk (e.g., CTE and RME). It may be informative to note where on the risk distribution each of the point estimates lies. By understanding the assumptions regarding the inputs and modeling

approaches used to derive point estimates and probabilistic estimates of risk, a communicator will be better prepared to explain the significant differences in risk estimates that may occur. Special emphasis should be given to the model and parameter assumptions that have the most influence on the risk estimates, as determined from the sensitivity analysis (see Appendix A).

6.4 GRAPHICAL PRESENTATION OF PRA RESULTS TO VARIOUS AUDIENCES

Graphics can be an effective tool for communicating concepts in PRA. As the old adage goes, "A picture is worth a thousand words." A graphic usually can be most easily understood by a diverse audience when it conveys a single message. It is generally a good idea to keep the graphics simple so that the message is clear. In general, each graphic should be developed and modified depending on the type of presentation and the intended audience.

☞ *The key to presenting graphics in PRA effectively is to select a relatively small number of appropriate messages, and to find a balance between meaningful information and overwhelming detail.*

Points to consider when developing graphics for public meetings, senior staff, and the press are presented below. Certainly, recommendations for presenting clear and informative graphics are applicable to all three forums. Practical recommendations for graphical analysis techniques and tips for successful visual displays of quantitative information are given by Tufte (1983) and Helsel and Hirsch (1993).

6.4.1 PUBLIC MEETING

For a public availability session (or meeting), care should be taken to assure that the graphics are of appropriate size and the lettering is easy to read. For example, a graphic on an 8 ½ x 11 inch sheet of paper, or a font size smaller than 18 pt in a computer presentation, may not be easily seen from the back of a large auditorium. It may be appropriate to present information using large posters, spaced so that the audience may move among them and discuss the posted results with the risk assessor or RPM. Handouts and a glossary of terms may also be used. Using slides with too much text should be avoided, since the information may be difficult to read and understand. Pre-planning and pilot testing the graphics before the presentation may be helpful in assuring that the message is accurately portrayed to the community.

Consistent with EPA's guidance on risk characterization, the CTE and RME cancer risks and noncancer hazards, and EPA's decision point should be highlighted on graphics. The discussions accompanying the graph should emphasize that these values represent risks to the average and high-end individuals, respectively, and serve as a point of reference to EPA's decision point. The distribution of risks should be characterized as representing variability among the population based on differences in exposure. Similarly, graphics that show uncertainty in risk can be described using terms such as "confidence interval", "credible interval", or plausible range. The graphics need not highlight all percentiles. Instead, selected percentiles that may inform risk management decisions (such as the 5^{th},50^{th}, 90^{th}, 95^{th}, and 99^{th} percentiles) should be the focus. Figure 6-1 shows an example of a PDF for variability in risk with an associated text box for identifying key risk percentiles.

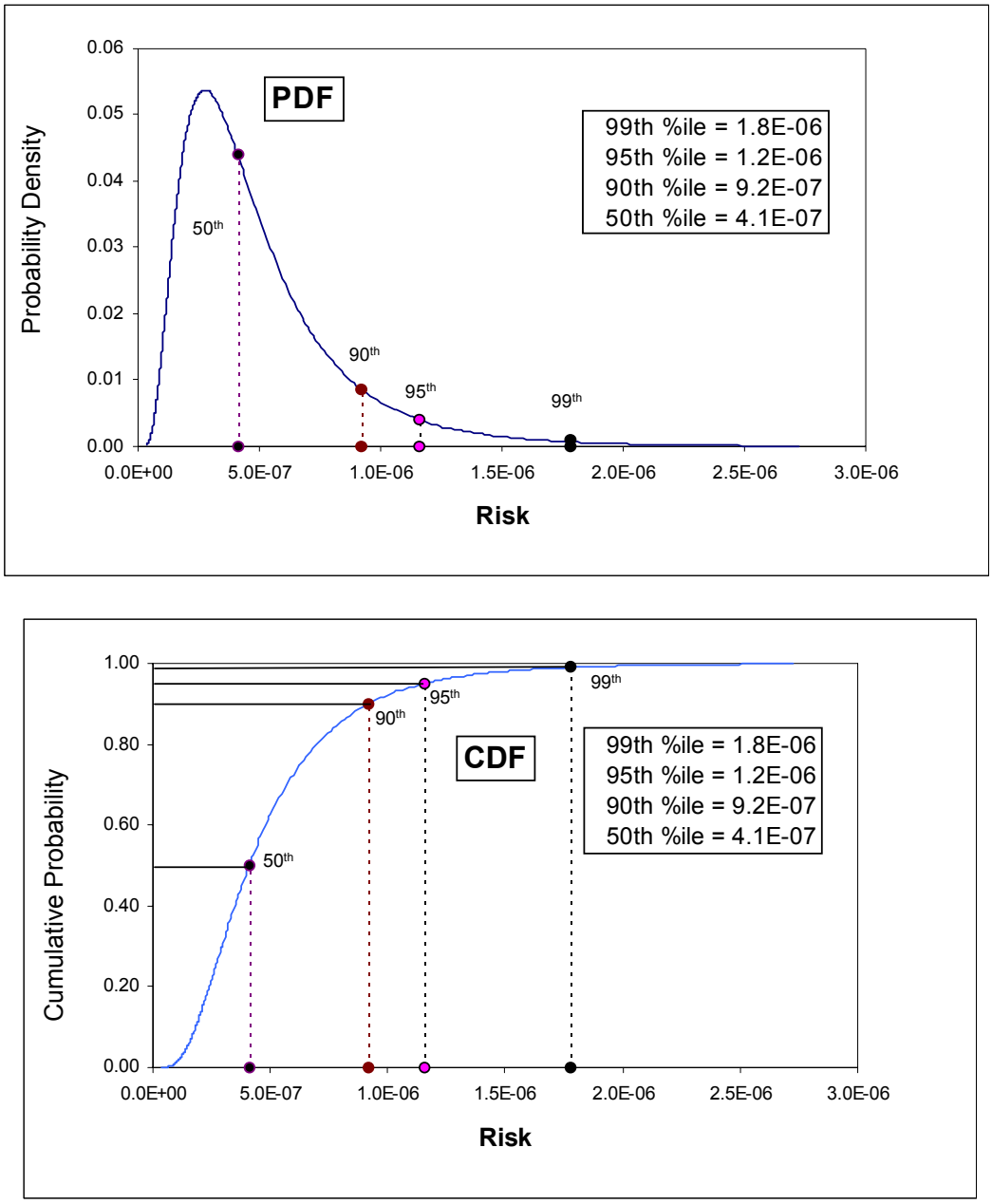


Figure 6-1. Hypothetical PRA results showing a probability density function (PDF) (top panel) for cancer risk with selected summary statistics for central tendency and high-end percentiles. This view of a distribution is useful for illustrating the shape of the distribution (e.g., slightly right-skewed) and explaining the concept of probability as the area under a curve (e.g., most of the area is below 1E-06, but there is a small chance of 2E-06). Although percentiles can also be overlayed on this graphic, a cumulative distribution function (CDF) (bottom panel) may be preferable for explaining the concept of a percentile.

Figure 6-2 gives two examples of graphics that can be used to display results of a sensitivity analysis from a Monte Carlo Analysis (MCA). While both graphics are likely to be understood by non-technical audiences, the pie chart may be more familiar. The pie chart (Figure 6-2A) suggests that the results should sum to 1.0, which may not be true if there are correlations among one or more variables, or if only a subset of the variables are displayed (e.g., those that contribute at least 1%). The available data can be normalized so that the squared correlation coefficients do sum to 100%, and this approach has been adopted by some commercial software available to run Monte Carlo simulations (e.g., *Crystal Ball*® by Decisioneering, www.decisioneering.com). The benefit of showing the squared correlation coefficient (r^2 or *r-square*, also called the coefficient of determination), rather than the correlation coefficient (r) is that *r-square* is proportional to the total variation in risk associated with specified input variable. Therefore, one can use the *r-square* to describe, in quantitative terms, the contribution of the input variable to the total variance in the risk distribution. In this example, exposure duration (ED) contributes approximately two-thirds (64%) to the total variance in risk.

A more technical graphic is the tornado plot (Figure 6-2B). In addition to showing the relative magnitude of the correlations (*r-square*), it illustrates the direction of influence a specific variable has on the final risk estimate. Bars that extend to the right indicate a positive correlation (e.g., high risk estimates correspond with high values for the variable), whereas bars that extend to the left indicate a negative correlation (e.g., high risk estimates correspond with low values for the variable.) In this example, the exposure duration (ED) has the largest positive correlation with risk, while body weight (BW) has the largest negative correlation with risk.

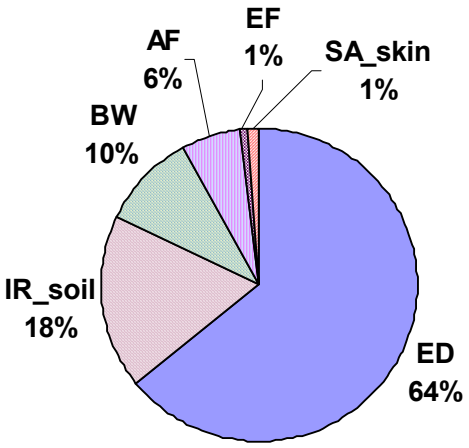
The graphics shown in this chapter are a small fraction of the graphics that might be used to communicate concepts related to PRA. Numerous additional examples are given throughout this guidance document. Table 6-1 provides a summary of cross references to other figures that were developed for this guidance document to convey specific concepts regarding variability and uncertainty.

Table 6-1. Examples of Graphics for Communicating PRA Concepts in this Guidance Document.

General PRA Topic Area	Location	Variability	Uncertainty
Conceptual Diagrams for Fundamental Concepts			
Monte Carlo Analysis	Figure 1-2	x	
Tiered process for PRA	Figure 2-1, 2-2	x	x
PDFs and CDFs			
Input variable(s)	Figure 1-1, 4-4, 4-5, 4-6	x	
Risk distribution with selected percentiles highlighted	Figure 6-1		x
Comparing RME risk (e.g., 95th percentile) with risk level of concern	Figure 1-3, 4-3, 7-2,	x	
Selecting and Fitting Probability Distributions			
Fitting distributions - frequency distribution overlaid by a PDF	Figure 3-1	x	
Lognormal probability plot	Figure 5-2	x	
Sensitivity Analysis			
Sensitivity analysis - tornado plot of Spearman rank correlations	Figure 3-6, 6-2b	x	
Sensitivity analysis - pie chart	Figure 6-2a	x	
Joint probability curve	Figure 4-8	x	
Variability in toxicity			
Species sensitivity distribution	Figure 4-7	x	
Iterative Simulations			
CDFs from multiple 1-D MCA simulations to convey uncertainty in the risk distribution	Figure 3-3		x
PRG Selection			
Estimation from best-fit line for RME risk and EPC	Figure 5-1	x	
RME risk ranges corresponding to alternative choices of PRG	Figure 7-4	x	
90% credible interval for RME risk (95th percentile) corresponding to alternative choices of PRG	Figure 7-5		x

Bi-model distribution for concentration showing pre-remediation EPC, post-remediation EPC, remediation action level, and uniform distribution for clean fill	Figure 5-3	**X**	**X**
2-D MCA Results			
Illustration of tabular and graphic outputs of a 2-D MCA	Figure 4-9		**X**
Confidence intervals (or credible intervals) on a risk distribution	Figure 1-4, 4-10, 4-11, 4-12		**X**
Box-and-whisker plot for results of 2-D MCA	Figure 3-4, 7-3		**X**
Horizontal box-and-whisker plots with multiple CDFs	Figure 6-3	**X**	**X**

A. Pie Chart



B. Tornado Plot

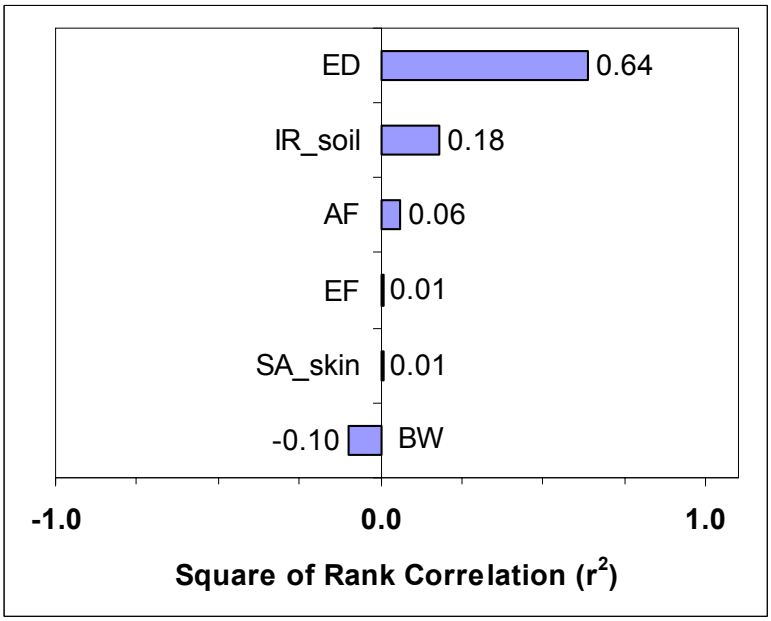


Figure 6-2. Results of a sensitivity analysis shown as a pie chart **(A)** and tornado plot **(B)**. Both graphics illustrate the concept of the relative contribution to variance for exposure variables that contribute at least 1% to the variance in risk. The pie chart suggests that the sum of the squared rank correlations equals 1.0, which is true only if the results are normalized to 100%. The tornado plot gives both the magnitude and direction (positive or negative) of the correlation. ED=exposure duration, IR_soil=soil ingestion rate, AF=absorption fraction, EF=exposure frequency, SA_skin=surface are of skin, and BW=body weight.

6.4.2 EPA SENIOR STAFF

For communicating PRA with EPA's senior risk managers (e.g., EPA Section Chiefs, EPA Branch Chiefs, or EPA Division Directors), an executive summary or executive briefing package may be appropriate. This presentation should highlight major findings, compare point estimate and probabilistic results, provide sensitivity analysis results, and state uncertainties addressed in the PRA.

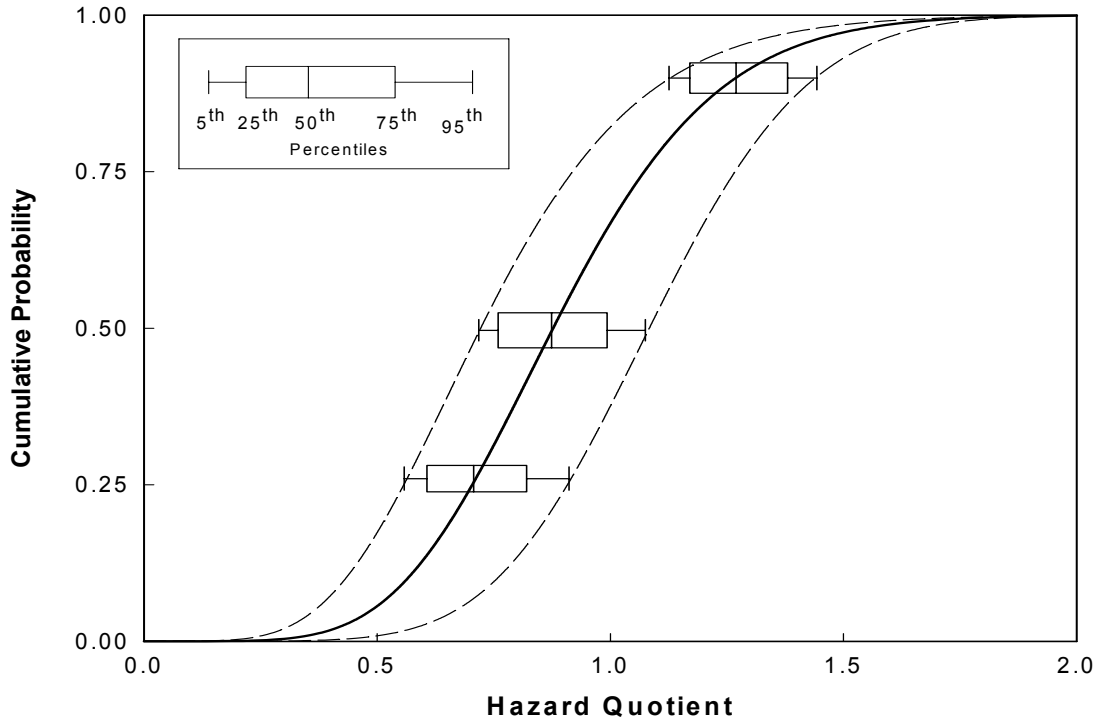


Figure 6-3. The results of a 2-D MCA. The graphic shows a method of presenting variability as a cumulative distribution function and uncertainty as box plots at the 25th, 50th, and 95th percentiles of variability. The CDF of the 50th percentile is represented by the solid line and the CDFs given by the dotted lines represent the 5th and 95th percentiles of uncertainty for each percentile of variability.

EPA senior level risk managers would generally be most interested in the risk estimates at the 50th, 90th, 95th, and 99.9th percentiles (i.e., a CTE risk estimate and the RME risk range). EPA senior managers may also wish to know the uncertainty surrounding each of the percentiles of risk. This uncertainty can be described in a table (e.g., confidence intervals around the 95th percentile risk) or a graphic (e.g., box-and-whisker plots). It is advisable for the risk assessor to have this information on hand during the briefing to respond to questions. Presenting distributions of uncertainty along with distributions of variability can create a very busy figure or table—it is best to keep things simple.

Figure 6-3 shows cumulative distribution functions (CDFs) for the Hazard Quotient (HQ) for a single chemical, representing variability in HQ. One method of displaying uncertainty is to use box-and-whisker plots. In this example, the horizontal box and whiskers represent uncertainty around selected percentile estimates of variability. Specifically, the three box-and-whisker plots correspond to the 25th, 50th, and 95th percentiles of the distribution for variability in HQ. The box shows the 25th and 75th percentiles (i.e., interquartile range) of uncertainty, whereas the whiskers show the 5th and 95th percentiles of uncertainty. In this example, uncertainty in the 95th percentile HQ is quantified by the box-and-whiskers plot in which the 5th percentile of uncertainty is 1.1, the 50th percentile is 1.3, and the 95th percentile is 1.4. This suggests that despite the uncertainty in the estimate of the 95th percentile of variability, an HQ of 1.0 is likely to be exceeded. Sometimes such results are said to describe the 90% *confidence interval* in the 95th percentile HQ. The term "confidence interval" is used loosely in this context to convey information about uncertainty; however, it is not the same as a statistical confidence limit that one might obtain by estimating a population parameter from a sample. An alternative term that may be more appropriate in this case is "credible interval".

The three curves represent similar information on uncertainty across the complete range of percentiles for variability. The solid line shows the CDF for all of the 50th percentiles of uncertainty, whereas the dotted lines show the 5th and 95th percentiles of uncertainty.

The box-and-whisker plot is simple to produce, conveys information about the symmetry and width of the confidence interval, and is easy to interpret (Tufte, 1983). In general, box-and-whisker plots are useful for summarizing results from two-dimensional Monte Carlo (2-D MCA) simulations. The methods and inferences associated with 2-D MCAs are discussed further in Appendix D. The results of a 2-D Monte Carlo simulation represent a range of possible estimates for the percentile given one or more sources of uncertainty that were included in the simulation. If the target audience for this graphic has a greater understanding of statistics, it may be less confusing if alternative phrases are used to describe the results, such as "credible interval" or "probability band".

Graphics that show probability density functions for uncertainty (PDFu's) are generally more meaningful to a technical audience of risk assessors and uncertainty analysts. Alternative graphics may be needed to communicate other sources of uncertainty in risk estimates (e.g., use of alternative probability models for exposure variables, effect of changes in the model time step, application of spatial weighting to concentration data, etc.). Additional information on communicating risks to senior EPA managers is given by Bloom et al. (1993).

The results from the sensitivity analysis may be useful to the senior managers in deciding whether additional sampling is necessary. One issue that may be important to address with risk managers and senior staff is that the width of the credible interval (e.g., 5th to 95th percentiles of uncertainty) will be determined in part by the number of sources of uncertainty that are quantified. As additional sources of uncertainty are quantified and included in the model, the interval around the risk distribution will tend to widen. This situation may appear to be counterintuitive for those managers who expect confidence to increase as uncertainty is quantified. However, by uncovering and quantifying the sources of uncertainty, the benefits in the risk communication and decision-making process should become clear. The results of the sensitivity analysis should help to focus discussions, data collection efforts, and analyses on the more significant sources of uncertainty. In addition, by developing estimates of credible intervals of uncertainty in risk estimates, the decision-making process using the tiered approach may become more transparent.

6.4.3 PRESS RELEASES

For a press briefing presentation, care should be given to identify messages and develop publication quality graphics with clear descriptions that can be provided in press packages. It is usually a good idea to provide the graphics in both color and black and white so that the press can choose the most appropriate presentation style for the story. The RPMs generally should work with the CIC, the press staff in the Communication Division, and senior managers to develop press materials. Adequate time should be left for the preparation of materials and internal Agency review and approval before information is released.

6.5 PERCEPTION OF RISK AND UNCERTAINTY

The purpose of this section is to present current thinking about how people view risk and uncertainty. This section should provide useful information for planning risk communication and addresses the first step in the seven step process (Section 6.2.2), "Identify the Audience."

There are many individual differences in the way people regard the risks and hazards that are present in modern life. These differences have their roots in the differences in perception of risk and uncertainty of the individual human mind (Slovic, 1986). The risk assessor and/or risk communicator should keep in mind the general perceptions about risk held by different groups. Communications should be tailored to the specific audience. This section summarizes some of the criteria used to judge risks in the absence of scientific data and the direction of the potential bias that may be expected by applying these criteria. Additional publications on this issue are identified in the reference section at the end of this chapter.

In the absence of scientific data, the general public evaluates risks using inferences of judgment as described below (Slovic et al., 1979):

- **Availability**: People tend to judge risks as more likely if they are easy to recall.
- **Overconfidence**: People tend to be overconfident about the judgments they make based on the use of heuristics.
- **Desire for Certainty**: People tend to misgauge risk/benefit conflicts in favor of the benefits as a result of a desire for certainty and anxiety about uncertainty.

Slovic et al. (1979) identified nine characteristics of risk that may influence perceptions. These nine dimensions may provide a perspective on whether a health risk is perceived as "more risky" or "less risky", as described in the table below.

Dimension of Risk	More Risky	Less Risky
Voluntariness	Involuntary	Voluntary
Immediacy of the effect	Delayed	Immediate
Exposed persons' knowledge about risk	Low	High
Sciences' knowledge about risk	Low	High
Control over risk	Low	High
Newness	Unfamiliar or New	Familiar
Chronic/Catastrophic	Catastrophic	Chronic
Common/Dread	Dreaded	Common
Severity of the consequences	High	Low

The presentation of uncertainty in a risk estimate can be interpreted with vastly different conclusions depending on the audience and their perceptions. For example, a thorough scientific account of multiple sources of uncertainty presented to a group of interested risk assessors and environmental scientists may be clearly understood. Such a group will likely conclude that the assumptions made in the risk assessment were appropriate and that the results can be used with confidence as a decision support tool. In contrast, a similar scientific presentation given to the community may be misunderstood, and the perceived risk may be greater. Citizens are often more concerned about the potential impact to their personal situation, than to the uncertainty in the risk estimate. Consequently, the community may react negatively to a long, highly scientific presentation on uncertainty. A good rule of thumb is to limit the presentation to no more than 15 minutes.

Focusing heavily on uncertainty may cause citizens to conclude that the risk must be high. They may also conclude that the presenter is incompetent because he or she is not sure of anything, or that the presenter is trying to hide something by cloaking the information in technical jargon, or even that the presenter is intentionally avoiding the public's issues of concern. To the extent possible, technical jargon during the presentation should be avoided or explained.

A helpful presentation generally should incorporate the following steps: (1) present information about the conclusions that can be drawn from the risk assessment; it is extremely frustrating for decision-makers to receive detailed information on uncertainty without conclusions (Chun, 1996); (2) describe the certainty of the information that supports these conclusions; (3) address the uncertainty and its implications for the conclusions; and (4) present the information without jargon and in a frank and open manner. Section 6.4 provides examples of graphics that may be useful in presentations of PRA.

6.6 TRUST AND CREDIBILITY

The single most important quality a presenter may need to possess in order to communicate to others is a sense of trust and credibility. Trust and credibility are based on working with the community and providing thoughtful, accurate responses to questions and concerns raised by the community. Building trust and credibility is important, whether communicating to a high-level technical audience, a RPM/decision-maker who wishes to have the "big picture," or the public.

Credibility can best be established through a long history of frank and open discussions with the community. In addition, a presenter can gain credibility if he or she has the ability to restate the available information so that it addresses the concerns and interests of an audience. The ability to garner trust and credibility comes from knowing the audience, respecting their opinion, and communicating at an appropriate level (U.S. EPA, 1994).

6.7 COMMUNICATION ISSUES FOR RPMS

Following the RPM's decision to conduct a site-specific PRA, the level of stakeholder involvement in the development and review of the PRA should be evaluated. Establishing the appropriate level of stakeholder involvement may include input from the CIC, risk assessor and appropriate senior managers (e.g., Section Chief, Branch Chief, etc.). The level of stakeholder involvement may vary depending on the site complexity and the interest of the community. As an initial step, it may be appropriate to conduct an exploratory session where letters are sent to various stakeholders (e.g., environmental groups, CAG, etc.) inviting their participation in a general meeting on the topic of PRA. If there is a strong interest among the stakeholders, then a more involved communication plan may be appropriate including, but not limited to the following steps:

- Providing stakeholders with an introduction to the principles of PRA in an informal session (e.g., public availability session).

- Providing a draft Scope of Work (SOW) to interested stakeholders followed shortly thereafter by an availability session to discuss comments on the document.

- Providing a period of time for the stakeholders to review and comment on the selected distributions, including an availability session for discussions with EPA staff where the community may help to identify key site-specific information such as exposure factors and receptor behavior.

- Providing the opportunity for EPA risk assessor to meet with the TAG grantee (if appropriate) and stakeholders to ask questions regarding the SOW.

- Providing a revised SOW including a response to stakeholder comments.

- Providing an overview of the final PRA at a public meeting and providing appropriate supporting PRA documents in the repositories for stakeholder review and comment. This session may be part of the general session regarding the remedial investigation when the risk assessment is discussed. Based on the complexity of the PRA, it may be appropriate to hold a public availability session where the stakeholders (including the TAG grantee), if appropriate, are able to meet with EPA staff to ask questions and offer suggestions regarding the document.

- Providing a response to comments from stakeholders regarding the PRA.

If the level of interest is low, then a less extensive CIP may be appropriate. In this case, fact sheets (in plain language) describing the general principles of PRA to the stakeholders and the key findings of the PRA may be provided (U.S. EPA, 2000a). At public meetings where the risk assessment is discussed, a short discussion of the PRA findings and their significance may be appropriate. The PRA document should be made available in the repositories for review and comment by the stakeholders.

For sites with medium interest, a combination of the activities identified above may be appropriate. For example, it may be appropriate to have a public availability session on the principles of PRA and then make the documents available for review and comment.

The RPM should consider a number of administrative issues in developing the plan for involving the stakeholders in the PRA. Issues to consider include: staff resources, funds for obtaining meeting space, availability of contractor support, significance of PRA in decision making, and the length of time required to complete the RI/FS. To aid in reducing costs, it may be appropriate to combine meetings regarding PRA and point estimate risk assessment based on the close links between the documents.

REFERENCES FOR CHAPTER 6

Bloom, D.L. et al. 1993. Communicating Risk to Senior EPA Policy Makers: A Focus Group Study. U.S. EPA Office of Air Quality Planning and Standards.

Chun, A. 1996. Strategies for Communicating Uncertainty to the Public. IBM Risk Conference Proceedings, October 31.

Helsel, D.R. and R.M. Hirsch. 1993. *Statistical Methods in Water Resources*. Elsevier Science. Amsterdam.

Slovic, P., B. Fischoff, and S. Lichtenstein. 1979. Rating the Risks. *Environment* 21(3):14–20 and 36–39.

Slovic, P. 1986. Informing and Educating the Public About Risk. *Risk Anal*. 6(4):403–415.

Tufte, E.R. 1983. The Visual Display of Quantitative Information. Graphics Press. Chesire, CT.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment)*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1–89/002. NTIS PB90-155581.

U.S. EPA. 1991a. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume I: Human Health Evaluation Manual* (HHEM), Part B, Development of Risk-Based Preliminary Remediation Goals. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333.

U.S. EPA. 1991b. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9355.0-30.

U.S. EPA. 1994. *Seven Cardinal Rules of Risk Communication*. Office of Policy Analysis. Washington, DC. EPA/OPA/87/020.

U.S. EPA. 1998. *Superfund Community Involvement Handbook and Toolkit*. Office of Emergency and Remedial Response, EPA 540-R-98-007.

U.S. EPA. 1999a. *Risk Assessment Guidance for Superfund: Volume I–Human Health Evaluation Manual*. *Supplement to Part A: Community Involvement in Superfund Risk Assessments*. EPA/540/R-98/042, March.

U.S. EPA. 1999b. *Superfund Risk Assessment and How You Can Help: An Overview*. Videotape. September 1999 (English version) and August 2000 (Spanish version). English Version: EPA-540-V-99-003, OSWER Directive No. 9285.7-29B. Spanish Version (northern Mexican): EPA-540-V-00-001, OSWER Directive No. 9285.7-40. Available through NSCEP: 800.4909.198 or 513.489.8190.

U.S. EPA. 2000a. *El Superfund Hoy Día. La Estimación de Reisgos: Cómo Lograr La participatión del la Comunidad. ¿Qué es la Estimación del Riesgo para la Salud Humana?* OSWER Directive No. 9200.2-26K. Enero. (fact sheet)

U.S. EPA. 2000b. *Superfund Risk Assessment and How You Can Help*. Videotape. (English version only). EPA-540-V-99-002, OSWER Directive No. 9285.7-29A. Available through NSCEP: 800.4909.198 or 513.489.8190, September.

U.S. EPA. 2001. *Early and Meaningful Community Involvement*. Office of Solid Waste and Emergency Response. Washington, DC. OSWER Directive No. 9230.0-99. October 12.

Supplemental References Regarding Risk Communication and Public Perception

Connelly, N.A. and B.A. Knuth. 1998. Evaluating Risk Communication: Examining Target Audience Perceptions About Four Presentation Formats for Fish Consumption Health Advisory Information. *Risk Anal*. 18:649–659.

Covello, V.T. 1987. Decision Analysis and Risk Management Decision Making: Issues and Methods. *Risk Anal*. 7(2):131–139.

Deisler, P.E. 1988. The Risk Management-Risk Assessment Interface. Last in a Five-Part Series on Cancer Risk Assessment. *Environ. Sci. Technol.* 22:15–19.

Fischhoff, B. 1995. Risk Perception and Communication Unplugged: Twenty Years of Process. *Risk Anal.* 15(2):137–145.

Fischhoff, B. 1998. Communicate unto others. *Reliab. Eng. Syst. Saf*. 59:63–72.

Fischhoff, B., A. Bostrom and M.J. Quadrel. 1997. Chapter 34. Risk Perception and Communication. In: *Oxford Textbook of Public Health*, Vol. 2, pp 987–1002. London: Oxford Univ. Press (Ed. R. Defels, et al.).

Hora, S.C. 1992. Acquisition of Expert Judgment: Examples from Risk Assessment. *J. Energy Eng.* 118(2):136–148.

Ibrekk, H. and M.G. Morgan. 1987. Graphical Communication of Uncertain Quantities to Non-Technical People. *Risk Anal*. 7:519–529.

Johnson, B.B. and P. Slovic. 1995. Presenting Uncertainty in Health Risk Assessment: Initial Studies of its Effects on Risk Perception and Trust. *Risk Anal.* 15:485–494.

Kaplan, S. 1992. 'Expert Information' Versus 'Expert Opinions.' Another Approach to the Problem of Eliciting/Combining/Using Expert Knowledge in PRA. *Reliab. Eng. Syst. Saf.* 35:61–72.

Morgan, M.G., A. Bostrom, L. Lave and C. J. Atman. 1992. Communicating Risk to the Public. *Environ. Sci. Technol.* 26(11):2048–2056.

Ohanian, E.V., J.A. Moore, J.R. Fowle, et al. Workshop Overview. 1997. Risk Characterization: A Bridge to Informed Decision Making. *Fundam. Appl. Toxicol.* 39:81–88.

Thompson, K.M. and D.L. Bloom. 2000. Communication of Risk Assessment Information to Risk Managers. *J. Risk Res.* 3(4):333–352.

CHAPTER 7

ROLE OF THE PRA IN DECISION MAKING

7.0 INTRODUCTION

When deciding whether or not to remediate a hazardous waste site, the risk manager needs to know if an unacceptable risk is present, and if so, what cleanup level to apply to the contaminated media. For this information, the risk manager should turn to the risk assessor for help in interpreting the results of the risk assessment. This chapter provides guidance on how to interpret the results of a probabilistic risk assessment (PRA) to help determine if an unacceptable risk is present, and the criteria to consider when deriving a risk-based preliminary remediation goal (PRG) and a final remedial goal.

7.1 GENERAL PRINCIPLES OF RISK-BASED DECISION MAKING IN SUPERFUND

Under Agency policy, an individual with reasonable maximum exposure (RME) will generally be the principal basis for evaluating potential human health risks at Superfund sites (see *Risk Assessment Guidance for Superfund* (Section 6.1.2 of U.S. EPA, 1989) and the National Contingency Plan's (NCP) Preamble (U.S. EPA, 1990)). The RME is defined as the highest exposure that is reasonably expected to occur at a site, and is intended to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures. In general, where cumulative carcinogenic risk to the RME individual is less than 1E-04, and the non-carcinogenic Hazard Index (HI) is less than or equal to 1, remedial action is not warranted under Superfund unless there are adverse environmental impacts, or the applicable or relevant and appropriate requirements (ARARs) are not met. As discussed in Section 7.2.4, the RME receptor is often (although not always) an appropriate basis for evaluation of risks to ecological receptors, as well.

Once a determination of unacceptable risk to humans and/or ecological receptors has been made, the risk managers will typically ask the risk assessor to develop site-specific PRGs. PRGs are generally defined as health-based chemical concentrations in an environmental media for which the risks (cancer or noncancer) to the RME receptor would not exceed some specified target level. For systemic or noncarcinogenic toxicants, the target risk level is generally a HI of unity (1). This is considered to be a threshold concentration to which the human population (including sensitive subgroups) and ecological receptors may be exposed without adverse effect during less-than-lifetime (i.e., chronic, subchronic, or short-term) exposures. For carcinogens, the target risk level used to derive the PRG typically represents a cumulative lifetime cancer risk to an individual of between 1E-06 and 1E-04 (equivalently expressed as 10^{-6} and 10^{-4}). For carcinogenic risks, less-than-lifetime exposures are converted to equivalent lifetime values (U.S. EPA, 1989). The 1E-06 risk level is specified in the NCP as a point of departure for determining remediation goals when ARARs are not available or not sufficiently protective. It is important to remember that risk-based PRGs are initial guidelines and do not represent final cleanup or remediation levels. Remediation levels are finalized after appropriate analysis in the remedial investigation/feasibility study (RI/FS) and record of decision (ROD). A final cleanup level may differ from a PRG based on the risk manager's consideration of various uncertainties in the risk estimate, the technical feasibility of achieving the PRG, and the nine criteria outlined in the NCP (see Chapter 1, Exhibit 1-2).

EXHIBIT 7-1

DEFINITIONS FOR CHAPTER 7

Applicable or Relevant and Appropriate Requirements (ARARs) - Federal or state environmental standards; the NCP states that ARARs should be considered in determining remediation goals. ARARs may be selected as site-specific cleanup levels.

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in a population, usually considered to be the mean or median of the distribution.

Confidence Interval - A range of values that are likely to include a population parameter. Confidence intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). When used to characterize uncertainty in a risk estimate, it is assumed that methods used to quantify uncertainty in the model inputs are based on statistical principles such as sampling distributions or Bayesian approaches. For example, given a randomly sampled data set, a 95% confidence interval for the mean can be estimated by deriving a sampling distribution from a Student's t distribution.

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

Hazard Index (HI) - The sum of more than one Hazard Quotient for multiple substances and/or multiple exposure pathways. The HI is calculated separately for chronic, subchronic, and shorter-duration exposures.

Hazard Quotient (HQ) - The ratio of a single substance exposure level over a specified time period (e.g., subchronic) to a reference dose (or concentration) for that substance derived from a similar exposure period.

Preliminary Remediation Goal (PRG) - Initially developed chemical concentration for an environmental medium that is expected to be protective of human health and ecosystems. PRGs may be developed based on applicable or relevant and appropriate requirements, or exposure scenarios evaluated prior to or as a result of the baseline risk assessment. (U.S. EPA, 1991a, 1991b).

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

Remedial Investigation/Feasibility Study (RI/FS) - Studies undertaken by EPA to delineate the nature and extent of contamination, to evaluate potential risk, and to develop alternatives for cleanup.

RME Range - The 90th to 99.9th percentiles of the risk distribution generated from a PRA, within which an RME risk value may be identified. The 95th percentile is generally recommended as the starting point for specifying the RME risk in a Superfund PRA.

RME Risk - The estimated risk corresponding to the reasonable maximum exposure.

7.2 INTERPRETING A RISK DISTRIBUTION

7.2.1 WHAT IS A DISTRIBUTION OF RISK AND WHAT DOES IT LOOK LIKE?

In the traditional point estimate risk assessment approach, risks to the RME individual are characterized as single point values (e.g., HI=2, or cancer risk=1E-05). In the PRA approach, the output of the risk assessment is an estimate of the distribution of risks across all members of the population. An example is shown in Figure 7-1.

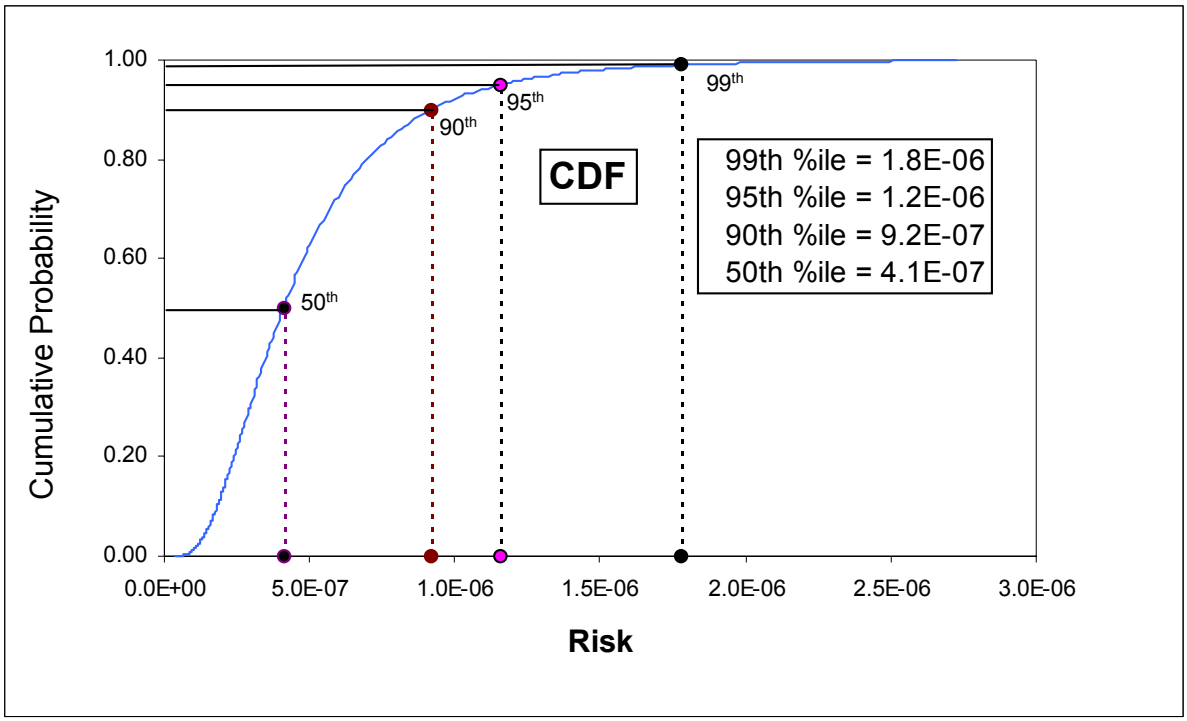


Figure 7-1. Hypothetical PRA results showing a cumulative distribution function (CDF) for lifetime excess cancer risk.

In this example, the x-axis of Figure 7-1 represents the excess lifetime cancer risk level and the y-axis represents the cumulative probability of the cancer risk level within the hypothetical population. The graph also shows various landmarks along the distribution curve such as the 50th percentile, the 90th, 95th, etc. In this illustration, the 95th percentile corresponds to a cancer risk of 1.2E-06.

7.2.2 WHAT IS THE RME RANGE?

Given a risk distribution such as shown in Figure 7-1, what part of the risk distribution should a risk manager be concerned about? As explained above, the risk to the RME receptor is a key factor in making decisions regarding the need for action at a Superfund site. EPA's *Guidelines for Exposure Assessment* (U.S. EPA, 1992) states that the "high-end" (or RME) of exposure for a population occurs between the 90th and 99.9th percentiles, with the 99.9th percentile considered a bounding estimate. Similarly, PRAs developed to support RME risk estimates for Superfund should reflect this approach.

> ☞ *In this guidance, the 90th to 99.9th percentiles of the risk distribution are collectively referred to as the **recommended RME range**.*

In utilizing PRA results to determine if an unacceptable risk is present and to develop a PRG which is sufficiently protective, risk managers should address two questions:

(1) What percentile of the risk distribution will be selected to represent the RME receptor?
(2) How will information on uncertainty in the high-end risk estimates be used in this process?

The risk manager may consider a number of factors in choosing a specific percentile to represent the RME individual. This may include both quantitative information and professional judgment. In particular, risk managers may need to understand what sources of variability and uncertainty are already explicitly accounted for by the modeling approach and inputs (i.e., point estimates and/or probability distributions) used to estimate the risk distribution, and what sources may be present but are not quantified. Approaches for selecting an appropriate percentile in human health and ecological risk assessments are described below.

7.2.3. RELATING THE RISK DISTRIBUTION TO THE RISK MANAGEMENT GOAL FOR HUMAN HEALTH

In most cases, a recommended starting point for risk management decisions regarding the RME is the 95th percentile of the risk distribution. The 95th percentile for the risk distribution is an appropriate description of high-end exposure as identified by the Presidential/Congressional Commission on Risk Assessment and Risk Management (1997).

> ☞ *In human health PRA, a recommended starting point for risk management decisions regarding the RME is the 95th percentile of the risk distribution.*

Figure 7-2 illustrates this approach for a site where cancer risks are the risk driver. Assume the risk manager has selected an excess cancer risk of 1E-05 as the risk management goal, and the 95th percentile as the definition of the RME. If line B on the graph represents a 1E-05 probability of cancer, a no-action decision may be warranted because the 95th percentile of the risk distribution is below the cancer risk level of concern. Conversely, if we were to assume that the 95th percentile is above the risk level of concern (i.e., line A on the graph represents 1E-05), remedial action may be warranted.

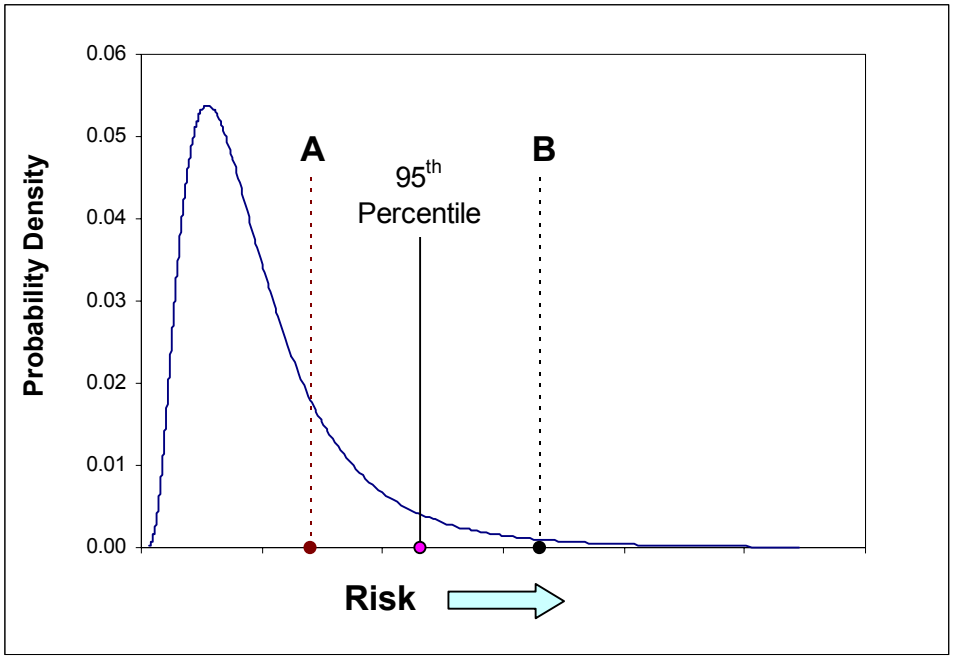


Figure 7-2. Example of a probability distribution for risk illustrating the 95th percentile and two different risk levels of concern (A and B). Assuming the 95th percentile corresponds to the RME, the need for remedial action depends on how the RME risk compares with the risk level of concern. For Case A (RME > level of concern), remedial action may be warranted. For Case B (RME < level of concern), remedial action may be unnecessary.

Although the 95th percentile is recommended as a starting point for defining the RME in the majority of human health risk assessments conducted within the Superfund program, the risk manager may use discretion in selecting a different percentile within the RME range (90th to 99.9th percentiles). In situations where the risk manager believes that a sufficient amount of site-specific information has been collected to indicate that the risk estimates are much more likely to be high (e.g., overestimated due to multiple health protective inputs), the risk manager may choose a lower percentile within the recommended RME risk range (e.g., the 90th) as the most representative of the RME estimate at the site. Conversely, when the risk manager believes that the risk estimates may tend to underestimate true risks, or if there is substantial uncertainty in the accuracy of the risk estimates, the risk manager may choose a percentile higher than the 95th in the recommended RME risk range (e.g., the 98th or the 99th). There are a variety of factors that can be considered when making this decision, such as the qualitative and quantitative uncertainty in the exposure assessment calculations, the uncertainty in the toxicity values, and the presence of biological or measured data (in contrast to modeled data). These factors are discussed below in Section 7.3. It is highly recommended that the risk manager consult with the site risk assessor when applying these factors to determine an appropriate percentile in the RME risk range.

7.2.4 RELATING THE RISK DISTRIBUTION TO THE RISK MANAGEMENT GOAL FOR ECOLOGICAL RISK ASSESSMENT

For ecological risk assessments, the choice of the percentile of the variability distribution for exposure or risk that will be protective depends on the receptor that is being considered as well as the nature of the endpoint used to establish the level of concern. For most species, the risk management objective will generally be to ensure population sustainability, even if some individual members of the population (those at the upper end of the exposure or risk distribution) may experience a higher risk of adverse effects. The risk management goal of population stability does not necessarily correspond to protection of the central tendency receptor at or below the regulatory level of concern.

As indicated in Chapter 4, without knowledge of the proportion of the local population that must survive and reproduce for the population to be stable, the choice of the central tendency exposure (CTE) receptor as the basis of the risk management goal may not be protective. Sustainability of a local population often depends upon the amount of "reserves" within that subpopulation to fill in ecological niches left voided by toxicologically impaired individuals. At a very small number of sites, a population biologist may be able to provide information about the level of effect associated with a decrease in population sustainability. At the majority of sites, the use of the CTE receptor by risk management as the basis for adequate protection of local populations of ecological receptors cannot be supported. Therefore, in the absence of such species-specific (trophic level) information, it is prudent and appropriate to base PRGs and cleanup levels on the upper end of the distribution of variability in the Hazard Quotient (HQ) to provide greater confidence that the receptor population of concern will be protected.

For threatened or endangered species, it will normally be appropriate to provide protection to as high a percentile of the distribution (i.e., the RME receptor) as is practicable (e.g., high-end of the RME range of 90^{th} to 99.9^{th} percentiles), since injury to even a single individual is undesirable.

7.3 FACTORS TO CONSIDER IN CHOOSING THE PERCENTILE FOR THE RME

Risk assessments (both point estimate and PRA) should be based on the best quality data available. A key component of the risk management process is a careful review and evaluation of the potential limitations in the quality and relevance of the data that are used in the risk assessment (i.e., qualitative and quantitative uncertainties) in order to evaluate the strengths and weaknesses of the assessment (U.S. EPA, 1993). Communication between risk managers, risk assessors, and other technical team members is vital at this stage. The main question to be answered is, "How well do the inputs to the risk assessment represent exposure pathways and behaviors at a given site?" The answer to this question can be expressed qualitatively (e.g., high, medium, or low) or quantitatively (e.g., confidence intervals or credible intervals). Some examples of these types of evaluation are illustrated below.

Use of Default Exposure Distributions

When site-specific data are not available, the best available information on some exposure parameters most likely will be from studies at other sites (e.g., in other parts of the country). In both point estimate risk assessment and PRA, the use of surrogate data to support input parameters raises questions about representativeness for both current and future land use scenarios. A specific example of potentially poor representativeness would be the use of national data for estimating the exposure frequency of adult workers when the receptor of concern is a railroad worker. Railroad workers may typically be on the site for only 100 days/year. If the risk assessment were based on the national default assumption of 250 days/year, this choice would give a high bias to the risk estimate.

Another example of a site-specific exposure factor that may vary considerably among different locations is fish ingestion rates. At sites where ingestion of fish contaminated with metals poses a concern, tissue concentrations from fish fillets collected on site are often used to determine the concentration term. However, a cultural practice of people harvesting fish on site may include consuming some of the internal organs of the fish in addition to the fillets. If the metal contaminants selectively accumulate in the internal organs instead of the fillet tissues, use of data only on fillets contaminants would give a low bias to the risk estimate.

Other Factors that Influence Site-Specific Exposures

Exhibits 7-2 and 7-3 list other types of factors that may be important to consider when evaluating the representativeness of an exposure or risk model. Given the source of the available data, the risk assessor should identify potential uncertainties and discuss the likelihood that the values used may under- or overestimate actual site-specific exposures. The risk manager should consider this information in decision making throughout the tiered process for PRA (see Chapter 2).

EXHIBIT 7-2

EXAMPLES OF DEMOGRAPHIC, CULTURAL, AND BEHAVIORAL FACTORS THAT CAN AFFECT EXPOSURE

- Subsistence fishing, hunting, or ingestion of home-grown produce
- Exposures to cultural foods or medicines that contain contaminants
- Preparation of foods in containers that contain contaminants that may leach out into food or beverage
- Hobbies and other personal practices resulting in exposure to contaminants
- Age of the population (e.g., children may have greater exposure and susceptibility than adults (U.S. EPA, 1995b, 1996)

EXHIBIT 7-3

EXAMPLES OF PHYSICAL OR GEOGRAPHICAL FACTORS THAT CAN AFFECT EXPOSURE

- Geographical features that limit or enhance accessability (e.g., slopes, valleys, mountains)
- Land use, including where exposure occurs within the exposure unit, and the current or future manner in which the receptor contacts the contaminated media
- Availability of contaminated medium for exposure (e.g., grass vs. bare soil)
- Depth of contamination (e.g., surface soil is of greatest concern for direct contact)
- Bioavailability of contaminant from media or water (e.g., physiochemical factors that enhance or reduce absorption)
- Water quality and distribution systems, including water hardness and use of lead-soldered pipes
- Temporary barriers (e.g., fences, ground cover, and concrete) that affect current (but not necessarily future) exposures

For example, the features of a potentially exposed population and the physical and geographical factors at a site can increase or decrease exposure to contaminated media. These factors should be considered in defining exposure pathways and characterizing exposure variables in the risk assessment. Such site-specific information may support a decision to evaluate the entire RME range (90^{th} to 99.9^{th} percentile) before selecting the percentile that represents RME risk. A departure from the 95^{th} percentile would depend on whether or not qualitative or quantitative factors suggest an increased or decreased exposure, and hence, risk. In practice, multiple and sometimes competing factors may need to be balanced in order to determine an appropriate percentile for the RME risk (see hypothetical example in Section 7.5).

Subpopulations may be at increased risk from chemical exposures due to increased sensitivity, behavior patterns that result in high exposures, and/or current or past exposures from other sources. Environmental health threats to children are a particular concern (U.S. EPA, 1995b, 1996). Once identified, a subgroup can be treated as a population in itself, and characterized in the same way as the larger population using similar descriptors for population and individual risk (U.S. EPA, 1995a). This principle applies to both point estimate risk assessments and PRA.

Use of Biological Data

Biological monitoring data and/or other biomarker data can be useful sources of information for evaluating uncertainty in an exposure or risk assessment. These data can provide an indication of the magnitude of current or past exposures and the degree to which the exposures are correlated with contaminated site media. Examples of biological data that are useful in human health assessments include lead in blood, trichloroethylene and its metabolites in blood or urine, arsenic or methyl parathion metabolites in urine, and polychlorinated biphenyls (PCBs) or dioxins in blood or fat tissue. Tissue burdens of contaminants are also widely useful as biomarkers of exposure in ecological risk assessments. Just as air or groundwater monitoring data can provide increased (or decreased) confidence in the results of predictive air or groundwater models, biomarkers can be used in a similar manner to evaluate how much confidence should be placed in predictive exposure assessment models. Biological data can be subject to the same shortcomings as other exposure data in terms of data quality and representativeness. The design and performance of the biological data collection effort generally should be carefully evaluated for these factors (e.g., low, medium, and high quality or confidence; low or high bias, etc.) before using the results in the risk decision. Currently, collection of biological monitoring data is limited at Superfund sites and requires coordination with appropriate agencies outside of EPA.

Issues Related to Toxicity Factors

A variety of factors may affect the magnitude of adverse responses expected to occur in similarly exposed individuals such as age, physiological status, nutritional status, and genotype. In general, these sources of inter-individual variability, and related uncertainties, are taken into account in the derivation of toxicity values (e.g., reference concentration (RfC), reference dose (RfD), and carcinogenic slope factor (CSF)) used in human health risk assessments. Thus, human health toxicity values usually are derived to be health-protective for the most sensitive populations.

> ☞ *Sources of variability or uncertainty are often accounted for in the derivation of toxicity values. The level of protectiveness afforded by the toxicity value may be an important factor in deciding on the appropriate RME risk percentile to use.*

Risk managers, in collaboration with risk assessors, should carefully consider whether the toxicity value is representative of the population of concern. For example, the toxicity value may be based on oral exposures to drinking water, whereas exposure to a site population being evaluated may be via soil ingestion. Similarly, the toxicity value could be based on effects in a healthy worker population, whereas the site population encompasses all ages and a range of individual health conditions. Uncertainty in toxicity values may reflect insufficient data to evaluate developmental toxicity concerns or to account for *in utero* exposures. Also, it may be unclear whether the population of concern has similar characteristics to the sensitive population accounted for in the derivation of the toxicity value. This determination may require coordination with a toxicologist to review the basis for the derivation of the toxicity values in question. Even then, in most cases, the determination will be very difficult, because our understanding of human variability in toxicologic responses is very limited for many chemicals. When data are insufficient to support a more quantitative representation of these sources of inter-individual variability an uncertainty factor may be used in the derivation of non-cancer human health toxicity values (RfD, RfC).

Some of the same factors that should be considered when employing toxicity values to estimate risk are also relevant to the use of toxicokinetic and toxicodynamic modeling in risk assessment. For example, a toxicity assessment for methylmercury used a technique called benchmark dose modeling (BMD) to relate the levels in maternal blood to adverse developmental effects, based on data from a large epidemiology study of Faroes Islanders (Grandjean et al., 1997; Budtz-Jørgensen et al., 2000). The RfD determined is well-supported by the other large human studies from the Seychelles (Davidson et al., 1995, 1998) and New Zealand (Kjellstrom et al., 1986, 1989) as well as a physiologically-based pharmacokinetic (PBPK) model based on the Seychelles data (Clewell et al., 1999). The RfD obtained with benchmark dose modeling (BMD) was 1E-04 mg/kg-day. The PBPK model incorporated variability in toxicokinetics to obtain a range of acceptable intakes of methylmercury between 1E-04 and 3E-04 mg/kg-day. Although the PBPK model was not used in the derivation of the benchmark dose value, it was used to support the choice of uncertainty factors in the derivation of the RfD.

At the time this guidance was finalized, the understanding of this type of toxicity information (i.e., human variability) was not well developed. Although such information was not used to characterize variability in human health risks, the estimates of variability from the PBPK model did provide additional information on uncertainty. For decision makers, the toxicity data and the choice of the endpoint (e.g, neurodevelopmental effects in the case of methylmercury) can guide qualitative risk management choices regarding the percentile representing the RME (within the 90^{th} to 99.9^{th} percentile range) and/or the appropriate level of confidence in the RME estimate. Exhibit 7-4 lists some of the issues to consider when evaluating the uncertainty in a toxicity value.

EXHIBIT 7-4

EXAMPLES OF TOXICITY CONSIDERATIONS

- How severe is the effect?
- Is the effect reversible?
- How steep is the slope of the dose-response curve at low dose?
- Is the contaminant persistent in the environment or in receptors?
- Does the contaminant bioconcentrate as it moves through the food chain?
- How bioavailable is the contaminant?

Use of Quantitative Uncertainty Estimates

PRA methods such as a two-dimensional Monte Carlo analysis (2-D MCA) may be used to quantify the uncertainty or confidence surrounding risk estimates, and this information may be helpful in selecting the RME risk percentile. Figure 7-3 provides hypothetical results of a 2-D MCA where a credible interval has been quantified for a 95th percentile of variability in noncancer HI. In exposure units (EU) 1 and 3, the credible intervals for the 95th percentile are fairly narrow, which suggests a high degree of confidence that the risks in EU1 are negligible and that the risks in EU3 are unacceptable. Conversely, the relatively wide credible intervals in EU2 and EU4 give less confidence in the results, but suggest that the 95th percentiles likely exceed a target HI of 1 in both cases. Further efforts to reduce or characterize uncertainties may affect the risk management decision in these two areas.

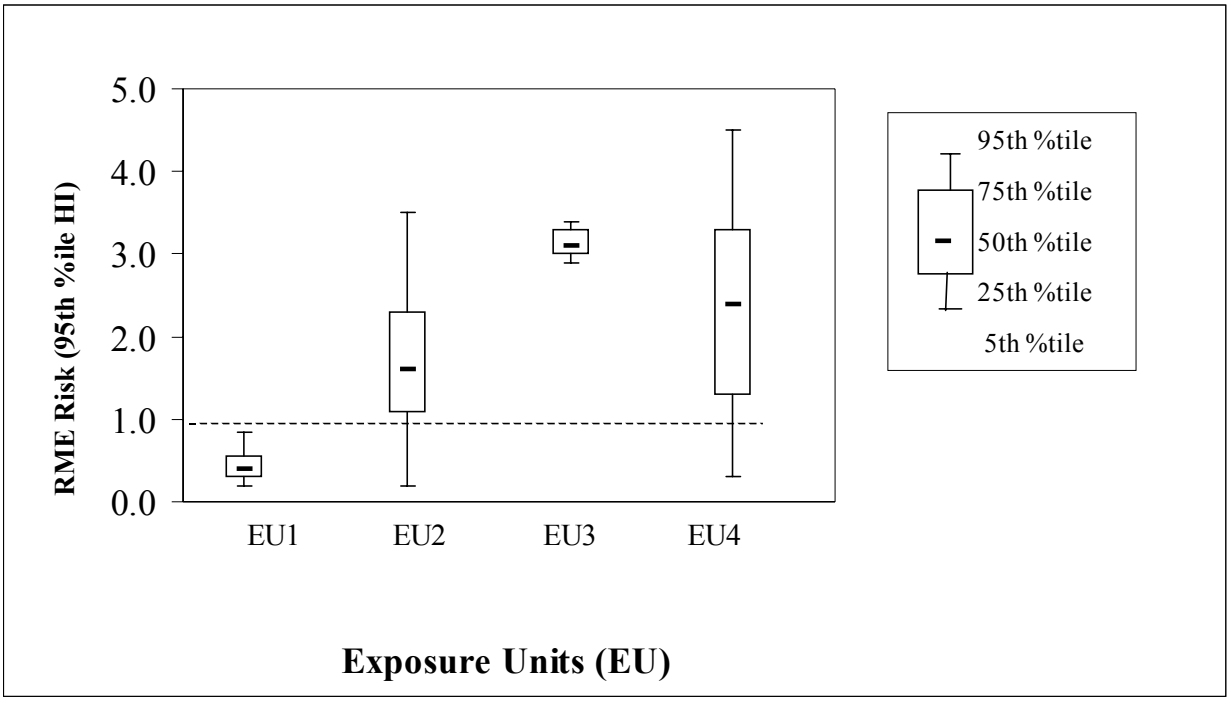


Figure 7-3. Box and whisker plots characterizing uncertainty in the RME risk estimates (95th percentile of the Hazard Index) at four locations. The box represents the inter-quartile range (25th to 75th percentiles) while the whiskers represent the 90% credible interval (5th to 95th percentiles).

Summary: Multiple Criteria Form the Basis of the Remedial Decision

Final risk management decisions should be based on a weighted consideration of all of the relevant factors that influence confidence in the risk distribution. For example, a risk manager may be presented with a risk assessment for a heavy metal in residential soil in which the distribution of cancer risk estimates in the RME range (i.e., 90th to 99.9th percentiles) overlaps the risk range of concern (1E-06 to 1E-04). The risk manager then should proceed with the site technical team to evaluate the data available to define inputs for the risk assessment, as well as the site-specific factors, and the available biological monitoring data. Assume that several factors that are likely to increase the confidence in the risk estimates were noted: (1) the soil collection and analysis effort was well-designed; (2) the predominant chemical and physical forms of the metal in the soil are characterized by relatively low bioavailability; (3) all of the yards in the residential neighborhood are covered with grass lawns, a feature generally expected to reduce direct exposure to soil; and (4) biomonitoring data from the site are all within normal physiological ranges, suggesting little, if any, excess contaminant exposure occurred at the site. In addition, generic national data were used in the absence of site-specific information on two input variables that ranked highest in the sensitivity analysis, thereby reducing confidence in the risk estimates. In this example, the consideration of these factors collectively suggests that the results of the risk assessment are likely biased towards an overestimate of risk, and this information may be used in a risk management selection of a percentile of the risk distribution to represent the RME receptor (e.g., less than or equal to the 95th percentile).

7.4 UNCERTAINTY ASSOCIATED WITH THE USE OF THE 99.9TH PERCENTILE

As previously stated, this guidance adopts the 90th to 99.9th percentiles of the risk distribution as the recommended RME risk range for decision-making purposes, consistent with EPA's *Guidelines for Exposure Assessment* (U.S. EPA, 1992). A cautionary note should be added about the selection of the higher percentiles within that range, especially the 99.9th percentile. The extreme percentiles ("tails") of an input distribution are understandably the most uncertain part of a PDF, since the number of data values in these ranges are less abundant than in the center of the range. This uncertainty in the tails of the input distributions leads in turn to greater uncertainty in the tails of the calculated exposure or risk distribution, and the magnitude of this uncertainty increases rapidly at the very high percentiles. In many cases, estimates at the extreme tails, such as the 99.9th percentile, may be neither accurate nor plausible. For that reason, great care should be taken when evaluating an RME risk in the upper percentiles of the risk range.

7.5 MOVING FROM A PRG TO A REMEDIAL GOAL

As discussed above, where an unacceptable risk is identified, the risk assessor is typically asked to develop site-specific PRGs (see Chapter 5 for discussion on derivation of PRGs). PRGs may be developed using a probabilistic approach much in the same manner as they are developed using a point estimate approach. The target risk level should be set for a specified percentile (corresponding to the RME receptor), and the concentration in contaminated media which corresponds with that target risk level should be calculated. It is important to understand that the PRG is an early step, not the last step, in the selection of a final cleanup level. During the RI/FS, the risk manager should evaluate the remedial alternatives using the nine criteria described in the NCP (U.S. EPA, 1990) (Chapter 1, Exhibit 1-2). Achieving a target level of protection for human and/or ecological receptors is one of the primary factors, but this objective should be balanced by criteria such as feasibility, permanence, state and community acceptance, and cost. Indeed, there may be times when a purely risk-based PRG may be impracticable as a final cleanup goal. In cases such as this, it is important to remember that the RME is not a single, fixed percentile on the risk distribution, but instead represents the portion of the risk distribution curve between

the 90th and 99.9th percentiles. Depending on the specific exposure and toxicity information available at a site, a PRG developed using the 90th percentile of risk may be sufficient to protect the reasonably maximum exposed individual. Alternatively, at some sites, the risk manager may feel that a PRG developed using even the 95th percentile of risk is not sufficiently protective of the RME individual and thus may choose to develop a PRG using a higher percentile.

> ☞ *Selection of final remediation or cleanup levels during the RI/FS and ROD may be an iterative process, and may consider a range of factors in addition to the initial PRG estimate.*

For example, at a former nuclear energy site, a PRG of 200 picocuries/gram (pCi/g) was developed for plutonium in soil based on a one-dimensional Monte Carlo analysis (1-D MCA) and the recommended starting point of the 95th percentile for the RME individual. At this particular site, the surrounding communities were strongly opposed to this PRG as a cleanup level. They felt it was not adequately protective, and as a result, limited progress occurred in remediating the site over the years. The communities pointed out to the risk manager that many of the exposure assumptions used in the PRA were not site-specific, and some members of the community felt that exposures occurred more often (i.e, with higher frequency) and for a longer period of time (i.e., for a greater duration) than were assumed. Based on the exposure parameters recommended by the community, the PRG would have been 75 pCi/g. At this point, the risk manager could have chosen to either go back and collect sufficient site-specific demographic and exposure data to refine the risk calculations and the PRG derivation, or evaluate the feasibility of a PRG associated with higher percentiles on the risk distribution curve (e.g., 99th percentile). In this particular example, the risk manager compared the costs associated with the cleanup that would be required to satisfy the community concerns with the costs associated with collection of additional data and recalculation of the risk and PRG. The risk manager decided that the additional cost of cleanup was manageable and expected that the PRG based on the 99th percentile would be accepted by the community. In addition, remedial activity could begin quickly without more investigation. When the risk manager presented these findings to the community, the citizens quickly agreed with this approach and remediation activities moved forward.

How does Variability and Uncertainty in Risk Relate to the Choice of a PRG?

An effective approach for communicating the results of a probabilistic analysis to risk managers is to develop graphics that relate variability and uncertainty in risk to the choice of a PRG. Two graphics are illustrated in Figures 7-4 and 7-5, based on the concept of iterative simulations presented in Chapter 5 (Section 5.5). Continuing the PRG example discussed above, assume that multiple 1-D MCA simulations are run with PRGs for plutonium ranging from 25 pCi/g to 250 pCi/g in increments of 25 pCi/g. As the concentration term is changed to correspond with a PRG, each Monte Carlo simulation yields a different distribution of risk. Figure 7-4 focuses on the RME range of percentiles from the risk distribution (i.e., 90th - 99.9th percentiles). A risk manager might use this graphic to evaluate how the PRG could change based on the choice of the percentile used to represent the RME. A hypothetical risk level of concern of 1E-05 corresponds with the 90th percentile at a PRG of approximately 125 pCi/g, whereas 1E-05 intersects the 95th percentile line at a PRG of approximately 75 pCi/g. Therefore, when variability in risk is the focus of the decision, the difference between an RME set at the 95th percentile instead of the 90th percentile is 50 pCi/g.

Figure 7-5 presents information on uncertainty, rather than variability. This graphic could be used to summarize results of a 2-D MCA (see Appendix D), or a series of 1-D MCA simulations (see

Chapter 3, Section 3.4) applied to the same range of PRGs evaluated in Figure 7-4. In this case, the results yield a 90% credible interval (CI) for the risk distribution. Figure 7-5 highlights the 90% CI for the 95th percentile, assuming that a risk manager selects the 95th percentile to represent the RME risk, and she is interested in the uncertainty in the risk estimates. Using the same hypothetical risk level of concern (1E-05), the 90% upper CI for the 95th percentile corresponds with 1E-05 at a PRG of approximately 25 pCi/g. The risk manager may need to consider the cost and feasibility of achieving a PRG as low as 25 pCi/g. In addition, the 90% lower CI corresponds to a PRG of 250 pCi/g. The risk manager may determine that this range of uncertainty (i.e., an order of magnitude) is too wide to set a PRG, and that further steps are needed to reduce identify the major sources (i.e., sensitivity analysis).

Variations on Figures 7-4 and 7-5 can be developed to focus on different percentiles of the risk range. This information, together with the results of the sensitivity analysis which highlights the major sources of variability and uncertainty, should help to guide the selection of final remediation or cleanup levels, or continued data collection and analysis following the tiered process for PRA.

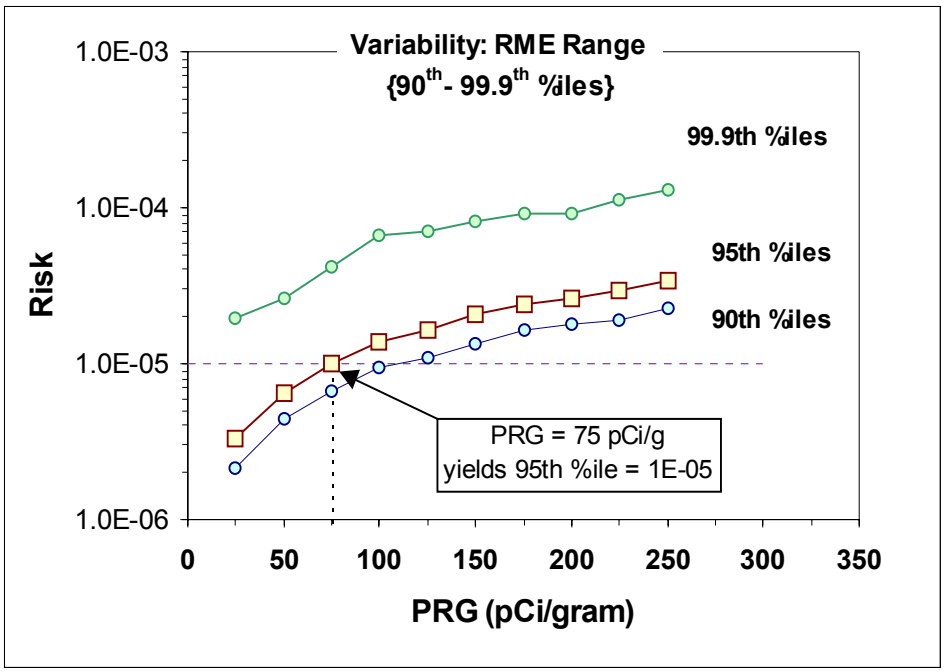


Figure 7-4. Example of graphic showing variability in risk (i.e., RME range, or 90th to 99.9th percentiles) associated with different choices of PRG for plutonium in soil (pCi/g). The hypothetical risk level of concern (1E-05) corresponds to a 90th percentile risk at a a PRG of ~ 100 pCi/g, and a 95th percentile at a PRG of ~ 75 pCi/g. In this example, all of the 99.9th percentiles exceed 1E-05, leaving no choices for PRG at the high end of the RME range.

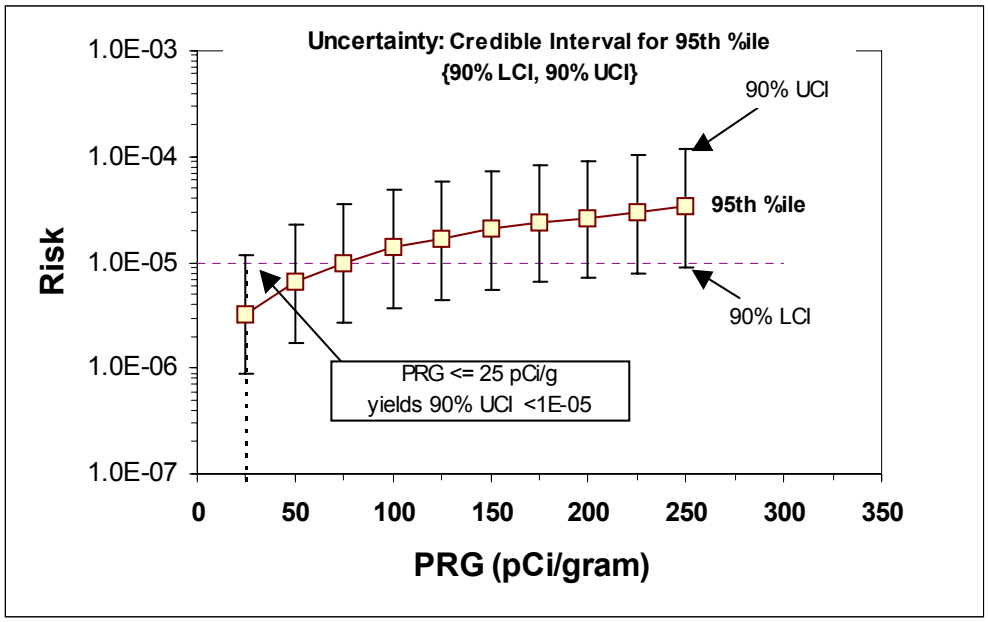


Figure 7-5. Example of graphic showing uncertainty in 95th percentile risk associated with the same choices of PRGs given in Figure 7-4. Uncertainty is given by the 90% upper and lower credible interval (CI). The hypothetical risk level of concern (1E-05) corresponds with the 90% upper CI at a PRG of ~ 25 pCi/g, and the 90% lower CI at a PRG of ~ 250 pCi/g.

REFERENCES FOR CHAPTER 7

Budtz-Jørgensen, E., P. Grandjean, N. Keiding, et al. 2000. Benchmark Dose Calculations of Methylmercury-Associated Neurobehavioral Deficits. *Toxicol. Lett.* 112–113:193–199.

Clewell, H.J., J.M. Gearhart, P.R. Gentry, et al. 1999. Evaluation of the Uncertainty in an Oral Reference Dose for Methylmercury Due to Interindividual Variability in Pharmacokinetics. *Risk Anal.* 19:547–558.

Davidson, P., G. Myers, C. Cox, et al. 1995. Longitudinal Neurodevelopmental Study of Seychellois Children Following in Utero Exposure to Methylmercury from Maternal Fish Ingestion: Outcomes at 19 and 29 Months. *NeuroToxicology* 16:677–688.

Davidson, P.W., G.J. Myers, C. Cox, et al. 1998. Effects of Prenatal and Postnatal Methylmercury Exposure from Fish Consumption on Neurodevelopment: Outcomes at 66 Months of Age in the Seychelles Child Development Study. *JAMA* 280:701–707.

Grandjean, P., P. Weihe, R. White, et al. 1997. Cognitive Deficit in 7-year-old Children with Prenatal Exposure to Methylmercury. *Neurotoxicol. Teratol.* 20:1–12.

Kjellstrom, T., P. Kennedy, S. Wallis, et al. 1986. Physical and Mental Development of Children with Prenatal Exposure to Mercury from Fish. Stage 1: Preliminary Test at Age 4. *Natl. Swed. Environ. Protec. Bd.*, Rpt 3080 (Solna, Sweden).

Kjellstrom, T., P. Kennedy, S. Wallis, et al. 1989. Physical and Mental Development of Children with Prenatal Exposure to Mercury from Fish. Stage 2: Interviews and psychological tests at age 6. *Natl. Swed. Environ. Prot. Bd.*, Rpt 3642 (Solna, Sweden).

Presidential/Congressional Commission on Risk Assessment and Risk Management. 1997. *Risk Assessment and Risk Management in Regulatory Decision Making*. Final Report, Volume 2.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991a. *Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9355.0-30.

U.S. EPA. 1991b. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume I: Human Health Evaluation Manual* (HHEM), Part B, Development of Risk-Based Preliminary Remediation Goals. Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333.

U.S. EPA. 1992. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. *57 Federal Register*, 22888-22938, May 29.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund*. Office of Solid Waste and Emergency Response. Washington, DC.

U.S. EPA. 1995a. *Memorandum from Carol Browner on Risk Characterization*. Office of the Administrator. Washington, DC. February 22.

U.S. EPA. 1995b. *Memorandum from Carol Browner on Policy on Evaluating Health Risks to Children*. Office of the Administrator. Washington, DC. October 20.

U.S. EPA. 1996. *Memorandum from Carol Browner on EPA's Report, Environmental Health Threats to Children*. Office of the Administrator. Washington, DC. September.

APPENDIX A

SENSITIVITY ANALYSIS: HOW DO WE KNOW WHAT'S IMPORTANT?

A.0 INTRODUCTION

Sensitivity analysis, as it is applied to risk assessment, is any systematic, common sense technique used to understand how risk estimates and, in particular, risk-based decisions, are dependent on variability and uncertainty in the factors contributing to risk. In short, sensitivity analysis identifies what is "driving" the risk estimates. It is used in both point estimate and probabilistic approaches to identify and rank important sources of variability as well as important sources of uncertainty. The quantitative information provided by sensitivity analysis is important for guiding the complexity of the analysis and communicating important results (see Chapter 6). As such, sensitivity analysis plays a central role in the tiered process for PRA (see Chapter 2). This Appendix focuses on a set of graphical and statistical techniques that can be used to determine which variables in the risk model contribute most to the variation in estimates of risk. This variation in risk could represent variability, uncertainty, or both, depending on the type of risk model and characterization of input variables.

There is a wide array of analytical methods that may be referred to as sensitivity analysis, some of which are very simple and intuitive. For example, a risk assessor may have two comparable studies from which to estimate a reasonable maximum exposure (RME) for childhood soil ingestion. One approach to evaluating this uncertainty would be to calculate the corresponding RME risk twice, each time using a different plausible point estimate for soil ingestion rate. Similarly, in a probabilistic model, there may be uncertainty regarding the choice of a probability distribution. For example, lognormal and gamma distributions may be equally plausible for characterizing variability in an input variable. A simple exploratory approach would be to run separate Monte Carlo simulations with each distribution in order to determine the effect that this particular source of uncertainty may have on risk estimates within the RME range (90^{th} to 99.9^{th} percentile, see Chapter 1).

Sensitivity analysis can also involve more complex mathematical and statistical techniques such as correlation and regression analysis to determine which factors in a risk model contribute most to the variance in the risk estimate. The complexity generally stems from the fact that multiple sources of variability and uncertainty are influencing a risk estimate at the same time, and sources may not act independently. An input variable contributes significantly to the output risk distribution if it is both highly variable *and* the variability propagates through the algebraic risk equation to the model output (i.e., risk). Changes to the distribution of a variable with a high sensitivity could have a profound impact on the risk estimate, whereas even large changes to the distribution of a low sensitivity variable may have a minimal impact on the final result. Information from sensitivity analysis can be important when trying to determine where to focus additional resources. The choice of technique(s) should be determined by the information needs for risk management decision making.

This appendix presents guidance on both practical decision making and theoretical concepts associated with the sensitivity analysis that are commonly applied in risk assessment. An overview of the type of information provided by sensitivity analysis is presented first, followed by guidance on how to decide what method to use in each of the tiers. A straightforward example of applications of Tier 1 and Tier 2 sensitivity analysis methods is shown, followed by a more detailed discussion of the theory and equations associated with the different methods.

EXHIBIT A-1

DEFINITIONS FOR APPENDIX A

Continuous Variables - A random variable that can assume any value within an interval of real numbers (e.g., body weight).

Correlation - A quantitative expression of the statistical association between two variables; usually represented by the Pearson correlation coefficient for linear models, and the Spearman rank correlation coefficient (see below) for nonlinear models.

Discrete Variables - A random variable that can assume any value within a finite set of values (e.g., number of visits to a site in one year) or at most a countably infinite set of values, meaning that you can count observations, but there is no defined upper limit. An example of countably infinite would be the number of dust particles in a volume of air (a Poisson distribution), whereas *uncountably* infinite would be the number of points in a line segment.

Local Sensitivity Analysis - Evaluation of the model sensitivity at some nominal points within the range of values of input variable(s).

Monte Carlo Analysis (MCA) or Monte Carlo Simulation - The process of repeatedly sampling from probability distributions to derive a distribution of outcomes. MCA is one of several techniques that may be used in PRA.

Multiple Regression Analysis - A statistical method that describes the extent, direction, and strength of the relationship between several (usually continuous) independent variables (e.g., exposure duration, ingestion rate) and a single continuous dependent variable (e.g., risk).

Nonparametric Tests - Statistical tests that do not require assumptions about the form of the population probability distribution.

Range Sensitivity Analysis - Evaluation of the model sensitivity across the entire range of values of the input variable(s).

Rank - If a set of values is sorted in ascending order (smallest to largest), the rank corresponds to the relative position of a number in the sequence. For example, the set {7, 5, 9, 12} when sorted gives the following sequence {5, 7, 9, 12} with ranks ranging from 1 to 4 (i.e., rank of 5 is 1, rank of 7 is 2, rank of 9 is 3, and rank of 12 is 4).

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis attempts to provide a ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic r that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by least-squares regression on the other variable, also called the coefficient of determination..
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Sensitivity Score - A sensitivity ratio that is weighted by some characteristic of the input variable (e.g., variance, coefficient of variation, range).
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic r that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

A.1.0 UTILITY OF SENSITIVITY ANALYSIS

As highlighted in Exhibit A-2, sensitivity analysis can provide valuable information for both risk assessors and risk management decision makers throughout the tiered process for PRA. By highlighting important sources of variability and uncertainty in the risk assessment, sensitivity analysis is generally an important component of the overall uncertainty analysis. For example, methods that quantify parameter uncertainty and model uncertainty may yield different estimates of the RME risk. This information can be used to guide the tiered process by supporting decisions to conduct additional analyses or prioritize resource allocations for additional data collection efforts. Results of sensitivity analysis can also facilitate the risk communication process by focusing discussions on the important features of the risk assessment (e.g., constraints of available data, state of knowledge, significant scientific issues, and significant policy choices that were made when alternative interpretations of data existed).

EXHIBIT A-2

UTILITY OF SENSITIVITY ANALYSIS

- **Decision making with the tiered approach-** *e.g., After quantifying parameter uncertainty, we are 95 percent confident that the RME risk is below the risk level of concern— no further analysis is needed.* Also—*selection of a beta distribution over a lognormal distribution for ingestion rate changes the 95th percentile of the risk distribution by a factor of 10—further evaluation may be needed.*
- **Resource allocation** - *e.g., Two of the 10 exposure variables contribute 90 percent of the variability in the risk estimate.*
- **Risk communication** - *e.g., For input variable X, if we were to use a distribution based on site-specific data instead of a national survey, we would expect a minimal change in the RME risk estimate.*

Decision Making with the Tiered Approach

In general, the type of information provided by a sensitivity analysis will vary with each tier of a PRA. Table A-1 provides an overview of the methods that may be applied in each tier based on the type of information needed. In Tier 1, sensitivity analysis typically involves changing one or more input variables or assumptions and evaluating the corresponding changes in the risk estimates. Ideally, the results for Tier 1 would be useful in deciding which exposure pathways, variables, and assumptions are carried forward for further consideration in subsequent tiers of analysis. By identifying the variables that are most important in determining risk, one can also decide whether point estimates, rather than probability distribution functions (PDFs), can be used with little consequence to the model output. This information is important not only for designing 1-D MCA models of variability, but also for designing more complex analyses of uncertainty discussed in Appendix D (e.g., 2-D MCA models, geostatistical analysis, Bayesian analysis). Section A.2.2 provides an overview of the Tier 1 methods and some insights regarding their limitations. Methods associated with Monte Carlo simulations used in Tiers 2 and 3 can take advantage of the ability to vary multiple inputs simultaneously and account for correlations. Sections A.2.3 and A.3 provide an overview of the sensitivity analysis methods that can be applied in a probabilistic analysis.

Table A-1. Overview of Sensitivity Analysis Methods Applicable in Tiers 1, 2, and 3 of a PRA.

Tier	Goal	SA Method(s)	What to Look For	Rationale
1	Quantify contributions of each exposure pathway to risk, identify major and minor pathways	Calculate % of total risk from each exposure pathway	Exposure pathways that contribute a very small percentage (e.g., < 5%) to total risk	Good preliminary step in Tier 1 for reducing the number of exposure variables to focus on in subsequent tiers.
			Exposure variables that appear in multiple exposure pathways	Risk estimates are likely to be more sensitive to variables that appear in multiple exposure pathways.
1	Identify the form of the dose equation for key pathways	Inspection	Equation is multiplicative or additive	SR values can be determined with minimal effort (see Table A-3). For multiplicative equations, SR=1.0 for all variables in the numerator, and SR is a function of the percent change for all variables in the denominator.
			Equation contains variables with exponents (e.g., powers, square roots)	Output is likely to be more sensitive to variables with exponents greater than 1.0.
1	Quantify contributions of each exposure variable to total risk, identify major and minor variables	Sensitivity Ratio (SR), unweighted	SR = 1.0, or SR is the same for multiple variables	It's likely that this is a multiplicative equation (see above), and the SR approach will not be effective at discriminating among relative contributions. Explore sensitivity further with other methods.
			SR ≠ 1.0	SR may vary as a function of the % change in the input variable. In this situation, it can be informative to explore small deviations (± 5%) and large deviations (min, max) in the input variables.
			SR < 1.0	Implies an inverse relationship between the input and output variables (e.g., inputs in the denominator of a risk equation).

Table A-1. Overview of Sensitivity Analysis Methods Applicable in Tiers 1, 2, and 3 of a PRA.

Tier	Goal	SA Method(s)	What to Look For	Rationale
			SR=0	Variable probably appears in both the numerator and denominator and, therefore, cancels out of the risk equation. Examples include exposure duration (ED) in noncancer risk equations, and body weight (BW) if ingestion rate is expressed as a function of body weight.
1	(cont'd) Quantify contributions of each exposure variable to total risk	Sensitivity Ratio (SR), weighted—also called Sensitivity Score	Differences in SR based on the weighting factor	A more informative approach than unweighted SR value for those variables that have sufficient information to define a weighting factor (e.g., coefficient of variation or range).
2	Quantify relative contributions of exposure pathways to risk	1-D MCA for variability or uncertainty, with outputs specifying % contribution of exposure pathways	Compare mean with high- and low-end percentiles of % contribution to risk	The % contribution of each exposure pathway will vary as a function of the variability (or uncertainty) in the inputs; exposure pathways that appear to be relatively minor contributors on average, or from Tier 1 assessment, may in fact be a major contributor to risk under certain exposure scenarios. The likelihood that a pathway is nonnegligible (e.g., > 5%) can be useful information for risk managers.
2	Quantify relative contributions of exposure variables to risk	1-D MCA for variability or uncertainty, Graphical analysis— scatterplots of inputs and output	Nonlinear relationship	Easy and intuitive approach that may identify relationships that other methods could miss. May suggest transformations of input or output variables (e.g., logarithms, power transformations) that would improve correlation and regression analyses.
		1-D MCA, Correlation Analysis using Pearson and /or Spearman Rank	Very high or low correlation coefficients Differences between relative rankings based on Pearson and Spearman	Easy to implement with commercial software; rank orders the variables based on the *average* contribution to variance. Differences in magnitude of coefficients are expected between Pearson and Spearman rank approaches, but relative order of importance is likely to be the same.

Table A-1. Overview of Sensitivity Analysis Methods Applicable in Tiers 1, 2, and 3 of a PRA.

Tier	Goal	SA Method(s)	What to Look For	Rationale
		1-D MCA, Multiple Linear Regression Analysis (e.g., stepwise)	Very high or low regression coefficients R^2 and adjusted R^2 for total model	Easy to implement with commercial software; gives contribution to reduction in residual sum of squares (RSS) For risk equations with large sets of input variables, a small subset of inputs may be able to explain the majority of the variance.
2	Quantify relative contributions of exposure variables to RME risk range	1-D MCA; same as previous step, but for subset of risk distribution (e.g., > 90^{th} percentile)	Difference in relative contributions for entire risk distribution and the RME range of the risk distribution	Variables may contribute differently to the high-end of the risk distribution, especially if the input variables are highly skewed. This situation would warrant a closer look at the assumptions regarding the estimate of the variance, differences in the upper tail (high-end percentiles) for alternative choices of probability distributions, and assumptions associated with truncation limits.
		1-D MCA, Goodness-of-fit, K-S or Chi-square; Sort output as above; perform GoF on input distribution only, comparing subset of input values corresponding with high-end risk to subset corresponding with remainder of risk distribution	GoF result—rejection of null (distributions are the same) suggests the variable may be an important contributing factor to the RME risk estimate	A second method for identifying variables that contribute differently at the high-end of the risk distribution. GoF test results should be interpreted with caution because a Monte Carlo simulation will generally yield large sample sizes (e.g., n=5,000 iterations), which is more likely to result in a positive GoF test (i.e., rejection of the null).
3	Quantify relative contributions of exposure pathways and variables to variability and uncertainty in risk	2-D MCA, same sensitivity analysis methods as Tier 2	For variability, evaluate inner loop values; for parameter uncertainty, evaluate outer loop values	The results of a sensitivity analysis depend on the question that is being asked about the risk estimate—are we interested in variability or uncertainty? The major sources of variability in risk may point to a different set of input variables than the major sources of uncertainty in risk.

Resource Allocation

Decisions regarding allocation of future resources and data collection efforts to reduce lack of knowledge generally should take into consideration the most influential input factors in the model, and the cost of gaining new information about the factors. Sensitivity analysis is a key feature of determining the expected value of information (EVOI) (see Appendix D). Once a sensitivity analysis is used to identify an input variable as being important, the source of its variability generally should be determined. If an input factor has a significant uncertainty component, further research and/or data collection can be conducted to reduce this uncertainty. Reducing major sources of uncertainty, such as the most relevant probability model for variability or the parameter estimates for the model, will generally improve confidence in the model output, such as the estimated 95th percentile of the risk distribution. An input factor may contribute little to the variability in risk, but greatly to the uncertainty in risk (e.g., the concentration term). Likewise, a variable may contribute greatly to the variability in risk, but, because the data are from a well characterized population, the uncertainty is relatively low (e.g., adult tap water ingestion rate).

An example of the output from a 2-D MCA of uncertainty and variability (see Appendix D) is shown in Figure A-1. Assume for this example that the decision makers choose the 95th percentile risk as the RME risk, and that a sensitivity analysis is run to identify and quantitatively rank the important source(s) of parameter uncertainty. The bar chart (top panel) in Figure A-1 indicates that the mean soil concentration contributes most to the uncertainty in the 95th percentile risk estimate. In addition, the mean exposure frequency is a greater source of uncertainty than the standard deviation exposure frequency. Since both the sample size and variance impact the magnitude of the confidence limits for an arithmetic mean soil concentration, one way to reduce the confidence limits (i.e., the uncertainty) would be to collect additional soil samples. As shown by the box-and-whisker plots (bottom panel) in Figure A-1, increasing the sample size (from n=25 to n=50) reduced the 90% confidence limits for the 95th percentile risk to below 1E-05, assuming the additional observations support the same estimate of the mean and standard deviation as the original sample.

Although the uncertainty in a risk estimate can be reduced by further data collection if the sensitive input distribution represents uncertainty, this is not necessarily true for input distributions that represent variability. For example, variability in the distribution of body weights can be better characterized with additional data, but the coefficient of variation (i.e., standard deviation divided by the mean) will not in general be reduced.

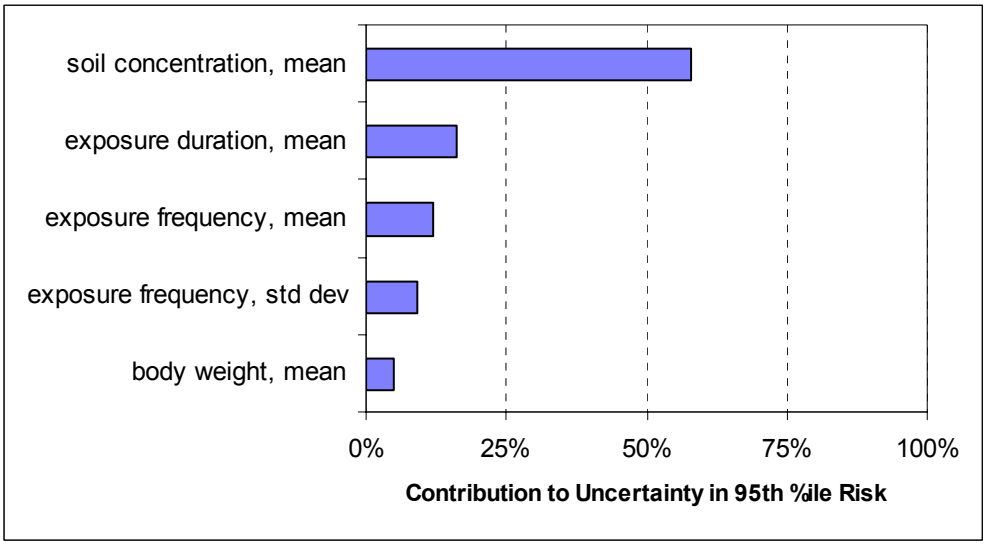
Risk Communication

Even if additional data are not collected to reduce uncertainty, identifying the exposure factors that contribute most to risk or hazard may be useful for risk communication. For example, assume that the input for exposure frequency has the strongest effect on the risk estimate for a future recreational open space. Further examination of this exposure variable reveals that the wide spread (i.e., variance) of the PDF is a result of multiple users (e.g., mountain bikers, hikers, individuals who bring picnics, etc.) of the open space who may spend very different amounts of time recreating. As a result of this analysis, the decision makers and community may decide to focus remediation efforts on protecting the high-risk subpopulation that is expected to spend the most time in the open space.

After determining which contaminants, media, and exposure pathways to carry into a PRA, numerical experiments generally should be performed to determine the sensitivity of the output to various distributions and parameter estimates that may be supported by the available information. Variables that

do not strongly affect the risk estimates may be characterized with point estimates without significantly altering the risk estimates. This guidance document does not recommend a quantitative metric or rule of thumb for determining when a variable strongly affects the output; this would generally be determined on a case-by-case basis. A qualitative or quantitative analysis may be used depending on the complexity of the risk assessment at this point. For example, incidental ingestion of soil by children is often an influential factor in determining risk from soil, a factor recognized by risk assessors. This recognition is a *de facto* informal sensitivity analysis. An array of quantitative techniques is also available, ranging from something as simple as comparing the range of possible values (i.e., maximum-minimum) for each variable, to more complex statistical methods such as multiple regression analysis. Several of these methods are discussed in more detail in this appendix.

Often, sufficient information is available to characterize a PDF for a minor variable without significant effort. This situation raises a question of whether the variable should be characterized with a point estimate or a PDF. The results of sensitivity analysis should be viewed as supplemental information, rather than an absolute rule for determining when to use a PDF. There are at least two issues to consider related to risk communication. First, the risk communication process may be facilitated by narrowing the focus of the evaluation to the key factors. More attention can be given to the discussion of key variables quantified by PDFs by describing the minor variables with point estimates. However, the decision to use a point estimate should be balanced by considering a second issue regarding perception and trust. There may be a concern that by reducing sources of variability to point estimates, there would be a reduction (however small) in the variability in risk, especially if multiple small sources of variability add up to a nonnegligible contribution. To address these concerns, it may be prudent to leave the PDFs in the calculations despite the results of a sensitivity analysis.



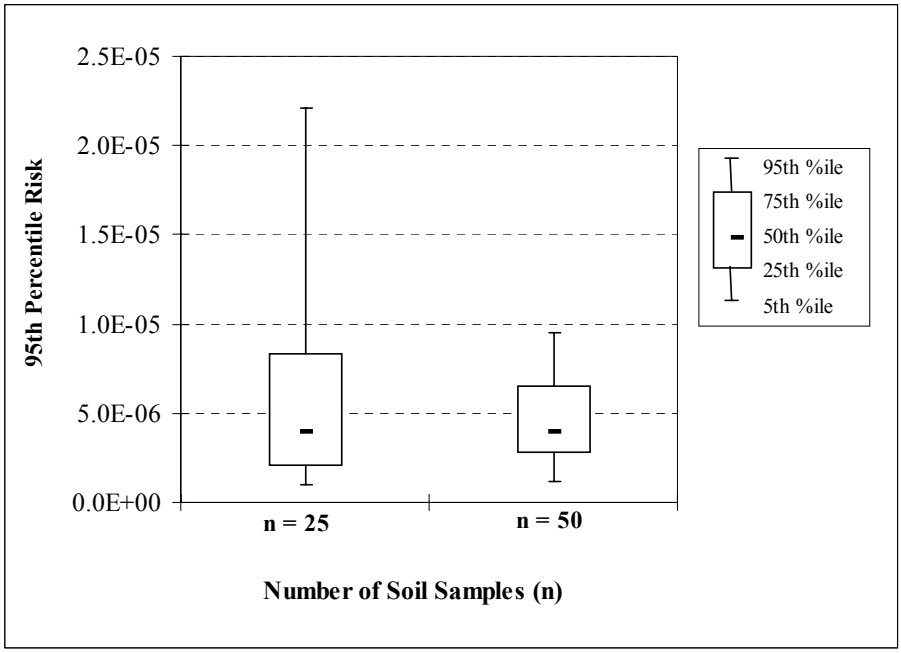


Figure A-1. Results of 2-D MCA in which parameters of input distributions describing variability are assumed to be random values. Results of a sensitivity analysis (top graph) suggest that more than 50% of the uncertainty in the 95th percentile of the risk distribution is due to uncertainty in the arithmetic mean concentration in soil. The bottom graph gives box-and-whisker plots for the 95th percentile of the risk distribution associated with Monte Carlo simulations using different sample sizes (n=25 and n=50). For this example, the whiskers represent the 5th and 95th percentiles of the distribution for uncertainty, otherwise described as the 90% confidence interval (CI). For n=25, the 90% CI is [1.0E-06, 2.2E-05]; for n=50, the 90% CI is reduced to [1.2E-06, 9.5E-06]. While increasing n did not change the 50th percentile of the uncertainty distribution, it did provide greater confidence that the 95th percentile risk is below $1x10^{-5}$.

A.2.0 COMMON METHODS OF SENSITIVITY ANALYSIS

Of the numerous approaches to sensitivity analysis that are available (see Exhibit A-3), no single approach will serve as the best analysis for all modeling efforts. Often, it will make sense to apply multiple approaches. The best choice(s) for a particular situation will depend on a number of factors, including the nature and complexity of the model and the resources available. A brief description of the more common approaches is provided in this appendix. Sensitivity analysis need not be limited to the methods discussed in this guidance, which focuses on the more common approaches. A large body of scientific literature on various other methods is available (e.g., Iman et al., 1988, 1991; Morgan and Henrion, 1990; Saltelli and Marivort, 1990; Rose et al., 1991; Merz, Small, and Fischbeck, 1992; Shevenell and Hoffman 1993; Hamby, 1994; U.S. EPA, 1997). Any method used, however, generally should be documented clearly and concisely. This includes providing all information needed by a third party to repeat the procedure and corroborate the results. Relevant information might include the following: exposure pathways and equations; a table with the input variables with point estimates, probability distributions and parameters; and tables or graphs giving the results of the sensitivity analysis and description of the method used. A hypothetical example is presented in this appendix to illustrate how to apply and present the results of selected approaches to sensitivity analysis.

EXHIBIT A-3

SOME KEY INDICES OF SENSITIVITY ANALYSIS

- Relative contribution of exposure pathways
- Inspection of risk equation
- Sensitivity ratios (i.e., elasticity)
- Sensitivity scores (i.e., weighted sensitivity ratios)
- Graphical techniques with results of Monte Carlo simulations (e.g., scatter plots)
- Correlation coefficient (or coefficient of determination, r^2) (e.g., Pearson product moment, Spearman rank)
- Normalized multiple regression coefficient
- Goodness-of-fit test for subsets of the risk distribution

Hypothetical Example of a Noncancer Risk Equation

To illustrate the application of sensitivity analysis concepts to Tier 1 and Tier 2, a hypothetical risk assessment is presented based on the general equation for Hazard Index (HI) given by Equation A-1. Note that HI is equal to the sum of the chemical-specific Hazard Quotient (HQ) values, so technically, this example reflects exposures from a single chemical.

$$HI = \frac{C_i \times I_i \times AF_i \times EF \times ED}{BW \times AT} \times \frac{1}{RfD} \qquad \text{Equation A-1}$$

The terms in Equation A-1 can be defined as follows: concentration in the i^{th} exposure medium (C_i), ingestion or inhalation rate of the i^{th} exposure medium (I_i), absorption fraction of chemical in the i^{th} exposure medium (AF_i), exposure duration (ED), exposure frequency (EF), body weight (BW), averaging time (AT=ED x 365 days/year), and reference dose (RfD).

For this example, HI is calculated as the sum of the exposures to adults from two exposure pathways: tap water ingestion and soil ingestion. Equation A-2 gives the equation for HI while Table A-2 gives the inputs for a point estimate assessment and a probabilistic assessment of variability.

$$HI = \frac{((C_w \times I_w \times AF_w) + (C_s \times I_s \times AF_s)) \times EF \times ED}{BW \times AT} \times \frac{1}{RfD} \quad \text{Equation A-2}$$

Table A-2. Point estimates and probability distributions for input variables used in the hypothetical example of HI associated with occupational exposure via water and soil ingestion.

Input Variable in Equation A-2	Point Estimate		Probability Distribution		Units
	CTE	RME	Type	Parameters	
Concentration in Water (C_w)	40	40	point estimate	40	mg/L
Tap Water Ingestion Rate (I_w)	1.3	2.0	lognormal[1]	[1.3, 0.75]	L/day
Absorption Fraction Water (AF_w)	0.30	0.50	beta[2]	[2.0, 3.0]	unitless
Concentration in Soil (C_s)	90	90	point estimate	90	mg/kg
Soil Ingestion Rate (I_s)	0.05	0.10	uniform	[0, 0.13]	kg/day
Absorption Fraction Soil (AF_s)	0.10	0.30	beta[2]	[1.22 , 4.89]	unitless
Exposure Frequency (EF)	250	350	triangular	[180, 250, 350]	days/yr
Exposure Duration (ED)	1	7	empirical[3]	see below	years
Body Weight (BW)	75	75	lognormal[1]	[74.6, 12.2]	kg
Averaging Time (AT)	365	2555	empirical[4]	ED x 365	days
RfD_{oral}[5]	0.5	0.5	point estimate	0.5	mg/kg-day

[1]Parameters of lognormal distribution are [arithmetic mean, standard deviation].
[2]Parameters of beta distribution are [alpha, beta], with range defined by min=0 and max=1.0. Parameter conversions for arithmetic mean and standard deviation are given in Table A-7.
[3]Parameters of empirical cumulative distribution function (ECDF) for ED ~ [min, max, {x}, {p}] = [0, 30, {0.08, 0.18, 0.30, 0.44, 0.61, 0.84, 1.17, 1.72, 3.1, 6.77, 14.15, 23.94}, {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.975, 0.99}], where *x* is the array of values and *p* is the array of corresponding cumulative probabilities.
[4]AT=ED x 365 for noncarcinogenic risks (Hazard Index).
[5]For simplicity, RfD_{oral} is assumed to be applicable to the ingestion of the chemical in both water and soil.

A.2.1 TIER 1 APPROACHES

Approaches for sensitivity analysis in Tier 1 of a PRA are limited to calculations that are based on changing point estimates. They are generally easy to perform and to communicate. As given by Table A-1, goals for the sensitivity analysis in Tier 1 include quantifying the relative contributions of the exposure pathways, identifying potential nonlinear relationships that may exist between input variables and the risk estimate, and rank ordering the relative contribution of exposure variables to variability or uncertainty in the risk estimate. This last goal may be the most difficult to achieve due to the limitations associated with the point estimate methodology. Methods are applied to the hypothetical example presented above (Section A.2.0) in order to demonstrate the inherent limitations of the Tier 1 approaches in some situations.

A.2.1.1 PERCENTAGE CONTRIBUTION OF EXPOSURE PATHWAYS TO TOTAL RISK

For cancer and noncancer risk assessments central tendency exposure (CTE) and RME risk is typically calculated as the sum of risks from multiple exposure pathways. Risks may be dominated by one or two exposure pathways, which can be determined through a simple calculation as shown below. The relative contributions of exposure pathways are likely to differ between the CTE risk and RME risk.

The point estimates in Table A-2 were applied to Equation A-2 to obtain CTE and RME point estimates of HI. Table A-3 gives the percent contributions of soil ingestion and tap water ingestion using Equations A-3 and A-4. Tap water ingestion contributes at least 90% to HI, and the total HI is greater than 1.0 for both CTE and RME point estimates. If 1.0 is the level of concern for HI, and a decision was made to explore variability and uncertainty in a probabilistic analysis, this result might support prioritizing the evaluation of data and assumptions associated with the tap water ingestion pathway.

Table A-3. Percent contribution of exposure pathways to HI for the example in Section A.2.

Exposure Pathway	CTE Point Estimate		RME Point Estimate	
	HI	% of total[2]	HI	% of total
Soil Ingestion	0.02	6 %	0.15	13 %
Tap Water Ingestion	0.28	94 %	1.02	87 %
Total	0.30	100 %	1.17	100 %

[1]Equation A-3: $HI_{total} = HI_{soil} + HI_{water}$
[2]Example using Equation A-4: % of total RME HI for soil ingestion = (0.15 / 1.17) x 100% = 13%.

$$HI_{total} = \sum_{i=1}^{n} HI_i \qquad \text{Equation A-3}$$

$$Percent\ Contribution_i = \frac{HI_i}{HI_{total}} \times 100\% \qquad \text{Equation A-4}$$

In this example, the choice of CTE and RME point estimates reflects an effort to explore variability in HI, rather than uncertainty. Even if the concentration terms represent the upper confidence limit on the mean (e.g., 95% UCL), the point estimates chosen to represent the CTE and RME for other exposure variables reflect assumptions about the variability in exposures. There is uncertainty that the choices actually represent the central tendency and reasonable maximum exposures. To explore this uncertainty, alternative choices for CTE and RME may have been selected. This type of exploration of uncertainty in Tier 1 may also be viewed as a form of sensitivity analysis. The percent contribution of exposure pathways could be recalculated, and the sensitivity ratio approaches discussed below may also be applied.

A.2.1.2 INSPECTION OF RISK EQUATION

For many Superfund risk assessments, risk equations can be characterized as relatively simple algebraic expressions involving addition, multiplication, and division of input variables. The term "product-quotient" model is often applied to describe equations such as Equation A-1. For these risk equations, the input variables that are likely to contribute most to the variability or uncertainty in risk can be identified by inspection. In addition, inspection of the risk equation can help to identify which sensitivity analysis methods are unlikely to reveal the relative importance of the input variables. This concept is illustrated by comparing the results of the sensitivity ratio approach (Section A.2.1.3) with the Tier 2 approaches (Section A.2.2) applied to the hypothetical example in Section A.2.0.

Some risk equations can be more complex, involving conditional probabilities, or expressions with exponents (e.g., $y=x^2$, or $y=\exp(1- x)$). In these cases, the Tier 1 sensitivity analysis methods may be effective and highlighting the variables that contribute most to the risk estimates.

A.2.1.3 SENSITIVITY RATIO (SR)

A method of sensitivity analysis applied in many different models in science, engineering, and economics is the **Sensitivity Ratio** (SR), otherwise know as the *elasticity* equation. The approach is easy to understand and apply. The ratio is equal to the percentage change in output (e.g., risk) divided by the percentage change in input for a specific input variable, as shown in Equation A-5.

$$SR = \frac{\left(\frac{Y_2 - Y_1}{Y_1}\right) \times 100\%}{\left(\frac{X_2 - X_1}{X_1}\right) \times 100\%} \qquad \text{Equation A-5}$$

where, Y_1 = the baseline value of the output variable using baseline values of input variables
Y_2 = the value of the output variable after changing the value of one input variable
X_1 = the baseline point estimate for an input variable
X_2 = the value of the input variable after changing X_1

Risk estimates are considered most sensitive to input variables that yield the highest absolute value for SR. The basis for this equation can be understood by examining the fundamental concepts associated with partial derivatives (see Section A.3.2). In fact, SR is equivalent to the normalized partial derivative (see Equation A-12).

Sensitivity ratios can generally be grouped into two categories—local SR and range SR. For the local SR method, an input variable is varied by a small amount, usually ±5% of the nominal (default) point estimate, and the corresponding change in the model output is observed. For the range sensitivity ratio method, an input variable is varied across the entire range (plausible minimum and maximum values). Usually, the results of local and range SR calculations are the same. When the results differ, the risk assessor can conclude that different exposure variables are driving risk near the high-end (i.e., extreme tails of the risk distribution) than at the central tendency region.

Demonstration of the Limitations of SR Approach

Although SR is a relatively simple and intuitive approach, it does not provide useful information under certain conditions for the more common risk equations. To demonstrate the limitations, first Equation A-5 is applied to the hypothetical example given in Section A.2.0. The results are then extended to a more general case of any of the more common risk models that involve the products of terms (i.e., multiplicative model) or the sum of terms (i.e., additive model).

Table A-4 presents an example of the local SR and range SR approach applied to the set of RME inputs given in Table A-2. For the local SR, each input was increased by 5% (i.e., Δ=+5%), while for the range SR, each input was increased by 50%. Inputs for exposure frequency were truncated at the maximum value of 365 days/year, which represents a 4.29% increase over the nominal RME value of 350 days/year.

Table A-4. Results of the Sensitivity Ratio (SR) approach applied to the hypothetical example of RME HI given in Section A.2.0. Includes *both* soil ingestion and tap water ingestion pathways.

Input Variable , X in Equation A-2[1]	Nominal RME value (X_1)	Local SR (Δ = + 5.0%)			Range SR (Δ = + 50% or max)		
		X_2	Δ in HI (%)	SR	X_2	Δ in HI (%)	SR
Tap Water Ingestion Rate, I_w (L/day)	2.0	2.1	4.35	0.87	3.0	43.5	0.87
Absorption Fraction Water, AF_w (unitless)	0.50	0.525	4.35	0.87	0.75	43.5	0.87
Soil Ingestion Rate, I_s (kg/day)	0.100	0.105	0.65	0.13	0.150	6.5	0.13
Absorption Fraction Soil, AF_s (unitless)	0.30	0.315	0.65	0.13	0.45	6.5	0.13
Exposure Frequency, EF (days/yr)	350	365[2]	4.29	1.00	365[2]	4.29	1.00
Exposure Duration, ED (years)	7	7.35	0.00	0.00	10.5	0.00	0.00
Body Weight, BW (kg)	75	78.75	- 4.46	- 0.89	112.5	- 33.33	- 0.67

[1]Only input variables that represent variability are included. Concentrations are point estimates of uncertainty. Averaging time is a function of exposure duration. RfD is a fixed point estimate.
[2]Maximum EF of 365 days/yr represents a 4.29% change in the nominal RME value of 350 days/yr.

The following observations can be made from these results:

- In decreasing order of sensitivity:

  Local SR (Δ = 5%) rankings: EF > BW > I_w = AF_w > I_s = AF_s > ED

  Range SR (Δ = 50%) rankings: EF > I_w = AF_w > BW > I_s = AF_s > ED

- EF is the most sensitive variable with an SR value of 1.0. Since EF is a variable in the numerator for both exposure pathways, this result is to be expected, as will be explained below.

- ED yields an SR=0, suggesting it does not contribute to the HI estimate. Upon closer inspection of the risk equation, it is apparent that ED occurs in the numerator of Equation A-2, as well as in the denominator (AT=ED x 365). Thus, ED effectively cancels out of the product quotient model and does not effect the estimate of HI.
- BW, the only variable in the denominator of the risk equation, is also the only variable to yield a different SR value when comparing the local and range SR approaches. Thus, BW is the only variable for which SR depends on the percent change in the input (Δ).
- BW is the only negative SR value, indicating that HI and BW are inversely related. This is true in general for any variable in the denominator of a product quotient model.
- For variables unique to the water ingestion pathway (I_w, AF_w), SR=0.87. Similarly, for variables unique to the soil ingestion pathway (I_s, AF_s), SR=0.13 These SR values are exactly the same as the percent contributions of the tap water ingestion pathway and soil ingestion pathway to HI (see Table A-3).

Since tap water ingestion is the dominant pathway (i.e., 87% of RME HI), a reasonable strategy for the Tier 1 sensitivity ratio approach might be to limit the subsequent probabilistic analysis in Tier 2 to the tap water ingestion pathway; so that input variables unique to the soil ingestion pathway would be characterized by point estimates. For this relatively simple example, this would mean that soil ingestion rate (I_s) and absorption fraction from soil (AF_s) would be described by point estimates instead of PDFs. The question to address would then become—Of the exposure variables in the tap water ingestion pathway, which ones contribute most to HI? A sensitivity ratio approach was applied to the tap water ingestion pathway to address this question. The results are presented in Table A-5.

Table A-5. Results of the Sensitivity Ratio (SR) approach applied to the hypothetical example of RME HI given in Section A.2.0. Includes *only* tap water ingestion pathway.

Input Variable , X in Equation A-2[1]	Nominal RME value (X_1)	Local SR (Δ = + 5.0%)			Range SR (Δ = + 50% or max)		
		X_2	Δ in HI (%)	SR	X_2	Δ in HI (%)	SR
Tap Water Ingestion Rate, I_w (L/day)	2.0	2.1	5.0	1.00	3.0	50	1.00
Absorption Fraction Water, AF_w (unitless)	0.50	0.525	5.0	1.00	0.75	50	1.00
Exposure Frequency, EF (days/yr)	350	365[2]	4.29	1.00	365[2]	4.29	1.00
Exposure Duration, ED (years)	7	7.35	0.00	0.00	10.5	0.00	0.00
Body Weight, BW (kg)	75	78.75	- 4.46	- 0.89	112.5	- 33.33	- 0.67

[1]Only input variables that represent variability are included. Concentrations are point estimates of uncertainty. Averaging time is a function of exposure duration. RfD is a fixed point estimate.
[2]Maximum EF of 365 days/yr represents a 4.29% change in the nominal RME value of 350 days/yr.

The following observations can be made from these results:

- In decreasing order of sensitivity:

  Local SR (Δ = 5%) rankings: I_w = AF_w = EF > BW > ED

  Range SR (Δ = 50%) rankings: I_w = AF_w = EF > BW > ED

- SR values for variables in the numerator (I_w, AF_w, and EF) are all equal to 1.0, so the SR approach suggests that they contribute equally to the HI estimate.
- BW values are the same as in Table A-4. They are negative, and the values change as a function of the percent change in the nominal RME value (Δ).

Tables A-4 and A-5 suggest that the SR approach provides essentially the same information about sensitivity as other Tier 1 methods. Specifically, inspection of the risk equation reveals that ED does not contribute to HI. In addition, for pathway-specific variables in the numerator, like ingestion rates and absorption fractions, SR values are equal to the percent contributions of the exposure pathways. This actually reflects the fact that each factor in the numerator of a multiplicative equation has an SR of 1.0.

The results of the SR approach applied to the example above can be generalized to all multiplicative and additive risk equations, as discussed below.

Generalizing the Limitations of the SR Approach

In many cases, the general equation for SR (Equation A-5) will give values that can be determined *a priori*, without doing many calculations. To understand why this is true, it is useful to simplify the algebraic expression given by Equation A-5. Let Δ equal the percentage change in the input variable, X_1. For *SR* calculations, Δ may be either positive or negative (e.g., ±5% for local SR; ±100% for range SR), and the new value for the input variable (i.e., X_2) is given by Equation A-6.

$$\begin{aligned} X_2 &= X_1 + (X_1 \times \Delta) \\ &= X_1 \times (1 + \Delta) \end{aligned} \qquad \text{Equation A-6}$$

Therefore, the denominator in Equation A-5 reduces to Δ:

$$\frac{X_2 - X_1}{X_1} = \frac{X_1(1+\Delta) - X_1}{X_1} = \frac{(1+\Delta) - 1}{1} = \Delta$$

and Equation A-5 reduces to Equation A-7:

$$SR = \frac{\mathbf{1}}{\Delta} \times \left(\frac{\mathbf{Y_2 - Y_1}}{\mathbf{Y_1}} \right) \qquad \text{Equation A-7}$$

Equation A-7 can be used to evaluate SR for different types of exposure models in which the intake equation is generally expressed as a simple algebraic combination of input variables. Solutions to SR calculations for input variables in both multiplicative and additive equations are given in Table A-6. For any such risk equation, the solution will fall into one of the five categories given by Exhibit A-4.

EXHIBIT A-4

CATEGORIES OF SOLUTIONS FOR SENSITIVITY RATIOS OF MULTIPICATIVE OR ADDITIVE EQUATIONS

Case 1 SR is a constant (e.g., 1.0). SR is independent of the choice of nominal (default) values for input variables and the choice of Δ.

Case 2 SR is a constant determined only by the nominal values for the input variables. SR is independent of the choice of Δ.

Case 3 SR is constant determined only by the choice of Δ. SR is independent of the nominal values for the input variables.

Case 4 SR is a function of both the nominal values for the input variables and the choice of Δ.

Case 5 SR is 0. The variable does not contribute to the risk estimate.

Table A-6. Examples of algebraic solutions to Sensitivity Ratio calculations for additive and multiplicative forms of risk equations.[1, 2]

Equation Type (Output = Y, Inputs = A, B, C, D)		SR_A =	SR_B =	SR_C =	SR_D =
1) Additive in Numerator	$Y = \frac{A + B}{C}$	$\frac{A}{A + B}$	$\frac{B}{A + B}$	$-\frac{1}{1 + \Delta}$	NA[3]
2) Additive in Denominator	$Y = \frac{A}{C + D}$	1.0	NA	$-\frac{C}{C(1 + \Delta) + D}$	$-\frac{D}{D(1 + \Delta) + C}$
3) Multiplicative in Numerator	$Y = \frac{A \times B}{C}$	1.0	1.0	$-\frac{1}{1 + \Delta}$	NA
4) Multiplicative in Denominator	$Y = \frac{A}{C \times D}$	1.0	NA	$-\frac{1}{1 + \Delta}$	$-\frac{1}{1 + \Delta}$

[1]Sensitivity Ratio for input variable A for an equation that is additive in the numerator: $SR_A = A / (A + B)$.
[2]Δ=% change in input variable. For example, Δ for $C=[(C_2 - C_1)/C_1] \times 100\%$, where C_1=the original point estimate and C_2=the modified point estimate. Similarly, $C_2 = C_1 (1 + \Delta)$.
[3]NA=not applicable because the variable is not in the equation.

The following observations can be made for the four forms of the risk equation, based on one of the five cases described in Exhibit A-4:

(1) Additive in Numerator

- **Case 2**: SR values for variables in the numerator depend exclusively on the nominal point estimates for all variables in the numerator. The values are independent of the choice of percent change in the inputs (Δ).
- **Case 3**: SR values for variables in the denominator depend exclusively on Δ, and are negative (i.e., inversely related to the output). Also, the lower the choice for Δ, the higher the resulting SR values. Therefore, SR is somewhat arbitrary, especially for the range SR approach since input variables may have different plausible minimum and maximum values.

(2) Additive in Denominator

- **Case 1**: SR values for variables in the numerator are always equal to 1.0. Since they are independent of the nominal values and Δ, there is no way to distinguish the relative contributions to the output.
- **Case 4**: SR values for variables in the denominator are a function of both the nominal values of variables in the denominator and Δ.

(3) Multiplicative in Numerator and (4) *Multiplicative in Denominator*

- **Case 1**: SR values for variables in the numerator are always equal to 1.0. Since they are independent of the nominal values and Δ, there is no way to distinguish the relative contributions to the output.
- **Case 3**: SR values for variables in the denominator depend exclusively on the Δ, and are negative (i.e., inversely related to the output). Also, the lower the choice for Δ, the higher the resulting SR values. Therefore, SR is somewhat arbitrary, especially for range SR since input variables may have different plausible minimum and maximum values.

These generalized results highlight a major limitation in the use of the SR approach for obtaining information from sensitivity analysis. For simple exposure models in which the relationship between exposure and risk is linear (e.g., multiplicative), the ratio offers little information regarding the relative contributions of each input variable to the risk estimate. In many cases, all of the input variables will have the same constant, either equal to 1.0 (in the case of a single exposure pathway) or equal to the relative contributions of the exposure pathways. For more complex models that combine additive, multiplicative, and nonlinear relationships between inputs and outputs (e.g., environmental fate and transport models, pharmacokinetic models), the ratio is likely to be an effective screening tool for identifying potentially influential input variables and assumptions.

Another difficulty with the SR approach is that it generally requires an assumption that the input variables are independent. Two variables may actually be positively correlated (e.g., high values of X_1 correspond with high values of X_2) or negatively correlated (e.g., high values of X_1 correspond with low values of X_2). If input variables are correlated, holding the value for one variable fixed while allowing the other to vary may produce misleading results, especially with the range sensitivity ratio approach. For example, it may not be realistic to hold body weight fixed at a central tendency while allowing skin surface area to vary from the minimum to maximum values. An improvement over the sensitivity ratio approach would be to allow correlated input variables to vary simultaneously.

A.2.1.4 SENSITIVITY SCORE

A variation on the sensitivity ratio approach may provide more information from a Tier 1 sensitivity analysis, but it requires that additional information be available for the input variables. The *sensitivity score* is the SR weighted by a normalized measure of the variability in an input variable (U.S. EPA, 1999). Examples of normalized measures of variability include the coefficient of variation (i.e., standard deviation divided by the mean) and the normalized range (i.e., range divided by the mean), as given by Equation A-8.

$$Sensitivity\ Score = SR \times \frac{\sigma}{\mu} \quad or \quad SR \times \frac{(\max - \min)}{\mu} \qquad \text{Equation A-8}$$

By normalizing the measure of variability (i.e., dividing by the mean), this method effectively weights the ratios in a manner that is independent of the units of the input variable, and provides a more robust method of ranking contributions to the risk estimates than the SR alone. This approach does require that the coefficient of variation or range can be calculated for each variable. Tables A-7 and A-8 present the results of the sensitivity scores based on the CV applied to the hypothetical example from Section A.2.0.

Table A-7. Calculation of coefficient of variation (CV = SD / Mean) for the hypothetical example of RME HI given in Section A.2.0.

Input Variable , X in Equation A-2[1]	Probability Distribution[2]	Mean[3]	SD[3]	CV = SD/Mean
Tap Water Ingestion Rate, I_w (L/day)	lognormal (1.3, 0.75)	1.3	0.75	0.58
Absorption Fraction, Water, AF_w (unitless)	beta (2.0, 3.0)	0.4	0.2	0.50
Soil Ingestion Rate, I_s (kg/day)	uniform (0, 0.13)	0.065	0.038	0.58[2]
Absorption Fraction, Soil, AF_s (unitless)	beta (1.22, 4.89)	0.20	0.15	0.75
Exposure Frequency, EF (days/yr)	triangular (180, 250, 350)	260	35	0.13[3]
Exposure Duration, ED (years)	empirical CDF (see Table A-2 for parameters)	1.75	3.86	2.21
Body Weight, BW (kg)	lognormal (74.6, 12.2)	74.6	12.2	0.16

[1]Only input variables that represent variability are included. Concentrations are point estimates of uncertainty. Averaging time is a function of exposure duration. RfD is a fixed point estimate.
[2]Beta (a, b): mean=a / (a+b) and SD = ((a x b) / [(a + b)^2 x (a+b+1)])^0.5)
Uniform (min, max): mean = (min + max)/2 and SD = ((1/12)^0.5) x (max - min) = 0.289 x (max - min)
Triangular (min, mode, max): mean = (min + mode + max)/3 and SD = (1/18) x (min^2 + mode^2 + max^2 - min x max - min x mode - mode x max)
Empirical CDF ({x}, {p}): mean and SD were estimated by Monte Carlo simulation.
[3]Mean=arithmetic mean; SD=arithmetic standard deviation

Table A-8. Results of the Sensitivity Score (Score) approach applied to the hypothetical example of RME HI given in Section A.2.0. Calculations for Sensitivity Ratio (SR) and Coefficient of Variation (CV) are given in Table A-4 and Table A-7, respectively.

Input Variable , X in Equation A-2[1]	Nominal RME value (X_1)	CV (Table A-7)	Local SR (Δ = + 5%)		Range SR (Δ = + 50%)	
			SR (Table A-4)	Score[2]	SR (Table A-4)	Score[2]
Tap Water Ingestion Rate, I_w (L/day)	2.0	0.58	0.87	**0.50**	0.87	**0.50**
Absorption Fraction, Water, AF_w (unitless)	0.50	0.50	0.87	**0.44**	0.87	**0.44**
Soil Ingestion Rate, I_s (kg/day)	0.100	0.58	0.13	**0.06**	0.13	**0.06**
Absorption Fraction, Soil, AF_s (unitless)	0.30	0.75	0.13	**0.10**	0.13	**0.10**
Exposure Frequency, EF (days/yr)	350	0.13	1.00	**0.13**	1.00	**0.13**
Exposure Duration, ED (years)	7	2.21	0.00	**0**	0.00	**0**
Body Weight, BW (kg)	75	0.16	- 0.89	**- 0.14**	- 0.67	**- 0.11**

[1]Only input variables that represent variability are included. Concentrations are point estimates of uncertainty. Averaging time is a function of exposure duration. RfD is a fixed point estimate.
[2]Score=SR x CV (see Equation A-8)

The following observations can be made from these results:

- In decreasing order of sensitivity:

  - Score based on local SR (Δ = 5%): I_w > AF_w > BW > EF > AF_s > IR_s > ED

  - Score based on range SR (Δ = 50%): I_w > AF_w > EF > BW > AF_s > IR_s > ED

- Compared with the SR approach alone in which sensitivity can only be expressed for exposure pathways, the sensitivity score approach provides a measure of sensitivity for exposure variables within each exposure pathway.
- Although ED has the highest CV, it continues to have no contribution to the HI.
  - If Tier 1 sensitivity analysis is based on the sensitivity score, the highest ranked variables are generally those with the highest CV in the exposure pathway that contributes the most to the total risk (HI). For this hypothetical example, I_w and AF_w are the two highest ranked variables.

A.2.2 TIER 2 APPROACHES

Approaches for sensitivity analysis in Tier 2 of a PRA utilize the results of Monte Carlo simulations, which allows multiple input variables to vary simultaneously. The methods are relatively simple to perform with spreadsheets or commercial statistical software. The results are generally easy to communicate, although the details of the methodology are more complex than Tier 1 approaches. As given by Table A-1, goals for the sensitivity analysis in Tier 2 are the same as Tier 1:quantifying the relative contributions of the exposure pathways, identifying potential nonlinear relationships that may exist between input variables and the risk estimate, and rank ordering the relative contribution of exposure variables to variability or uncertainty in the risk estimate. In addition, since the output is a distribution, Tier 2 sensitivity analysis methods can also utilize graphical techniques to observe nonlinear relationships, as well as evaluate potential changes in relative importance of variables and assumptions for risks in the RME risk range. Methods are applied to the hypothetical example presented in Section A.2.0 in order to demonstrate the advantages over the Tier 1 methods.

A.2.2.1 GRAPHICAL TECHNIQUES

Simple scatter plots of the simulated input and output (e.g., risk vs. exposure frequency, or risk vs. arithmetic mean soil concentration) can be used to qualitatively and quantitatively evaluate influential variables. A "tight" best-fit line through the scatter plot, as indicated by the magnitude of the r^2, suggests that a variable may significantly influence the variance in risk. Hypothetical scatter plots used to identify sensitive and insensitive variables are shown in Figure A-2. Another method for visualizing the relationship between all of the inputs and outputs is to generate a scatterplot matrix (Helsel and Hirsch, 1992). This graphic shows both histograms and scatter plots for all variables on the same page.

Figure A-3 illustrates scatter plots for the 1-D MCA simulations associated with the example from Section A.2.0. Based on the r^2 values (i.e., coefficient of determination for simple linear regression analysis), the relationship between HI and I_w is very strong ($r^2 = 0.47$) while the relationship between HI and I_s is very weak ($r^2 < 0.01$), suggesting that HI is more sensitive to variability in I_w than I_s.

A.2.2.2 CORRELATION COEFFICIENTS

The variance in a risk estimate from a Monte Carlo simulation is due to the variance in the probability distributions used in the risk equation. It is commonly said that a Monte Carlo model propagates sources of variability simultaneously in a risk equation. Numerous statistical techniques, known collectively as correlation analysis and regression analysis, can be applied to a linear equation to estimate the relative change in the output of a Monte Carlo simulation based on changes in the input variables. Examples of metrics of sensitivity include the simple correlation coefficient, the rank correlation coefficient, and a variety of coefficients from multiple regression techniques. The underlying assumptions associated with these approaches are discussed in greater detail in Section A.3. As explained in Section A.3.3.1, correlation coefficients and regression coefficients are based on different interpretations of the input variables, but they can be calculated with similar equations.

When the output distribution is compared with the distribution for one input variable at a time, two of the more common approaches are to calculate the Pearson product moment correlation and the Spearman rank correlation. Correlation analysis with one input variable will generally yield reasonable results when the input variables are sampled independently in a Monte Carlo simulation. Some statistical packages offer the correlation coefficient as an index of sensitivity, so it is important to identify which

coefficient is being calculated. *Crystal Ball*® and *@Risk* can be used to calculate the Spearman rank correlation, which tends to be more robust when the relationships between inputs and outputs are nonlinear. If the relationships are linear, such as with the product quotient models presented in this appendix, the two metrics of correlation will yield similar rankings of input variables. Rank correlation coefficients shown in *Crystal Ball*® and *@Risk* are calculated by the standard method provided in most statistics texts. *Crystal Ball*® also indicates that sensitivity can be determined as contribution to variance. This is not the relative partial sum of squares techniques discussed in Section A.3.3.2 (Equation A-19). Instead, *Crystal Ball*® calculates the contribution to the variance by squaring the rank correlation coefficients and normalizing them to 100%. Many other commonly used commercial software packages will perform Spearman rank correlation. Pearson product moment correlations (r) can be calculated in Microsoft Excel using the trendline feature in a scatter plot chart, or by using the function *Correl(X array, Y array)*, where *X array* corresponds with the Monte Carlo simulation of an input variable, and *Y array* corresponds with the output of the simulation.

Figure A-4 illustrates results of the correlation analysis for the 1-D MCA simulations associated with the example from Section A.2.0. The graphics were generated using *Crystal Ball*® 2000. The results are summarized in Table A-9. If the model output variable (e.g., HI) and input variable are highly correlated, it means that the output is sensitive to that input variable. By squaring the coefficient, the results can be expressed in terms of the percentage contribution to variance in the output (Figure A-4, top panel). To determine if the correlation is positive or negative, the correlation coefficient should not be squared (Figure A-4, bottom panel). For risk equations, in general, variables in the numerator of the equation (ingestion rate, absorption fraction, exposure frequency, etc.) will tend to be positively correlated with risk, while variables in the denominator (body weight) will tend to be negatively correlated with risk. The greater the absolute value of the correlation coefficient, the stronger the relationship.

Table A-9. Results of Tier 2 sensitivity analyses applied to hypothetical example in Section A.2.0: Pearson product moment correlations and Spearman rank correlations.[1]

Exposure Variable	Product Moment Correlation		Spearman Rank Correlation[2]		
	r	r^2 x 100%	r	r^2 x 100%	normalized r^2 x 100%
Tap Water Ingestion Rate, I_w (L/day)	0.644	41.4	0.603	36.3	39.5
Absorption Fraction Water, AF_w (unitless)	0.583	34.0	0.666	44.4	48.3
Body Weight, BW (kg)	- 0.216	4.7	- 0.229	5.2	5.7
Exposure Frequency, EF (days/yr)	0.174	3.0	0.167	2.8	3.0
Absorption Fraction Soil, AF_s (unitless)	0.109	1.2	0.149	2.2	2.4
Soil Ingestion Rate, I_s (g/day)	0.061	0.4	0.099	1.0	1.1
Exposure Duration, ED (years)	0.010	0.0	0.010	0.0	0.0

[1]Monte Carlo simulation using *Crystal Ball® 2000*, Latin Hypercube sampling, and 5000 iterations.
[2]*Crystal Ball® 2000* output includes Spearman rank correlations, *r*, and *normalized* r^2 values, calculated by dividing each r^2 value by the sum of all the r^2 values (i.e., 0.920 in this example). Figure A-4 illustrates the *r* and normalized r^2 values for the Spearman rank correlation analysis.

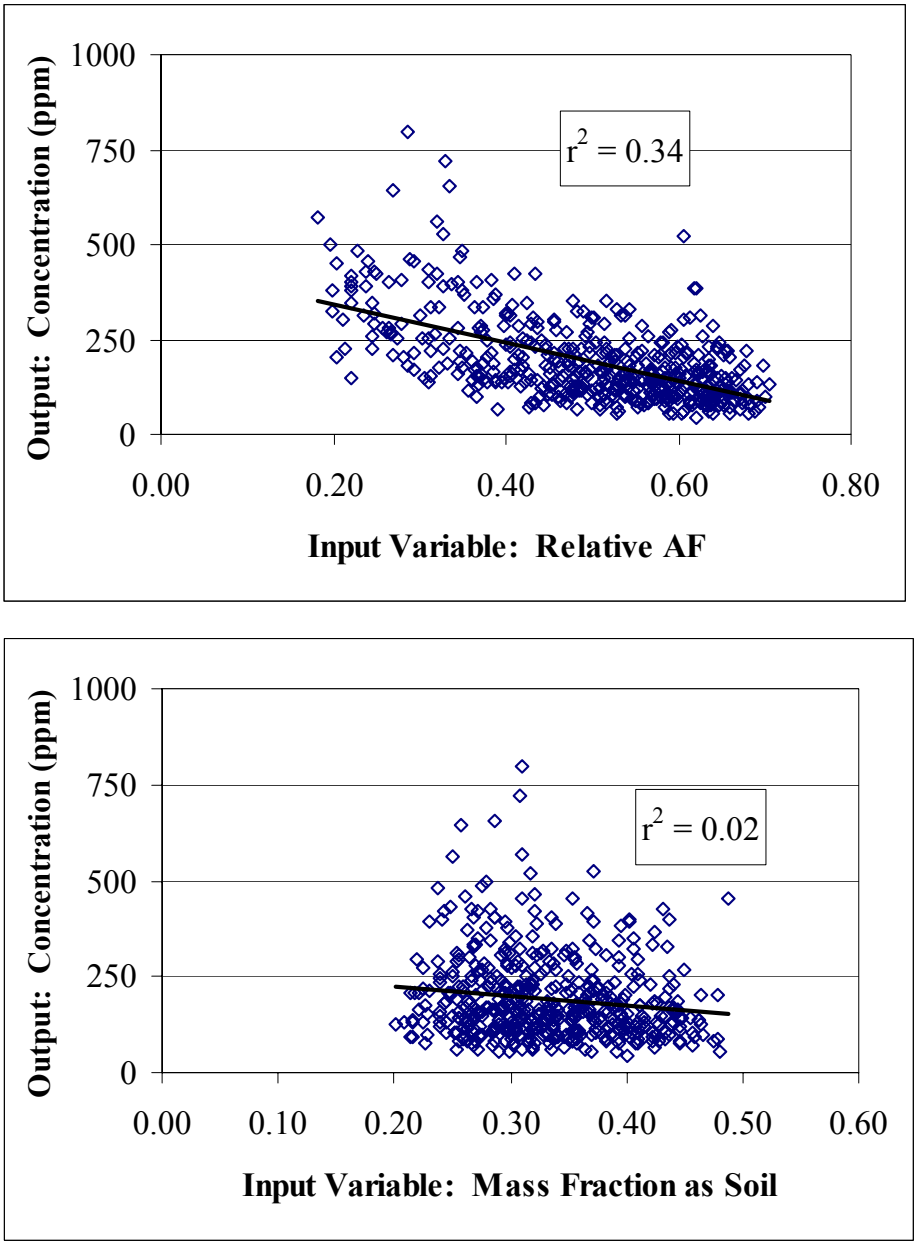


Figure A-2. Scatterplots of simulated random values from a 1-D MCA of variability. The output from the model is a contaminant concentration in soil (C) that corresponds with a prescribed (fixed) level of risk for a hypothetical population (based on Stern, 1994). For each iteration of a 1-D MCA simulation, random values were simultaneously selected for all model variables and the corresponding concentration (C) was calculated. Inputs were simulated as independent random variables. Scatterplots of 500 consecutive random values and estimates of C are shown for two input variables: relative absorption fraction, RAF (top graph); and mass fraction of dust as soil, F (bottom graph). There is a moderate, indirect relationship between C and RAF (r^2=0.34), compared with the weak relationship between C and F (r^2=0.02), suggesting that the model output (C) is more sensitive to variability in RAF than F.

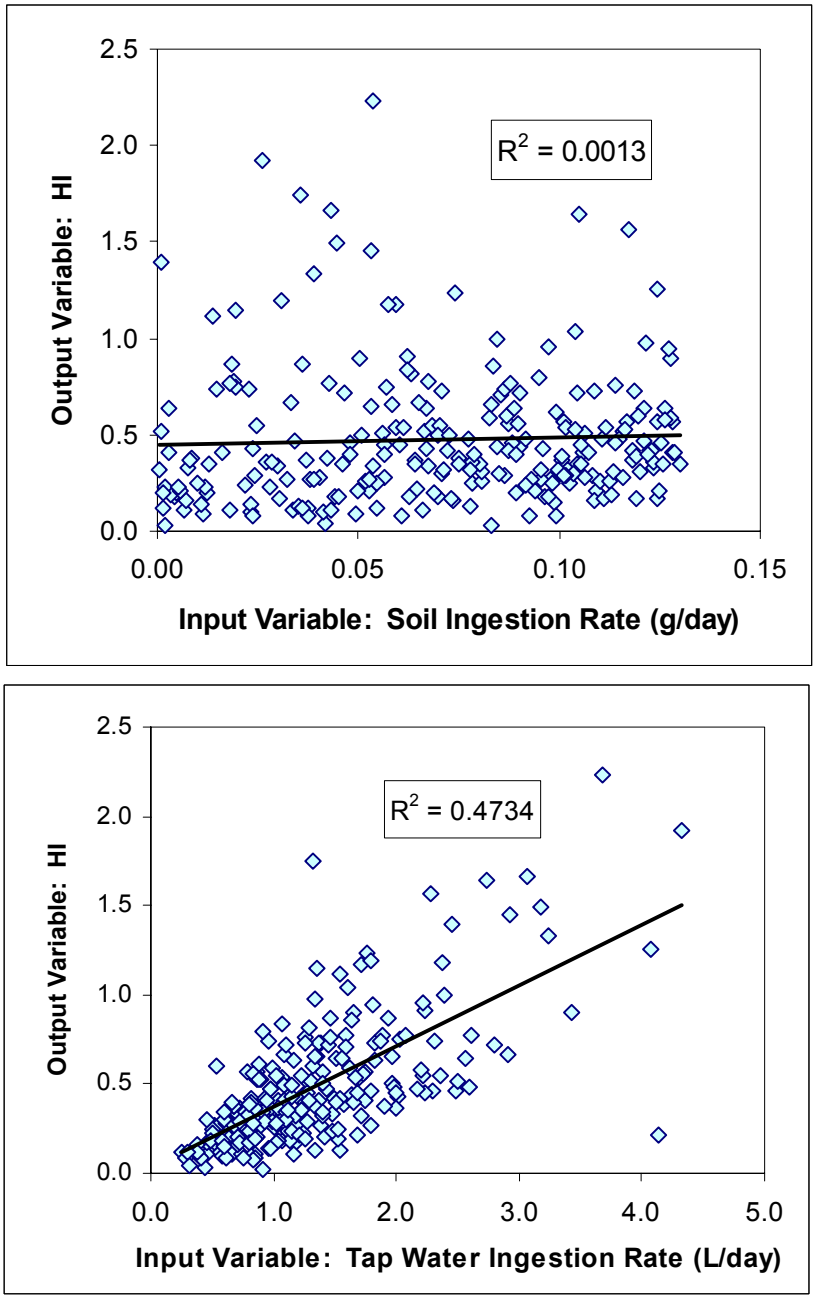


Figure A-3. Scatterplots of simulated random values from a 1-D MCA of variability for example in Section A.2.0. The output from the model is HI. For each iteration of a 1-D MCA simulation, random values were simultaneously selected for all model variables and the corresponding HI was calculated. Inputs were simulated as independent random variables. Scatterplots of 250 consecutive random values and estimates of HI are shown for two input variables: soil ingestion rate, I_s (top graph); and tap water ingestion rate, I_w (bottom graph). There is a negligible relationship between HI and I_s ($r^2 < 0.01$), compared with the strong relationship between HI and I_w (r^2=0.47), suggesting that the model output (HI) is more sensitive to variability in I_w than I_s. Best-fit lines were generated with the Simple Linear Regression in Microsoft Excel's trendline option for scatterplots; r^2 values represent the coefficient of determination (see Section A.3).

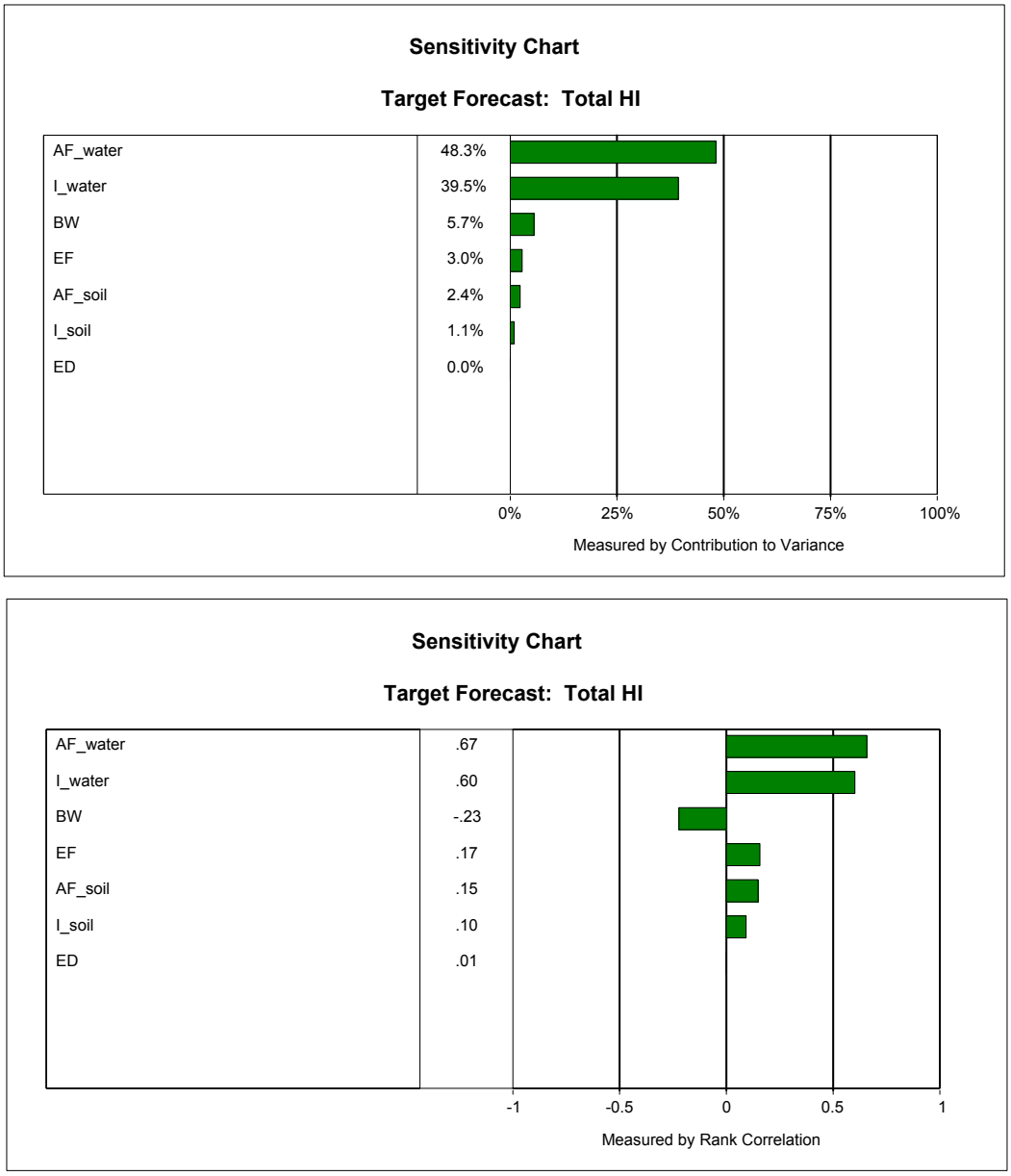


Figure A-4. Top panel - bar graph showing the r^2 values (square of Spearman rank correlation coefficient), a metric for the dependence of HI on exposure factors based on 1-D MCA for variability. Bottom panel - bar graph, sometimes referred to as "tornado plot", showing rank correlation coefficient. This graph is effective for showing both the relative magnitude and direction of influence (positive or negative) for each variable. Abbreviations for input variables are given in Table A-4. In this example, the variable with the greatest effect on HI is the absorption fraction in water (AF_w), followed by the water ingestion rate (I_w). Concentration does not influence variability because, in this example, long-term average concentration is characterized by a point estimate (i.e., 95% UCL), rather than a probability distribution. Exposure duration does not influence variability because variability in ED is expressed in both the numerator (ED) and denominator (AT=ED x 365 for noncarcinogenic effects), and cancels out. Output was generated with *Crystal Ball*®, which calculates the contribution to variance by squaring the rank correlation coefficient and normalizing to 100%.

In this example, seven exposure variables are used to characterize variability in HI. The remaining variables in the risk equation (i.e., concentration terms, and RfD) are characterized by point estimates. Because point estimates do not vary in a Monte Carlo simulation, they do not contribute to the variance in the output. This result does not mean that concentration is an unimportant variable in the risk assessment. Concentration may still contribute greatly to the uncertainty in the risk estimate. A sensitivity analysis of parameter uncertainty in a risk equation can be explored using iterative simulations, such as with 2-D MCA.

Results of the Pearson correlation and Spearman rank correlation give similar rankings of the input variables, with absorption fraction of water (AF_w) and tap water ingestion rate (I_w) being the two dominant exposure variables. Pearson correlations suggest that I_w is the most sensitive variable (r =0.644), whereas the highest Spearman rank correlation is for AF_w (r = 0.603). This may reflect the fact that I_w is characterized by an untruncated lognormal distribution, whereas AF_w is bounded between 0 and 1.0. The effect on the correlations of the occasional high-end value for I_w generated from random sampling of the lognormal distribution will tend to be expressed by Pearson correlations, but muted by the Spearman rank correlations.

A comparison of the Tier 1 and Tier 2 results is given below:

- **Tier 1, Sensitivity Ratios:**

  - Local SR (Δ = 5%) rankings: EF > BW > I_w = AF_w > I_s = AF_s > ED

  - Range SR (Δ = 50%) rankings: EF > I_w = AF_w > BW > I_s = AF_s > ED

- **Tier 1, Sensitivity Scores:**

  - Score based on local SR (Δ = 5%): I_w > AF_w > BW > EF > AF_s > IR_s > ED

  - Score based on range SR (Δ = 50%): I_w > AF_w > EF > BW > AF_s > IR_s > ED

- **Tier 2, Correlation Coefficients:**

  - Pearson: I_w > AF_w > BW > EF > AF_s > IR_s > ED

  - Spearman Rank: AF_w > I_w > EF > BW > AF_s > IR_s > ED

The Tier 1 sensitivity scores and Tier 2 correlation coefficients yield similar results, suggesting that, if sufficient information is available to estimate the coefficient of variation in the input variables, a Tier 1 analysis can help to focus efforts on the variables that contribute most to the variance in risk. By contrast, the Tier 1 sensitivity ratio approach suggested that EF was the most influential variable, when in fact it contributes less than 5% to the variance in the HI. These results suggest that Tier 1 sensitivity ratios are best applied to identify dominant exposure pathways, rather than dominant exposure variables in the risk equation.

A.2.2.3 FOCUSING ON THE RME RANGE OF THE RISK DISTRIBUTION

Monte Carlo methods can also be used to determine the sensitivity over a subset of the output distribution, such as the RME range (i.e., 90th to 99.9th percentiles). For some exposure models, the relative contribution of exposure variables may be different for the high-end exposed individuals than for the entire range of exposures. The general strategy for exploring sensitivity over subsets of risk estimates is to first sort the distribution of simulated output values in ascending (or descending) order, and then apply a sensitivity analysis to the subset of interest (e.g., > 90th percentile). For the hypothetical example presented in this appendix, there was no difference in the relative rankings of inputs in the RME range.

A.2.2.4 INSPECTION

With Monte Carlo analysis, the probability distributions assumed for the various input variables are used to generate a sample of a large number of points. Statistical methods are applied to this sample to evaluate the influence of the inputs on the model output. A number of different "indices" of sensitivity can be derived from the simulated sample to quantify the influence of the inputs and identify the key contributors. Most of these are based on an assumption that the model output *Y* varies in a monotonic, linear fashion with respect to various input variables (X_1, X_2, etc.). For example, an estimate of average daily intake (mg/kg-day) from multiple exposure pathways is linear with respect to the intake from each pathway. Since most risk models are linear with respect to the input variables, the output distribution (particularly its upper percentiles) tends to be dictated by the input variables with the largest coefficient of variation (CV), or the ratio of the standard deviation to the mean. For example, Equation A-9 represents a simple expression for intake rate as a function of random variables X_1 and X_2 :

$$Y = X_1 + X_2 \qquad \text{Equation A-9}$$

where X_1 and X_2 may represent dietary intake associated with prey species 1 and 2, respectively. If the same probability distribution was used to characterize X_1 and X_2, such as a lognormal distribution with an arithmetic mean of 100 and standard deviation of 50 (i.e., CV=50/100=0.5), each variable would contribute equally to variance in Y. If, however, X_2 was characterized by a lognormal distribution with an arithmetic mean of 100 and standard deviation of 200 (i.e., CV=200/100=2.0), we would expect Y to be more sensitive to X_2. That is, X_2 would be a greater contributor to variance in Y.

While the coefficient of variation may be a useful screening tool to develop a sense of the relative contributions of the different input variables, a common exception is the case when X_1 and X_2 have different scales. For example, Equation A-10 is an extension of Equation A-9:

$$Y = a_1 X_1 + a_2 X_2 \qquad \text{Equation A-10}$$

where a_1 and a_2 are constants that may represent the algebraic combination of point estimates for other exposure variables. If the means of X_1 and X_2 are equal, but $a_1 >> a_2$, then X_1 would tend to be the dominant contributor to variance, regardless of the CV for X_2. This concept was demonstrated by the sensitivity score calculations given in Table A-8. Water ingestion rate (I_w) and soil ingestion rate (I_s) had the same CV (0.58), but I_w was the dominant variable because tap water ingestion contributed approximately 90% to the HI.

The most influential random variables generally have the highest degrees of skewness or are related to the output according to a power function (Cullen and Frey, 1999). For example, Equation A-11 presents an extension of Equation A-10 in which there is a power relationship between X_2 and Y. In this

$$Y = a_1 X_1 + a_2 X_2^{\theta}$$ Equation A-11

example, assume Y represents the total dietary intake rate of cadmium for muskrats, X_1 and X_2 represent the dietary intake rate associated with prey species 1 and 2, respectively, a_1 and a_2 represent additional point estimates in the equation, and θ is the power exponent. In general, for $\theta > 1$, the total dietary intake rate (Y) will be more sensitive to the intake rate associated with species 2 (X_2) than species 1. Assume (hypothetically) that the power relationship stems from the fact that there is a direct relationship between availability of prey species X_2 and chemical body burdens of prey species X_2 because individuals that are more accessible to the muskrat also happen to frequent areas of the site with higher concentrations.

A.3.0 ADVANCED CONCEPTS IN SENSITIVITY ANALYSIS

This section provides additional information on the underlying principles of sensitivity analysis, although it is not a comprehensive summary and is not intended to substitute for the numerous statistical texts and journal articles on sensitivity analysis. Section A.3.1 begins with a general framework for relating model output to model input. Section A.3.2 explains the sensitivity ratio approach and highlights some of its limitations. Section A.3.3 reviews some of the metrics reported by the commercial software that report results of sensitivity analysis following Monte Carlo simulations (e.g., *Crystal Ball®*, *@Risk*). While statistical software for MCA provides convenient metrics for quantifying and ranking these sources, it is strongly recommended that risk assessors and risk managers develop an understanding of the underlying principles associated with these metrics.

A.3.1 RELATING THE CHANGE IN RISK TO THE CHANGE IN INPUT VARIABLE X

For purposes of discussion, let *Y* denote a model output (e.g., risk) and suppose that it depends on the input variable *X*. In general, a risk assessment model may use any number of inputs; however, for purposes of illustrating concepts, it is convenient to restrict this discussion to one variable. The model relates the output *Y* to values of *X* (i.e., x_0, x_1, ♦♦♦, x_n) based on the function expressed as *Y=F(x)*. The sensitivity of *Y* to *X* can be interpreted as the slope of the tangent to the response surface *F(X)* at any point x_i. This two-dimensional surface can be a simple straight line, or it may be very complex with changing slopes as shown in Figure A-5a. The sensitivity, therefore, may depend on both the value of *X* and the amount of the change Δx about that point. This concept can be extended to two input variables, X_1 and X_2, where the response is characterized by a three-dimensional surface. The shape may be a simple plane (Figure A-5c) or it may be very complex with many "hills" and "valleys" depending on the defining function $F(X_1, X_2)$. In a typical risk assessment with ten or more variables, the surface can be very complex, but the shape is likely to be dominated by a small subset of the input variables.

A sensitivity analysis based on a relatively small deviation about the point may be referred to as a local sensitivity analysis, while a large deviation may be referred to as range sensitivity analysis. In either case, the objective is to evaluate the sensitivity at some nominal point (X_1*, X_2*) such as the point defined by the mean or median of X_1 and X_2. At any point, the sensitivity of the model output, $Y^* = F(X_1^*, X_2^*)$, to one of the inputs (X_1 *or* X_2), is represented by the rate of change in *Y* per unit change in *X*. This is the slope of the surface at that nominal point in the direction of *X* and is expressed as $\partial Y/\partial X_i$, the *partial derivative* of *Y* with respect to *X*.

$$Partial\ Derivative = \frac{\partial Y}{\partial X} \approx \frac{\Delta Y}{\Delta X}$$

If the function $F(X_1, X_2)$ is known explicitly, it may be possible to determine the partial derivatives analytically. This is not a requirement, however, because an estimate can be obtained by incrementing X_i by a small amount, ΔX_i, while keeping the other inputs fixed and reevaluating the model output Y. The resulting change in Y divided by ΔX_i will approximate $\partial Y/\partial X_i$ at the nominal point. In practice, analytical solutions can be approximated using Monte Carlo techniques. This information is presented to highlight the fundamental concepts of sensitivity analysis. The partial derivative, *per se*, would typically not be one of the methods of sensitivity analysis used in a PRA. However, all of the approaches that are presented in this appendix are variations on this concept.

One drawback to using the partial derivative to quantify the influence of X_i is that the partial derivative is influenced by the units of measurement of X_i. For example, if the measurement scale for X_i is changed from grams to milligrams, the partial derivative $\partial Y/\partial X_i$ will change by a factor of 1,000. Therefore, it is necessary to **normalize the partial derivative** to remove the effects of units (see Section A.3.2).

If the relationship between Y and all of the inputs is linear, then the response surface is a flat plane and each of the partial derivatives at each point, (X_i, Y), will remain constant regardless of where the point is in the surface (Figure A-5b). In this case, it is a simple matter to determine the relative influence that the various inputs have on the model output. When the relationship is nonlinear, however, the situation is more complex because the influence of a particular input may vary depending on the value of that input.

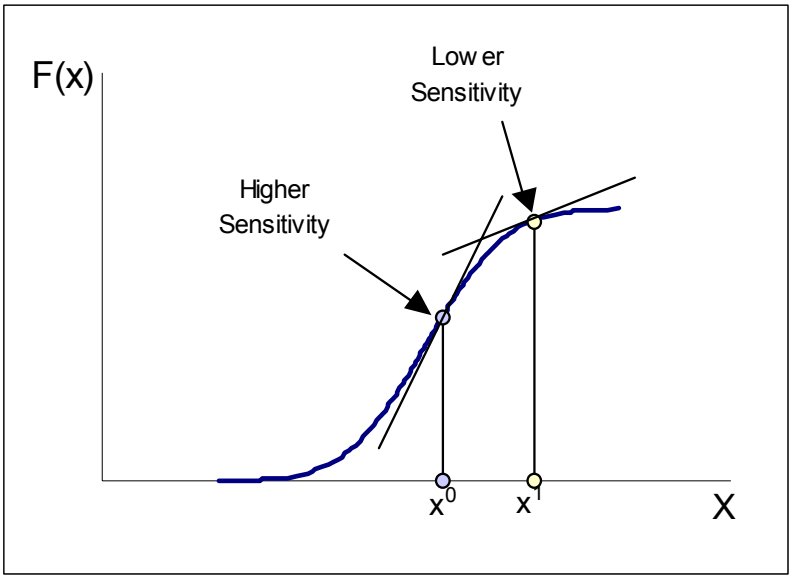


Figure A-5a. Hypothetical 2-D response surface for Y given one input variable: $Y=F(X)$. The sensitivity of Y with respect to X is calculated as the slope at a specific point on the surface (x^0, x^1), or the partial derivative, $\partial Y/\partial X_i$.

$$slope = \frac{\partial Y}{\partial X_i}$$

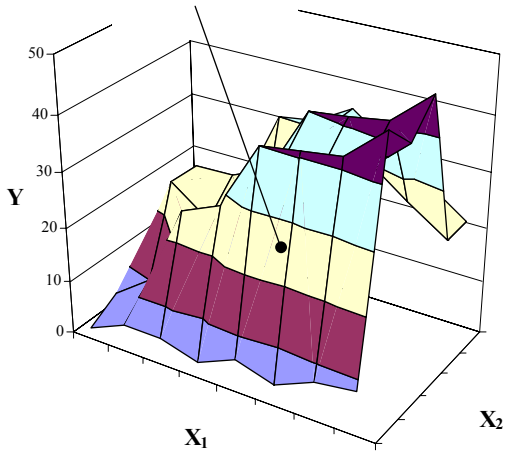


Figure A-5b. Hypothetical 3-D response surface for *Y* given two input variables: Y = $f(X_1, X_2)$. The sensitivity of *Y* with respect to X_i is calculated as the slope at a specific point on the surface, or the partial derivative, $\partial Y/\partial X_i$.

$$slope = \frac{\partial Y}{\partial X_i} = constant$$

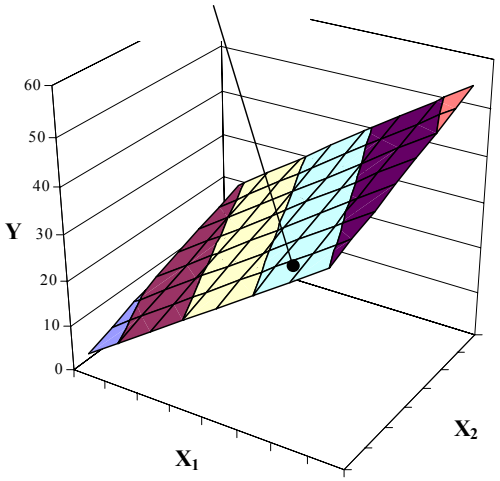


Figure A-5c. Hypothetical 3-D response surface when *Y* is a linear function of two input variables: Y=$f(X_1, X_2)$. The slope (i.e., the partial derivative, $\partial Y/\partial X_i$) is constant for any point (X_i, *Y*) on the surface in the direction of X_i. In this case, $\partial Y/\partial X_1$=5 while $\partial Y/\partial X_2$=2.

A.3.2 NORMALIZED PARTIAL DERIVATIVE

Classical sensitivity analysis methods use estimates of the partial derivatives of the model output with respect to each variable. For the purpose of evaluating the relative influence of the various input variables on the model output at a single point, the **normalized partial derivative** provides a useful index.

If the input variables are all discrete and take on a small number of values, then it is possible to evaluate the influence of the various input variables at each of the points defined by considering all possible combinations of the inputs. Then the influence can be evaluated for each input by computing normalized partial derivatives at each point. This approach is limited to situations where the number of inputs as well as the number of possible values for each input is relatively small; otherwise, the number of combinations to be evaluated will be unmanageable. Furthermore, when evaluating the influence at different points on the input-output surface simultaneously, it is important to take into account the probability associated with each of those points. For example, the fact that a particular input has a large influence on the model output at a particular point would be discounted if the probability associated with that particular point is very low.

A similar approach may be used to analyze inputs that are continuous variables if a few points representing the range of values are selected. For example, low, medium (or nominal), and high values may be selected for each of the continuous input variables and then the relative influence of each of the input variables can be computed as in the case of discrete inputs. One limitation of this approach, however, is that the continuous nature of the inputs makes it impossible to calculate an exact probability for each of the points. Generally, in a PRA, many if not all of the inputs will be random variables described by probability distributions and it will be necessary to quantify the influence of each input, X_i, over the entire range of X_i.

An estimate of the partial derivative can be obtained by incrementing X_i by a small amount, say ΔX_i while keeping the other inputs fixed and reevaluating the model output Y. The resulting change in Y divided by ΔX_i will approximate $\partial Y/\partial X_i$ at the nominal point.

$$\text{Partial Derivative} = \frac{\partial Y}{\partial X} \approx \frac{\Delta Y}{\Delta X}$$

As previously noted, one complication to using the partial derivative to quantify the influence of X_i is that the partial derivative is influenced by the units of measurement of X_i. One way this is accomplished is to divide the partial derivative by the ratio of the nominal point estimates, Y^* / X_i^* (or equivalently multiply by X_i^* / Y^*). An approximation of the normalized partial derivative is given by Equation A-12.

$$\text{Normalized Partial Derivative} \approx \frac{\Delta Y}{\Delta X} \times \frac{X_1}{Y_1} = \frac{\left(\frac{Y_2 - Y_1}{Y_1}\right)}{\left(\frac{X_2 - X_1}{X_1}\right)} \quad \text{Equation A-12}$$

This is the same as the equation for calculating sensitivity ratios (Section A.2.1.3), or elasticity (see Equation A-5). As with the SR approach, the normalized partial derived can be weighted by characteristics of the input variable (Section A.2.1.4). One approach is to divide by the ratio of standard deviations (σ_Y / σ_X), where σ_Y is the standard deviation of Y and σ_X is the standard deviation of X. This method requires that the standard deviations be known, or that a suitable estimate can be obtained.

As previously noted, if the relationship between Y and all of the inputs is nonlinear, the influence of a particular input may vary depending on the value of that input. One approach to this problem is to consider a range of values for the input and to examine the influence over that range. If the input is considered to be a random variable following some specified probability distribution, then it may be desirable to look at the influence that the random input has on the model output across the distribution of input values. This can be accomplished with a Monte Carlo approach. Another technique that addresses nonlinearities is to calculate contributions to variance using input variables that are transformed (e.g., lognormal or power transformation).

A.3.3 REGRESSION ANALYSIS: R^2, PEARSON R, AND PARTIAL CORRELATION COEFFICIENTS

In order to understand R^2, it is necessary to first understand simple and multiple linear regression. In regression analysis, we are interested in obtaining an equation that relates a dependent variable (Y) to one or more independent variables (X):

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad \text{Equation A-13}$$

where β_0 and β_1 are regression coefficients, and ε is called a random error. Equation A-13 is the general equation for a simple linear regression, because there is only one Y and one X variable, and their relationship can be described by a line with intercept β_0 and slope β_1.

Note that *linear* regression refers to the linear relationship between parameters (β_0, β_1), not X and Y. Thus, the equation $Y = \beta_0 + \beta_1 X_1^2 + \varepsilon$ is considered linear. *Multiple* linear regression involves more than one X related to one Y $\left[Y = \beta_0 + \beta_1 X_1 + \beta_2 X \cdots\right]$, while *multivariate* regression involves more than one Y to more than one X.

EXHIBIT A-5

SIMPLIFYING ASSUMPTIONS IN REGRESSION ANALYSIS

- Y is a linear function of the unknown coefficients (β_i)
- Successive values of Y are uncorrelated
- Variance of Y is constant for all values of inputs (X_i)

The random error, ε, represents the difference between an observed Y value (calculated from the observed input variables), and a Y value predicted by the regression line ($\hat{y}$). It is also called

the *residual* (i.e., ε=y–ŷ). The random error takes into account all unpredictable and unknown factors that are not included in the model. Exhibit A-5 gives some of the simplifying assumptions that apply to regression analysis. Assumptions about ε are that the random error has mean = 0 and constant variance, and is uncorrelated among observations. One method of finding the best regression line is to minimize the residual sum of squares (i.e., least-squares method), also called the sum of squares due to error (SSE).

In terms of sensitivity analysis, we are interested in how much of the variation in *Y* can be explained by the variation in *X*, and how much is unexplained (due to random error). If a scatter plot of paired observations (*x*, *y*) shows that our regression line intersects all of the observations exactly, then all of the variation in *Y* is explained by *X*. Another way of stating this is that the difference between the mean output ($\bar{y}$) and an observed *y* (y_i), or (y_i - $\bar{y}$), is equal to the difference between the mean output and a predicted *y* or ($\hat{y} - \bar{y}$).

In general, the total deviation of y_i from $\bar{y}$ is equal to the sum of the deviation due to the regression line plus the deviation due to random error:

$$(y_i - \bar{y}) = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})$$
$$\sum (y_i - \bar{y})^2 = \sum (y_i - \hat{y}_i)^2 + \sum (\hat{y}_i - \bar{y})^2 \quad \text{Equation A-14}$$
$$SST = SSE + SSR$$

Thus, the total sum of squares (SST) equals the sum of squares due to error (SSE) plus the sum of squares due to regression (SSR).

A.3.3.1 CALCULATIONS OF R^2 AND ADJUSTED R^2

The R^2 term is a measure of how well the regression line explains the variation in *Y*, or:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$
$$R = \sqrt{\frac{\text{variation explained by regression}}{\text{total variation in Y}}} \quad \text{Equation A-15}$$

where R^2 is called the *coefficient of multiple determination* and *R* is called the *multiple correlation coefficient*. If R^2=0.90 for a certain linear model, we could conclude that the input variables (X_1, X_2,...X_k) explain 90% of the variation in the output variable (*Y*). R^2 reduces to the *coefficient of determination* r^2 for simple linear regression when one independent variable (*X*) is in the regression model. The *sample correlation coefficient*, *r*, is a measure of the association between *X* and Y, and calculated by Equation A-16. It is also referred to as the Pearson product moment correlation coefficient.

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\left[\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2\right]^{0.5}} \qquad \text{Equation A-16}$$

In addition, r is an estimate of the unknown population parameter, ρ, defined by Equation A-17:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad \text{Equation A-17}$$

where σ_X and σ_Y denote the population standard deviations of the random variables X and Y, and where σ_{XY} is called the covariance between X and Y. The covariance σ_{XY} is a population parameter describing the average amount that two variables "covary". Thus, another way of thinking about a correlation coefficient (R) is that it reflects the ratio of the covariance between two variables divided by the product of their respective standard deviations; and the value always lies between -1 and +1. *@Risk* and *Crystal Ball®* provide both the R^2 for the entire model, as well as the correlation coefficients for each input variable (or regressor). The higher the value of R_i for X_i, the more sensitive the output variable is to that input variable.

Although the calculations are the same, there is a subtle conceptual difference between the coefficient of determination (r^2) from regression, and the square of the correlation coefficient. When evaluating two variables *(X, Y)*, the key is whether X is interpreted as a "fixed" quantity (i.e., an explanatory variable), or a random variable just like Y. In regression analysis, r^2 measures how well the regression line explains the variation in Y given a particular value for X (Equation A-15). Correlation requires that X be considered a random variable, typically having a bivariate normal distribution with Y (see Appendix B).

One artifact of regression analysis is that R^2 increases as you add more and more input variables to your model; however, the increased fit of the model due to one or more of the input variables may be insignificant. Sometimes an adjusted R^2 is calculated to take into account the number of input variables (called regressors) in the model (k) as well as the number of observations in the data set (n):

$$R^2_{adj} = \frac{(n-1)R^2 - 1}{n - k - 1} \qquad \text{Equation A-18}$$

While R^2 gives the proportion of the total *variation* of Y that is explained, R^2_{adj} (Equation A-18) takes into account the degrees of freedom (*df*), and gives the proportion of the total *variance* of Y that is explained (variance = variation /df); or stated simply, R^2_{adj} is the R^2 corrected for *df*, where *df* is described by [1 - k/(n-1)].

- If the relationship between an input variable and an output variable is strong, but nonlinear, the R^2 statistic will be misleadingly low.
- If the means of the sampling data are used rather than the individual observations for each variable, R^2 will be misleadingly high. This is because taking the mean of a sample reduces the fraction of the

total variation due to *random* variation (see discussion of random error above). This is an important consideration when trying to interpret the results of regression analyses that incorporate data averaged over different spatial scales (e.g., regression of PbB on soil lead concentrations taken at the city block level may give an inflated R^2 value if the sampling data are averaged over a larger spatial scale, such as the census tract level).

A multiple regression analysis can also be performed to estimate the **regression coefficients** (see Appendix A.3.3). Each coefficient essentially represents an "average" value of the partial derivative across the entire distribution of the input. The regression coefficient, like the partial derivative, depends on the units of measurement so, as in the case of the partial derivative, it must be normalized. This can be accomplished by multiplying the regression coefficient by the ratio of estimated standard deviations s_y/s_x.

A convenient way to carry out a sensitivity analysis is to perform a stepwise regression analysis. Some statistical software packages (e.g., SAS, SPSS) offer a variety of different approaches for this; however, in general, they can be classified into two general categories: forward selection and backward elimination. In the forward selection, the inputs are added to the model one by one in the order of their contribution. In the backward elimination, all of the inputs are used in the model initially and then they are dropped one by one, eliminating the least important input at each step. A true stepwise procedure is a variation on the forward selection approach where an input can drop out again once it has been selected into the model if at some point other inputs enter the model that account for the same information.

A.3.3.2 RELATIVE PARTIAL SUM OF SQUARES (RPSS)

The **relative partial sum of squares (RPSS)** measures the sensitivity of the model output to each of the input variables by partitioning the variance in the output attributable to each variable using multiple regression techniques (Rose et al., 1991). The RPSS is presented as a percentage reflecting the proportion of influence a given variable has on risk. The results of RPSS are intuitive and generally easy to understand.

Briefly, the RPSS represents the percentage of the total sum of squares attributable to each of the variables. To calculate RPSS for variable V_i, the difference between the regression sum of squares (RSS) for the full model and the regression sum of squares for the model with V_i missing (RSS_{-i}) is divided by the total sum of squares (TSS) and expressed as a percentage:

$$RPSS_i = \frac{100\,(RSS - RSS_{-1})}{TSS} \qquad \text{Equation A-19}$$

This procedure can be thought of as analogous to least squares linear regression, but performed in the *n*-dimensional parameter space of the risk equation. Since this approach depends on the adequacy of the linear regression model between the output variable (e.g., risk) and all the variables, an additional diagnostic is to check how close R^2 is to 1.0. For equations with more than three parameters (such as those used in Superfund risk assessments), the computational overhead of this process is large and requires specific computer programs. The software program *Crystal Ball*® does not perform this calculation, but it can be determined with most standard statistical software packages that perform multiple regression. *@Risk* performs a calculation similar to this called multivariate stepwise regression that yields correlation coefficients in lieu of percent contributions to output variance.

A.3.3.3 SPEARMAN'S RANK CORRELATION COEFFICIENT (RHO)

The validity of using indices such as regression coefficients, correlation coefficients, and partial correlation coefficients depends on the assumptions of the underlying linear model being met. If there is any doubt that a data set satisfies the model assumptions, a nonparametric measure of correlation based on the rank orders of the inputs and associated outputs can be used. The Spearman Rank correlation coefficient is a nonparametric statistic; it measures an association between variables that are either count data or data measured on an ordinal scale, as opposed to data measured on an interval or ratio scale. An example of an ordinal scale would be the ranking of sites based on their relative mean soil concentrations. For example, if there are four categories of soil contaminant concentrations, sites with the highest concentrations may receive a rank of 1 while sites with lowest concentrations may receive a rank of 4. Ordinal scales indicate relative positions in an ordered series, not "how much" of a difference exists between successive positions on a scale.

To calculate the Spearman rank correlation coefficient, assign a rank to each of the input variables (X_j) and output variables (Y_k). For each ranked pair (X_j, Y_k), calculate the difference, d, between the ranks. For example, if the first observation for variable X has a ranking of 5 (relative to all of the observations of X), and the corresponding value of Y has a ranking of 3 (relative to all of the observations of Y), the difference (d) is equal to 5–3=2. Spearman rho (r_s) is calculated as:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{(n^3 - n)} \qquad \text{Equation A-20}$$

Hence ($-1 \le r_s \le 1.0$), and r_s=-1 describes a perfect indirect or negative relationship between ranks in the sense that if an X element increases, the corresponding Y element decreases. Similarly, r_s=0 suggests that there is no relationship between X and Y.

The Pearson product moment correlation coefficient is equal to the Spearman rank correlation coefficient when interval/ratio values of the measured observations (X, Y) are replaced with their respective ranks.

REFERENCES FOR APPENDIX A

Cullen, A.C. and H.C. Frey. 1999. Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

Hamby, D.M. 1994. A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models. *Environ. Monit. and Assess.* 32:135–154.

Helsel, D.R. and R.M. Hirsch. 1992. Statistical Methods in Water Resources. Elsevier Science B.V.

Iman, R.L. and J.C. Helton. 1988. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Anal.* 8:71–90.

Iman, R.L. and J.C. Helton. 1991. The Repeatability of Uncertainty and Sensitivity Analyses for Complex Probabilistic Risk Assessments. *Risk Anal.* 11:591–606.

Merz, J., M.J. Small, and P. Fischbeck. 1992. Measuring Decision Sensitivity: A Combined Monte Carlo-Logistic Regression Approach. *Medical Decision Making*, 12: 189–196.

Morgan, M.G. and M. Henrion.1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press.

Palisade Corporation. 1994. *Risk Analysis and Simulation Add-In for Microsoft Excel or Lotus 1-2-3*. Windows Version Release 3.0 User's Guide, Palisade Corporation, Newfield, NY.

Rose, K.A., E.P. Smith, R.H. Gardner, A.L. Brenkert, and S.M. Bartell. 1991. Parameter Sensitivities, Monte Carlo Filtering, and Model Forecasting Under Uncertainty. *J. Forecast* 10:117–133.

Saltelli, A and J. Marivort. 1990. Non-Parametric Statistics in Sensitivity analysis for Model Output: A Comparison of Selected Techniques. *Reliab. Engin. Syst. Saf.* 28:299–253.

Shevenell, L. and F.O. Hoffman. 1993. Necessity of Uncertainty Analyses in Risk Assessment. *J Hazard Mater*. 35:369–385.

Stern, A.H. 1994. Derivation of a Target Level of Lead in Soil at Residential Sites Corresponding to a *de minimis* Contribution to Blood Lead Concentration. *Risk Anal.* 14:1049–1056.

U.S. EPA. 1997. *Guiding Principles for Monte Carlo Analysis*. Risk Assessment Forum and National Center for Environmental Assessment. EPA/630/R-97/001.

U.S. EPA. 1999. *TRIM, Total Risk Integrated Methodology, TRIM FATE Technical Support Document Volume I: Description of Module*. Office of Air Quality Planning and Standards. EPA/43/D-99/002A.

APPENDIX B

SELECTION AND FITTING OF DISTRIBUTIONS

B.0 INTRODUCTION

An important step in Monte Carlo analysis (MCA) is to select the most appropriate distributions to represent the factors that have a strong influence on the risk estimates. This step in the development of a Monte Carlo model can be very challenging and resource intensive.

> ☞ *Specifying probability distributions for all of the input variables and parameters in a probabilistic risk assessment (PRA) will generally not be necessary.*

If the sensitivity analysis indicates that a particular input variable does not contribute significantly to the overall variability and uncertainty, then this variable may be represented as a point estimate. As discussed in Appendix A, however, different approaches to sensitivity analysis may be applied throughout the tiered approach (e.g., sensitivity ratios, correlation analysis), and the ability to reliably identify variables as being minor or major can vary. Sometimes it can be helpful to develop probability distributions based on preliminary information that is available from Tier 1 in order to explore alternative options for characterizing variability and uncertainty. Likewise, sometimes the important "risk drivers" are apparent, and resources can be allocated to fully characterize the variability and uncertainty in those input variables. Therefore, the process of selecting and fitting distributions may also be viewed as a tiered approach. This appendix reviews the methods available to select and fit distributions and provides guidance on the process for determining appropriate choices depending on the information needed from the assessment and the information available to define the input variables.

In PRA, there are some important distinctions in the terminology used to describe probability distributions. A probability density function (PDF), sometimes referred to as a probability model, characterizes the probability of each value occurring from a range of possible values. Probability distributions may be used to characterize variability (PDFv) or uncertainty (PDFu). One advantage of using a PDFv and PDFu is that distributions represent a large set of data values in a compact way (Law and Kelton, 1991). For example, a lognormal distribution provides a good fit to a large data set of tap water ingestion rates (n=5,600) among children ages 1 to 11 years (Roseberry and Burmaster, 1992). Therefore, the distribution type (lognormal) and associated parameters (mean and standard deviation) fully describes the PDFv for intake rates, from which other statistics of interest can be calculated (e.g., median, and 95th percentile). Reducing a complex exposure model to a series of representative and well-fitting distributions can facilitate both the quantitative analysis and the communication of the modeling methodology. Alternatively, a PDFu may be specified to characterize parameter uncertainty. For example, the sample mean ($\bar{x}$) is generally an uncertain estimate of the population mean (μ) due to measurement error, small sample sizes, and other issues regarding representativeness (see Section B.3.1). A PDFu can be used to represent the distribution of possible values for the true, but unknown parameter. Understanding whether uncertainty or variability is being represented by a PDF is critical to determining how the distribution and parameters should be specified and used in a PRA.

EXHIBIT B-1

DEFINITIONS FOR APPENDIX B

Bayesian Analysis - Statistical analysis that describes the probability of an event as the degree of belief or confidence that a person has, given some state of knowledge, that the event will occur. Bayesian Monte Carlo combines a prior probability distribution and a likelihood function to yield a posterior distribution (see Appendix D for examples). Also called subjective view of probability, in contrast to the frequentist view of probability.

Bin - Regarding a histogram or frequency distribution, an interval within the range of a random variable for which a count (or percentage) of the observations is made. The number of bins for a histogram is determined on a case-by-case basis. In general, equal interval widths are used for each bin; however, in some cases (e.g., Chi-square test), individual bin widths are calculated so as to divide the distribution into intervals of equal probability.

Countably Infinite - Used to describe some discrete random variables, this term refers to a set of numbers that can be counted with integers (e.g., one, two, three) and that has no upper limit. Examples include the number of tosses required for a coin to show a head—we can count each toss, but it is possible that at least one more toss is needed. The number of dust particles in a volume of air is another example. Countably finite implies there is an upper limit (e.g., days of work per year).

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF, gives the cumulative probability of occurrence for a random independent variable. Each value c of the function is the probability that a random observation x will be less than or equal to c.

Empirical Distribution Function (EDF) -The EDF, also called the empirical CDF (ECDF), is based on the frequency distribution of observed values for a random variable. It is a stepwise distribution function calculated directly from the sample, in which each data point is assigned an equal probability.

Frequency Distribution or Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

Goodness-of-Fit (GoF) Test - A method for examining how well (or poorly) a sample of data can be described by a hypothesized probability distribution for the population. Generally involves an hypothesis test in which the null hypothesis H_0 is that a random variable X follows a specific probability distribution F_0. That is, H_0: $F=F_0$ and H_a: $F \neq F_0$.

Independence - Two events A and B are independent if whether or not A occurs does not change the probability that B occurs. Likewise, knowing the value of B does not affect the value of A. Input variables, X and Y, are independent if the probability of any paired values (X, Y) is equal to the probability of X multiplied by the probability of Y. In mathematical terms, X and Y are independent if f(X, Y)=f(X) x f(Y). Independence is not synonymous with correlation. If X and Y are independent, then their correlation is zero, Cor(X, Y)= 0. But, the converse is not always true. There may be a nonlinear relationship between X and Y that yields Cor(X, Y)=0, but the variables are highly dependent.

Nonparametric Method - Also called a *distribution-free* method, a procedure for making statistical inferences without assuming that the population distribution fits a theoretical distribution such as normal or lognormal. Common examples are the Spearman rank correlation, (see Appendix A) and the bootstrap-t approach..

Parameter - In PRA, a parameter is a quantity that characterizes the probability distribution of a random variable. For example, a normal probability distribution may be defined by two parameters (e.g., arithmetic mean and standard deviation).

Parametric Distribution - A theoretical distribution specified by a distribution type and one or more parameters. Examples include the normal, Poisson, and beta distributions.

EXHIBIT B-1 —*Continued*
DEFINITIONS FOR APPENDIX B

Probability Density Function (PDF) - A function representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Probability Distribution - The mathematical description of a function that associates probabilities with specified intervals or values for a random variable. A probability distribution can be displayed in a graph (e.g., PDF or CDF), summarized in a table that gives the distribution name and parameters, or expressed as a mathematical equation. In PRA, the process of selecting or fitting a distribution that characterizes variability or uncertainty can also be referred to as applying a *probability model* to characterize variability or uncertainty. In this guidance, the probability model is considered to be one source of model uncertainty.

Step Function - A mathematical function that remains constant within an interval, but may change in value from one interval to the next. Cumulative distribution functions for discrete random variables are step functions.

Z-score - The value of a normally distributed random variable that has been standardized to have a mean of zero and a SD of one by the transformation $Z=(X-\mu)/\sigma$. Statistical tables typically give the area to the left of the z-score value. For example, the area to the left of z=1.645 is 0.95. Z-scores indicate the direction (+/-) and number of standard deviations away from the mean that a particular datum lies assuming X is normally distributed. Microsoft Excel's *NORMSDIST(z)* function gives the probability p such that $p=Pr(Z \leq z)$, while the *NORMSINV(p)* function gives the z-score z_p associated with probability p such that $p=Pr(Z \leq z_p)$.

B.1.0 CONCEPTUAL APPROACH FOR INCORPORATING A PROBABILITY DISTRIBUTION IN A PRA

Often, more than one probability distribution may appear to be suitable for characterizing a random variable. A step-wise, tiered approach is recommended for incorporating probability distributions in a PRA. This appendix provides guidance on selecting and fitting distributions for variability and parameter uncertainty based on the overall strategy given in Exhibit B-2. Many of the same principles of selecting and fitting distributions are also given in EPA's *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999a).

EXHIBIT B-2

GENERAL STRATEGY FOR SELECTING AND FITTING DISTRIBUTIONS

(1) Hypothesize a family of distributions
(2) Assess quality of fit of distribution
(3) Estimate distribution parameters
(4) Assess quality of fit of parameters

Probability distributions may be developed to characterize variability or uncertainty. Example flow charts for specifying a PDFv and PDFu are given in Figures B-1 and B-2, respectively. Both approaches outline an iterative process that involves three general activities: (1) identify potentially important sources of variability or uncertainty to determine if a PDF may be needed; (2) apply the general strategy given in Exhibit B-1 and evaluate plausible alternatives for distributions and parameter estimates; and (3) document the decision process. The flowcharts provide a general outline of the process and contain terms which are explained in subsequent sections. Just as with the point estimate approach, different sites may require different probability distributions for input variables, depending on the unique risk management issues and sources of uncertainty.

B.2.0 PRELIMINARY SENSITIVITY ANALYSIS

Selecting and fitting probability distributions for *all* of the input variables can be resource intensive and is generally unnecessary. Ideally, a subset of variables could be identified that contribute to most of the variability and uncertainty in a risk estimate. Sensitivity analysis can play an important role in helping to identify and quantitatively rank the major exposure pathways and variables. Since the information obtained from a sensitivity analysis may vary, depending on the approach(es) used and the information available to characterize the input variables, risk assessors should understand inherent limitations of each approach. A variety of approaches that are common for Tier 1 and 2 analyses are described and applied to a hypothetical example in Appendix A.

In a Tier 1 assessment, sensitivity analysis is typically limited to exploring the effect of alternative point estimates on the risk estimate. These methods can be helpful if additional information regarding the variability in the input variables is incorporated into the analysis (i.e., sensitivity scores). Alternatively, a reasonable approach is to specify preliminary probability distributions for one or more inputs in order to maximize the advantages of probabilistic methods. The difference between a preliminary distribution and a subsequent distribution reflects the level of effort invested in characterizing variability and uncertainty. If a robust data set is available in Tier 1 to define point estimates, then a preliminary distribution may, in fact, fully characterize variability with very high confidence. For other variables, summary statistics, rather than sample data, may be available, allowing for estimates of central tendency or plausible ranges. The use of preliminary distributions reflects an effort to employ more robust sensitivity analysis techniques without expending the effort and resources that might otherwise be applied to a PRA in Tier 2. The goal of the preliminary analysis would not be necessarily to evaluate risks and/or develop a PRG; rather, the focus would be on identifying input variables that may be important to explore more fully. Preliminary sensitivity analysis can provide insight into the importance of selecting among alternative probability distributions and exposure scenarios.

One-dimensional Monte Carlo simulations with preliminary (or screening-level) distributions can be run prior to engaging in a more involved process of selecting and fitting distributions. The distributions can be selected based on knowledge regarding the mechanisms that result in variability, and information already available for determining point estimates (e.g., summary statistics, U.S. EPA guidance, etc.). Table B-1 provides examples of preliminary distributions that might be selected based on the type of information available, sometimes referred to as the *state of knowledge*. In many cases, the distribution is intended to estimate the plausible bounds of a variable, while requiring no additional data collection effort. For example, given estimates of a lower bound [min], upper bound [max], and the assumption that each value is equally likely, a uniform distribution would be used to represent variability (or parameter uncertainty). If no mechanistic basis for selecting a distribution exists, then the preliminary distribution would be chosen based on the available information. For example, given the estimates of the arithmetic mean [μ] and a percentile value [a] for a random variable, an exponential distribution might be recommended with $\lambda=1/\mu$.

Guidance on matching the choice of the distribution to the state of knowledge is extended to a more diverse array of scenarios later in this appendix (see Table B-4).

Table B-1. Examples of Preliminary Distributions Based on Information Available[1], [2]

Information / Constraints	Distribution Shape
[a, b]	uniform
[a, m, b]	triangular
[a, b, α_1, α_2, β]	beta
[μ, σ]	normal
γ	exponential
[a, b, μ, σ]	Johnson Sb, Lognormal
[α, β]	gamma

a=minimum, b=maximum, m=mode, α=shape parameter, μ=mean, σ=standard deviation, γ=average rate of occurrence of events, β=scale,

It may be informative to explore alternative choices for distributions applied to the same variable. For example, a simple yet informative approach is to run two 1-D MCA simulations for variability with an input variable characterized first by a Johnson Sb (i.e., a four-parameter lognormal distribution; Hahn and Shapiro, 1967) and then by a normal distribution. The difference in the risk distribution, especially at the percentile that is relevant to the risk management decision (e.g., 95th percentile), may offer insights regarding the importance of the shape of the PDFv.

B.3.0 WHAT DOES THE DISTRIBUTION REPRESENT?

Distributions may be specified to characterize variability or uncertainty. Often, a Monte Carlo simulation of variability will focus on describing differences between individuals in a population (i.e., inter-individual variability). In this case, the goal is to select a distribution that is representative of the *target* population—the set of all receptors that are potentially at risk. There may be uncertainty that the choice of PDFv reflects variability in the target population. In general, risk assessors should fully disclose uncertainties in the PDFv, especially because the use of a distribution instead of a point estimate may inappropriately suggest that there is a greater state of knowledge. Following the tiered process (see Chapter 2, Figure 2-1), there are multiple opportunities to consider consequences of alternative modeling approaches early in the process of developing a probabilistic model. The importance of relating the distribution to the *target* population, clearly distinguishing between variability and uncertainty, and evaluating data representativeness is emphasized in Sections B.3.1, B.3.2 and B.4.

[1]The preliminary distributions are based in part on maximum entropy concepts. Maximum entropy is a technique for determining the distribution that represents the maximum uncertainty allowed by the available information and data (Vose, 1996). Although the approach can be used to quickly define distributions that maximize uncertainty, the credibility of the distribution depends on the use of accurate, unbiased information.

[2]See Table B-2 for more detailed descriptions of selected distributions.

B.3.1 CONCEPTS OF POPULATION AND SAMPLING

The distinction between a *target* population, a *sampled* population, and a *statistical* population should be considered carefully when evaluating information for use in both Tier 1 and Tier 2 of a PRA. The *target* population is often considered to be the "population of concern". A risk assessor is often interested in quantifying specific attributes of the population (e.g., exposure duration, exposure frequency, etc.). A *sampled* population is the set of receptors available for selection and measurement. For purposes of this appendix/guidance, the *sampled* population may be the *target* population or it may be a different population that is thought to be representative of the *target* population. For purposes of this guidance, a *statistical* population is an approximation of the *target* population based on information obtained from the *sampled* population.

Distributions are generated from representative *sample* populations to make inferences about the *target* population. Ideally, a *sampled* population should be a subset of a *target* population and should be selected for measurement to provide accurate and representative information about the exposure factor being studied. However, defining representative samples is a matter of interpretation.

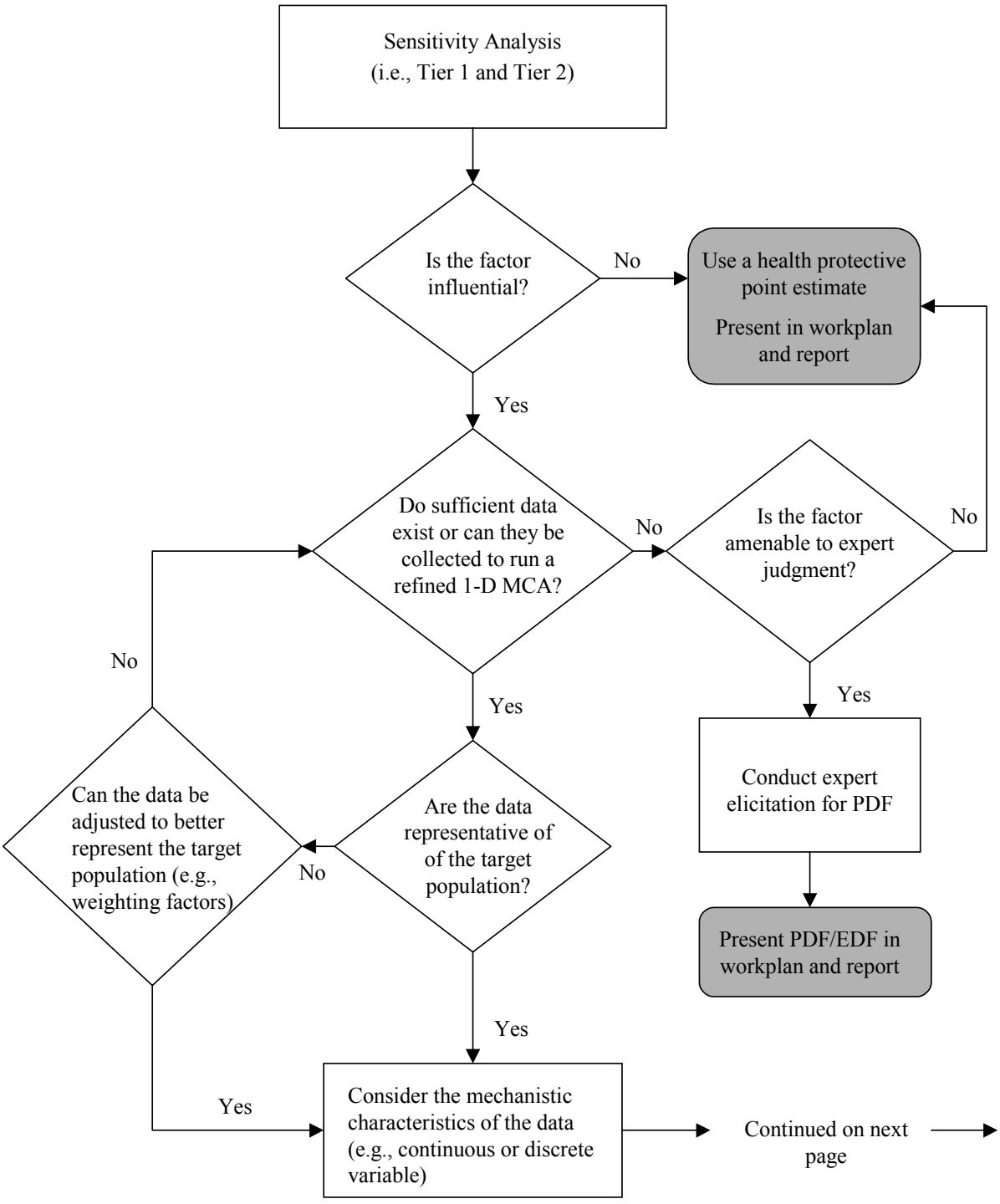


Figure B-1 (page 1 of 2). Conceptual approach for incorporating probability distributions for variability in PRA.

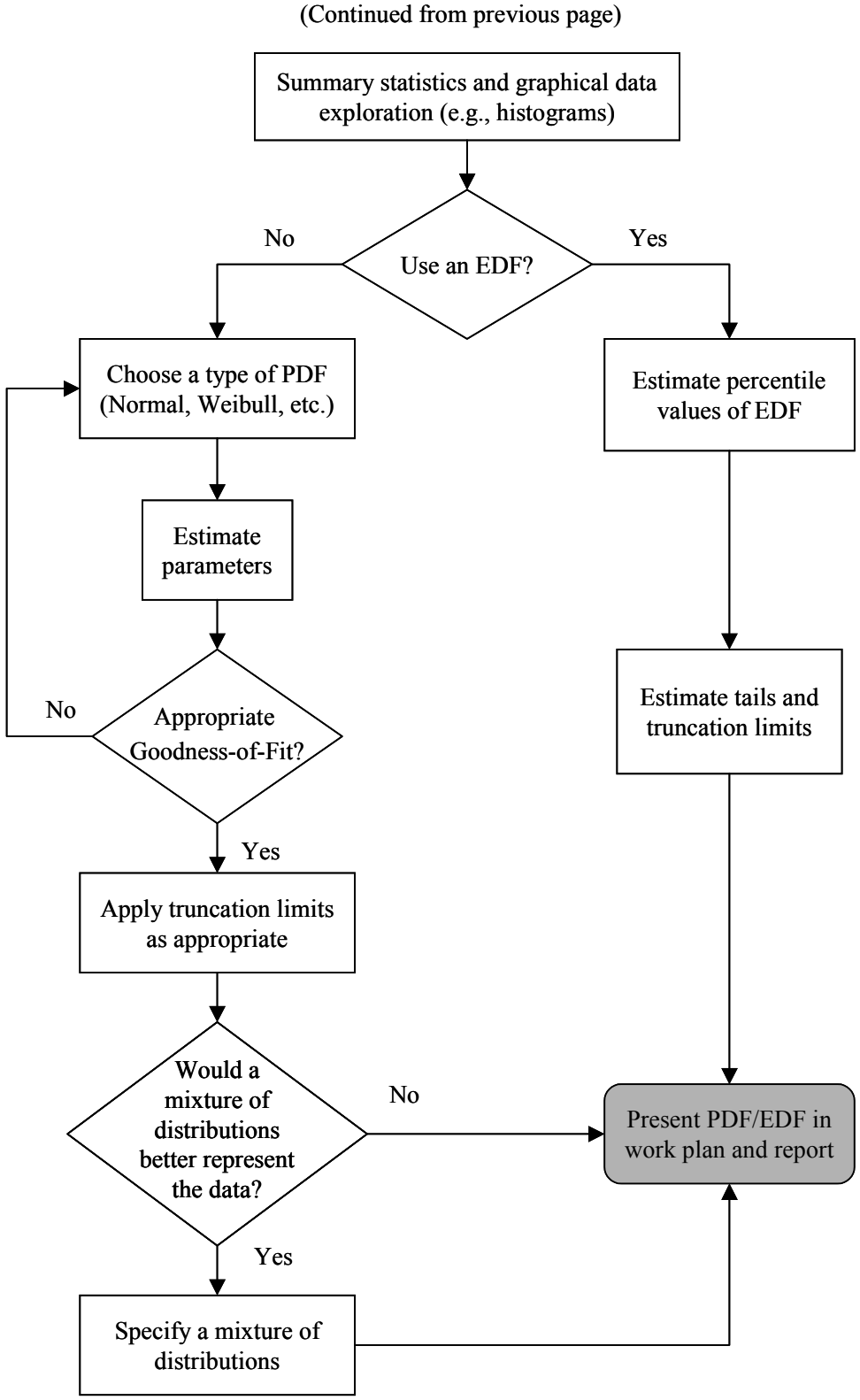


Figure B-1 (page 2 of 2). Conceptual approach for incorporating probability distributions for variability in PRA.

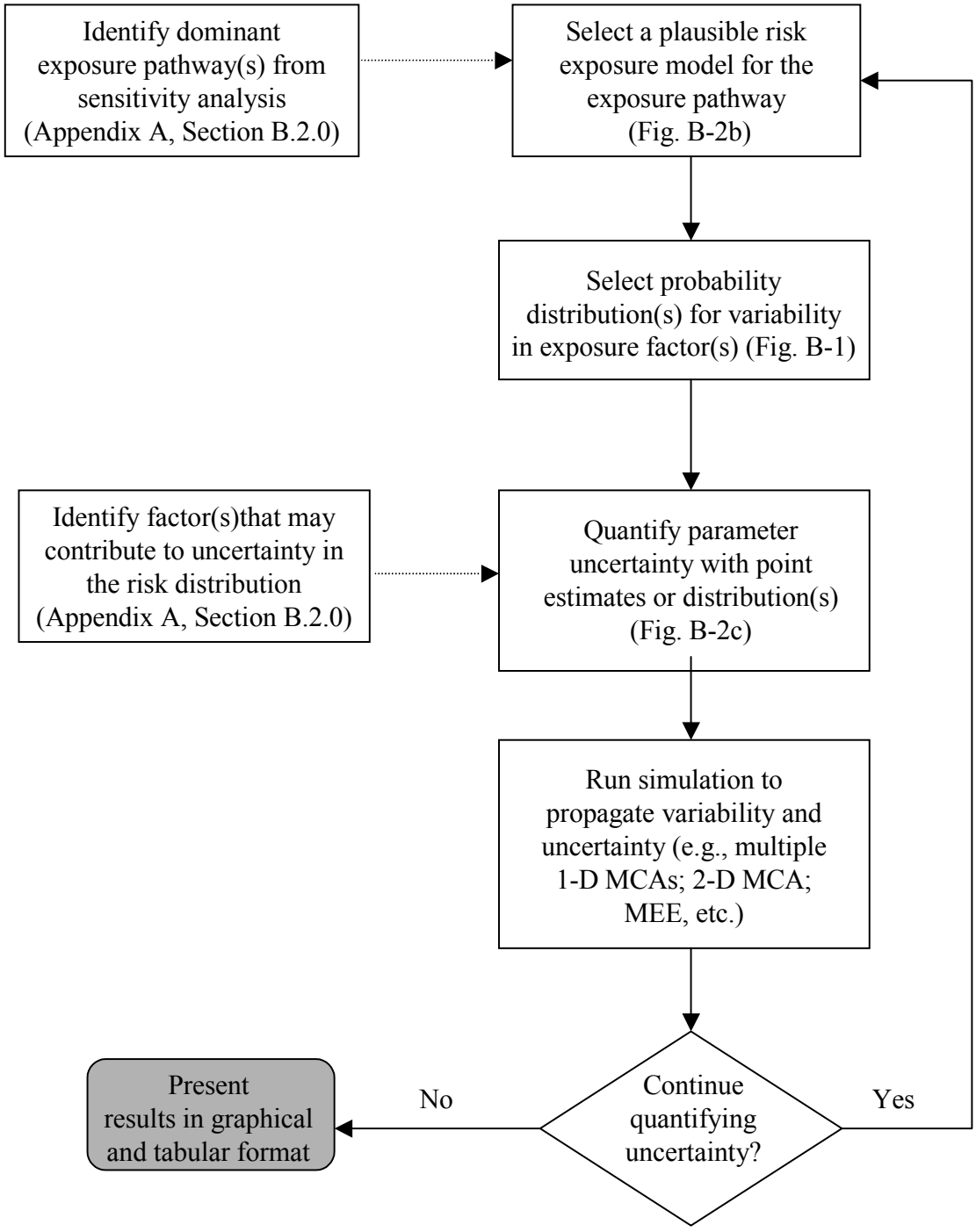


Figure B-2a (page 1 of 3). Conceptual approach for quantifying model and parameter uncertainty in PRA.

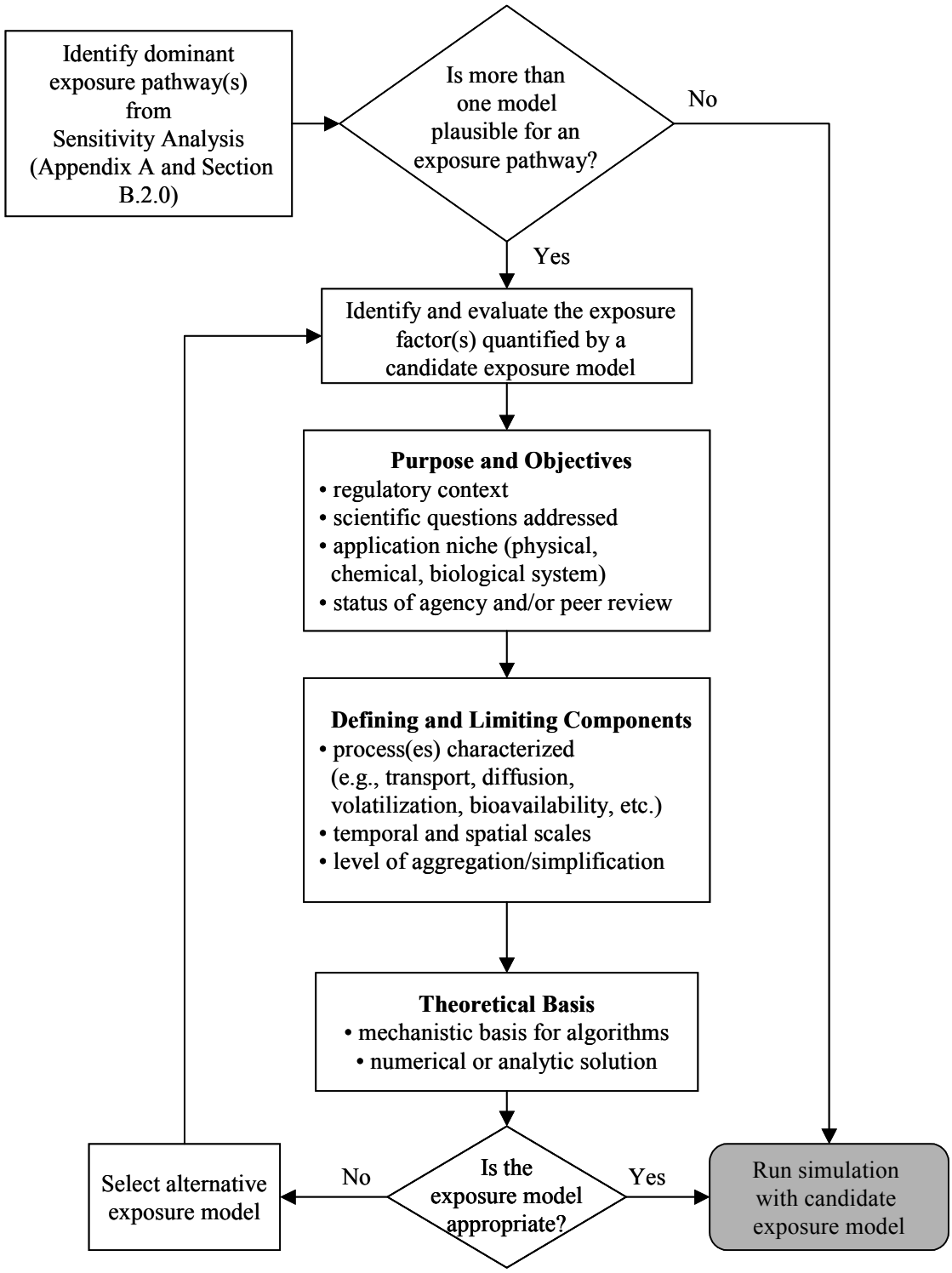


Figure B-2b (page 2 of 3). Detailed conceptual approach for incorporating model uncertainty in PRA.

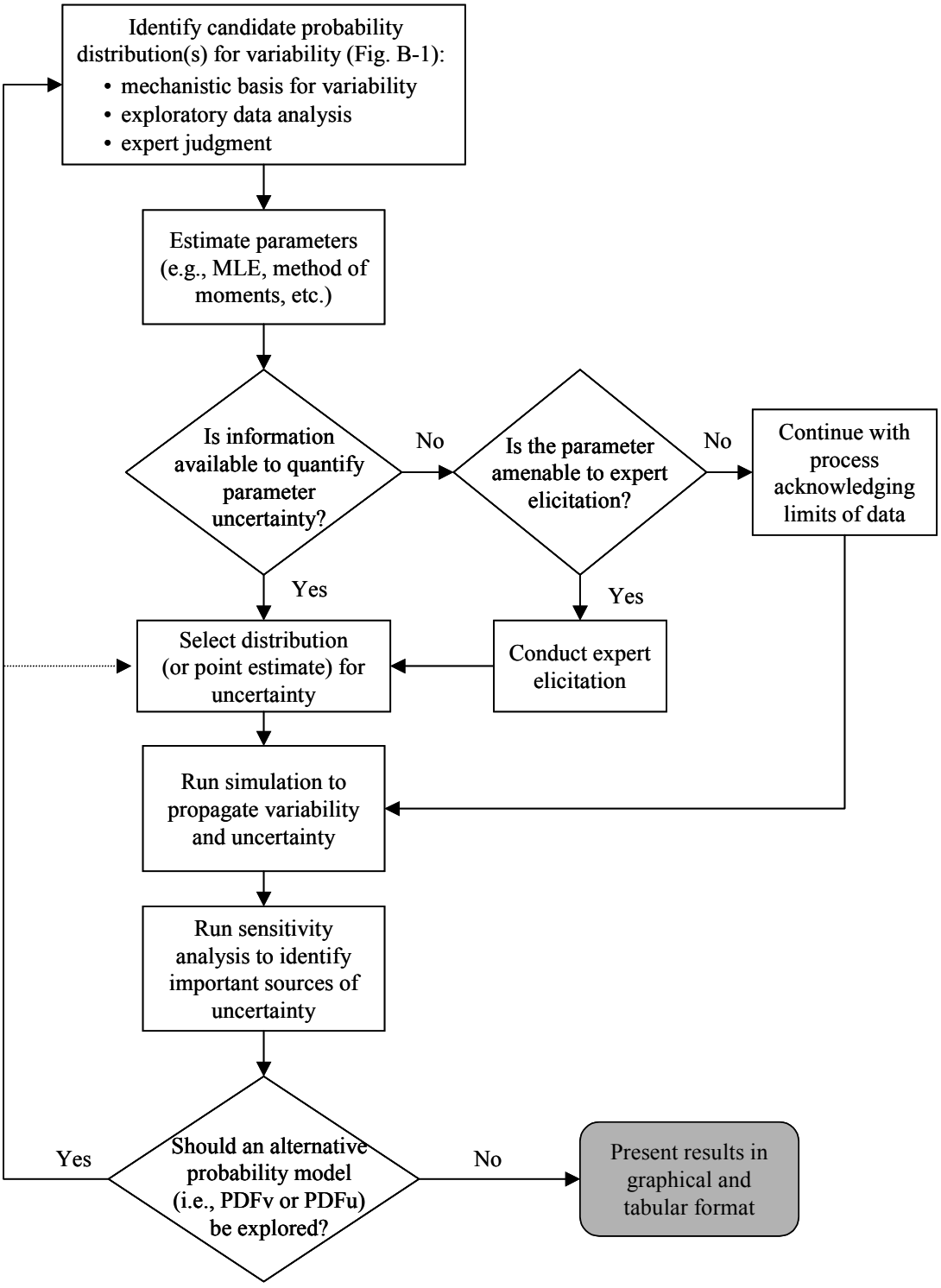


Figure B-2c (page 3 of 3). Detailed conceptual approach for incorporating parameter uncertainty in PRA.

B.3.2 CONSIDERING VARIABILITY AND UNCERTAINTY IN SELECTING AND FITTING DISTRIBUTIONS

Multiple probability distributions may be used to describe variability and uncertainty in an input variable. For example, a normal probability distribution may be selected to characterize variability in body weight, whereas a uniform distribution may selected to characterize uncertainty in the estimate of the arithmetic mean of the normal distribution. The appropriate interpretation and analysis of data for an exposure variable will depend on whether one is specifying a PDFv or PDFu. Figure B-1 outlines one useful process for selecting distributions for variability, whereas Figure B-2 (three pages) outlines a useful process for quantifying both model and parameter uncertainty.

Variability generally refers to observed differences attributable to true heterogeneity or diversity in a population (U.S. EPA, 1997b). Variability results from natural random processes. Inter-individual variability may stem from environmental, lifestyle, and genetic differences. Examples include human physiological variation (e.g., natural variation in body weight, height, breathing rates, drinking water intake rates), changes in weather, variation in soil types, and differences in contaminant concentrations in the environment. Intra-individual variability may reflect age-specific changes (e.g., body weight and height). Variability is not reducible by further measurement or study. A PDF for variability can usually be obtained by fitting a distribution to the sample measurements.

Sources of Uncertainty

Uncertainty generally refers to the lack of knowledge about specific factors, parameters, or models (U.S. EPA, 1997b). Although uncertainty in exposure and risk assessment may be unavoidable due to the necessary simplification of real-world processes, it generally can be reduced by further measurement and study. Parameter uncertainty may stem in part from measurement errors, sampling errors, or other systematic errors in the collection and aggregation of data. Model uncertainty may reflect the simplification of a complex process, a mis-specification of the exposure model structure, a misuse or misapplication of an exposure model, use of the wrong distributional model, and the use of surrogate data or variables. Scenario uncertainty may reflect uncertainty in an exposure model, such as the relevance of specific exposure pathways to the target population. A conceptual exposure model can be used to provide direction in specifying a probability distribution for uncertainty. For example, the concentration term in a Superfund risk assessment typically represents the long-term average concentration to which a receptor is exposed (see Chapter 5). An uncertainty distribution for the concentration term could be developed in part from ideas about the statistical uncertainty of estimating the long-term average from a small sample, and the assumption of random movement of the receptors within a defined exposure unit.

Probability Distributions and Model Uncertainty

This appendix primarily focuses on methods for quantifying uncertainty associated with both the selection of a variability distribution, and estimating parameters of a distribution. A probability distribution can be referred to as a type of model in the sense that it is an approximation, and often a simplified representation of variability or uncertainty that combines both data and judgment. A broader use of the term model refers to a representation of a chemical, physical, or biological process. In risk assessment, many different models have been developed, with varying objectives, major defining and limiting components, and theoretical basis. Figure B-2b provides a general process for exploring model uncertainty of this type. This figure reflects the concepts and spirit of the *Agency Guidance for Conducting External Peer Review of Environmental Regulatory Modeling* (U.S. EPA, 1994). In general, EPA regional risk assessors should be consulted in order to determine the types of exposure and risk models that may be plausible for quantifying exposure at a particular site.

Parameter Uncertainty

Quantifying parameter uncertainty in a probabilistic model typically requires judgment (see Appendix C). When data are uncertain due to, for example, small sample sizes or questionable representativeness (Section B.3.1), Monte Carlo simulation can be a useful tool for demonstrating the effect of the uncertainty on the risk estimates. It is most important to model uncertainty when the sensitive input variables are uncertain. Uncertainty can be quantified in both the point estimate approach (e.g., a range of possible central tendency exposure values) or a probabilistic approach (e.g., a range of possible values for the arithmetic mean of a distribution). While a quantitative uncertainty analysis may complicate a risk management decision by suggesting that risk estimates are highly uncertain, this information can be helpful by focusing additional efforts towards collecting data and reducing uncertainty in the most sensitive input variables. Likewise, if an estimated risk is below a regulatory level of concern, even after quantifying highly uncertain inputs to the exposure model, the risk manager may be more confident in a decision. As emphasized in Figures B-2a, B-2b, and B-2c, risk assessors should generally refrain from setting *ad hoc* probabilities to different candidate distributions in a single Monte Carlo simulation. Instead, this guidance strongly recommends exploring model or parameter uncertainty by running a separate simulation with each candidate model. For example, rather than randomly assigning a beta distribution or a lognormal distribution to an exposure variable for each iteration of a simulation, separate simulations should be run with the candidate probability distributions. Similarly, if a range of temporal or spatial scales is plausible for quantifying exposure, multiple simulations should be designed to demonstrate the importance of these assumptions on the risk estimates.

Uncertainty in parameter estimates may be characterized using a variety of methods. Similar to a PDF for variability, a PDF for parameter uncertainty may be represented by a probability distribution with a unique set of parameters. Sometimes the distribution for uncertainty can be specified by knowing (or assuming) a distribution for variability. For example, if X is a normally distributed random variable, the Student's t distribution and the Chi-square (χ^2) distribution can be used to develop PDFu's for random measurement error uncertainty in the sample mean and variance, respectively. The PDFu for both the Student's *t* and Chi-square distributions is determined by the sample size (*n*). If a PDFu cannot be determined from the PDF for variability, or assumptions regarding the underlying distribution for variability are not supportable, nonparametric or "distribution free" techniques may be used (e.g., bootstrapping). Both parametric and nonparametric techniques may yield confidence intervals for estimates of population parameters.

B.4.0 DO DATA EXIST TO SELECT DISTRIBUTIONS?

Developing site-specific PDFs for every exposure assumption (or toxicity value, in the case of ecological risk) can be time and resource intensive, and in many cases, may not add value to the risk management decision. For those exposure variables that do exert a significant influence on risk, a PDF may be developed from site-specific data, data sets available in the open literature (e.g., EPA's *Exposure Factors Handbook*, U.S. EPA 1997a), or from existing PDFs in the literature (e.g., Oregon DEQ, 1998).

At Superfund sites, perhaps the most common exposure variable that will be described by site-specific data will be the media concentration term. The sample (i.e., collection of empirical measurements) will most often be used to estimate either a point estimate of uncertainty (e.g., an upper confidence limit for the arithmetic mean concentration—the 95% UCL), or a distribution that characterizes the full distribution of uncertainty in the mean. Exposure variables such as ingestion rates, exposure duration, and exposure frequency will most likely be derived from existing PDFs or data sets in the open literature. The Agency supports the development PDFs that may be generally applicable to

different sites (e.g., body weight, water intake, and exposure duration) (U.S. EPA, 1999b, 2001). Until final recommendations of PDFs are available for the more generic exposure variables, PDFs for exposure variables that lack adequate site-specific data will typically be selected from: (1) existing PDFs; (2) data on the entire U.S. population; or (3) data on subsets of the U.S. population that most closely represent the target population at a site. If risks to a sensitive subpopulation, such as young children, elderly adults, ethnic groups, or subsistence fishermen, are a concern at a site, then existing PDFs or data sets that best characterize these subpopulations would be preferable to national distributions based on the entire U.S. population. If adequate site-specific data are available to characterize any of the exposure variables, distributions can be fit to those data.

Uncertainty Associated with Sample Size

An appropriate question to consider when evaluating data sets for use in exposure and risk assessment is, "What sample size is sufficient?" Generally, the larger the sample size (n), the greater one's confidence in the choice of a probability distribution and the corresponding parameter estimates. Conversely, for small n, Goodness-of-fit (GoF) tests (see Section B.6.2) will often fail to reject many of the hypothesized PDFs. In general, there is no rule of thumb for the minimum sample size needed to specify a distribution for variability or uncertainty. Increasing a sample size may be an appropriate option to consider when evaluating risk management strategies to reduce uncertainty.

Statistical sampling, in general, is important to consider when estimating parameters of a probability distribution. One rule of thumb is that the parameters that reflect the central tendency of a distribution (e.g., arithmetic mean, median, mode) can be estimated with greater confidence than parameters that reflect the extremes of the distribution (e.g., 95th percentile). When deciding on appropriate truncation limits (minimum and maximum values), it is unlikely that the statistical sample actually includes the plausible bounds. See Section B.5.7 for more detailed guidance on specifying truncation limits for probability distributions.

B.4.1 WHAT ARE REPRESENTATIVE DATA?

The question, "What is a representative sample?", is important to address when selecting and fitting distributions to data. Many of the factors that may determine representativeness (e.g., sample size and the method of selecting the target, and sample population (Section B.3.1)) are relevant to both point estimate and PRA. EPA's *Guidance for Data Usability in Risk Assessment, Part A* (U.S. EPA, 1992) describes representativeness for risk assessment as the extent to which data define the true risk to human health and the environment.

The goal of representativeness is easy to understand. However, evaluating data to determine if they are representative is more difficult, especially if the problem and decision objectives have not been clearly defined.

The importance of representativeness also varies with the level of complexity of the assessment. If a screening level assessment is desired, for example, to determine if concentrations exceed a health protective exposure level, then representativeness may not be as important as health protectiveness. However, if a complete baseline risk assessment is planned, the risk assessor should generally consider the value added by more complex analyses (e.g., site-specific data collection, sensitivity analysis, and exposure modeling). A tiered approach for making these decisions for a PRA is presented in Chapter 2, and examples of more complex analyses are presented in Appendix D. In addition, the Agency (U.S.

EPA, 1999a) summarizes the advantages and weaknesses of proposed checklists for risk assessors to evaluate representativeness of exposure factors data.

For purposes of this guidance, a surrogate study is one conducted on a sampled population that is similar to, but not a subset of, the target population. When using surrogate data, the risk assessor should generally exercise judgment about the representativeness of the data to the target population. For example, the distribution of body weights of deer mice from two independent samples from similar ecosystems may differ depending on the age structure, proportion of males and females, and the time of year that the samples were obtained. When in doubt about which study results to use in defining a probability distribution, one option is to develop a distribution and calculate risks with each sample independently, and compare the results. This approach can be a simple, but effective type of uncertainty analysis. At a minimum, uncertainties associated with the use of surrogate studies should be discussed in the assessment.

In many cases, the surrogate population shares common attributes with the target population, but is not truly representative. The risk assessor should then determine the importance of the discrepancies and whether adjustments can be made to reduce those differences. There are a wide variety of methods that can be used to account for such discrepancies, depending on the available information. Summary statistics (e.g., as presented by the *Exposure Factors Handbook*, U.S. EPA, 1997a) can be used to estimate linear characteristics of the target population from the sample population. For example, if the mean, standard deviation, and various percentiles of the sample population are known, then the mean or proportion exceeding a fixed threshold can be calculated using a simple weighted average. Adjustment options are more numerous if the risk assessor has access to the raw data. Adjustments for raw data include: weighted averages, weighted proportions, transformations, and grouping of the data based on the available information (e.g., empirical data, and professional judgment).

In most cases, the evaluation of data representativeness will necessarily involve judgment. The workplan should generally include a description of the data, the basis for the selection of each distribution, and the method used to estimate parameters (see Chapter 2). Empirical data (i.e., observations) are typically used to select distributions and derive parameter estimates. However, it may be necessary to use expert judgment or elicitation in cases where the quality or quantity of available data are found to be inadequate.

B.4.2 THE ROLE OF EXPERT JUDGMENT

Expert judgment refers to inferential opinion of a specialist or group of specialists within an area of their expertise. When there is uncertainty associated with an input variable, such as a data gap, expert judgment may be appropriate for obtaining distributions. Note that distributions elicited from experts reflect individual or group inferences, rather than empirical evidence. Distributions based on expert judgment can serve as Bayesian priors in a decision-analytic framework. The distributions and Bayesian priors can be modified as new empirical data become available. There is a rich literature base regarding the protocol for conducting expert elicitations and using the results to support decisions (Morgan and Henrion, 1990). Elicitation of expert judgment has been used to obtain distributions for risk assessments (Morgan and Henrion, 1990; Hora, 1992; U.S. EPA, 1997b) and for developing air quality standards (U.S. EPA, 1982).

Bayesian analysis is a statistical approach that allows the current state of knowledge, expressed as a probability distribution, to be formally combined with new data to reach an updated information state. In PRA, Bayesian Monte Carlo analysis (Bayesian MCA) can be used to determine the reduction in

uncertainty arising from new information. When combined with techniques from decision analysis, Bayesian MCA can help to determine the type and quantity of data that generally should be collected to reduce uncertainty. The benefits and limitations of expert elicitation, Bayesian statistics, Bayesian MCA, and decision analysis (i.e., value of information [VOI]), as applied to PRA, are discussed in greater detail in Appendix D.

EXHIBIT B-3

FACTORS TO CONSIDER IN SELECTING A PROBABILITY DISTRIBUTION*

- *Is there a mechanistic basis for choosing a distributional family?*
- *Is the shape of the distribution likely to be dictated by physical or biological properties or other mechanisms?*
- *Is the variable discrete or continuous?*
- *What are the bounds of the variable?*
- *Is the distribution skewed or symmetric?*
- *If the distribution is thought to be skewed, in which direction?*
- *What other aspects of the shape of the distribution are known*?
- *How well do the tails of the distribution represent the observations?*

*Source: U.S. EPA, 1997b

B.5.0 FITTING DISTRIBUTIONS TO DATA

Sometimes more than one probability distribution may adequately characterize variability or uncertainty. The choice of a distribution should be based on the available data and on knowledge of the mechanisms or processes that result in variability. In general, the preferred choice is the simplest probability model that adequately characterizes variability or uncertainty and is consistent with the mechanism underlying the data. For example, a log-logistic distribution would not necessarily be selected over a 2-parameter lognormal distribution simply because it was ranked higher in a GoF test by a statistical software package. Some distributions (e.g., normal, lognormal) are well known among risk assessors. The statistical properties for these distributions are well understood and the formal descriptions can often be brief.

Important factors to consider in selecting a PDF are described in Exhibit B-3. An initial step in selecting a distribution should be to determine if the random variable is discrete or continuous. Continuous variables take any value over one or more intervals and generally represent measurements (e.g., height, weight, concentration). For a continuous variable, a mathematical function generally describes the probability for each value across an interval. Discrete variables take either a finite or *countably infinite* number of values. Unique probabilities are assigned to each value of a discrete variable. The number of rainfall events in a month is an example of a discrete random variable, whereas the amount of rainfall is a continuous variable. Similarly, the number of fish meals per month is a discrete variable, whereas the average size (mass) of a fish meal is continuous.

Another important consideration is whether there are plausible bounds or limits for a variable. For example, it is highly unlikely that an American adult will weigh less than 30 kg or more than 180 kg. Most exposure variables may assume any nonnegative value within a plausible range. Therefore, distributions will generally be truncated at a minimum of zero (or higher), or a probability distribution that is theoretically bounded at a nonzero value may be specified (see Table B-3). A more detailed discussion of factors to consider in selecting a PDF and specifying parameter values is provided below.

B.5.1 CONSIDERING THE UNDERLYING MECHANISM

There may be mechanistic reasons depending on known physical or biological processes that dictate the shape of the distribution. For example, normal distributions result from processes that sum random variables whereas lognormal distributions result from multiplication of random variables. A Poisson distribution is used to characterize the number of independent and randomly distributed events in a unit of time or space. An exponential distribution would describe the inter-arrival times of independent and randomly distributed events occurring at a constant rate. If, instead, the elapsed time until arrival of the k^{th} event is of interest, then the appropriate probability distribution would be the gamma distribution (Morgan and Henrion, 1990).

> ☞ *In all cases, it is incumbent on the risk assessor to explain clearly and fully the reasoning underlying the choice of a distribution for a given exposure variable—primarily from a mechanistic standpoint if possible.*

Table B-2 lists some of the probability distributions that may commonly be used in PRA. This is not an exhaustive list, and the scientific literature contains numerous examples with alternative distributions. Where practicable, a mechanistic basis is presented for the choice of the distribution. For some distributions, such as beta, triangular, and uniform, a mechanistic basis is not offered because it is unlikely that a chemical or biological process will yield a random variable with that particular shape. Nevertheless, such distributions may be appropriate for use in PRA because they reflect the extent of information that is available to characterize a specific random variable. Preliminary distributions are discussed in Section B.2.0 and Table B-4. Because many of the distributions given in Table B-2 can assume flexible shapes, they offer practical choices for characterizing variability.

Table B-2 also illustrates probability distributions (both PDFs and CDFs) commonly used in PRA. While intuitively appealing, identifying a mechanistic basis for a distribution can be difficult for many exposure variables; however, it may be relatively apparent that the variable is bounded by a minimum (e.g., ingestion rate $\geq$ 0 mg/day) and a maximum (e.g., absorption fraction $\leq$ 100%), or that the relevant chance mechanism results in a discrete distribution rather than a continuous distribution, as described above.

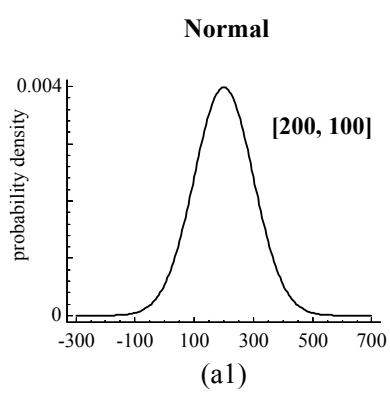
For each distribution, one or more examples with different parameter estimates are given to demonstrate the flexibility in the shape of the PDF. In addition to the descriptions of the distributions in Tables B-2, Table B-3 provides a summary of the parameters and theoretical bounds that define the PDFs. For a further discussion of characteristics of PDFs see Thompson, 1999. Figures (a-h) immediately following Table B-2 present examples of PDFs and the corresponding CDFs for distributions commonly used in PRA.

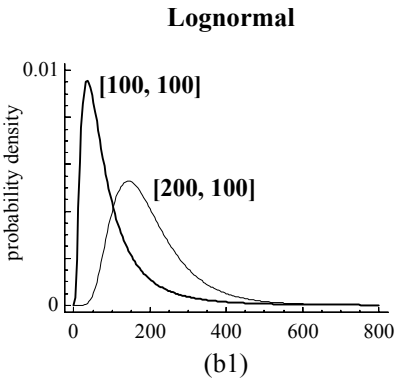
Table B-2. Examples of Selected Probability Distributions for PRA.

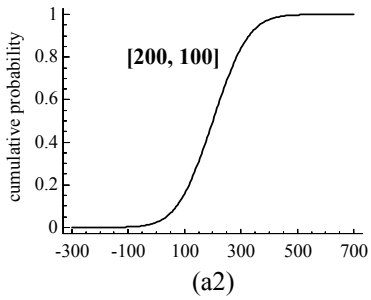
Distribution	Mechanistic Basis	Example(s)
Beta Figure (e)	Describes a continuous random variable with finite upper and lower bounds. This distribution can take on very flexible shapes, but generally does not have a mechanistic basis.	Absorption fraction bounded by 0 and 100%; fraction of time an individual spends indoors.
Binomial	Describes a discrete random variable produced by processes that: (1) occur in a fixed number *n* of repeated independent "trials"; (2) yield only one of two possible outcomes (e.g., "success" or "failure") at each trial; and (3) have constant probability *p* of "success". A binomial distribution is characterized by parameters *n*, *p*, and *x*, representing the number of trials, the probability of success of each trial, and the number of successes, respectively.	The number of animals with tumors (or some other quantitative outcome) in a chronic animal bioassay.
Exponential Figure (h)	If instead of counting the number of events in the Poisson process (below), one measures the time (or distance) between any two successive, random, independent events.	The length of time between two radiation counts; length of time between major storm events; distance between impact points of two artillery shells.
Gamma Figure (g)	Similar to exponential except that time until occurrence of the k^{th} event in the Poisson process is measured (rather than time between successive events). Reduces to exponential when k=1.	Time until k^{th} radiation count; elapsed time until k^{th} major storm event.
Lognormal Figure (b)	Multiplication of a large number of random variables, or equivalently adding the logarithms of those numbers, will tend to yield a distribution with a lognormal shape.	Chemical concentrations in environmental media; media contact rates; rates and flows in both fate and transport models. Because the basic risk equation is multiplicative, distributions of risk are generally lognormal. In practice, lognormal distributions often provide good fits to data on chemical concentrations in a variety of media (Gilbert, 1987; Ott, 1990).
Normal Figure (a)	Addition of independent random variables, with no one variable contributing substantially to the total variation of the sum, will tend to yield a distribution with a normal shape. This result is established by the central limit theorem.	The "Gaussian Plume Model" for the dispersion of air pollutants is based on the idea that, at a micro level, individual parcels of air, or molecules of pollutants, are subject to many random collisions from other molecules that act together as if a large number of random numbers were being added/subtracted from an initial 3-dimensional description of a position.
Poisson	Observed when counting the frequency of discrete events, where the events are independent of one another, and randomly distributed in space or time. Approximates the binomial distribution when sample size, *n*, is large and probability, *p*, is small.	The number of counts of radiation that occur in a particular time interval; the release of synaptic transmitter from nerve cells; the number of artillery shells falling within a fixed radius; the occurrence of major storm events in a month; number of leaks in average length of pipe.

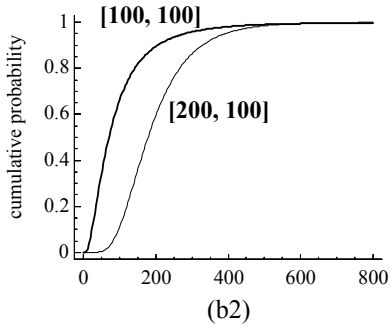
Table B-2. Examples of Selected Probability Distributions for PRA.

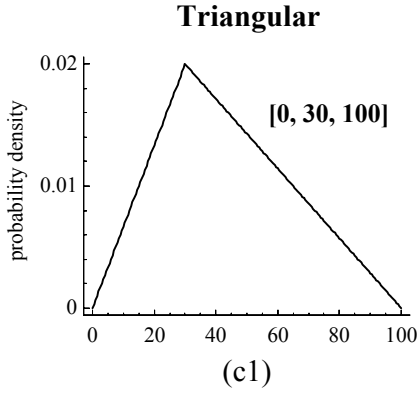
Distribution	Mechanistic Basis	Example(s)
Triangular Figure ©)	The PDF is shaped like a triangle, with parameters representing plausible bounds and a most likely value (i.e., mode). This is a "rough" probability model that generally describes a random variable based on limited information rather than mechanistic basis.	Variability in shower droplet diameter. Uncertainty in the mean air exchange rate in a shower.
Uniform Figure (d)	The PDF is shaped like a rectangle, with parameters representing plausible bounds. This is a "rough" probability model that generally describes a random variable based on limited information rather than a mechanistic basis.	Variability in the air ventilation rate in a house.
Weibull Figure (f)	Originated in reliability and (product) life testing as a model for time to failure or life length of a component when the failure rate changes with time. A very flexible model taking a wide range of shapes. If the failure rate is constant with time, the Weibull reduces to the exponential distribution.	Examples for exponential and gamma would also be appropriate for Weibull.

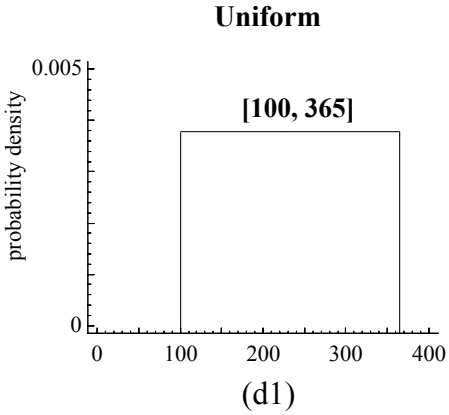

Normal
probability density
0.004
0
[200, 100]
-300 -100 100 300 500 700
(a1)

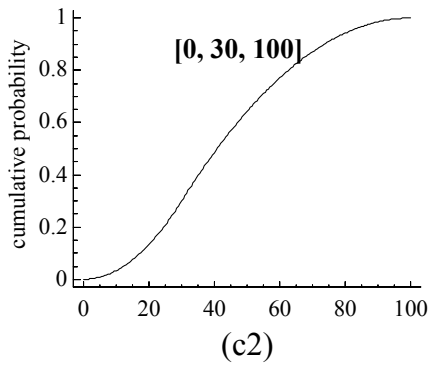

Lognormal
probability density
0.01
0
[100, 100]
[200, 100]
0 200 400 600 800
(b1)

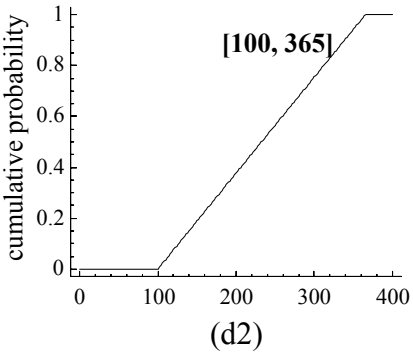

cumulative probability
1
0.8
0.6
0.4
0.2
0
[200, 100]
-300 -100 100 300 500 700
(a2)


cumulative probability
1
0.8
0.6
0.4
0.2
0
[100, 100]
[200, 100]
0 200 400 600 800
(b2)

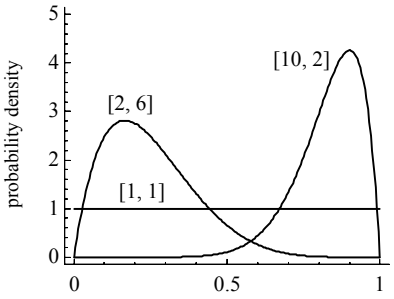

Triangular
probability density
0.02
0.01
0
[0, 30, 100]
0 20 40 60 80 100
(c1)


Uniform
probability density
0.005
0
[100, 365]
0 100 200 300 400
(d1)

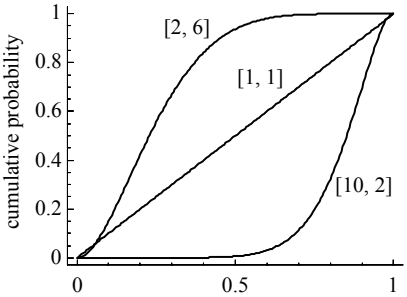

cumulative probability
1
0.8
0.6
0.4
0.2
0
[0, 30, 100]
0 20 40 60 80 100
(c2)


cumulative probability
1
0.8
0.6
0.4
0.2
0
[100, 365]
0 100 200 300 400
(d2)

Beta

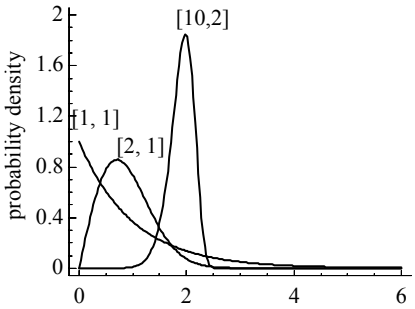


(e1)

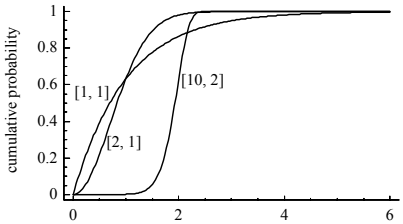


(e2)

Weibull

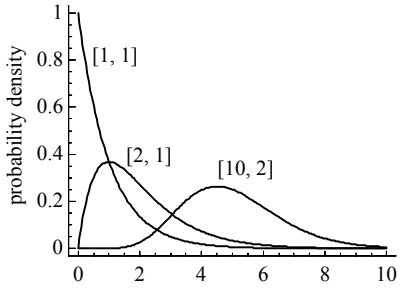


(f1)

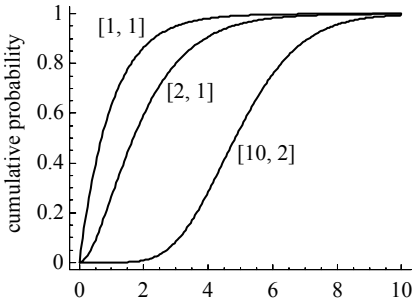


(f2)

Gamma

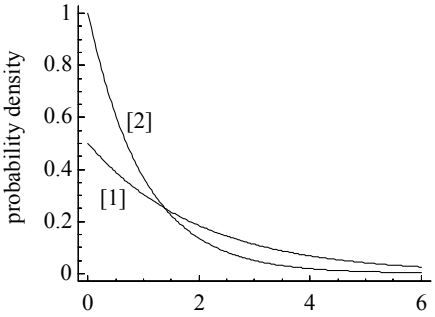


(g1)

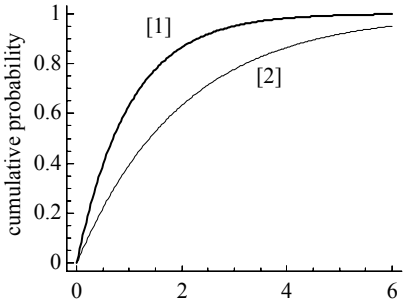


(g2)

Exponential



(h1)



(h2)

B. 5.2 EMPIRICAL DISTRIBUTION FUNCTIONS (EDFS)

In some cases, an empirical distribution function (EDF) may be preferred over fitting the data set to a hypothesized distribution. EDFs, also called empirical cumulative distribution functions (ECDF), provide a way to use the data itself to define the distribution of the relevant variable. Briefly, an EDF for a random variable is described by a step function based on the frequency distribution of observed values. An EDF for a continuous random variable may be linearized by interpolating between levels of the various bins in a frequency distribution. The CDF for a linearized EDF appears as a line, rather than steps. Example B-3 at the end of this Appendix illustrates an EDF, linearized EDF, and beta distribution (α_1=0.63, α_2=2.85, rescaled to min=0, max=364) fit to percentile data for soil ingestion rates in children (Stanek and Calabrese, 1995). A plausible range (i.e., minimum and maximum values) was imposed on the data set for this example.

EXHIBIT B-4

VARIATIONS OF THE EDF

Linearized - Linearly interpolates between two observations, yielding a linearized cumulative distribution pattern.

Extended - In addition to linearizing (see above), adds lower and upper bounds based on expert judgment.

Mixed Exponential - Adds an exponential upper and/or lower tail to the EDF.

EDFs provide a complete representation of the data with no loss of information. They do not depend on the assumptions associated with estimating parameters for theoretical probability models. EDFs are designed to provide direct information about the shape of the distribution, which reveals skewness, multimodality, and other features of the data set. However, EDFs may not adequately represent the tails of a distribution due to limitations in data acquisition. In the simplest case, an EDF is constrained to the extremes of the data set. This may be an unreasonable restriction if limiting the EDF to the smallest and largest sample values is likely to greatly underestimate the distributional tails. If this is an important source of uncertainty, the risk assessor may choose to extend the tails of the distribution to plausible bounds or to describe the tails with another distribution (see Exhibit B-4). For example, an exponential distribution may be used to extend the tails based on the last 5% of the data. This method is based on extreme value theory, and the observation that extreme values for many continuous, unbounded distributions follow an exponential distribution (Bratley et al., 1987). As with other probability models, uncertainty in the plausible bounds of an EDF may be reduced by obtaining additional information.

Advantages and disadvantages of using EDFs in PRA are discussed in detail in the *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999a).

B.5.3 GRAPHICAL METHODS FOR SELECTING PROBABILITY DISTRIBUTIONS

Graphical methods can provide valuable insights and generally should be used in conjunction with exploratory data analysis. Examples of graphical methods are frequency distributions (i.e., histograms), stem-and-leaf plots, dot plots, line plots for discrete distributions, box-and-whisker plots, and scatter plots (Tukey, 1977; Conover, 1980; Morgan and Henrion, 1990).

☞ *Graphical methods are invaluable for exploring a data set to understand the characteristics of the underlying population*.

Together with statistical summaries, graphical data summaries can reveal important characteristics of a data set, including skewness (asymmetry), number of peaks (multi-modality), behavior in the tails, and data outliers.

Frequency Distribution or Histogram

The frequency distribution, or histogram, is a graphical approximation of the empirical PDF. Frequency distributions can be plotted on both linear and log scales. The general strategy for selecting the number of bins to partition the data is to avoid too much smoothing and too much jaggedness. Equation B-1 (U.S. EPA, 1999a) provides a starting point for estimating the number of bins based on the sample size (n).

$$Number of\ Bins = 1 + 3.322 \log_{10} n \qquad \text{Equation B-1}$$

Probability Plotting

Another method that may be used to visualize distributions and estimate parameters is probability plotting, also referred to as linear least squares regression or regression on ordered statistics. This technique involves finding a probability and data scale that plots the CDF of a hypothesized distribution as a straight line. The corresponding linearity of the CDF for the sample data provides a measure of the GoF of the hypothesized distribution. The general approach involves sorting the sample data in ascending order and converting the ranks to percentiles. The percentile value for the i^{th} rank is calculated according to Gilbert (1987) as:

$$Percentile = 100 \times \frac{i - 0.5}{n} \qquad \text{Equation B-2}$$

An alternative formula is provided by Ott (1995):

$$Percentile = 100 \times \frac{i}{n + 1} \qquad \text{Equation B-3}$$

Plotting positions given by Equations B-2 and B-3 are special cases of the more general formula given by Equation B-4 (Helsel and Hirsch, 1992):

$$Percentile = 100 \times \frac{i - a}{n + 1 - 2a} \qquad \text{Equation B-4}$$

where a is a constant that varies from 0 (Equation B-3) to 0.5 (Equation B-2).

The percentiles are used to calculate the *z-scores*, which represent the number of standard deviations away from the mean that a particular datum lies assuming the data are normally distributed. For normal distributions, the data are plotted against the z-scores; for lognormal distributions, the data are log-transformed and plotted against the z-scores. In both cases, parameters of the distribution can be estimated from the least-squares regression line. When the hypothesized distribution is a poor fit to the data, p-plots can yield misleadingly low estimates of the standard deviation (Cullen and Frey, 1999). Both Gilbert (1987) and Ott (1995) provide excellent descriptions of the use of probability plotting to derive parameter estimates for a given distribution. Probability plotting techniques with best-fit lines have been used to estimate parameters for a wide variety of distributions, including beta, Weibull, and gamma.

Cullen and Frey (1999) point out that probability plotting may not be a primary choice for selecting a fitting distributions because the method violates an important assumption of least squares regression—independence of the observations (see Appendix A, Exhibit A-5). This is because the rank-ordered data are no longer independent. Nevertheless, this approach may yield good results when the fit is good and the choice of distributions is somewhat subjective.

B.5.4 PARAMETER ESTIMATION METHODS

As a rule, there are often a number of different methods available for estimating a given parameter. The most appropriate method to apply may require judgment, depending on the relative difficulty in applying a method for a particular parameter, as well as the desired statistical properties of the method. The following simple example provides a useful analogy. Suppose that the parameter of interest, A, is the total area of an approximately square exposure unit. If the exposure unit is a perfect square, and the length of one side (L_1) is known, the area would be equal to L_1^2 (i.e., for a square, $A=L_i^2$). Suppose L is unknown, but two independent measurements, X_1 and X_2, are available to estimate the length (see Exhibit B-5). If it is assumed that the random variable, L, has a probability distribution with mean μ, then the area of the square piece of property is $A=\mu^2$. What is a reasonable estimate of the area (i.e., $\hat{A} = \hat{\mu}^2$) based on X_1 and X_2? Three plausible methods for calculating $\hat{\mu}^2$ are given below.

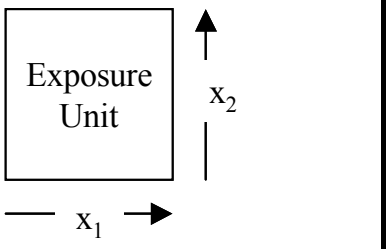
1. $\hat{\mu}_a^2 = \left(\frac{X_1 + X_2}{2}\right)^2$

2. $\hat{\mu}_b^2 = \frac{X_1^2 + X_2^2}{2}$

3. $\mu_c^2 = X_1 \times X_2$

EXHIBIT B-5

ESTIMATING THE AREA OF A HYPOTHETICAL EXPOSURE UNIT



Because these three estimators will, as a rule, give different answers, it may be useful to set criteria for selecting which one gives the "best" answer. Some of the statistical criteria that are used for this purpose are *consistency*, *efficiency, robustness, sufficiency*, and *unbiasedness* (see Exhibit B-6). It turns out, each method is relatively easy to implement, but the third method is preferred because it is a more efficient estimator.

In many cases, particularly if a model is complex, potential estimators of the unknown parameters are not readily apparent. To assist in developing estimators, several general methods have been developed. Exhibit B-7 lists some of the more common parameter estimation methods.

Perhaps the simplest method is the method of matching moments (MoMM), also called the method of moments. MoMM is appropriately named, as it involves expressing the unknown parameters in terms of population moments and then "matching", or equating the sample moments to the population

moments. For example, the sample mean ($\bar{x}$) and standard deviation (*s*) are estimators for the corresponding population parameters (μ and σ).

Maximum Likelihood Estimation (MLE) is a commonly applied method, that is often thought of as a parameter estimate for which the observed data are most "likely". The likelihood function is defined for independent continuous random variables as follows:

$$L(\theta_1, \theta_2,...\theta_k) = \prod_{I=1}^{n} f(x_1|\theta_1, \theta_2, ..., \theta_k)$$

The likelihood function is evaluated based on the product of the PDF for each value of x. The parameters of the probability model, (θ_k), are chosen to maximize the likelihood function value and thereby are most likely to produce the sample data set (Cullen and Frey, 1999).

EXHIBIT B-6

CRITERIA FOR EVALUATING PARAMETER ESTIMATION METHODS*

Consistency	A consistent estimator converges to the "true" value of the parameter as the number of samples increases.
Efficiency	An efficient estimator has minimal variance in the sampling distribution of the estimate.
Robustness	A robust estimator is one that works well even if there are departures from the assumed underlying distribution.
Sufficiency	A sufficient estimator is one that makes maximum use of information contained in a data set.
Unbiasedness	An unbiased estimator yields an average value of the parameter estimate that is equal to that of the population value.

*Source: Cullen and Frey, 1999

It has also been demonstrated that MLE yields estimators that generally have good properties when evaluated by the criteria listed above. In some cases (e.g., for smaller sample sizes), these estimators are not *unbiased*; however, this can often be accounted for by "adjusting" the estimator. A familiar example of this adjustment is in estimation of the variance of a normal distribution. The MLE for the variance is biased by a factor of ((*n*-1)/*n*), but this is easily corrected by multiplying the MLE by (*n*/(–1)). For some distributions, calculations of the MLE are straightforward. For example, MLE for parameters of a normal distribution are given by the mean and standard deviation of the sample data, the same as MoMM. MLE for parameters of a lognormal distribution are given by the mean and standard deviation of the log-transformed data, which is different from MoMM. In general, MLE calculations are complex, and commercial software such as *@Risk* and *Crystal Ball*® may be used. A more detailed discussion of the derivation and properties of MoMM and MLE can be found in the statistics literature (e.g., Chapter 5 of Mood and Graybill, 1963; Chapter 9 of Mendenhall and Scheaffer, 1973; Section 6.5 of Law and Kelton, 1991; Section 5.6 of Cullen and Frey, 1999).

EXHIBIT B-7

PARAMETER ESTIMATION METHODS

- Method of Matching Moments
- Maximum Likelihood
- Minimum Chi-Square
- Weighted Least-Squares

B.5.5 DEALING WITH CORRELATIONS AMONG VARIABLES OR PARAMETERS

Correlations between exposure variables or between parameters of the probability distribution may be important components of a probabilistic model. Correlation is a measure of association between two quantitative random variables. Two random variables may either be positively or negatively correlated. A positive correlation exists between two variables if the value of X_1 increases as the value of X_2 increases. For example, higher hand dust lead levels have been associated with higher pediatric blood lead levels (Charney et al., 1980). A negative correlation exists between two variables if the value of X_1 increases as the value of X_2 decreases. For example, studies suggest the ingestion of soil and dust particles increases as particle size decreases (Calabrese et al., 1996).

A first step in identifying correlations is to assess the possible physical and statistical relationships that exist between variables. In an ecological risk assessment (ERA), for example, the largest surf scoter (diving duck) does not consume the least amount of food, nor does the smallest surf scoter consume the greatest amount of food. Random sampling of body weight and ingestion rate as separate parameters, however, allows for these two possibilities. Neglecting a correlation between two variables may restrict (underestimate) the tails of the ecological Hazard Quotient (HQ) for each chemical of concern (COC), which are frequently the areas of the distribution of most interest.

The degree to which correlations affect the output of a risk model depends on: (1) the strength of correlations between the two variables, and (2) the contribution of the correlated variables to overall variance in the output (Cullen and Frey, 1999). Therefore, it is useful to conduct a preliminary sensitivity analysis to assess the impact of alternative correlation assumptions on the model output. If the impact is significant, correlations should be identified and accounted for in the PRA.

There are several approaches to account for dependencies in MCA including: (1) modifying the model to include the correlation; and (2) simulating dependence between variables for sample generation (Cullen and Frey, 1999). Modifying the model is preferred as simulation techniques cannot capture the full complexity between model inputs. However, when this is not possible, dependencies between variables can be simulated and approximated by correlation coefficients and bivariate normal distributions.

Correlation coefficients are a numerical measure of the strength and direction of the relationship between two variables. Sample correlation coefficients measure the linear relationship between variables. However, if two variables are from different probability distributions, it is unlikely that they are linearly related. Consequently, simulation software programs such as *Crystal Ball*® and *@Risk* can be used to calculate and employ the nonparametric statistic, Spearman's Rank Correlation Coefficients (Rho) in simulating correlation between inputs. Rank Correlation Coefficients measure the linear dependence not of the data values themselves, but of the rank value of the data. The ranks indicate relative positions in an ordered series, not the quantitative differences between the positions. The disadvantage of losing information by using the rank values (rather than the actual values) is offset by the ability to correlate random variables from different distribution types (See Appendix A).

Exhibit B-8 gives an example of a straightforward approach to specifying a rank correlation between two input variables in a one-dimensional Monte Carlo analysis (1-D MCA) for variability. A range of correlations is explored as a form of uncertainty analysis on the distribution of intakes given a fish advisory of 7.0 μg/day for a chemical.

EXHIBIT B-8

CORRELATION OF INPUT VARIABLES FOR 1-D MCA OF VARIABILITY

Intake Equation Intake = (CF x IR x FI x EF x ED)/(BW x AT)

Variables	Description and Units	Units	Point Estimate or PDFv
CF	concentration in fish	ug/kg	25
IR	fish ingestion rate	kg/meal	lognormal (0.16, 0.07)[1]
FI	fraction ingestion from source	unitless	1.0
EF	exposure frequency	meals/yr	lognormal (35.5, 25.0)[1]
ED	exposure duration	years	30
BW	body weight	kg	70
AT	averaging time	days	10950

[1]Lognormal PDF parameters: arithmetic mean, standard deviation

- Correlation between IR and EF is suggested by Burger et al. (1999) study of 250 anglers on the Savannah River, South Carolina. Moderate correlation (Kendall's tau=0.17, p=0.04)
- Uncertainty Analysis: 1-D MCA simulations of variability correlating IR and EF using *Crystal Ball® 2000* (5,000 iterations, Latin Hypercube sampling). Spearman rank correlations: 0.10, 0.50, 0.90

Statistics of PDFv for Intake (ug/day) compared to Fish Advisory of 7.0 ug/day

Rank Correlation (r)	0.10	0.50	0.90
Intake Statistics (ug/day)			
mean	1.6	1.8	2.0
50th percentile	1.1	1.1	1.1
95th percentile	4.4	5.4	6.5
97.5th percentile	5.7	7.0	9.0

- For this example, only IR and EF are characterized by PDFs. They contribute approximately equally to the distribution of intakes. Positive rank correlations have little effect on the median (50th percentile) of the output distribution, but tend to widen the tails of the distribution. Increasing the correlation from 0.10 to 0.90 increases the 90th percentile from 4.4 to 6.5 ug/day, and the 97.5th percentile from 5.7 to 9.0 ug/day.
- If the fish advisory is 7.0 ug/day, uncertainty in the correlation coefficient may have important consequences for the risk management decision.

Correlations may also be specified for parameters of a probability distribution. This is an important concept when designing a two-dimensional Monte Carlo analysis (2-D MCA) in which parameters of the same PDFv might be otherwise be described by independent PDFu's. A common approach for correlating two parameters is to specify a bivariate normal distribution (Nelsen, 1986, 1987; Brainard and Burmaster, 1992). A bivariate normal distribution allows for the distribution of one variable to be sampled conditional on the other. This is a special case of a joint distribution in which both x and y are random variables and normally distributed (as the conditional distribution of x or of y is always normal) (Wonnacott and Wonnacott, 1981). Example B-4 further explains bivariate normal distributions and demonstrates this approach applied to coefficients of a simple linear regression model that relates contaminant concentrations in soil and dust.

The results of correlation analysis should be interpreted with caution. Two variables may be associated due to: (1) a dependency between the two variables; (2) chance (two independent variables appear dependent due to chance in the sampling procedure); and (3) variables not included in the analysis (lurking variables) are affecting the two variables being analyzed. Likewise, a low correlation measure does not necessarily mean the two variables are independent. As a lurking variable may cause the appearance of an association between the two independent variables, it may also mask the association between two dependent variables.

☞ *Correlation describes a degree of mathematical association, not a causal relationship between the two variables.*

Efforts to extrapolate or predict correlations outside the range of observed values should also be done with caution. For example, there may be a strong linear relationship between age and height in children; however, it would be inappropriate to apply this correlation to adults. Additional caution is needed when correlating more than two factors at a time. In general, because of the complexity of specifying a valid covariance matrix when correlating more than two factors at a time, risk assessors may need to consult a statistician to avoid generating misleading risk estimates.

B.5.6 CENSORED DATA

In order to define the exposure point concentration, estimates of summary statistics representative of the entire distribution of data are needed (Helsel and Hirsch, 1992). Censored data complicate the process of selecting and fitting PDFs and estimating parameter estimates. A censored data set is a data set for which measurements above or below a certain threshold are not available. Left censored data occurs frequently at Superfund sites, where samples for a number of chemicals are often below the reporting limit. A censored datum (often denoted by ND) commonly represents a value of half of the laboratory reporting limit.

Three general methods for estimating summary statistics for left censored data sets include: (1) simple substitution; (2) distributional methods; and (3) robust methods (Helsel and Hirsch, 1992). These methods may be evaluated based on the root mean squared error (RMSE) estimate, a measure of the difference between the sample statistic (e.g., the sample mean, $\bar{x}$) and the true population parameter (e.g., population mean, μ).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \frac{(\bar{x} - \mu)^2}{\mu}}{N}}$$

Methods which yield estimates closer to the true parameter value have lower bias, higher precision, and lower RMSEs.

Simple Substitution Methods

Simple substitution methods entail substituting values equal to or lower than the reporting limit in the data set. These surrogate values are then included in the calculation of the summary statistics and in determining the distributional shape of the data set. Although this method is frequently used, it is important to understand its limitations; depending on the surrogate value used (e.g., half the reporting limit) the simple substitution method may yield biased parameter estimates (e.g., low estimates of the mean) and may yield misleading distributional shapes. Studies such as those reported by Gilliom and Helsel (1986) have determined, in terms of the RMSE, that simple substitution methods perform more poorly than the distributional and robust methods described below.

Distributional Methods

With distributional methods, the entire data set is assumed to follow a theoretical distribution (e.g., normal distribution). Assuming a theoretical distribution, MLE and probability plotting (p-plot) methods provide summary statistics that best match the reported values of the data and the percentage of samples below the threshold value. If the data fit the theoretical distribution exactly, or if the sample size is large, both MLE and p-plots are unbiased methods. Often, however, the sample size is small and the distribution deviates from a theoretical distribution. In this case, the MLE and p-plot methods may yield biased and imprecise methods (Hesel and Hirsch, 1992).

Robust Methods

With robust methods, a theoretical distribution is needed. A theoretical distribution is fit to the data above the detection limit by MLE or p-plot methods. Based on this assumed PDF, the value of the data points below the detection limit are extrapolated and used in the summary statistics calculation. Unlike the simple substitution method, these extrapolated values are not estimates for the data points; rather, they are only used jointly to calculate summary statistics (Hesel and Hirsch, 1992). The method is considered robust as it uses the actual values of the sample data, rather than the distribution above the detection limit.

B.5.7 TRUNCATION

Truncation refers to imposing a minimum and/or maximum value on a probability distribution. The main purpose of truncation is to constrain the sample space to a set of "plausible values". For example, a probability distribution for adult body weight might be truncated at a minimum value of 30 kg and a maximum value of 180 kg in order to avoid the occasional selection of an unlikely value (e.g., 5 or 500 kg). Given the subjectiveness involved in selecting truncation limits, such choices should clearly be made with caution, and involvement of stakeholders who may be aware of site-specific circumstances.

For example, there may well be individuals who weigh more than 180 kg and less than 30 kg. The purpose for truncating the tails of a distribution is to confine each risk estimate of a Monte Carlo simulation to a combination of plausible input values. The advantage of truncating unbounded probability distributions in PRA is that central tendency and high-end risk estimates will not be biased by unrealistic values. The disadvantage is that the original parameter estimates of the nontruncated distribution are altered by constraining the sample space. The bias in the parameter estimates increases as the interval between the minimum and maximum truncation limit is reduced. For example, a normal distribution with an arithmetic mean of 100 may be fit to a data set; imposing a truncation limit of 300 may result in a truncated normal distribution with an arithmetic mean of 85. The relationship between the truncated and nontruncated parameter estimates can be determined analytically (Johnson et al., 1995) or approximated using Monte Carlo simulations under both truncated and nontruncated scenarios.

Table B-3. Theoretical bounds and parameter values for selected distributions.

Probability Distribution	Parameters[1]	Theoretical Bounds
Normal	(μ, σ)	$(-\infty, +\infty)$
Lognormal	(μ, σ)	$[0, +\infty)$
Weibull	(α, β)	$[0, +\infty)$
Exponential	(β)	$[0, +\infty)$
Gamma	(α, β)	$[0, +\infty)$
Beta	$(\alpha_1, \alpha_2, a, b)$	[a, b]
Uniform	(a, b)	[a, b]
Triangular	(a, m, b)	[a, b]
Empirical (bounded EDF)	(a, b, {x}, {p})	[a, b]

[1]a=minimum, b=maximum, μ=mean, σ=standard deviation, m=mode,
α=shape parameter, β=scale parameter, x=value, p=probability

Truncation is typically considered when using unbounded probability distributions (e.g., normal, lognormal, gamma, Weibull) to characterize variability. Table B-3 gives the theoretical bounds for selected probability distributions that may be more commonly used in PRA. Truncating the minimum value may also be appropriate for distributions whose minimum is defined as zero (e.g., lognormal, gamma, Weibull). Truncation is generally less important when a PDF is used to characterize uncertainty in a parameter estimate (e.g., arithmetic mean), since distributions for uncertainty are often bounded by definition (e.g., triangular, uniform). Bounded continuous distributions, such as the beta distribution or empirical distribution (see Section B.5.2) are not subject to the parameter bias of truncation, although plausible minimum and maximum values must still be identified.

Identifying appropriate truncation limits that reflect "plausible bounds" for an exposure variable will often require judgment. Given that most data sets represent statistical samples of the target population, it is unlikely that the minimum and maximum observed values represent the true minimum and maximum values for the population. However, there may be physiological or physical factors that can aid in setting plausible truncation limits. For example, the maximum bioavailability of chemicals in the gastrointestinal (GI) tract is 100%. Similarly, the solubility of chemicals in aquatic environments

(accounting for effects of temperature) will generally be less than the chemical solubility in water free of particulates.

In general, sensitivity analysis can be used to determine if truncation limits are an important source of parameter uncertainty in risk estimates. For exposure variables in the numerator of the risk equation, the maximum truncation limit is of greatest concern. For exposure variables in the denominator of the risk equation, the minimum truncation limit is of greatest concern. Details regarding the fit of the tails of the probability distribution and the effect of truncation on the parameter estimates should generally be included in the workplan.

B.6.0 ASSESSING QUALITY OF THE FIT

The quality of the fit of a distribution may be evaluated in several ways. Standard statistical approaches are available to test the fit of a theoretical distribution to a data set (i.e., GoF tests). In addition, alternative choices for distribution shapes and plausible bounds might be explored as a form of sensitivity analysis. Together with graphical exploration (Section B.5.3), this information may be useful when deciding whether or not to incorporate a specific type of distribution for an exposure variable into a PRA.

☞ *GoF tests are one tool among several to assess the quality of a distribution.*

Although GoF testing is a necessary part of distribution fitting, and tests are readily available with commercial software, it is less important than mechanistic considerations or graphical data exploration for choosing a candidate distribution. Examples of GoF tests are discussed below, and cautions regarding GoF are outlined in Section B.6.3.

B.6.1 WHAT IS A GOODNESS-OF-FIT TEST?

Goodness-of-fit (GoF) tests are formal statistical tests of the hypothesis that the data represent an independent sample from an assumed distribution. These tests involve a comparison between the actual data and the theoretical distribution under consideration.

In statistical hypothesis testing the null hypothesis (H_0) is assumed to be true unless it can be proven otherwise. The "evidence" upon which we base a decision to reject or not to reject H_0 is a random sample. Typically, we seek to reject H_0 in favor of H_a. For example, with the two sample *t*-test, the null hypothesis is that the means of two populations are equal (not different) and the alternative is that they are different. This is expressed as:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

Most often, the hypothesis test is used to show that the means are not equal (i.e., reject H_0 in favor of H_a) in order to state that there is a significant difference between the two populations at a specified significance level (e.g., α=0.05). Thus, the hypothesis test is often referred to as a significance test.

The *p-value* in a statistical test is calculated from a sample and represents the probability of obtaining a value of the test statistic as extreme or more extreme as the one observed if H_0 is in fact true. When the *p-value* is small it means either the null hypothesis is not true, or that we have witnessed an

unusual or rare event (by chance we drew an unusual sample that resulted in the extreme value of the test statistic). Often a value of 0.05 or 0.01 is designated as a cutoff, or significance level α. If the *p-value* is (e.g., $p < 0.05$), the null hypothesis is rejected in favor of the alternative, and we state that the test result is statistically significant at level α. This does not mean that we have proven H_a is true. Rather, we are saying that based on our sample results, it is unlikely that H_0 is true.

In a GoF test, the hypothesis test is set up the same way as a "traditional" hypothesis test, but the outcome is viewed a little differently. In GoF tests, we generally seek to *fail* to reject H_0 because the null hypothesis states that the data were obtained from a population described by the specified distribution (F_0). The alternative hypothesis is that the data were obtained from a population described by a different distribution. In most applications of GoF techniques, the alternative hypothesis is composite—it gives little or no information on the distribution of the data, and simply states that H_0 is false (d'Agostino and Stephens, 1986). This can be expressed as:

$$H_0 : F = F_0$$

$$H_a : F \neq F_0$$

where F_0 is a specific continuous distribution function, such as the CDF for a normal distribution.

> ☞ *GoF tests do not prove that the population is described by the specified distribution, but rather that this assumption could not be rejected.*

In general, *p-values* provide one metric of evaluating the fit of the distribution. For example, a *p-value* of 0.06 indicates that the null hypothesis (i.e., the assumption of a specified distribution) cannot be rejected at $\alpha=0.05$. Larger *p-values* indicate a better fit and stronger evidence that the distribution specified by the null hypothesis may be appropriate. This guidance does not recommend an arbitrary cutoff for the *p-value*. A risk assessor performing a GoF test generally should report the *p-value* and whether the fit is considered "good" or "poor".

B.6.2 WHAT ARE SOME COMMON GOODNESS-OF-FIT TECHNIQUES?

The following GoF tests can also be found in most general statistical and spreadsheet software. Both *Crystal Ball*® and *@Risk* software present the results of chi-square, K-S, and Anderson-Darling tests in their fitting routines.

Shapiro-Wilk Test

The most widely used GoF test in risk assessment is the Shapiro-Wilk test for normality (Gilbert, 1987). This simple hypothesis test can determine whether or not a small data set ($n \leq 50$) is normally distributed. The test can also be run on log-transformed data to assess whether the data are lognormally distributed. D'Agostino's test may be used for samples sizes larger than those accommodated by the Shapiro-Wilk test (i.e., $n > 50$) (d'Agostino and Stephens, 1986). In addition, Royston (1982) developed an extension of the Shapiro-Wilk test for n as large as 2000 (Gilbert, 1987).

Probability Plot Correlation Coefficient Test

The correlation coefficient r (or the coefficient of determination, r^2) between the data and the z-scores of a normal probability plot (Filliben, 1975; Helsel and Hirsch, 1992) is similar to the W statistic

of the Shapiro-Wilk test. A detailed comparison of the Shapiro-Wilk test and the product correlation coefficient test is given by Filliben (1975) and d'Agostino and Stephens (1986). Helsel and Hirsch (1992) summarize critical r* values derived by Looney and Gulledge (1985) for the probability plot correlation coefficient test.

Chi-Square Test

The chi-square test is a general test that may be used to test any distribution (continuous or discrete), and for data that are ordinal (e.g., categories such as high/medium/low). Chi-square is a measure of the normalized difference between the square of the observed and expected frequencies. For example, by constructing a frequency distribution of the data with k adjacent bins, j=1...k, the number of data points in the j^{th} bin can be compared with the expected number of data points according to the hypothesized distribution. Note that in the case of continuous, unbounded distributions (e.g., normal), the first and last intervals may include $[-\infty, a_1]$ or $[a_k, +\infty]$ (Law and Kelton, 1991). The chi-square test is very sensitive to the chosen number and interval width of bins—different conclusions can be reached depending on how the intervals are specified. Strategies for selecting bins (e.g., setting interval widths such that there are no fewer than 5 data points expected per bin) are given in the statistical literature (d'Agostino and Stephens, 1986; Law and Kelton, 1991). The test statistic is compared with a value of the chi-square distribution with (k - r - 1) degrees of freedom, where k is the number of sample values and r is the number of parameters of the hypothesized distribution. As described in Section B.6.1, in general, higher *p-values* suggest better fits.

Kolmogorov-Smirnov (K-S) Test

The K-S test is a nonparametric test that compares the maximum absolute difference between the step-wise empirical CDF and the theoretical CDF. Because the maximum discrepancy is compared with the test statistic, K-S is sometimes referred to as a *supremum* test (Cullen and Frey, 1999). In general, lower values of the test statistic indicate a closer fit. The K-S test is most sensitive around the median of a distribution, and, hence, it is of little use for regulatory purposes when the tails of distributions are most generally of concern (U. S. EPA, 1999a). Although it does not require grouping data into bins like the chi-square test, critical values for the K-S test depend on whether or not the parameters of the hypothesized distribution are estimated from the data set (Gilbert, 1987; Law and Kelton, 1991). The Lilliefors test was developed to surmount this problem when the hypothesized distribution is normal or lognormal (Gilbert, 1987).

Anderson Darling Test

The Anderson-Darling test assesses GoF in the tails (rather than the mid-ranges) of a PDF using a weighted average of the squared differences between the observed cumulative densities. The Anderson-Darling test is sometimes referred to as the *quadratic* test (Cullen and Frey, 1999). The test statistic should be modified based on sample size prior to comparison with the critical value. Like the K-S test, in general, lower values of the test statistic indicate a closer fit (i.e., if the adjusted test statistic is greater than the modified critical value for a specified α, the hypothesized distribution is rejected). The Anderson-Darling test may be particularly useful because it places more emphasis on fitting the tails of the distribution.

B.6.3 CAUTIONS REGARDING GOODNESS-OF-FIT TESTS

There are many statistical software programs that will run GoF tests against a long list of candidate distributions. It is tempting to use the computer to make the choice of distribution based on a test statistic. However, GoF tests have low statistical power and often provide acceptable fits to multiple distributions. Thus, GoF tests are better used for rejecting poorly fitting distributions than for ranking good fits. In addition, for many distributions, GoF statistics lack critical values when the parameters are unknown (i.e., estimated from the data). In practice, this limitation is often discounted and the critical values are interpreted as a semi-quantitative measure of the fit. It is most appropriate to form an idea of the candidate distributions based on some well reasoned assumptions about the nature of the process that led to the distribution, and then to apply a GoF test to ascertain the fit (U.S. EPA, 1999a). Whenever possible, mechanistic and process (i.e., phenomenologic) considerations should inform the risk assessor's choice of a particular distribution rather than the results of a comparison of GoF tests (Ott, 1995). In addition, the value of graphical evaluations of the fit cannot be overstated.

B.6.4 ACCURACY OF THE TAILS OF THE DISTRIBUTION

The tails of a distribution (e.g., $< 5^{th}$ and $> 95^{th}$ percentiles) for an input variable are often of greatest interest when characterizing variability in risk. Distributions fit to data may not characterize the tails of the distribution in a way that represents the target population. In general, the importance of uncertainty in the fit of the tails of particular distributions should be determined on a site-specific basis. For exposure variables in the numerator of the risk equation, the upper tail is of greatest concern. For exposure variables in the denominator of the risk equation, the lower tail is of greatest concern.

The tails of the input PDFs generally have a significant influence on the tails of the risk distribution, especially for those variables that are ranked highest in a sensitivity analysis. Different distributions may share the same mean and variance, but assume very different shapes. Experiments with Monte Carlo simulations have demonstrated that the shape of the input PDFs may have a minimal effect on the risk estimates in the tails of the probability distribution when the mean and variance of the input PDFs are held constant (Hoffman and Hammonds, 1992; Finley and Paustenbach, 1994). Nevertheless, it is generally a good practice in PRA to demonstrate that alternative choices of PDFs do not have a significant effect on percentiles in the RME risk range.

A common question when developing and evaluating Monte Carlo models is, “How many iterations is enough?”. Since Monte Carlo sampling is approximately random, no two simulations will yield the same results (unless the same starting point, or seed, of the random number generator is used). A rule of thumb is that the stability of the output distribution improves with increasing numbers of iterations, although there will always remain some stochastic variability. The stability is generally better at the central tendency region of the output distribution than at the tails; therefore, more iterations may be needed when the risk management decision is associated with the higher percentiles (e.g., $> 95^{th}$ percentile). Risk assessors are encouraged to run multiple simulations (with the same inputs) using different numbers of iterations in order to evaluate the stability of the risk estimate of concern. The results of such an exercise should generally be reported to the Agency when submitting a PRA for review. Note that while the speed of modern computers has essentially eliminated the issue for 1-D MCA (e.g., 10,000 iterations of most 1-D MCA models can be run in less than 1 minute), it may still be an important issue for more complex modeling approaches such as Microexposure Event analysis (MEE) and 2-D MCA (see Appendix D).

B.7.0 SELECTING PROBABILITY DISTRIBUTIONS BASED ON STATE OF KNOWLEDGE

Table B-4 summarizes preliminary strategies for proceeding with a PRA based on the amount of available information. Recommended starting points for each of the three steps in the general process are provided. This table provides guidance on candidate distributions that are consistent with the available information, however, it is not intended to discourage the use or exploration of alternative choices.

> ☞ *Table B-4 provides recommended preliminary strategies, not steadfast rules. As an analyst works through the PRA, alternative distributions, estimation methods, consideration of mechanism, and GoF tests may better guide the selection process.*

Case 1 represents the best scenario, in which the analyst has access to the raw data and a sufficiently large sample size (or $\geq$ 6 percentiles). In this case, the analyst has a variety of choices for distribution fitting and estimating parameters. However, frequently raw data are inaccessible to the analyst. Cases 2 and 3 have limited information available (i.e., mean and upper percentile) and, therefore, have a narrower set of starting points. Case 4 is the most extreme scenario of data availability requiring expert judgment on selecting and fitting distributions.

Table B-4. Strategies for conducting PRA based on available information. Preferred methods in Case 1 (most information) are identified by an asterisk (*).

Evaluation Step	Case 1	Case 2	Case 3	Case 4
	Decreasing Information →			
Data Availability	raw data of sufficiently large sample size *or* six or more percentiles	three to five statistics	two statistics	one statistic
Selection of Distribution Type	**Nonnegative Continuous** any in this category **Bounded** beta, Johnson's SB	**Nonnegative Continuous** lognormal, gamma, Weibull **Bounded** beta, Johnson's SB		case-by-case basis using expert judgment
Selection of Parameter Estimation / Fitting Method	maximum likelihood* regression methods matching moments	minimize average absolute percent error (MAAPE) for available statistics	exact agreement between 2-parameter PDF and available statistics	
Assessment of Quality of Fit	**Graphical Assessment** P-log Q plot*, P-Q plot* residual % error plot* P-P plot, Q-Q plot **GoF Tests** Anderson-Darling* K-S Chi-square	**Graphical Assessment** P-log Q plot, P-Q plot **GoF Test** Chi-square, Estimate *p*-value for MAAPE using parametric bootstrap (if sample size is known)	**Graphical Assessment** judgment based on comparative analysis of PDFs and CDFs	
Estimation of Parameter Uncertainty	**Large Sample** asymptotic normality assumption **Medium Sample** nonparametric bootstrap **Small Sample** parametric bootstrap	**Parametric bootstrap** generate random samples using the fitted distribution (if sample size is known)		

EXAMPLES OF FITTING DISTRIBUTIONS USING GRAPHICAL METHODS, GOODNESS-OF-FIT, AND PARAMETER ESTIMATION

Example B-1. Empirical Distribution Function (EDF) for Soil Ingestion Rates

This hypothetical example illustrates how graphical methods can be used to select probability distributions for variability based on percentile data reported in the literature. Table B-5 gives the summary statistics that are reported by Stanek and Calabrese (1995) for average daily soil ingestion rates among young children. Three options are explored for selecting a distribution: (1) empirical distribution function (EDF) represented by a step function; (2) linearized and extended EDF; and (3) continuous parametric distributions (beta and lognormal).

In order to specify an EDF, a plausible range (minimum and maximum) must be inferred using judgment. Exposure factors such as ingestion rate are nonnegative variables (i.e., minimum ≥ 0); given the relatively low value for the 25th percentile (10 mg/day), it is assumed that 0 mg/day is a reasonable minimum value for this example. If children with pica for soil are excluded from the population of concern, the maximum value may be inferred from the relatively shallow slope at the high-end of the distribution. That is, the 90th percentile is reported as 186 mg/day while the 99th percentile is 225 mg/day, an increase of only 39 mg/day; it is assumed that 300 mg/day is a plausible maximum value for this example. Commercial software such as *Crystal Ball*® and *@Risk* can be used to input EDFs. Figure B-3 illustrates the basic step-wise EDF represented by the reported percentile values, as well as the "linearized, extended EDF" (i.e., linear interpolation between reported values and extended lower and upper tails).

An alternative to relying on a linear interpolation between the percentile values is to fit a continuous probability distribution to the reported percentiles. Since the original data are unavailable, standard GoF tests for the EDF, such as K-S and Anderson-Darling (d'Agostino and Stephens, 1986), cannot be applied. Note that computer software (e.g., *Crystal Ball*®, *@Risk*) will provide test statistics and corresponding *p-values*, however, these results will (inappropriately) reflect the number of percentile values reported rather than the sample size of the original data. Nevertheless, graphical methods may be employed to assess the adequacy of the fit of various PDFs. In this example, a beta distribution and lognormal distribution were fit to the EDF using *Crystal Ball*®. Figure B-4 illustrates the selected statistics for both distributions.

The beta distribution appears to more closely match the reported percentile values, especially at the upper tail of the distribution. The lognormal distribution has an unbounded maximum that, for this example, results in an extreme overestimate of the 95th and 99th percentiles. The beta distribution, by definition, is bounded at 0 and 1, and rescaled in this example to a maximum of 364 mg/day. This analysis would support the use of a beta distribution in a Monte Carlo simulation.

Table B-5. Selected statistics for reported and fitted distributions for ingestion rate (mg/day).

Summary Statistic	Reported Values	Linearized, Extended EDF	Beta Distribution[1]	Lognormal Distribution[2]
minimum	--	0	0	0
25th percentile	10	10	13	11
50th percentile	45	45	44	31
75th percentile	88	88	100	86
90th percentile	186	186	165	216
95th percentile	208	208	205	375
99th percentile	225	225	322	3346
maximum	--	300	364	$+\infty$

[1]Parameters of best-fit beta distribution: α_1=0.63, α_2=2.85, min=0, max=364.
[2]Parameters of best-fit lognormal distribution: μ=97.6, σ=291.8.

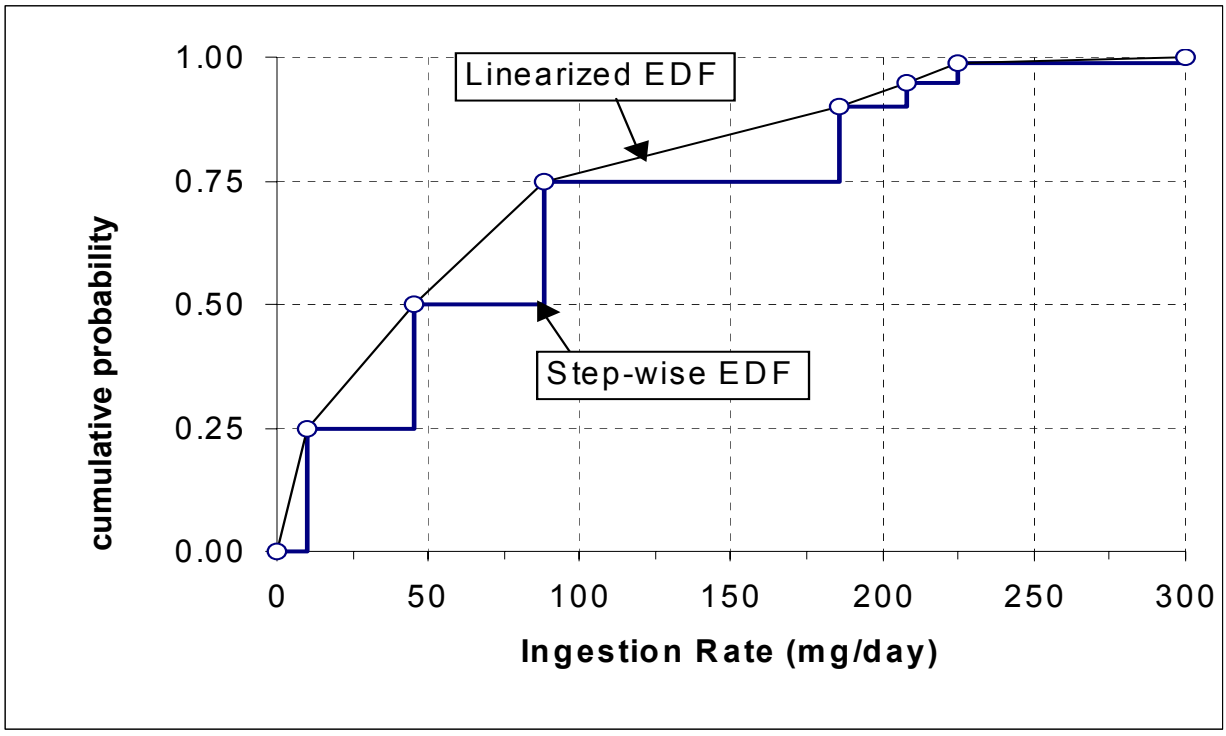


Figure B-3. Comparison of step-wise EDF and linearized EDF for ingestion rate. The upper and lower tails of both distributions are extended to a plausible range of [0, 300] mg/day.

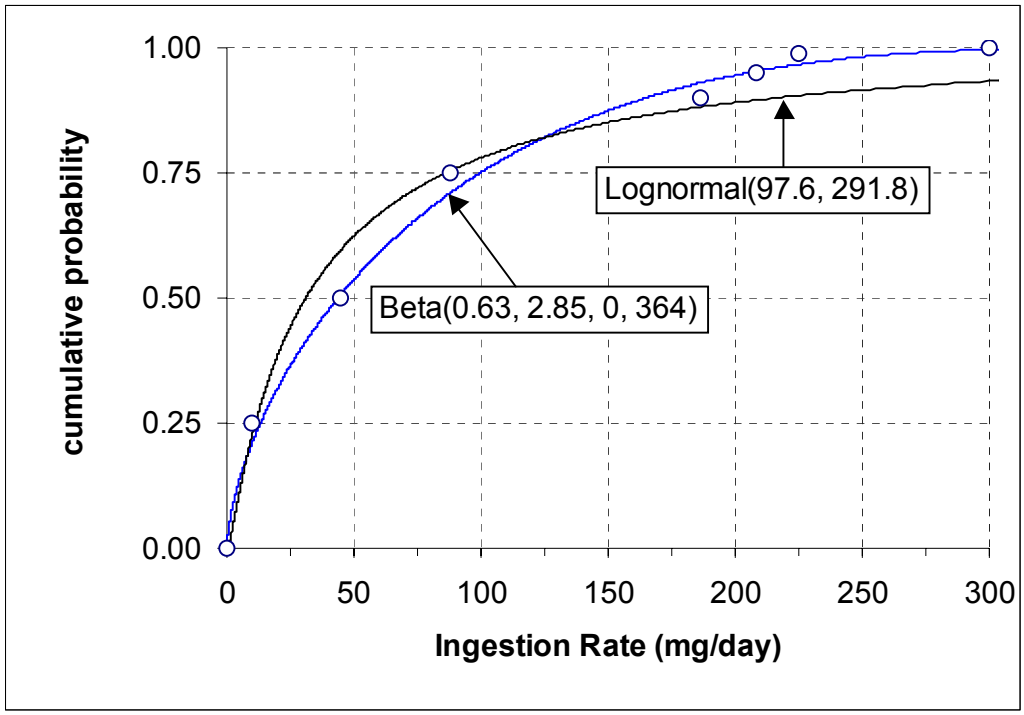


Figure B-4. Graphical assessment of beta and lognormal distributions fit to the cumulative distribution reported in the literature (circles). The beta distribution provides a closer fit to the percentile values in this example.

Example B-2. Variability in Lead Concentrations in Quail Breast Tissue

This hypothetical example demonstrates how the combination of graphical methods, GoF tests, and parameter estimation techniques provides strong evidence for selecting and fitting a lognormal distribution. Assume lead concentration in quail is an important variable for a food web model. Site-specific data (n=62) are used to estimate inter-individual variability in concentration (Table B-6). The histograms in Figure B-5 show lead concentrations in quail breast tissue collected near a settling pond at a plating works. Equation B-1 indicated that 7 bins is an appropriate starting point. The result (top left panel, Figure B-5) suggests that approximately 80% of the values are < 200 ppm and that the probability distribution for variability may be described by a nonnegative, right-skewed distribution (e.g., exponential, Weibull, lognormal, etc.). However, additional bins are needed to better understand the low-end of the distribution. After increasing the number of bins from 7 to 16 (top right panel, Figure B-5), graphical evaluation continues to suggest that the distribution is unimodal right skewed. The bottom panel of Figure B-5 illustrates that increasing the number of bins would not provide better resolution of the low-end of the distribution. For these data, 16 bins appear to provide a reasonable balance between too much smoothing and too much jaggedness.

Probability plots can be used to visually inspect the GoF of a specified distribution to the data, and, because the hypothesized distribution yields a straight line, the plots are particularly useful for evaluating deviations at the tails. In addition, parameter estimates can be obtained from the regression lines fit to the data, as discussed below. For this example, two lognormal probability plots are explored to evaluate how well the data can be described by a lognormal distribution (Figure B-6). The top panel gives the *z-score* on the abscissa (the "x" axis) and ln[concentration] on the ordinate (the "y" axis), while the bottom panel gives ln[concentration] on the abscissa and *z-score* on the ordinate. Plotting positions for both methods were calculated using Equation B-2. Equally plausible parameter estimates can be obtained from regression lines using either plotting method; however, the approach shown in the top panel may be easier to implement and interpret.

Despite the relatively large sample size of n=62, GoF tests generally fail to reject lognormality (i.e., normality of the log-transformed data) in this example. For the probability plot correlation coefficient test (Filliben, 1975; Looney and Gulledge, 1985), if $r < r^*$ (the value for r at a specified α), normality is rejected. For this example, r is 0.988, and r^* is between 0.988 and 0.989 for n=62 and α=0.25; therefore, the *p-value* for the concentrations is approximately 0.25 and one fails to reject lognormality at $\alpha \leq 0.25$. D'Agostino's test yields essentially the same conclusion, with a calculated Y value of -1.9166. For this data set, with n=62 and α=0.10, one rejects normality if $Y < -2.17$ or $Y > 0.997$ (see Table 9.7 in d'Agostino and Stephens, 1986); therefore, since Y is within this interval, one fails to reject the normal distribution. However, for α=0.20, the rejection criteria is [$Y < -1.64$ or $Y > 0.812$], Y falls outside the low-end of the interval, resulting in a rejection of the normal distribution. For this data set, the *p-value* associated with d'Agostino's test is slightly less than 0.20 and one fails to reject normality at $\alpha < 0.20$.

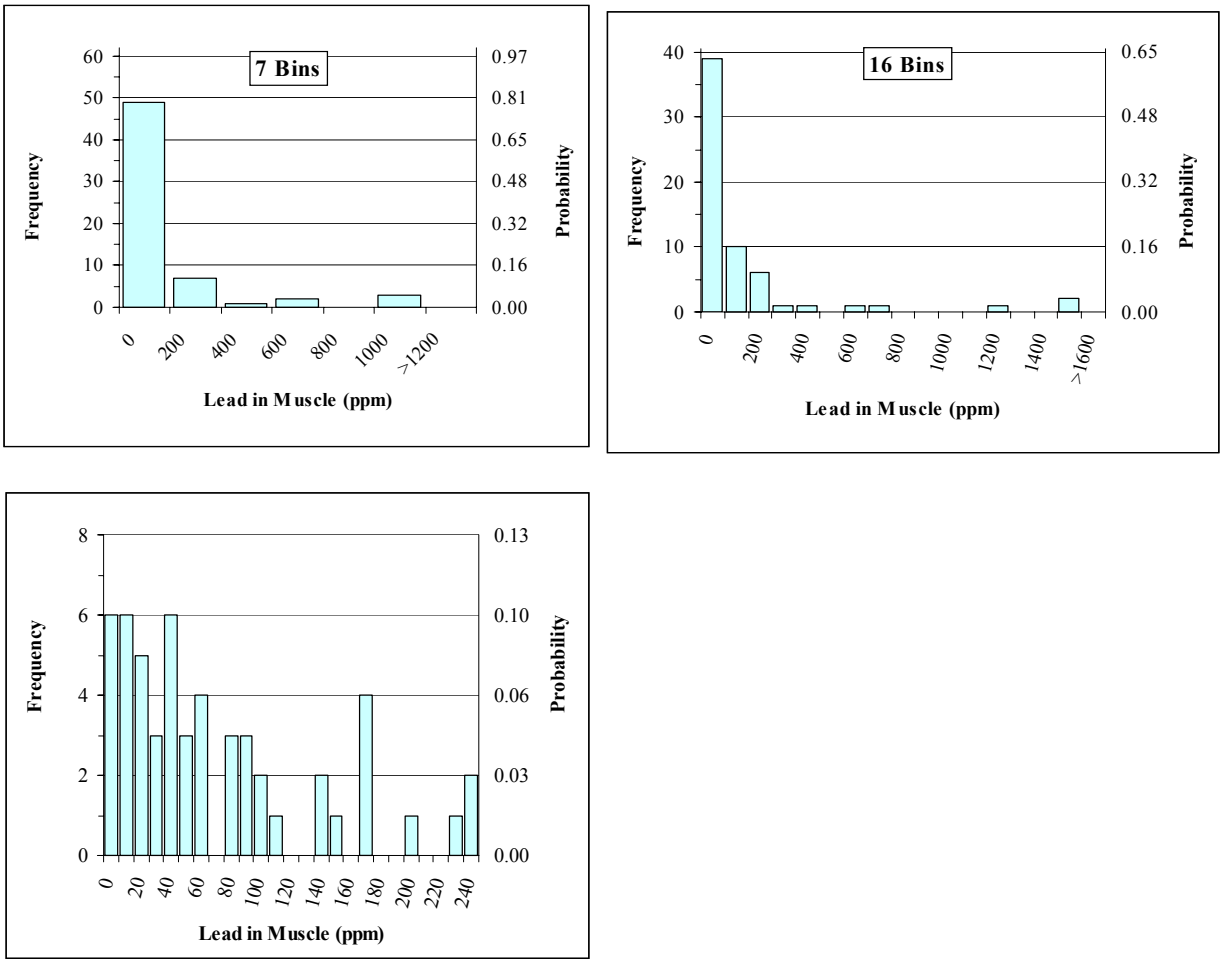


Figure B-5. Histograms of lead concentrations in quail breast muscle (*n*=62). The top left panel shows the result with seven bins; the top right panel shows the result with sixteen bins; the bottom panel uses bin widths of 10 ppm to highlight the lower tail (< 250 ppm) of the distribution.

Table B-6. Sample values of lead concentration (ppm) in quail breast muscle (*n*=62).

0.45	15.8	36.6	57	91	173	265
2.1	16	40	59.6	94.2	175.6	322
5.4	16.7	40.1	61.4	99	176	490
7.8	21	42.8	62	107	177	663.4
7.8	23	44	64	109	205	703
8.8	24	46	64	111	239	1231
11.8	24.8	47	84.6	149	241	1609
12	29.2	49	86.6	149	245	1634
15	35.5	53	86.8	154	264	

Different methods for obtaining the parameter estimates for the lognormal distribution can be explored in this example. For the lognormal distribution, MLE and MoMM simply require calculating the mean and standard deviation of the log-transformed sample data. For the lognormal probability plot method, the parameters can be obtained directly from the least squares regression line expressed as follows:

$$\ln(x) = [slope]z + [intercept] \qquad \text{Equation B-5}$$

such that exponentiating the intercept will give the geometric mean (GM) and exponentiating the slope will give the geometric standard deviation (GSD) (see Footnote 3 of Table B-7). Both the MLE and MoMM estimates will generally match the arithmetic mean of the log-transformed data (i.e., intercept) determined from lognormal probability plots; however, estimates of the standard deviation (i.e., slope) will vary (Cullen and Frey, 1999). In general, the probability plot method yields estimates of the standard deviation that are less than or equal to that of MoMM and MLE, and the results yield closer estimates as the correlation coefficient of the probability plot increases (Cullen and Frey, 1999). Table B-7 summarizes the parameter estimates using MLE, MoMM, and the two lognormal probability plotting techniques described above. The corresponding parameter estimates for the untransformed data are also presented.

In this example, the strong linearity of the probability plots (r^2=0.98) shown in Figure B-6 is an indication that a lognormal distribution is a reasonable model for describing variability in concentrations. The tails of the distributions fit the data fairly well, although the bottom panel suggests that the lognormal distribution slightly overestimates the lower tail. Furthermore, the parameter estimates of the lognormal distribution using probability plotting closely match the estimates using MLE and MoMM.

Table B-7. Parameter estimates for lognormal distribution of lead concentrations (ppm).

Parameter Estimation Method	Log-transformed Data		Untransformed Data[3]	
	Arithmetic mean [$\hat{\mu}$]	Arithmetic stdev [$\hat{\sigma}$]	Arithmetic mean [$\hat{\mu}$]	Arithmetic stdev [$\hat{\sigma}$]
Maximum Likelihood Estimate (MLE)	4.175	1.522	207	626
Method of Matching Moments (MoMM)	4.175	1.522	207	626
Log Probability Plot[1]	4.175	1.507	203	597
Log Probability Plot[2]	4.175	1.543	214	670

[1]Least squares regression line for Figure B-6, top panel.
[2]Least squares regression line for Figure B-6, bottom panel.
[3]For a lognormal distribution, the following equations can be used to convert parameters of the normal distribution of log-transformed data to corresponding parameters of the lognormal distribution of untransformed data. Assume μ^* and σ^* are the arithmetic mean and standard deviation, respectively, for the normal distribution of log-transformed data.

$$geometric\ mean = \exp[\mu^*]$$

$$geometric\ standard\ deviation = \exp[\sigma^*]$$

$$arithmetic\ mean = \exp[\mu^* + 0.5\sigma^{*2}]$$

$$standard\ deviation = \exp[\mu^*]\left(\exp[\sigma^{*2}]\exp[\sigma^{*2} - 1]\right)^{0.5}$$

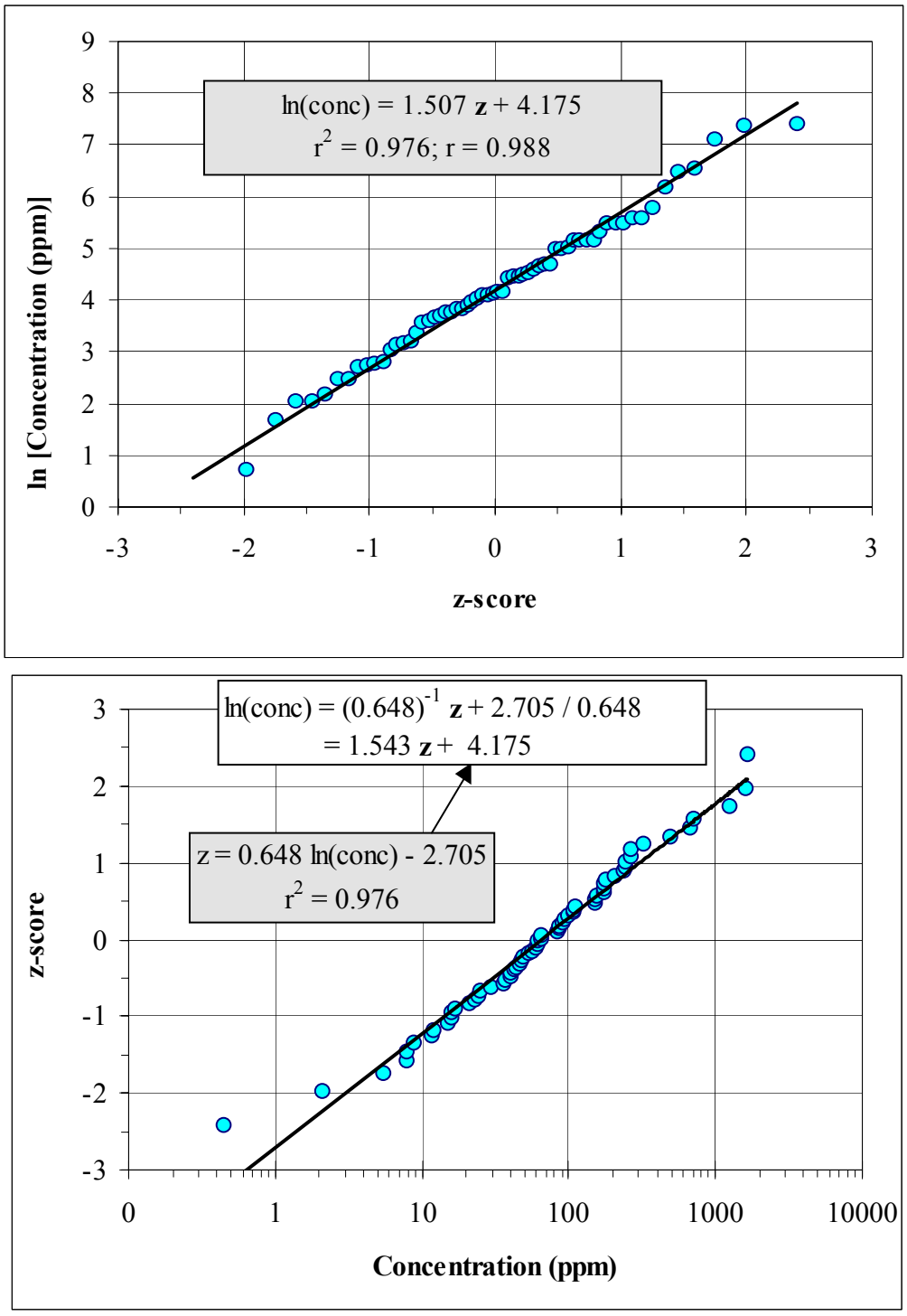


Figure B-6. Lognormal probability plots of lead in quail breast tissue. Top panel gives *z* on the abscissa and ln[concentration] on the ordinate. Bottom panel gives concentration (log scale) on the abscissa and *z* on the ordinate. Equally plausible parameter estimates can be obtained from regression lines using either plotting method. Bottom panel requires an additional step to express the equation that yields parameter estimates [ln(x)=(slope) z + (y-intercept)], where the slope estimates the standard deviation of ln(x) and the y-intercept (at z=0) estimates the arithmetic mean of ln(x).

Example B-3. Variability in Meal Sizes Among Consuming Anglers

A creel survey of anglers consuming contaminated fish was performed to estimate variability in fish meal sizes. The anglers were asked how many people would eat their fish. The lengths of the fish were measured and a regression equation was used to calculate the corresponding weights. The portion of the fish mass that is consumed was assumed to be 40% (e.g., fillets). Results given in Table B-8 are expressed in units of grams of fish per meal.

The appearance of the histograms (Figure B-7) suggests that the sample (n=52) may have been selected from a single distribution.

A normal probability plot of the meal sizes (Figure B-8) shows a departure from linearity. Specifically, there appears to be a "kink" in the probability plot at about 400 g/meal, suggesting that the sample may have been obtained from two unique distributions. Both the Filliben test and Shapiro-Wilk test indicated a significant departure from normality at α=0.01. Parameters may be read directly from the equations of the regression lines on the right hand panel of the graph. MoMM and MLE gave similar estimates.

Table B-8. Meal size (g/meal) (n=52).

65	182	310	405
74	208	314	415
74	221	318	416
77	226	318	477
90	241	327	531
110	248	332	572
111	253	336	608
133	260	337	745
143	261	350	831
150	281	351	907
163	303	360	1053
163	305	365	1189
174	305	390	1208

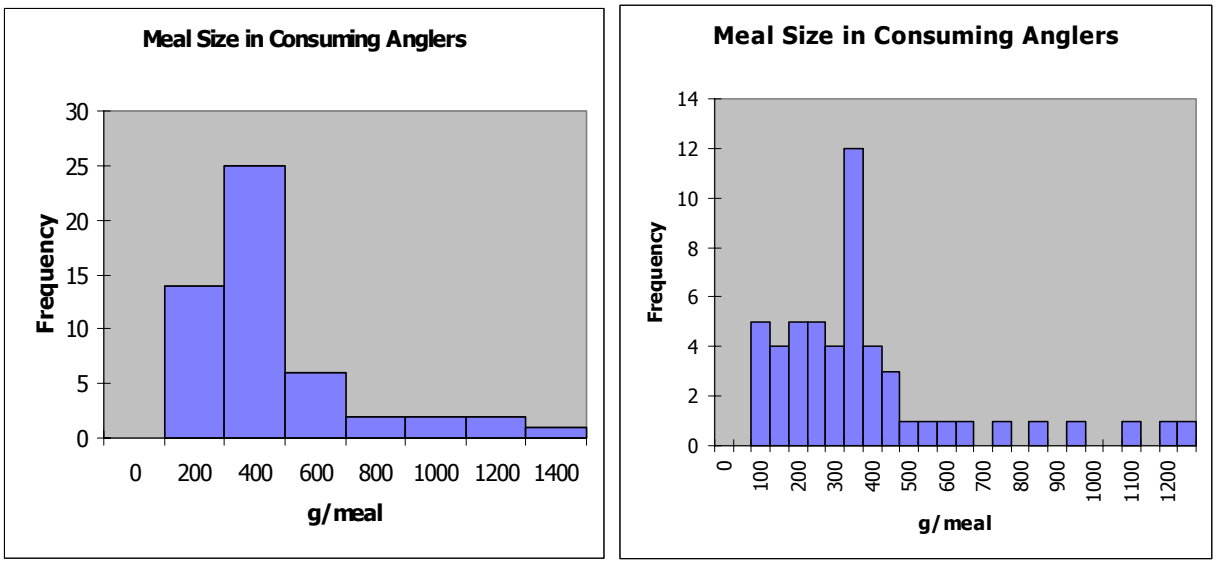


Figure B-7. Histograms of meal size (n=52) among consuming anglers. Left panel uses 7 bins, while the right panel uses 14 bins.

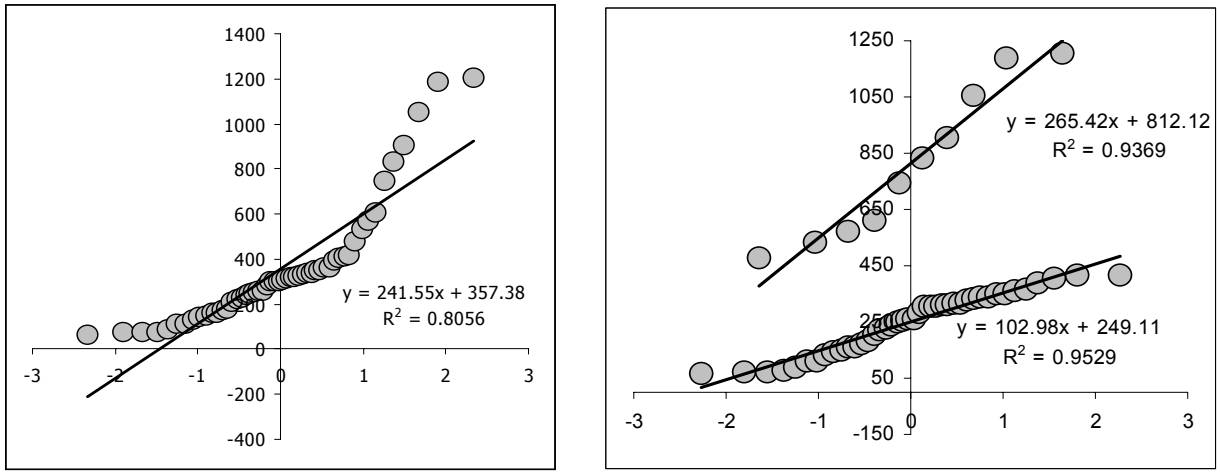


Figure B-8. Probability plot of meal size data from consuming anglers. The left panel shows the combined data, with a departure from linearity at ~ 400 g/meal. The right panel shows the data split between high consumers (top line) and low consumers (bottom line); note that separate lognormal probability plots were reconstructed for both subsets of the data. The point at which to "split" the distribution in the left panel is somewhat subjective. The break would be more obvious if the two distributions did not overlap.

Example B-4. Bivariate Normal Distributions

THIS EXAMPLE PRESENTS...

- Description of the assumptions associated with the bivariate normal distribution
- Guidance on simulating the bivariate normal distribution for two random variables
- Application of bivariate normal to a simple linear regression equation relating contaminant concentrations in soil and dust (see Figure B-9). Results are compared to the assumption of no correlation and perfect correlation

This example introduces the bivariate normal distribution to illustrate two concepts: (1) use of information on correlations in a Monte Carlo simulation; and (2) specifying distributions for uncertainty in parameter estimates. A brief explanation of the bivariate distribution is presented followed by an example comparing assumptions of no correlation and perfect correlation. A less complex example of a method for addressing correlations in PRA is given in Exhibit B-8.

Properties of a Bivariate Normal Distribution

One approach that can be used to correlate two random variables is to specify a bivariate normal distribution, which allows for the distribution of one variable to be sampled conditional on the other. A bivariate normal distribution is a special case of a joint distribution in which both x and y are random independent normally distributed variables. A bivariate normal distribution can be specified for all correlation coefficients including $\rho=0$, $\rho=1$, and $\rho=-1$. The bivariate distribution has a three dimensional shape and for $\rho=0$, from a bird's-eye view, is perfectly circular. As correlation increases (i.e. moves towards -1 or 1) this circle narrows and flattens to an elliptical shape, and finally for perfect correlation $\rightarrow=1$ and $\rho=-1$) becomes a straight regression line with a $r^2=1$. In three dimensional space the probability of obtaining measurement pairs (x, y) in the region is equal to the volume under the surface in that region. To completely specify the bivariate normal, estimates of the arithmetic mean and variance of the two parameters, as well as the correlation coefficient (μ_X and μ_Y, variances σ^2_X and σ^2_Y , and correlation coefficient ρ) are needed.

In a bivariate normal distribution, values of y corresponding to each value of x follow a normal distribution (Snedecor and Cochran, 1989). Analogously, the values of x corresponding to each value of y follow a normal distribution. Furthermore, if two random variables, *X* and *Y*, jointly follow a bivariate normal distribution, the marginal distribution of *X* is normal with mean μ_X and variance σ^2_X, and the marginal distribution of *Y* is normal with mean μ_Y and variance σ^2_Y.

Conditional Distributions

Assume we are interested in the conditional distribution of *X* given a certain value for *Y*. For example, if *X* and *Y* are positively correlated, we would expect that relatively high values of *X* tend to correspond with relatively high values of *Y*. The conditional distribution of *X* given that *Y*=y, where y represents a specific value for the random variable *Y*, is a normal distribution with:

$$mean = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \quad and$$
$$variance = \sigma_X^2 (1 - \rho^2) \qquad \text{Equation B-6}$$

Likewise, the conditional distribution of *Y* given that *X*=x, is also normal with:

$$mean = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X), \quad and$$
$$variance = \sigma_Y^2 (1 - \rho^2) \qquad \text{Equation B-7}$$

These general equations can be used to generate a correlated pair (*X*, *Y*), as described below.

*Note that the mean of the conditional distribution of *X* is a function of the given value of *Y* but the variance depends only on the degree of correlation.

General Approach for Correlating X and Y

To generate a correlated pair (*X*, *Y*), first generate *X* using a random value Z_1 from the standard normal distribution:

$$X = \mu_X + \sigma_X \times Z_1 \qquad \text{Equation B-8}$$

Next, express Y as a function of the conditional mean and variance of *Y* given *X* and a second standard normal variate Z_2:

$$Y = \mu_Y + \sigma_Y \times Z_2 \qquad \text{Equation B-9}$$

and generate a correlated Y by plugging Equation B-7 into Equation B-9. Using algebra, the combined equations yield the following simplified expression for generating Y:

$$Y = \mu_Y + \sigma_Y\left[(\rho \times Z_1) + \sqrt{1-\rho^2} \times Z_2\right] \qquad \text{Equation B-10}$$

The important component of this equation is that two random variates are needed (Z_1 and Z_2).

An alternative, but less general approach would be to obtain Y by first generating a normal variate X (Equation B-8) and then plugging that value into the regression equation of Y on X to obtain the associated value of Y. While this method maintains a correlation between X and Y, it will underestimate parameter uncertainty. The results are equal only for the special case of perfect correlation (ρ=1.0) between X and Y. Therefore, the more general bivariate normal distribution approach (given by Equations B-8 to B-10) is recommended for correctly correlating X and Y because it provides a more robust estimate of parameter uncertainty.

EXHIBIT B-9

STEPS FOR SIMULATING UNCERTAINTY IN LINEAR REGRESSION EQUATION USING A BIVARIATE NORMAL DISTRIBUTION TO CORRELATE PARAMETERS (B_0, B_1)

(1) Select Z_1 from a standard normal distribution $Z \sim N(0, 1)$

(2) Calculate β_0 using Equation B-8, where $X=\beta_0$, $\mu_x=\mu_{b0}$, and $\sigma^2_x=\sigma^2_{b0}$

(3) Select Z_2 from a standard normal distribution $Z \sim N(0, 1)$

(4) Calculate β_1 using Equation B-10, where $Y=\beta_1$, $\mu_y=\mu_{b1}$, $\sigma^2_y=\sigma^2_{b1}$, ρ=correlation between β_0 and β_1

Application of Bivariate Normal Distribution to Correlate Concentrations of Zinc in Soil and Dust

Assume random sampling of soil and dust zinc concentrations occurs in a residential area. Composite samples of soil and dust are collected from 21 locations such that samples are paired (i.e., each soil sample is co-located with a dust sample) (Table B-9). First the relationship between the zinc concentration in soil and dust is evaluated using simple least-squares regression. Next, the bivariate normal distribution for the slope (β_1) and intercept (β_0) is determined, yielding an arithmetic mean and standard deviation for each parameter (μ_{b0}, $\sigma^2_{b0,}$ μ_{b1}, and σ^2_{b1}), and correlation coefficient ρ between β_1 and β_0. In this context, the bivariate normal distribution may be considered a distribution for uncertainty in the parameter estimates.

Three simulation methods are employed to demonstrate the effect of assuming a bivariate normal distribution for parameters vs. perfect correlation, or independent parameters. Specifically:

(1) The slope and intercept of the regression line are described by a specific form of the bivariate normal distribution (i.e., follow *Steps 1, 2* in Exhibit B-9, and use Equation B-10 instead of *Step 4*).

(2) The slope and intercept of the regression line are described by a general form of the bivariate normal distribution (i.e., follow *Steps 1 to 4* in Exhibit B-9).

(3) The slope and intercept of the regression line are described by independent normal distributions (i.e., follow *Steps 1–4* in Exhibit B-9, but omit the correlation coefficient ρ in *Steps 2 and 4*).

For each approach, Monte Carlo simulations with I=5,000 iterations were run to determine the set

of parameter values (β_0, β_1) for a simple linear regression equation. Typically, the uncertainty in the parameter estimates is not accounted for when simple linear regression equations are used to relate to exposure variables in a model. Such an approach may fail to account for important sources of parameter uncertainty. Figure B-10 (middle panel) illustrates the preferred approach for characterizing parameter uncertainty based on the bivariate normal distribution. (Note that the correlation coefficient relating the intercepts and slopes generated from the simulation is consistent with the correlation coefficient that describes the bivariate normal distribution; this is a good check that the simulation was set up correctly and run for a sufficient number of iterations). These results are contrasted with results using a form of the bivariate normal (Equation B-10) that underestimates uncertainty (top panel) unless parameters are perfectly correlated. In addition, the simplistic approach of sampling from independent normal distributions (bottom panel), yields a "shot gun" scatter plot. Sampling from independent normal distributions results in unlikely extreme combinations of the slope and intercept more often than the correct bivariate normal approach; propagating this bias through a risk model may severely bias estimates of uncertainty in risk.

Table B-9. Zinc concentrations in paired (i.e., co-located) soil and dust samples (ppm) for *n*=21 locations.

Sample	Soil (X_i)	Dust (Y_i)	Sample	Soil (X_i)	Dust (Y_i)
1	120	216	12	560	200
2	190	149	13	560	256
3	270	83	14	720	496
4	285	508	15	800	239
5	310	215	16	880	203
6	340	219	17	910	757
7	350	203	18	1035	676
8	380	101	19	1445	426
9	440	178	20	1600	522
10	480	232	21	1800	276
11	560	199			

Bivariate Normal Distribution for Parameters of the Regression Equation		
B_0	mean	173.9
	variance	4162.2
B_1	mean	0.193
	variance	0.0063
s^2		27857.4
Cov (B_0, B_1)		-4.2428
r		-0.8254

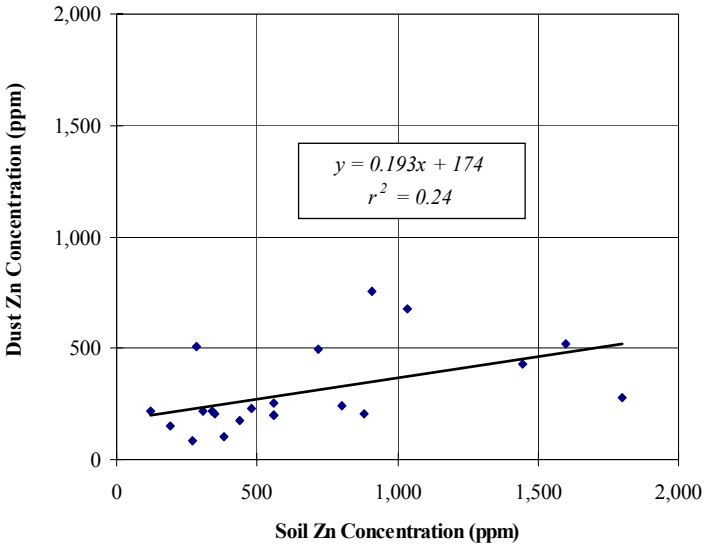


Figure B-9. Simple linear regression of zinc concentrations in soil and dust.

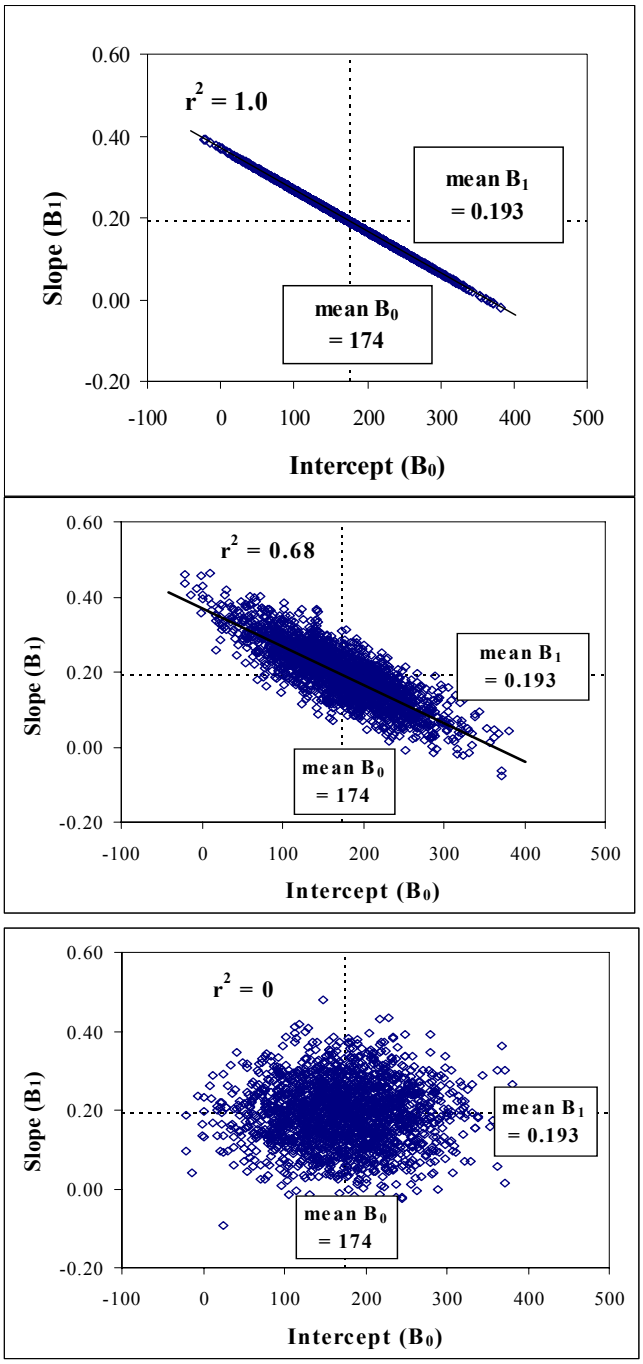


Figure B-10. Results of Monte Carlo simulation (*n*=5000 iterations) to estimate the slope and intercept of a regression equation. Top panel reflects the bivariate normal distribution for the special case that fails to capture the parameter uncertainty; middle panel reflects the preferred bivariate normal distribution with ρ=-0.825 based on empirical paired data; bottom panel reflects sampling from independent normal distributions.

REFERENCES FOR APPENDIX B

Brainard, J. and D.E. Burmaster. 1992. Bivariate Distributions for Height and Weight of Men and Women in the United States. *Risk Anal* 12(2):267–275.

Brately, P., B.L. Fox, and L.E. Schrage. 1987. *A Guide to Simulation*. Springer-Verlag, NY.

Burger, J., W. L. Stephens, Jr., C. S. Boring, M. Kuklinski, J.W. Gibbons, and M. Gochfeld. 1999. Factors in Exposure Assessment: Ethnic and Socioeconomic Differences in Fishing and Consumption of Fish Caught along the Savannah River. *Risk Anal*. 19(3):427–438.

Calabrese, E.J., Stanek, E.J., and Barnes R. 1996. Methodology to Estimate the Amount and Particle Size of Soil Ingested by Children: Implications for Exposure Assessment at Waste Sites. *Regul. Toxicol. Pharmacol*. 24:264–268.

Charney, E., J. Sayre, and M. Coulter. 1980. Increased Lead Absorption in Inner City Children: Where Does the Lead Come From? *Pediatrics* 65:226–231.

Conover, W.J. 1980. *Practical Nonparametric Statistics*. John Wiley & Sons, NY.

Cullen, A.C. and H.C. Frey. 1999. Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

d'Agostino, R.B. and M.A. Stephens. 1986. *Goodness-of-fit techniques*. Marcel Dekker, Inc, NY.

Filliben, J.J. 1975. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics* 17(1):111–117.

Finley, B.L. and D.J. Paustenbach. 1994. The Benefits of Probabilistic Exposure Assessment: Three Case Studies Involving Contaminated Air, Water and Soil. *Risk Anal* 14(1):53–73.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Hostrand Reinhold, NY.

Gilliom, R.J. D.R. Helsel. 1986. Estimation of Distributional Parameters for Censored Trace Level Water Quality Data, 1. Estimation Techniques. *Water Resources Research*. 22:135–146..

Hahn, G.J. and S.S. Shapiro. 1967. *Statistical Models in Engineering*. John Wiley & Sons, NY.

Helsel, D.R. and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. Elsevier. Amsterdam.

Hoffman, F.O. and J.S. Hammonds. 1992. *An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment.* ES/ER/TM–35. Martin Marietta.

Hora, S.C. 1992. Acquisition of Expert Judgment: Examples From Risk Assessment. *J. Energy Eng*. 118(2):136–148.

Johnson, N.L., S. Kotz, and N. Balakrishnan. 1995. *Continuous Univariate Distributions*. Volume 2, Second Ed. John Wiley & Sons, NY.

Law, A.M. and W.D. Kelton. 1991. *Simulation Modeling and Analysis*. McGraw-Hill, NY.

Looney, S.W. and T.R. Gulledge. 1985. Use of the Correlation Coefficient with Normal Probability Plots. *American Statist.* 39:297–303.

Mendenhall, W. and R.L. Scheaffer. 1973. *Mathematical Statistics with Applications*. Duxbury Press.

Mood, A.M. and F.A. Graybill. 1963. *Introduction to the Theory of Statistics*. Second Edition. McGraw-Hill, Inc.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.

Nelsen, R.B. 1986. Properties of a One-Parameter Family of Bivariate Distributions with Specified Marginals. *Comm. Stat. (Theory and Methods)* 15:3277–3285.

Nelsen, R.B. 1987. Discrete Bivariate Distributions with Given Marginals and Correlation. *Comm. Stat. (Simulation and Computation)* B16:199–208.

Oregon DEQ. 1998. *Guidance for the Use of Probabilistic Analysis in Human Health Exposure Assessments*. Waste Management and Cleanup Division. Interim Final. November.

Ott, W.R. 1990. A Physical Explanation of the Lognormality of Pollutant Concentrations. *J. Air Waste Manage Assoc.* 40(10):1378–1383.

Ott, W.R. 1995. *Environmental Statistics and Data Analysis*. CRC Press, Boca Raton.

Palisade Corporation. 1994. *Risk Analysis and Simulation Add-In for Microsoft Excel or Lotus 1-2-3*. Windows Version Release 3.0 User's Guide, Palisade Corporation, Newfield, NY.

Roseberry, A.M. and D.E. Burmaster. 1992. Lognormal Distributions for Water Intake by Children and Adults. *Risk Anal.* 12(1):99–104.

Royston, J.P. 1982. An Extension of Shapiro and Wilk's *W* test for Normality to Large Samples. *Appl. Stat*. 31:115–124.

Snedecor, G.W. and W.G. Cochran. 1989. *Statistical Methods*. Eighth Edition. Iowa State University Press, Iowa.

Stanek, E.J. and Calabrese, E.J. 1995. Daily Estimates of Soil Ingestion in Children. *Environ. Health Perspect.* 103:176–285.

Thompson, K. 1999. Developing Univariate Distributions from Data for Risk Analysis. *Hum. Eco. Risk Assess.* 5(4):755–783.

Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley, Boston.

U.S. EPA. 1982. *Air Quality Criteria for Particulate Matter and Sulfur Oxides*. ECAO, ORD. EPA 600/8–82-029.

U.S. EPA. 1992. *Guidance for Data Useability in Risk Assessment, Part* A. Office of Emergency and Remedial Response, Washington, DC. OSWER Directive No. 9285.7-09A.

U.S. EPA. 1994. *Guidance for Conducting External Peer Review of Environmental Regulatory Models.* Office of the Administrator, Washington, DC. EPA/100/B-94-001. July.

U.S. EPA. 1997a. *Exposure Factors Handbook.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa, Fb, and Fc.

U.S. EPA. 1997b. *Use of Probabilistic Techniques (Including Monte Carlo Analysis) in Risk Assessment*, Memorandum from Deputy Administrator Hansen and *Guiding Principles for Monte Carlo Analysis*. EPA/630/R-97-001.

U.S. EPA. 1999a. *Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments.* Risk Assessment Forum. EPA/630/R-98/004. January.

U.S. EPA. 1999b. *Options for Development of Parametric Probability Distributions for Exposure Factors.* Office of Research and Development. Research Triangle Institute Final Report. April 6.

U.S. EPA. 2001. *Development and Evaluation of Probability Density Functions for a Set of Human Exposure Factors.* Office of Emergency and Remedial Response. University of California Draft Report. May.

Vose, D. 1996. *Quantitative Risk Analysis: A Guide to Monte Carlo Modeling*. John Wiley & Sons, NY.

Wonnacott and Wonnacott. 1981. *Regression: A Second Course in Statistics*. John Wiley & Sons, NY.

APPENDIX C

CHARACTERIZING VARIABILITY AND UNCERTAINTY IN THE CONCENTRATION TERM

C.0 THE CONCENTRATION TERM AND THE EXPOSURE UNIT

Incomplete knowledge of the concentration of one or more chemicals in various exposure media is often the major source of uncertainty in Superfund risk assessments. In any risk assessment, the derivation of the concentration term will reflect assumptions about: (1) properties of the contaminant, (2) the spatial and temporal variability in contamination, (3) the behavior of the receptor, and (4) the time scale of the toxicity of the chemical(s). This appendix expands upon concepts introduced in Chapter 5. This appendix does not provide detailed equations for performing calculations, but instead refers the reader to other Environmental Protection Agency (EPA) guidance documents in which both the recommended approaches and calculations are provided.

The concentration term is linked to the concept of an exposure unit (EU). For Superfund risk assessments, an EU is the geographical area in which a receptor is randomly exposed to a contaminated medium for a relevant exposure duration. Environmental sampling provides information about the contamination within and around an EU. Multiple EUs may be defined at a site based on the choice of a receptor, the exposure medium, and the nature of contact with the medium. For example, residential exposures to children may involve exposures via soil and dust ingestion both at the primary residence and recreational areas at a day care facility. Site-specific information regarding the activities of receptors should guide assumptions about the receptor's contact with exposure media.

> ☞ *Defining the EU is critical to the success of the remedial strategy, as it affects the calculation of the concentration to which receptors are exposed.*

C.1.0 VARIABILITY IN PRA

In general, variability and uncertainty should be kept separate to the extent possible in any probabilistic risk assessment (PRA). For example, assume a one-dimensional Monte Carlo Analysis (1-D MCA) was developed to characterize variability in risk, but it combined a distribution for uncertainty in mean concentration with distributions for variability in exposure variables. The result would yield a single distribution for risk, however, each risk estimate would reflect both uncertainty and variability and distinguishing between the two would not be possible. Therefore, EPA's *Guiding Principles for Monte Carlo analysis* recommends against mixing distributions of variability and uncertainty in a 1-D MCA (U.S. EPA, 1997b) to avoid such ambiguities.

A fundamental concept in Monte Carlo analysis is that there is variability in exposure between receptors (inter-individual variability) as well as day-to-day variability for each individual (intra-individual variability). In most Tier 2 analyses (see Chapter 2), the goal of a 1-D MCA is to characterize inter-individual variability in exposure and risk. Typically, probability distributions for exposure represent variability (PDFv's) between individuals in the average value over the entire exposure duration. In this case, the exposure point concentration (EPC) should represent the average exposure concentration over the entire exposure duration. Because an EPC is calculated from a sample, there is uncertainty that

the sample mean equals the true mean concentration within the EU; therefore, to account for associated uncertainty, the 95% upper confidence limit for the mean (95% UCL) is generally used for Superfund risk assessments (U.S. EPA, 1992).

In a 1-D MCA, a point estimate for the EPC is combined with PDFv's for other variables to yield a probability distribution for risk. An alternative approach is to simulate long-term average exposures as a series of consecutive short-term exposure events. This approach is referred to as MicroExposure Event (MEE) Monte Carlo modeling, and is discussed in detail in Appendix D. In MEE modeling, the goal is to develop PDFv's for exposure variables that capture the event-to-event variability in exposures at the individual level. The concept of an averaging time still applies, but generally to a shorter time frame. For example, seasonal variability in exposure frequency might be expected among outdoor occupational workers so that different PDFv's are representative of inter-individual for each season. In this case, the EPC continues to represent an average concentration within the EU, but it would be linked to season-specific activity patterns. It may be important to develop two different weighted averages to reflect season-specific activity patterns and locations that are more frequently contacted in the summer compared with the winter, for example. As the time frame for the exposure scenario is shortened from the entire exposure duration, to a season, to a day, to an individual event, the concentration term should be reevaluated to assess the relevance of the assumption that concentrations contacted by the receptor are represented by the mean of the measured sample.

The following discussion introduces concepts of temporal and spatial variability as they apply to the estimate of the EPC for different exposure media and exposure scenarios. While the general rule of thumb applies to all Monte Carlo models—use a measure of the average concentration within the EU over the time frame of exposure—it is important to apply the site sampling data in a way that is consistent with the exposure scenario.

C.1.1 TEMPORAL VARIABILITY

Temporal variability in chemical concentrations may be an important consideration when developing a preliminary remediation goal (PRG) for any exposure medium (refer to Chapter 5 for a comprehensive discussion of using PRA to evaluate PRGs). For example, wind erosion may change chemical concentrations in surface soil over time; leaching may change concentrations in both subsurface soil and groundwater; and bioaccumulation may result in increasing concentrations in predatory fish with time. If possible, such factors should be considered early in the risk assessment process and included in the conceptual site model.

Development of the EPC normally will depend on the averaging time relevant to the exposure scenario and health endpoint of concern. In the shorter term, it may be unlikely that receptors are exposed throughout the entire EU due to temporal (and spatial) variability in the contaminant and inter-individual variability in activity patterns. Therefore, inter-individual variability in the EPC might be expected, and a distribution of EPCs may be developed to represent differences in exposure among the population. Variability in short-term exposure may be an important factor for assessing variability in acute toxicity. However, over time, short-term variability in the EPC will tend to smooth out and approach a long-term average concentration. A single estimate of the long-term average EPC may be reasonable to use in assessing risks to the receptor population. This is true regardless of the underlying distribution of the environmental sampling data (e.g., lognormal, normal, beta, etc.).

While most chemicals regulated by the Superfund program are based on concerns for chronic toxicity (e.g., lifetime cancer risk from exposure to a carcinogen for ten or more years), for some

chemicals, toxic effects occur with shorter exposure durations (e.g., nitrate in drinking water and methemoglobinemia in infants). Differences between acute and chronic health endpoints are important to consider for ecological receptors such as transient migratory species. Superfund guidance distinguishes between acute and chronic exposure to provide risk assessors the option of evaluating risk under different time frames. The EPC should be estimated within an EU during a period of time that has toxicological relevance for the exposed population.

☞ *The time scale of the concentration term should match the time scale of the toxicity criterion and exposure duration.*

C.1.2 SPATIAL VARIABILITY

Spatial variability in chemical concentrations is also an important property to consider when developing a PRG. Spatial variability arises from many factors, including the mechanism of contamination, physical and chemical dilution and transformation processes, and physical characteristics of the site (Cullen and Frey, 1999). Similarly, receptors may exhibit spatial variability in their contact with an exposure medium. In general, receptors are assumed to have equal access to all areas within an EU so that the concept of a long-term average concentration is applicable.

Often, the EPC is estimated without regard to the spatial patterns in contamination. The sampling design yields a measure of the variability in concentrations that is assumed to be representative of the receptor's contact with the exposure medium. However, even when the sampling design is representative (e.g., both are simple random samples within the EU), the concentrations may exhibit clear spatial patterns that could be used to reduce uncertainty in the EPC. Geostatistics (see Section C.5.2 and Appendix D) offers a wide range of techniques for incorporating spatial information into estimates of the EPC. These techniques are particularly useful when there is uncertainty in the representativeness of site sampling, due to a difference in scale between site sampling and the size of the EU, or the use of targeted sampling designs that oversample areas within an EU believed to contain the highest levels of contamination.

In point estimate risk assessments (Tier 1 of the PRA), the EPC is most often characterized by a point estimate of the mean concentration, typically given by the 95% UCL for the mean to account for uncertainty in the site characterization (U.S. EPA, 1992). Variability in concentrations is an important consideration for determining appropriate statistical methods used to estimate the 95% UCL. In addition, for some Monte Carlo models, a PDFv may be developed to determine the EPC for the exposure model. A PDFv for the EPC may be warranted in short-term exposure scenarios, particularly when the sampling density is relatively sparse in relation to the size of the EU (i.e., poor site characterization). For example, a risk assessment may include a future use residential scenario (e.g., currently the site is undeveloped) in which the EPC that is relevant to a potentially exposed population of children is the average concentration within a 0.5 acre lot. If the soil sampling yields 100 measurements, but a small subset of the samples (e.g., less than three) are available for any 0.5 acre area, the most appropriate measure of the average concentration for a hypothetical residence may be the maximum detected concentration or a single value from the PDFv in concentration among hypothetical receptors. In general, for any of the EU's that define a randomly located residence, the poor site characterization would be a source of uncertainty in both a point estimate and probabilistic risk assessment.

At the vast majority of sites, concentration data is the easiest data to obtain of all the exposure variables. In cases of poor site characterization, risk managers may opt to perform a point estimate risk assessment only using the maximum detected concentration and highly protective exposure assumptions.

In the scenario described above for 0.5 acre residential lots, it is possible that a residence would be located in an area in which the average concentration is represented by the maximum detected concentration in the sample. Should the risk manager opt for a Tier 1 point estimate risk assessment, the use of the maximum detected concentration of a chemical on the site should ensure the performance of a health-protective risk assessment within a smaller EU.

Consideration of variability is also warranted in short-term scenarios for ecological risk assessment (ERA) when the EU is much smaller than the site (see Section C.3.1.1). For example, the home range of the receptor populations may be relatively small in comparison to the spatial distribution of sampling locations (e.g., benthic invertebrates living in the sediment at the bottom of a river or soil invertebrates in a terrestrial habitat). In these cases, the receptor would be exposed to an area smaller than the sampling grid or measure of areal sampling density. A value from the PDFv that characterizes variability in the concentrations across a relatively large spatial scale may be used to define the EPC for a receptor population at a smaller scale. Again, risk assessors should take care in designing a 1-D Monte Carlo model when using a PDFv for the concentration term. It is unadvisable to mix a PDFv for the concentration term with PDFv's for other exposure scenarios when estimating risks within one EU. Use of the PDFv in this manner would incorrectly suggest that the mean concentration varied for each individual within the same EU according to the variability in concentration measured across a much larger area. A preferred approach is to use a PDFv to obtain a point estimate that represents the EPC, and then combine this point estimate with PDFv's for other variables in the Monte Carlo simulation to estimate risks in the small EU. If there are many EU's at a site, or if the boundaries of EUs are undefined, more advanced modeling approaches can be developed to efficiently run multiple scenarios. Methods for characterizing exposure point concentrations for ecological receptors are further discussed in Sections C.2 and C.3.

C.1.3 EXAMPLE OF TEMPORAL AND SPATIAL VARIABILITY

Exposure scenarios often require consideration of both temporal and spatial variability. The MEE might be used to assess temporal variability by simulating long-term intake as the sum of individual exposure events. The time step for MEE is an important consideration and will depend on the rate of change of the most rapidly changing exposure variable. In addition, there should be a correspondence between the time periods over which data were obtained and the time step used in the MEE model. For example, when a MEE is used for the risk assessment, the concentration term selected at each time period should match the "average" concentration within the EU appropriate for that particular time period. Assume that the receptor is a residential child, and the time period is a single day, and the child may contact only 1,000 square feet within the 0.5 acre (20,000 square feet) residential EU. The specific 1,000 square foot area may change with each day as the child chooses different areas in the yard to frequent. Hence, the variability in the sample may be a more appropriate measure of the concentration contacted by residential child receptor on a day-to-day basis than the long-term average within the 0.5 acre EU. Over the long-term, this receptor will be exposed to the entire EU and hence the average contaminant concentration within the 0.5 acre EU. Note that the day-to-day variability in concentration undergoes the familiar phenomenon of "regression to the mean" when considered over the long-term.

C.1.4 SPATIAL AND TEMPORAL VARIABILITY FOR DIFFERENT EXPOSURE MEDIA

C.1.4.1 VARIABILITY OF CONCENTRATIONS IN SOIL

Surface soil is subject to erosion by wind and surface water runoff. Over time, concentrations in surface soil may change, but generally at a slow rate relative to other media. The spatial variability of chemical contamination is most often due to the mechanism by which the contamination occurred. For example, particulate stack emissions will tend to fall in an even pattern downwind of the stack whereas over-application of pesticides and chemical spills can result in a patchy pattern of contamination.

Subsurface soil is not subject to wind erosion, so concentrations change mostly due to degradation processes or leaching of the contaminant to groundwater. At most Superfund sites, concentrations of chemicals in subsurface soil will remain relatively constant.

C.1.4.2 VARIABILITY OF CONCENTRATIONS IN GROUNDWATER

Exposure to groundwater contamination mostly occurs at a fixed point in space (e.g., the wellhead). Groundwater is subject to a variety of influences that can alter chemical concentrations within this medium such as aerobic and anaerobic biodegredation, volatization, and absorption. Due to these influences, monitored natural attenuation is an appropriate remedy under certain site conditions. If a risk assessor wishes to use a measure of the long-term average of a concentration in groundwater, a hydrogeologist should be consulted.

C.1.4.3 VARIABILITY OF CONCENTRATIONS IN SURFACE WATER

Concentrations in surface water can be very dynamic. Streams are constantly flowing and the effects of mixing, dilution and evaporation can change the chemical concentrations in surface water over relative short time periods. Any sampling of surface water is truly a “snapshot” in time. The sampling methods used to characterize spatial and temporal variability of concentrations in surface water will have a direct effect on the uncertainty in estimates of the average concentration over both short and long time frames.

C.1.4.4 VARIABILITY OF CONCENTRATIONS IN SEDIMENT

In some situations, sediment may be considered a relatively stable medium, similar to soil. Alternatively, sediment may be physically moved by currents, tides, the movement of ships and other events. Trend analysis may be used to establish the long-term average sediment transport at a site. This information could provide the basis for choosing a representative “average” concentration in the sediment available to ecological receptors (Piest and Miller, 1975; Van Sickel and Beschta, 1983; Walling, 1983; Meade et al., 1990).

C.1.4.5 VARIABILITY OF CONCENTRATIONS IN FISH

Concentrations in fish may vary due to a change in the availability of food and environmental conditions. Factors that may be used to model population dynamics may include intensity of angler harvest, death/attrition of the population, and the introduction of a predator species or a more adaptive species. In risk assessments that include a fish ingestion exposure pathway, the activities of the angler may be a more important factor in determining the EPC than the changes in concentrations in fish over time. For example, an avid recreational angler may harvest fish from different locations within a lake and

consume fish of different sizes and species. In this way, with the consumption of contaminated fish, both the contaminated medium and the exposure point change throughout the exposure duration.

Unless, samples of fish are collected over time, knowledge of these factors will generally be unknown. Concentrations of bioaccumulative chemicals in territorial fish (e.g., largemouth bass) obtained in different locations will generally reflect the concentrations in the sediment in the individual's home territory. Concentrations of bioaccumulative chemicals in migratory fish will be more difficult to predict as the fish will contact areas with varying sediment and surface water concentrations.

C.1.4.6 Examples of Temporal and Spatial Variability in the Concentration Term for Selected Exposure Media

Whatever medium is considered in the development of EPCs, the risk assessor should be aware that the EPC embodies aspects of both the spatial distribution of contamination, the movement of the receptor, and possibly the contaminated medium within the EU. Table C-1 presents examples of sources of temporal and spatial variability in the concentration term based on both the contamination in selected exposure media and the receptor.

Table C-1. Examples of temporal and spatial variability in selected media for the concentration term in common exposure scenarios.

Factor		Soil	Groundwater	Fish
Temporal Variability	Contaminant	• none, if contaminant source is inactive • aerial deposition from ongoing source emissions affected by wind patterns • degradation over time • volatilization • migration to groundwater • radioactive growth and decay	• seasonal fluctuation in groundwater table • migration of contaminant plume • natural attenuation	• seasonal changes in species availability • bioconcentration • long-term changes in population dynamics • fish tissue concentrations linked to temporal variability in water and sediment concentrations • physical and chemical processes
	Receptor	• changes in activity patterns and behaviors over time (e.g., with age)	• none, fixed location at specific wellhead • changes in well location over time	• dietary preferences for fish species • cooking practices
Spatial Variability	Contaminant	• heterogeneity in concentrations over a small area and with depth, including presence of hotspots • heterogeneity in soil properties that influence bioavailability	• migration of contaminant plume, based on hydrogeology and source emissions (e.g., bulk flow or continuous source)	• migration of fish • changes in fish population structure
	Receptor	• daily activity patterns involve contact with different areas of the EU	• none, fixed location at specific wellhead • changes in well location over time	• change in recreational habits, and areas fished

C.2.0 NONRANDOM EXPOSURES

As discussed in Section C.1.2, in the long-term it is generally assumed receptors exhibit random movement, such that there is an equal probability of contacting any area within the entire EU. Therefore, the long-term exposure concentration will most likely be the arithmetic mean of the concentration within the EU. However, in many situations, the assumption of random exposures in space may clearly be an oversimplification. People's behavior and preferences will cause them to access specific areas within an EU with greater frequency than others. The same is true in terms of ecological receptors with specific habitat preferences.

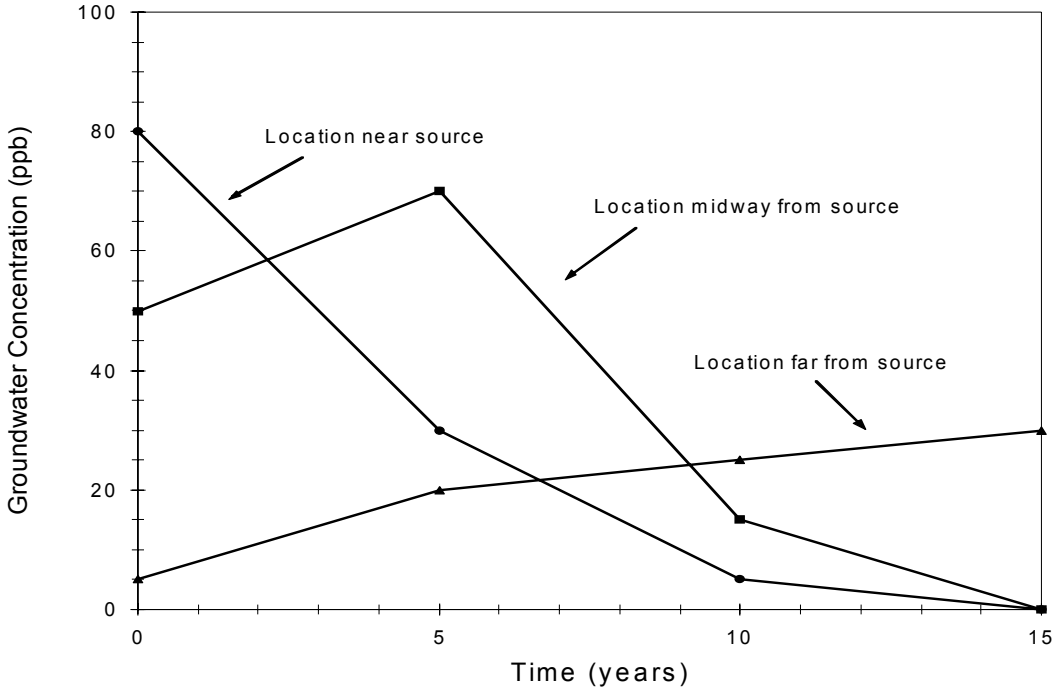


Figure C-1. Spatial and temporal variability in contaminant concentrations in groundwater.

For example, groundwater concentrations may show a large variation when sampled from wells in different locations (Figure C-1). Typically, residential receptors do not sample randomly from different wells, but draw chronically from individual wells. In such a case, the EU is a single wellhead. Fluctuations in the groundwater plume will depend on the hydrogeology of the site as well as the seasonal fluctuations in the water table. In this hypothetical example, concentrations are declining over time at distances nearest to the source, and concentrations are increasing as the plume moves farther from the source.

Incomplete information regarding the behavior patterns of people and environmental systems can be a large source of uncertainty in a risk assessment. Because of this, methods are being developed to model spatial relationships (between the contaminant and receptor) and nonrandom exposures. Recently, a quantitative technique to model nonrandom exposure has been proposed for ERA (Hope, 2000, 2001). Briefly, this technique divides the EU into smaller subunits and uses information about the attractiveness

of each subunit to assign a probability of the receptor occupying a given subunit for a period of time. Receptor movements are modeled stochastically and a time-weighted average of all the subunits provides a measure of the EPC. In some ecological risk assessments, telemetry data can be used to better characterize the areas of contamination that overlap with habitats of selected species. Hoff (1998) demonstrates an approach for American badgers (*Taxidea taxus*) in which telemetry data and geostatistical modeling provide an improved relationship between contaminant concentrations, tissue residues, and effects.

C.3.0 SOURCES OF UNCERTAINTY IN THE CONCENTRATION TERM

There are numerous potential sources of uncertainty in the estimate of the true mean concentration within an EU. As discussed in Chapter 5 (Section 5.1.1), sources of uncertainty can be grouped into four broad categories: sample data, location of the EU, behavior of the receptor, and from miscellaneous sources (e.g., physical and chemical processes). Development of an uncertainty distribution for the average concentration requires knowledge of the variability in chemical concentrations within the EU (unless distribution-free approaches are used), the toxicity of the chemicals, and the receptor's behavior. These distributions should be developed by risk assessors with the concept of the EU in mind. Differences in scale (e.g., small home range of an ecological receptor population relative to the site sampling design) can be a major source of uncertainty in ecological risk assessments. Methods for addressing such uncertainties in the concentration term are presented below. By incorporating these methods into the quantitative uncertainty analysis, risk managers may more effectively evaluate the importance of data-gaps and design subsequent rounds of site sampling to reduce the uncertainty in the EPC.

C.3.1 QUANTIFICATION OF UNCERTAINTY BASED ON THE SIZE OF THE EXPOSURE UNIT

Site characterization sometimes occurs before an EU has been defined. Therefore, an EU may be smaller than an entire site, equal to the site itself, or larger than the site. These three conditions lead to different conclusions and methods about the determination of the EPC. The most complex situation is when the EU is smaller than the site and the site can contain multiple EUs. For future scenarios in which the land use differs from the current land use, the difficulty in predicting the exact size and location of EUs necessitates accounting for the uncertainty in the EU.

Composite sampling is often used to maximize site information. However, it is important to note that the use of composite sampling influences the concentration term. If composite sampling is used exclusively at a site, the actual maximum concentration present or the best estimate of this maximum concentration will not be available. Depending on the time scale of the toxic effect or whether acute toxicity should be considered, this lack of knowledge of the maximum concentration present may be a large data gap. Risk assessors are urged to consider composite sampling and its ramifications for the concentration term.

C.3.1.1 WHEN THE EXPOSURE UNIT IS SMALLER THAN THE SITE

The size of the EU will be different depending on the length of exposure. A receptor can access a greater area if given more time. In almost all cases, the size of the EU for short-term exposure will be smaller than the EU for long-term exposure. Therefore, in addition to the uncertainty associated with sampling and analysis (which can be quantified with existing methods for calculating confidence intervals), there is uncertainty about the location of the EU within the site.

If contamination is evenly spread across the site, the location of the EU may not have any bearing on the EPC. In such a case, uncertainty may depend on the sample size or density of measurements within the EU relative to the entire site. In point estimate risk assessments, the concentrations of chemicals at the sampling location that poses the greatest risk may be considered as estimates of the EPC for this small EU. Using this "riskiest" sampling location as an estimate of the mean within an EU of unknown location accounts for both the uncertainty associated with limited sampling within a single EU and the uncertainty of the location within the site of the EU.

To express the uncertainty in location of the EU as a distribution, methods have been developed to place an EU of a given size randomly about a site (Burmaster and Thompson, 1997). A concentration term is developed for each of a large number of randomly located EUs. The distribution of these concentration terms will express the uncertainty in the location of the EU.

Risk assessors are cautioned to consider whether the statistical method used to estimate the EPC in an EU accounts for all sources of uncertainty in the concentration term. If only a few samples are used to characterize the average concentration within an EU, then the uncertainty in the EPC is large and should be presented in the risk characterization. These conditions may warrant additional sampling or the use of analytical methods that account for spatial variability within the entire site.

At some sites, geostatistical methods, pattern recognition, and geographical information systems (GIS) methods may provide additional insight and will aid in the development of the concentration term (see Section C.5.2). Although Table 3-1 shows several statistical methods for estimating both point estimates and distributions that encode uncertainty in the concentration term, a risk assessor's understanding of these uncertainties should be conceptual as opposed to purely statistical.

C.3.1.2 WHEN THE EXPOSURE UNIT IS THE SAME SIZE AS THE SITE

In this case, the entire environmental data set within the site boundaries can be used for the determination of the concentration term. Assuming the EU occupies the entire site, then the source of uncertainty associated with knowing the average concentration within the EU is the sampling and analytical uncertainty.

C.3.1.3 WHEN THE EXPOSURE UNIT IS LARGER THAN THE SITE

In this case, the EU extends beyond the site boundaries. Therefore, the entire environmental data set within the site boundaries can be used for determination of the concentration term. However, an additional term in the exposure assessment may be needed to account for the fraction of the exposures that are expected to occur off site. Essentially, the contribution of the chemical concentrations measured on and off site are weighted by the fraction ingested or contacted in each area. Similarly, the term "area use factor" is used in ecological risk assessments to refer to the percentage of time or area an animal inhabits a contaminated area. An exposure scenario in which the EU is defined by the multiple locations that may be visited would be a common extension of this concept. One reasonable assumption regarding off site exposures is that the concentrations would be equal to the "background" concentrations. If this assumption is made, a site risk assessor should be consulted to determine appropriate methods for incorporating background concentrations into the risk assessment. Alternatively, additional sampling at off site locations would be needed to estimate the concentrations.

C.4.0 SUMMARY OF RECOMMENDATIONS FOR THE CONCENTRATION TERM

Table C-2 presents general guidelines for establishing a concentration term in various media based on exposure time and the size of the EU. These general guidelines along with site-specific exposure conditions are the driving factors in risk assessment decision making for establishing the concentration term.

Table C-2. Summary of factors that may be considered in developing an EPC.

Medium	Exposure Time	Random	Non-Random	Size of EU relative to the site/sampling density	Recommendation (Human Health and Ecological)
Soil	Short-term		X	small	HH - consider variability in concentration relative to the time scale of toxicity. ECO - time weighted average of smaller subunits.
Soil	Long-term	X		variable	HH, ECO - consider uncertainty in the average concentration within an EU.
Fish	Short-term		X	variable	HH, ECO - consider variability in sample concentrations relative to the exposure time.
Fish	Long-term	X		variable	HH - consider uncertainty in the average concentration in consumed portion of fish. ECO - consider uncertainty in average concentration of whole fish.
Ground-water	Short-term		X	small - single well head	HH - consider either the highest detected concentration or uncertainty around the concentration at the center of the plume as a measure of a single well and relate to the time scale of the toxic effect. ECO - not applicable
Ground-water	Long-term		X	small - single well head	HH - consider variability among the higher concentration samples as a protective EPC. Alternatively, hydrogeologic modeling may be used to obtain a long-term average concentration in the most contaminated area. ECO - not applicable

C.5.0 METHODS FOR ESTIMATING UNCERTAINTY IN THE MEAN CONCENTRATION

Confidence intervals (CIs) and UCLs are computed to characterize uncertainty in a parameter estimate. CIs can be computed for any parameter. The general method for estimating confidence intervals is presented in equation C-1.

$$CI = \text{parameter estimate} \pm (\text{critical value}) \times SE \qquad \text{Equation C-1}$$

The parameter estimate is the estimated value for the unknown population parameter. The critical value is the number, z, with probability, p, lying to its right (for an upper critical value) or left (for a lower critical value). For a standard normal distribution (i.e., arithmetic mean=0, standard deviation=1), critical values are referred to as the *z-score* or *z-statistic*. These values are commonly given in statistics texts, and

may also be calculated using the Microsoft Excel function *Normsinv(p)*, where *p* corresponds to the probability lying to the right of the value. Distributions that characterize parameter uncertainty are sometimes referred to as sampling distributions. The standard error (SE) is the standard deviation of the sampling distribution for the parameter estimate. The confidence interval conveys two concepts: (1) an upper and lower confidence limit (for a 2-sided CI), and (2) a confidence level ($1-\alpha$), which gives the probability that the method yields an interval that encloses the parameter (Moore and McCabe, 1993). Methods for estimating SE vary for specific parameters. For example, the SE of a mean concentration may be calculated based on the sample variance and the sample size (due to Central Limit Theorem). Methods for calculating the SE for other parameters, such as the 95th percentile, are more complex, and may be estimated from a series of nested bootstrap simulations (Efron and Tibshirani, 1993; U.S. EPA, 2001a).

When comparing alternative approaches for quantifying parameter uncertainty, criteria that are important to consider include the variance of the original data set, and the bias and coverage of the CIs generated by each method. In statistics, a method is unbiased if the mean of the sampling distribution is equal to the true value of the parameter. Similarly, a method has accurate coverage if the probability *p* that a CI does not cover the true parameter is equal to the probability level used to construct the CI. For risk assessment, the most desirable method is one that deals well with high variance, yields CIs that are sufficiently wide (i.e., the CI does not underestimate the probability of enclosing the population parameter), and, more specifically, yields upper confidence limits that are not biased low. The choice of the most appropriate method will depend on the characteristics of the data set and a balance between two objectives: (1) the desire to be health protective and, therefore, have a low probability of underestimating the mean, and (2) a desire to be accurate, in the sense of choosing a method whose expected coverage equals the true coverage. As a general principle for quantitative uncertainty analysis, if alternative methods yield very different answers, it is helpful to explore the reasons for the differences. The objective is to explain why the estimates of the 95% UCL differ, and to determine if the differences are sufficiently great that they could alter the risk management decision or PRG. This information should be presented as part of the risk communication process associated with the scientific management decision points of the tiered process for PRA (see Chapter 2).

As discussed in Chapter 5, in Superfund risk assessment, the EPC is usually calculated as the 95% UCL for the mean to account for the uncertainty in estimating the average concentration within an EU. The 95% UCL is defined as a value that, when repeatedly calculated for randomly drawn subsets of size (*n*), equals or exceeds the true population mean 95% of the time. In other words, it is calculated and applied as a 1-sided confidence limit. The 95% UCL is one percentile on the probability distribution that characterizes uncertainty in the mean (i.e., the PDFu for the mean). It is equal to the 95th percentile of the sampling distribution for the mean. EPA's guidance on calculating the concentration term describes the rationale and methodology for selecting the 95% UCL as the point estimate for the concentration term (U.S. EPA, 1992).

Common methodologies for characterizing the 95% UCL for the arithmetic mean concentration include the following: (1) application of Equation C-1 using Student's t-statistic (for normal distributions), (2) Land method using H-statistic (for lognormal distributions) (Land 1971, 1975), and (3) bootstrap and Jacknife resampling techniques (Efron and Tibshirani, 1993). Details on these methods and on choosing an appropriate method are provided in the ORD/OSWER guidance bulletin, *Lognormal Distribution in Environmental Applications* (U.S. EPA, 1997a), and the more recent OSWER guidance bulletin, *Guidance on Calculation of UCLs at Superfund Sites* (U.S. EPA, 2001a). An overview of methods that may be used when data are not normal or lognormal is also provided by Schulz and Griffin (1999). It is the responsibility of the regional risk assessor to ensure that an appropriate method for

calculating a UCL or for developing an uncertainty distribution is chosen. Chapter 3 (Table 3-1) provides an overview of approaches for characterizing uncertainty in the concentration term in both 1-D MCA and 2-D MCA.

C.5.1 QUANTIFYING UNCERTAINTY WITHOUT INFORMATION ABOUT LOCATIONS OF SAMPLES AND RECEPTORS

Knowledge of both the sampling locations and the receptor's activity patterns with the EU can be used to derive a more representative estimate of the 95% UCL. If a risk assessor has access to an environmental data set without information about the sample locations, the risk assessor is forced to assume that the sample consists of a number of independent observations. The validity of this assumption depends on the unknown spatial variability of contamination at the site. The size and location of an EU, as well as the choice of a statistical method for estimating the distribution of uncertainty around the mean concentration will require often implicit (and possibly incorrect) assumptions about the spatial distribution of contamination. Similarly, if information regarding receptor activity patterns is unavailable, one must assume that any area within the EU is equally representative of potential exposures. The risk assessor is urged to explore the effects of these various assumptions and to make choices that are protective of human health and the environment.

C.5.2 QUANTIFYING UNCERTAINTY WITH INFORMATION ABOUT LOCATIONS OF SAMPLES AND RECEPTORS

In classical statistics, observations are assumed to be independent. This assumption is often invalid at contaminated sites where the method by which a chemical is released into the environment (e.g., deposition from airborne emissions; migration of contaminant plume from a point source) results in positive spatial autocorrelation. In other words, observations located next to each other tend to contain similar levels of contamination (i.e., redundant information) (Griffith and Layne, 1999). For example, the higher the spatial autocorrelation, the less incremental information is provided by adding observations in close proximity to existing observations. This decrease in the information content of a site sample is exacerbated by the tendency to choose sampling locations in the most contaminated areas rather than distributed at regular spatial intervals or specified using random sampling methodology.

At many hazardous waste sites, environmental sampling plans are designed with remedial actions rather than risk assessment in mind. Therefore, the risk assessor must establish a correspondence between the actual sampling locations and the locations a receptor would be expected to frequent. Geostatistics may provide information to establish this correspondence.

Geostatistics is a branch of spatial statistics that can be used to model spatial variability and parameter uncertainty. Geostatistics offers two fundamental contributions to risk assessment: (1) a group of methods to describe the spatial distribution of a contaminant in a quantitative fashion, and (2) the ability to maximize the information available in the data set (Deutsch and Journel, 1988; Isaacs and Srivastava, 1989).

Geostatistics is capable of using the information revealed by a correlation analysis of the data to estimate concentrations at unsampled locations. For example, geostatistics is able to use the spatial information contained in the data to model uncertainty in contaminant concentrations for areas where data are sparse, a situation commonly encountered in site assessment work. Using geostatistics, information from samples collected from outside an EU can be used to model the uncertainty in the mean concentration within an EU. Approaches that do not consider the geospatial information present in the

data are limited to the subset of samples within an EU. However, this ability to model uncertainty in areas where data are sparse is also limited, and a well characterized site is still the best path to understanding the risk at that site.

Geostatistical methods may be used to calculate a distribution of uncertainty in the mean of the concentration term for use in PRAs. In the past, geostatistics has not been widely applied to risk assessment, even though uncertainty in the exposure concentration is often a major source of uncertainty in risk estimates. Most risk assessors quantify uncertainty in the long-term average concentration without explicitly considering the spatial information present in data obtained from environmental sampling or knowledge of the receptor's movement and activities within the EU. When spatial information does not exist, the inherent assumption is that environmental sampling yields a data set that is representative of the spatial variability in concentrations encountered by a receptor. This assumption represents one source of uncertainty in the EPC. In addition, data collected outside an EU are often ignored in the analysis, even though they can provide a more comprehensive view of patterns of contamination across the site, including the EU of interest. Ignoring site-wide information may result in less informed estimates of risk and, therefore, less effective remedial designs (i.e., too little or too much remediation). In the past five years, with rapidly expanding software and hardware capabilities, some examples of the application of geostatistics can be found in exposure assessment and remedial design (e.g., Gomez-Hernandez, 1996; Goovaerts, 1996, 1997; Kriakidis, 1996; Ginevan and Splitstone, 1997; McKenna, 1997, 1998) as well as site assessment guidance (e.g., U.S. EPA, 2000).

A limit to applying geostatistics at hazardous waste sites is that the method is resource intensive and requires personnel experienced with the software and techniques. Risk assessors and risk managers should ensure that contractors and other personnel have the necessary capabilities before applying geostatistical methods to risk assessment or site cleanup. Geostatistics is a powerful tool, but it cannot incorporate quantitative knowledge regarding all sources of uncertainty. The risk assessor is cautioned to consider all possible sources of uncertainty as described in Chapter 5. As indicated previously, a full discussion of geostatistics is beyond the scope of this guidance, and interested readers are urged to consult the OSWER guidance document, *Guidance on Strategy for Surface Soil Cleanup at Superfund Sites* (U.S. EPA, 2001b).

EPA has produced several software packages used for geostatistical estimation. Among these are GEO-EAS and GEO-PACK. Expertise in geostatistics can be obtained from ORD/Las Vegas.

REFERENCES FOR APPENDIX C

Burmaster, D.E. and K.M. Thompson. 1997. Estimating Exposure Point Concentrations for Surface Soils for use in Deterministic and Probabilistic Risk Assessments. *Human Eco. Risk Assess*. 3(3): 363-84.

Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Plenum Press.

Deutsch, C.V. and Journel, A.G. 1998. *Geostatistical Software Library and User's Guid*e, *2nd E*d., Oxford University Press, NY.

Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, CRC Press.

Ginevan, M.E. and D.E. Splitstone. 1997. Improving Remediation Decisions at Hazardous Waste Sites with Risk-Based Geostatistical Analysis. *Environ. Science Tech.* 31(2):92A–96A.

Gomez-Hernandez, J.J. 1996. Issues on Environmental Risk Assessment. In: *Proceedings of the Fifth International Geostatistics Congres*s, Vol. 1. Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 15-26.

Goovaerts, P. 1996. Accounting for Local Uncertainty in Environmental Decision-Making Processes. In: *Proceedings of the Fifth International Geostatistics Congres*s, Vol. 2. Baafi, E.Y and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 929-940.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation.* New York: Oxford University Press.

Griffith, D.A. and L.J. Layne. 1999. *A Casebook for Spatial Statistical Analysi*s. Oxford University Press, NY.

Hoff, D.J. 1998. Integrated Laboratory and Field Investigations Assessing Contaminant Risk to American Badgers (*Taxidea taxus*) on the Rocky Mountain Arsenal National Wildlife Refuge. Ph.D. Dissertation, Clemson University, Clemson, S.C.

Hope, B.K. 2000. Generating Probabilistic Spatially-Explicity Individual and Population Exposure Estimates for Ecological Risk Assessment. *Risk Anal.* 20(5):575–590.

Hope, B.K. 2001. A Case Study Comparing Static and Spatially Explicit Ecological Exposure Analysis Methods. *Risk Anal.* 21(6):1001–1010.

Isaacs, E.H. and R.M. Srivastava. 1989. *An Introduction to Applied Geostatistic*s. Oxford University Press, NY.

Kriakidis, P.C. 1996. Selecting Panels for Remediation in Contaminated Soils via Stochastic Imaging. In: *Proceedings of the Fifth International Geostatistics Congres*s, Vol. 2. Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 973-983.

Land, C.E. 1971. Confidence Intervals for Linear Functions of the Normal Mean and Variance. *Ann. Math. Stat*. 42:1197–1205.

Land, C.E. 1975. Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. In: *Selected Tables in Mathematical Statistics*, Vol. 3. American Mathematical Society, Providence, RI.

McKenna, S.A. 1997. *Geostatistical Analysis of Pu-238 Contamination in Release Block D, Mound Plant, Miamisburg, Ohi*o. SAND97-0270, Sandia National Laboratories, Albuquerque, NM.

McKenna, S.A. 1998. Geostatistical Approach for Managing Uncertainty in Environmental Remediation of Contaminated Soils: Case Study. *Environmental and Engineering Geoscienc*e, 4(2):175-184.

Meade R.H., T.R. Yuzyk, and T.J. Day. 1990. *Movement and Storage of Sediment in Rivers of the United States and Canada*. In: Wolman et al. (eds) Surface Water Hydrology. The Geology of North America. Geological Society of America, Boulder, CO.

Moore, D.S. and G.P. McCabe. 1993. *Introduction to the Practice of Statistics*. W.H. Freeman and Company, NY.

Piest R.F. and C.R. Miller. 1975. *Sediment Yields and Sediment Sources*. In: Vanoni V.A. (ed.) Sedimentation Engineering, American Society of Civil Engineers, NY.

Schulz, T.W. and S. Griffin. 1999. Estimating Risk Assessment Exposure Point Concentrations when the Data are not Normal or Lognormal. *Risk Anal*. 19(4): 577– 584.

U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response. Washington, DC. OWSER Directive No. 9285.7-081.

U.S. EPA. 1997a. *Lognormal Distribution in Environmental Applications*. Office of Research and Development and Office of Solid Waste and Emergency Response. Washington, DC. EPA/600/R-97/006.

U.S. EPA. 1997b. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May 15.

U.S. EPA. 2000. *Statistical Estimation and Visualization of Ground-water Contamination Dat*a. Office of Research and Development, Washington, DC. EPA/600/R-00/034.

U.S. EPA. 2001a. *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites*. Office of Emergency and Remedial Response, Washington, DC.

U.S. EPA. 2001b. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Peer review draft. Office of Solid Waste and Emergency Response. Washington, DC. OSWER No. 9355.4-24. March.

Van Sickel J. and R.L. Beschta. 1983. Supply-Based Models of Suspended Sediment Transport in Streams. *Water Resour. Res.* 19:768–78.

Walling D.E. 1983. The Sediment Delivery Problem. *J. Hydrol*. 65:209–37.

APPENDIX D

ADVANCED MODELING APPROACHES FOR CHARACTERIZING VARIABILITY AND UNCERTAINTY

D.0 INTRODUCTION

This appendix briefly describes the following advanced modeling approaches that can be used in probabilistic risk assessment (PRA) to characterize variability and uncertainty: two-dimensional MCA (2-D MCA), microexposure event analysis (MEE), geospatial statistics, and Bayesian analysis. Except for 2-D MCA, these approaches can also be applied to point estimate risk assessment. The application of many of these approaches will require access to expertise in specialized areas of statistics and, in some cases, specialized or even custom-designed computer software. The intent here is to introduce some of the basic concepts and terminology, as well as to provide references where the reader can find more exhaustive coverage of these topics.

D.1.0 EXPRESSING VARIABILITY AND UNCERTAINTY SIMULTANEOUSLY

A Monte Carlo analysis that characterizes either uncertainty or variability in each input variable (see Chapter 1) can be described as a one-dimensional Monte Carlo analysis (1-D MCA). A 2-D MCA is a term used to describe a model that simulates both uncertainty and variability in one or more input variables. All probability distributions that are used to describe variability in a PRA model have a certain degree of associated uncertainty. For example, suppose variability in soil concentration (ppm) is estimated using a normal probability density function (PDF) defined by a mean (μ_{soil}=5) and standard deviation (σ_{soil}=1), and subjectively truncated (min, max) at (0, 50). Uncertainty in the parameter estimates can be represented in a PRA model by assuming both parameters are also random variables. To illustrate this concept, assume normal PDFs for *uncertainty* can be specified for both parameters. Uncertainty in the mean is described by the normal PDF with parameters (μ_{mean}=5, σ_{mean}=0.5); similarly, uncertainty in the standard deviation is described by the normal PDF with parameters (μ_{SD} =1, σ_{SD} =0.5). Model variables are represented in this manner when there is a compelling reason to believe that a unique probability distribution does not adequately describe one's knowledge of each variable in the model. A variable described in this way is called a second order random variable. Figure D-1 (Panel A) shows a collection of *n*=20 cumulative probability distributions (CDFs), each curve representing a unique set of (mean, SD) parameter estimates for the normal PDF for variability. Panel B shows the 90% *confidence interval*[1] based on 2,500 simulated CDFs. The 95% lower and upper bounds correspond to the distribution of 5th percentiles and 95th percentiles, respectively (i.e., CDF for 2,500 5th percentiles and CDF for 2,500 95th percentiles). The 90% credible interval (CI) for the 50th percentile is (3.4, 6.7).

[1]Note that the term "credible interval" may be more appropriate than "confidence interval" given that the range is based on subjective as well as statistical considerations. Brattin, Barry, and Chiu (1996) provide additional examples of uncertain PDFs that illustrate this concept.

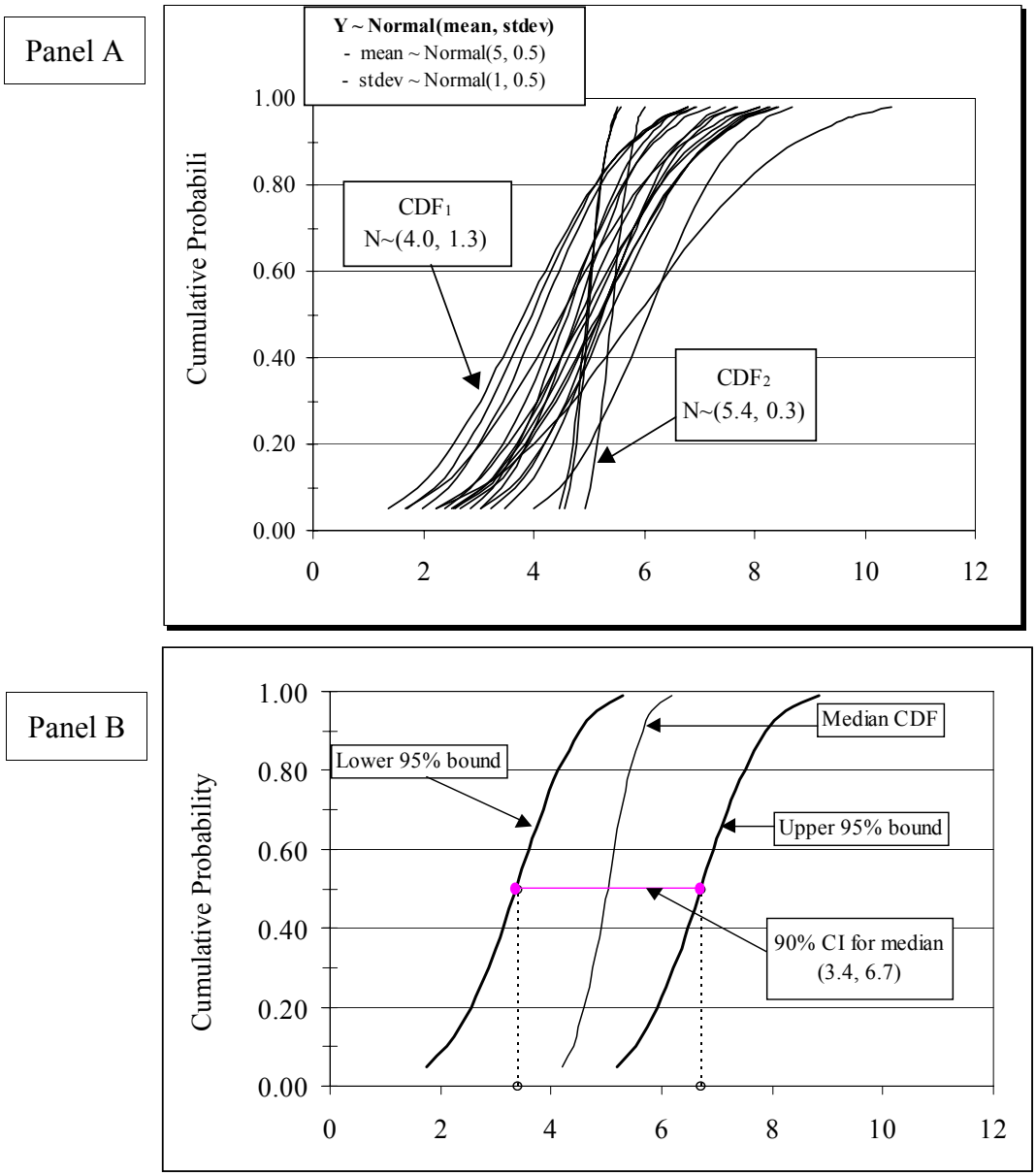


Figure D-1. Panel A shows a family of 20 CDFs for a hypothetical random variable, *Y* (e.g., concentration in units of ppm), characterized by a normal PDF where both the mean and SD are also random variables representing uncertainty in the parameter estimates: Mean~ Normal(5, 0.5), SD~ Normal(1, 0.5). Each CDF represents a single simulation of *n*=2500 iterations using a unique set of parameters. For example, CDF_1 represents N~(4.0, 1.3) while CDF_2 represents N~(5.4, 0.3). **Panel B** shows the "90% credible interval" for the CDF based on 2,500 *simulations*, each simulation using *n* = 2500 iterations (i.e., a 2-D MCA with 2,500 outer loop iterations and 2,500 inner loop iterations). Lower, median, and upper bounds represent the simulated 5th, 50th, and 95th percentiles, respectively. The 90% confidence interval for the estimate of the 50th percentile is: {3.4, 6.7}.

EXHIBIT D-1

DEFINITIONS FOR APPENDIX D

Bayesian Statistics - A specialized branch of statistics that views the probability of an event occurring as the degree of belief or confidence in that occurrence.
Geospatial Statistics - A specialized branch of statistics that explicitly takes into account the georeferenced context of data and the information (i.e., attributes) it contains.
Frequentist - A term referring to classical statistics in which the probability of an event occurring is defined as the frequency of occurrence measured in an observed series of repeated trials.
Image Analysis - A technique in geostatistics used to restore a degraded image or interpret images that have been contaminated by noise or possibly some nonlinear transformation.
Kriging - A geostatistical method of spatial statistics for predicting values at unobserved locations.
Likelihood Function - A Bayesian term referring to a probability distribution expressing the probability of observing a piece of new information given that a particular prior belief is true.
Location Tag - The spatial coordinates of a sampling location (e.g., longitude, latitude).
Microexposure Event Analysis (MEE) - An approach to modeling exposure in which long-term exposure of an individual is simulated as the sum of separate short-term exposure events.
Point Pattern Analysis - A technique in geostatistics of restricting the analysis to location information, ignoring attribute information, addresses two location problems: (1) describing points according to spacing, and (2) describing points according to density.
Posterior Distribution - A Bayesian term referring to a probability distribution that has been updated with new information.
Prior Distribution - A Bayesian term referring to the hypothesized, expected, or calculated probability distribution for an event prior to the collection of new information.
Spatial Autocorrelation - The tendency of data from locations that are relatively close together to be geographically correlated.
Thiessen (Voronoi) Polygon Analysis - A method of spatial statistics in which an area is subdivided into subregions, or polygons, in order to predict values at unobserved locations.
Time Step - A modeling term used to describe the time interval within which variable values do not change.
Two-Dimensional Monte Carlo analysis (2-D MCA) - Separate representation of variability and uncertainty in an MCA, usually accomplished using nested computation loops.

In the example shown in Figure D-1, the mean and standard deviation for soil concentration were allowed to vary independently. Thus, a distribution could be defined by a combination of a low mean and a high standard deviation, high mean and low standard deviation, or any other combination in between. The assumption of independence of variable parameters may not be valid in all cases. It may be unreasonable to assume that a high mean soil concentration would occur with a low standard deviation. An alternative assumption would be that the standard deviation of the mean is a constant proportion of the mean (i.e., a constant coefficient of variation). Correlations between parameters should be considered in the design of the PRA. One approach that is especially useful for characterizing relationships between the slope and intercept of a simple linear regression is to specify the bivariate normal distribution for the parameter estimates.

D.2.0 TWO-DIMENSIONAL MONTE CARLO ANALYSIS (2-D MCA)

Two-dimensional MCA is an approach for computing risk (or hazard) when combining distributions that represent variability and uncertainty. In 2-D MCA, distributions representing variability and uncertainty are sampled using nested computational loops (Figure D-2). The inner loop simulates variability by repeatedly sampling values for each variable from their defined probability distributions. With each circuit of the outer loop, new parameter values for each variable are selected, and the inner loop sampling is repeated. The result is a collection of inner loop simulations, one for each parameter value selected. If the inner loop samples 5,000 times, and the outer loop samples 1,000 times, then each

variable is sampled 5,000,000 times and 1,000 simulated probability distributions of risk are generated from the PRA model. These probability distributions can be analyzed to estimate the distributions for specific risk estimates. For example, confidence limits on the estimate of specific risk percentiles can be simulated using 2-D MCA (Figure D-3).

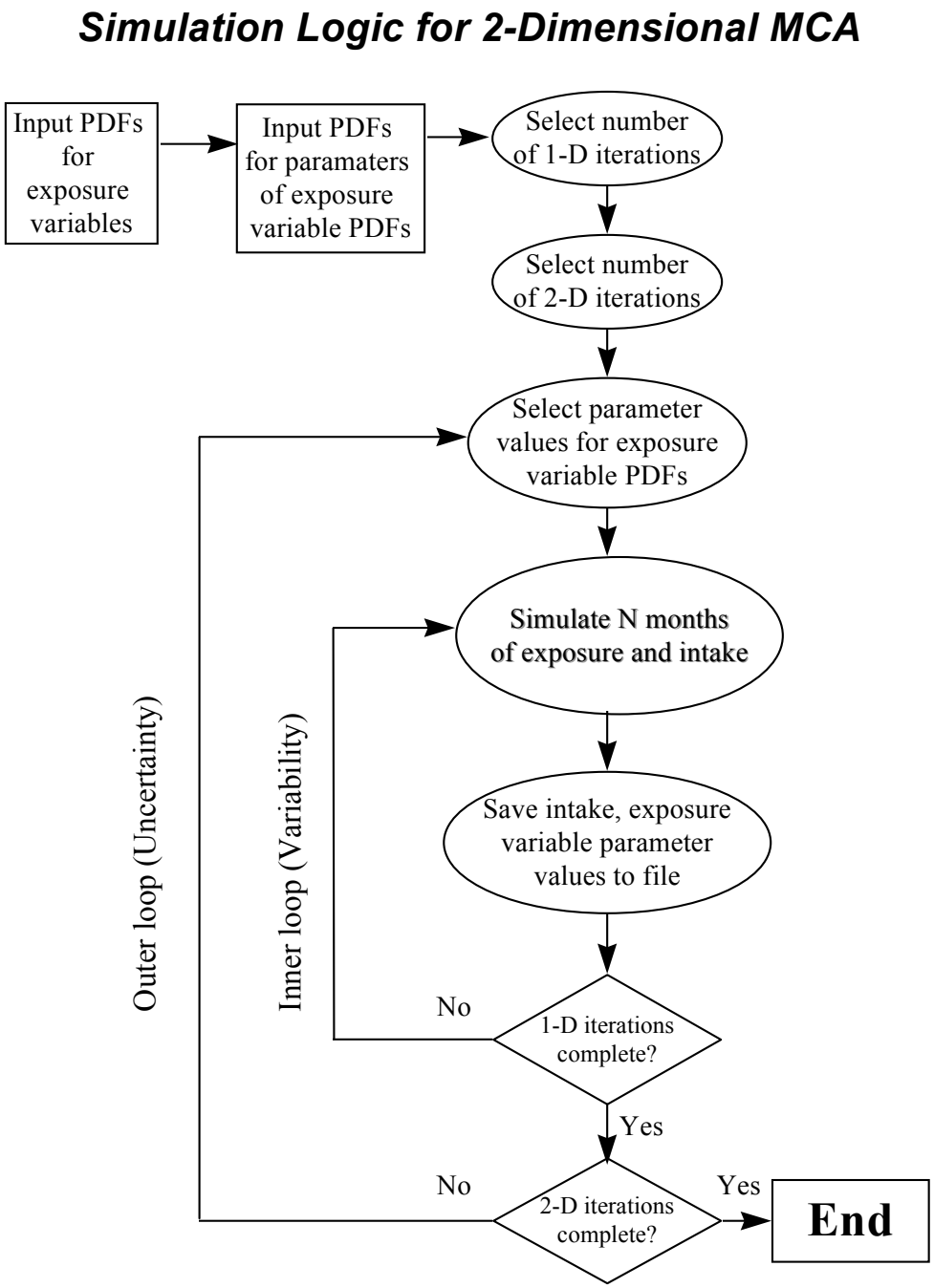


Figure D-2. Diagram showing of a 2-D Monte Carlo model in which the variability and uncertainty dimensions are computed in nested loops. In this example, values for exposure variables in the inner loop represent monthly averages.

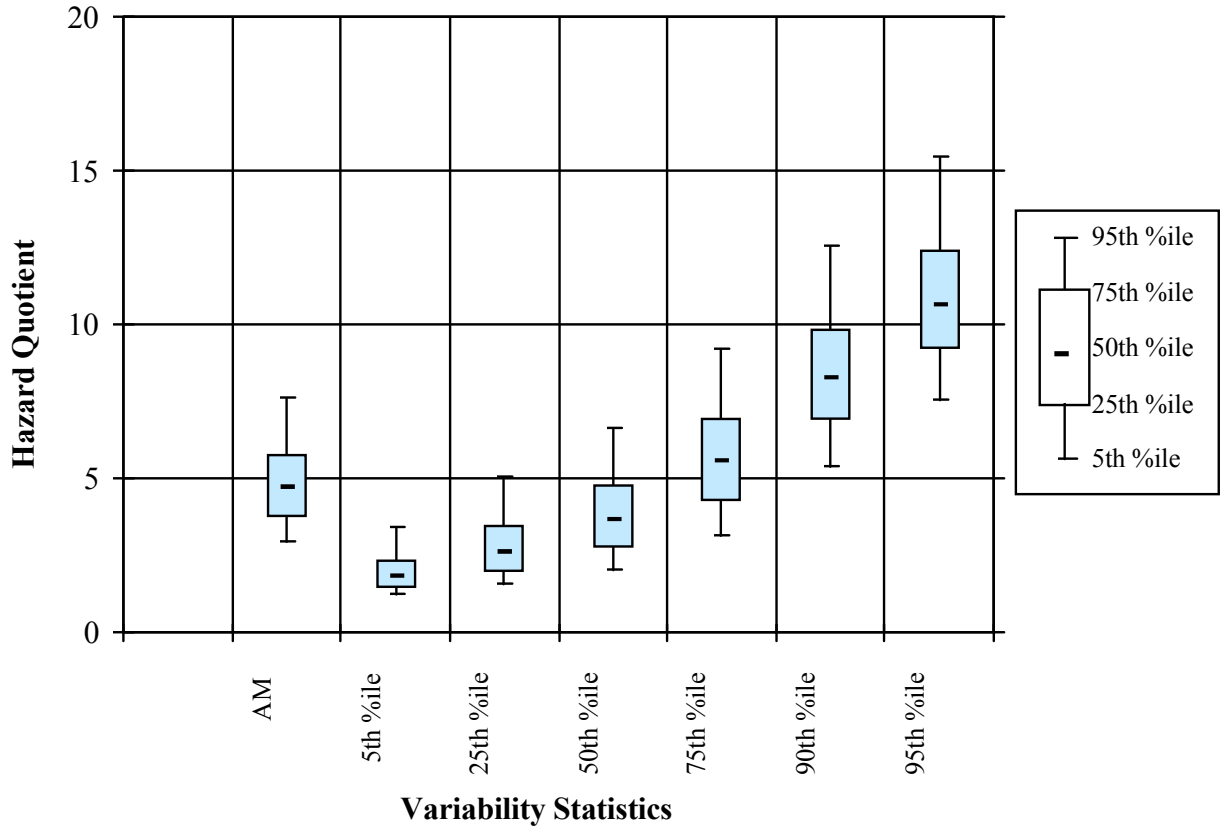


Figure D-3. Output from a 2-D MCA showing the estimated mean Hazard Quotient (HQ) and the 90% confidence interval for the arithmetic mean (AM) and selected percentiles of the HQ distribution. The 95th %ile HQ would be the reasonable maximum exposure (RME) risk estimate. The simulation suggests that there is a 95% probability that the RME HQ (95th percentile) is below 16.

D.3.0 MICROEXPOSURE EVENT ANALYSIS

The standard dose equation generally used in Superfund site risk assessments represents exposures averaged over a specified time period that is relevant to the health endpoint of concern (Equation D-1). If the risk assessment is directed at assessing life-time risk to humans, the averaging time used in Equation D-1 would generally be 70 years (i.e., estimated average human lifetime), and the calculated chemical intake would generally represent the life-time average daily dose (LADD). Where information is available to characterize variability on a smaller time scale than life-time, an alternative expression of dose that accommodates such variability may be desirable.

Standard Time-Averaging

$$DOSE = \frac{C \times IR \times EF \times ED}{BW \times AT} \qquad \text{Equation D-1}$$

Microexposure Event Modeling

$$DOSE = \frac{1}{AT} \sum_{j=1}^{ED} \frac{1}{BW_j} \sum_{i=1}^{Events_j} C_{ij} \cdot IR_{ij} \qquad \text{Equation D-2}$$

C = Concentration; I = exposure event; j = year of life
IR = Intake Rate
EF = Exposure Frequency
ED = Exposure Duration
BW = Body Weight
AT = Averaging Time

Concentrations in various environmental media can be expected to vary over time. For example, wind erosion may change chemical concentrations in surface soil. Leaching may change concentrations in both subsurface soil and groundwater. The change in the concentration term is most readily apparent when considering anglers harvesting fish. If an angler consumes a large amount of fish from a single location (e.g., a specific lake, pond, or river), then the average chemical concentration in the fish consumed by that angler can be expected to be similar to the average of the chemical concentration of fish in the population. However, if an angler consumes fish only occasionally, or harvests fish from different locations, there will be considerably more uncertainty in the concentration term. In addition, a harvesting angler may consume varying amounts of fish over the period of the exposure duration due to changing tastes, changes in the fish population size or other factors.

Daily activity patterns, food intake, soil ingestion and other behavioral factors are measured in a time period of less than a year. The extrapolation of these short term results to the chronic exposure situation is a source of uncertainty. Exposure events are real but unknowable, whereas data regarding the nature and magnitude of these events is known but its application to a real world situation is uncertain. Microexposure event analysis (MEE) attempts to explicitly quantify this uncertainty. Figure D-5 presents the general approach for MEE analysis. (Price et al., 1996, 2000). MEE modeling provides an alternative to the standard time-averaging approach represented by Equation D-1. In the MEE approach, long term intake is viewed as the sum of individual exposure events (Equation D-2). Implementing the MEE approach in a PRA requires dividing the exposure duration into short epochs, or time steps, within which the values assigned to exposure variables remain constant, but are allowed to vary from one time step to the next. In a PRA model, exposure variables are adjusted at each time step by selecting values from the probability distributions representing each variable (Figure D-4). Discussion of the implementation of

MEE analysis in risk assessment and its merits and limits can be found in Wallace et al. (1994), Price et al. (1996), Slob (1996), and Buck et al. (1997).

In MEE modeling, the time step becomes an important variable, with associated uncertainty. The time step should be selected based on information available to describe how exposures change over time. For example, a model of a moving plume of solvents in groundwater might suggest that chemical concentrations in a given location are dropping by between 16 and 25% quarterly. Several rounds of sampling may support this prediction. This rapid decline in concentrations suggests that an appropriate time step might be one quarter (i.e., three months).

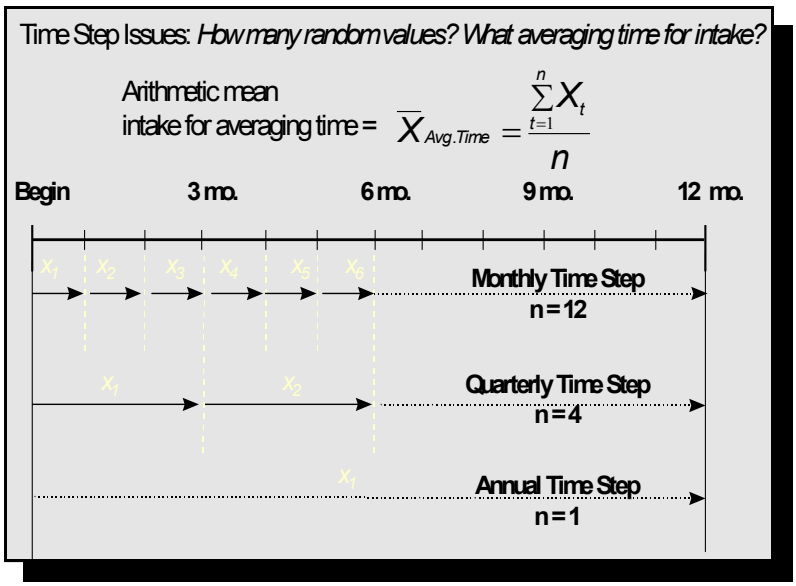


Figure D-4. Time Step for MEE.

On the other hand, where risk is being assessed for metals, dioxin, or PAHs in soil, the concentrations might be expected to change much more slowly, if at all, and the basis of the time step might be the increase in age and corresponding changes in behavior of the receptor. The time step may be global; that is, one time step may apply to all variables in the model. In this case, the same number of random values would be selected for each exposure variable in a Monte Carlo simulation. A more complex model may use different time steps for different variables, requiring some probability distributions to be sampled more often than others. The selection of a value for a time step implies that the value represents the average value for that variable during the time step.

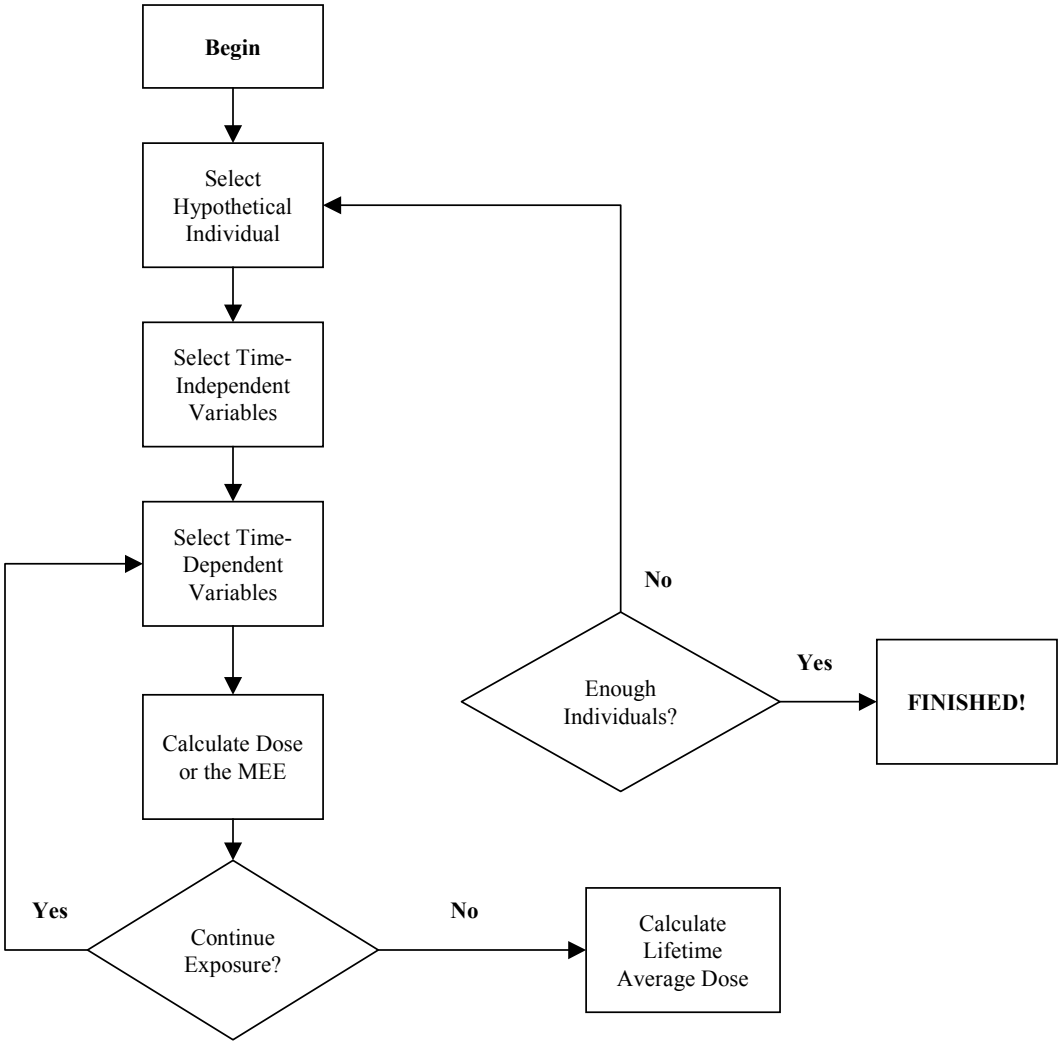


Figure D-5. Flowchart showing general approach for Microexposure Event (MEE) analysis.

Two important issues related to time step should be considered in implementing the MEE approach in PRA models. The first is the relationship between the length of the time step and the number of times random values are generated from a defined probability distribution. As the time step decreases, more time steps are needed to simulate exposures over a specified duration. For example, given a time step of one year and an exposure duration of 30 years, each random variable will be sampled 30 times (once per year); for a time step of one month and an exposure duration of 30 years, each random variable would be sampled 360 times (i.e., 12 months/year x 30 years). The Central Limit Theorem indicates that as *n* increases, the distribution of sample means is approximately normal, and the standard deviation of the sample distribution is inversely proportional to the square root of *n*. Thus a highly skewed input distribution (e.g., lognormal) may tend to become less skewed with increasing *n* (Figure D-6). A biased estimate of the RME risk in a PRA model may result if an inappropriately small or large time step is used in the model. This emphasizes the importance of having an empirical basis for selecting the time step and of exploring the time step as a variable in a sensitivity analysis of the model.

The second issue related to the time step concerns temporal correlations. Is it reasonable to assume that random values selected for consecutive time steps are completely independent? For example, consider body weight. The body weights of an individual measured at different times would be expected to show positive temporal autocorrelation; that is, body weight is likely to be similar (but not constant) from one time step to the next. For example, if an individual weighs 60 kg during one month, it is unlikely that they will weigh 80 kg the next month. If this scenario is accepted, then body weight should not be allowed to vary independently from one monthly time step to the next in the model. At shorter time steps, temporal correlation becomes more likely as a result of temporal autocorrelation. For example, one can expect a higher correlation between body weights on an individual measured on two successive days (one-day time step) than between weights measured at the midpoint of two successive years. Approaches to simulating temporal correlations in probabilistic models might include fixing an individual within a percentile range of a distribution (e.g., randomly assigned quartile) or using randomly assigned fluctuations (e.g., $BW_t = BW_{t-1} \pm x$).

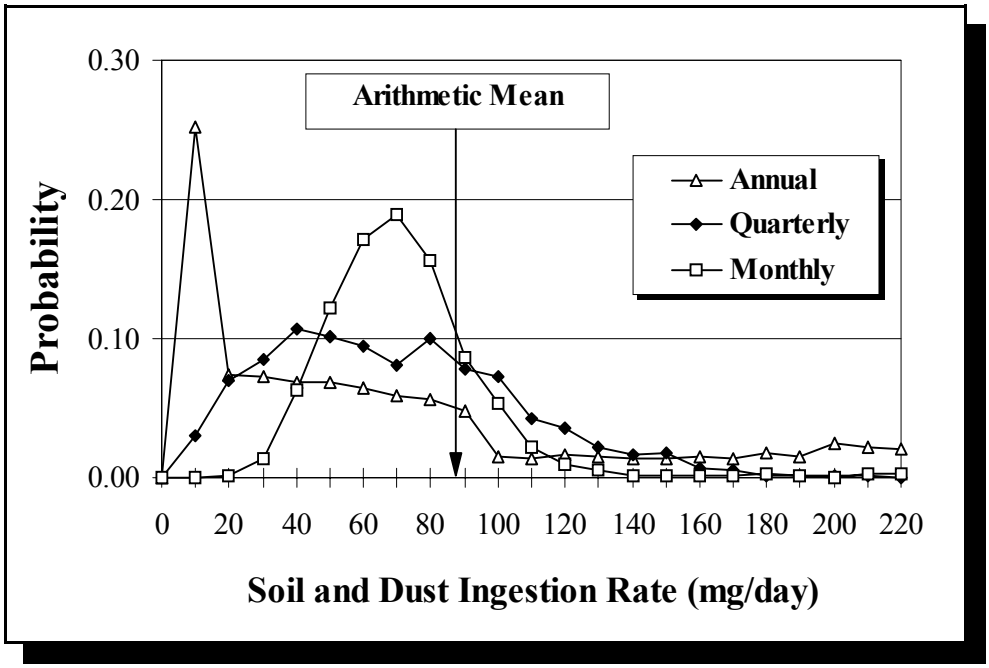


Figure D-6. Hypothetical example showing the effect of model time step on the probability distribution for soil and dust ingestion rate in children over a 1-year period. Number of samples (*n*) needed to simulate exposures: Annual (1), Quarterly (4), Monthly (12).

D.4.0 GEOSPATIAL STATISTICS

EXHIBIT D-2

POSITIVE SPATIAL AUTOCORRELATION

- Locations with a high value of Y tend to be surrounded by nearby high values of Y.
- Locations with a medium value of Y tend to be surrounded by nearby medium values of Y.
- Locations with a low value of Y tend to be surrounded by nearby low values of Y.

Spatial statistics is a specialized branch of statistics, falling under the heading of multivariate statistics, that explicitly takes into account the georeferenced or locational tagged context of data. Generally, environmental samples collected at Superfund sites have this geolocational information By acknowledging the geography of site chemicals, information about the spatial distribution of contamination can be incorporated into an exposure assessment. In addition, knowledge about a receptors home range or patterns of movement may also be incorporated into the definition of the exposure unit (see Appendix C, Section C.2.0). Explicitly accounting for spatial relationships may lead to a more accurate estimate of the confidence limits for the arithmetic mean concentration. Geospatial statistics quantifies the spatial autocorrelation (Exhibit D-2) of sample measurements and allows for the exploration of the spatial distribution of exposure and risk using techniques of map generalization. By recording locational tags for each sample, information about spatial patterns within an exposure unit (EU) can be exploited to estimate both pre- and post-remediation exposure and risk.

In the past five years, with rapidly expanding software and hardware capabilities, some examples of the application of geostatistics can be found in exposure assessment and remedial design (e.g., Gomez-Hernandez, 1996; Goovaerts, 1996, 1997; Kriakidis, 1996; Ginevan and Splitstone, 1997; McKenna, 1998; Hope, 2000; 2001) as well as site assessment guidance (e.g., U.S. EPA, 2000).

Several important risk assessment issues are closely linked to geospatial statistics, as described in Exhibit D-3. Geospatial statistics comprises:

- *spatial autoregression*
- *geostatistics*
- *point pattern analysis*
- *image analysis*

The first three of these subjects can contribute to spatial statistical support of site risk assessments. The key concept linking all three is spatial autocorrelation, which refers to covariation among samples for a single chemical, or the tendency of data from locations that are relatively close together to be geographically correlated. By analogy, classical statistics treats soil samples as though they are balls, each having a battery of attributes, that can be placed into an urn for statistical analysis; geospatial statistics treats soil samples as though they are clusters of grapes,

EXHIBIT D-3

EXAMPLES OF RISK ASSESSMENT ISSUES LINKED TO GEOSPATIAL STATISTICS

- Sampling tends to disproportionately represent "hot spots" (i.e., a relatively large portion of a data set with a small sample size (n) tends to be concentrated at "hot spots").
- The upper confidence limit (UCL) for the arithmetic mean exposure concentration (e.g., chemical concentrations in soil) depends on the sample size.
- Additional sampling may be needed, especially to better define the spatial patterns or the extent of contamination.
- There is uncertainty about locations not sampled at a site, as well as uncertainty regarding the representativeness of neighboring samples in nearby EUs.

with the branchy stems representing locational tags. Concentrations located on the same "branch" will be more strongly correlated than concentrations on different branches.

How is Geostatistics Different from Classical Statistics?

In general, geostatistics provides information beyond that provided by classical statistical techniques for at least two reasons. First, in classical statistics, observations are assumed to be independent. This assumption is often invalid at contaminated sites where the method by which a chemical is released into the environment (e.g., deposition form airborne emissions; migration of contaminant plume from a point source) often results in positive spatial autocorrelation (see Section D.4.1). In other words, observations located next to each other tend to contain similar levels of contamination (i.e., redundant information). For example, the higher the spatial autocorrelation, the less incremental information is provided by adding observations in close proximity to existing observations. This issue is compounded when the sample locations have been preferentially determined (e.g., "hot spot" sampling) rather than distributed at regular intervals or specified using random sampling methodology.

Second, geostatistics is able to use the geospatial information contained in the data to model uncertainty in contaminant concentrations for areas where data are scarce, a situation commonly encountered in site assessment work. Using geostatistics, information from samples collected from outside an EU can be used to model the uncertainty in the mean concentration within an EU. Approaches that do not consider the geospatial information present in the data are limited to the subset of samples within an EU.

D.4.1 CORRELATION AND SPATIAL AUTOCORRELATION

Several simple bivariate statistical approaches may be used to introduce the concept of spatial autocorrelation. Consider two variables, X and Y. For positive correlation there is a tendency for high values of X to be paired with the high values of Y, medium values of X to be with the medium values of Y, and low values of X with the low values of Y. The tendency is in the opposite direction for negative correlation; high values of X tend to be paired with low values of Y, and so on. Spatial autocorrelation, which virtually always is positive, directly parallels these definitions, but is written in terms of a single variable as shown in Exhibit D-2.

Just as the bivariate relationship between two variables, X and Y, can be portrayed by a scatter plot (Y versus X), the spatial autocorrelation relationship can be portrayed for a single variable, Y, (e.g., Y versus Y). A good example is the Moran scatterplot, which plots the sum or average of nearby values of Y versus Y. This plot is most effective when Y has been converted to z-scores. As shown in Figure D-7 and Section D.4.2, scatter plots can be used to illustrate some important issues related to sample size.

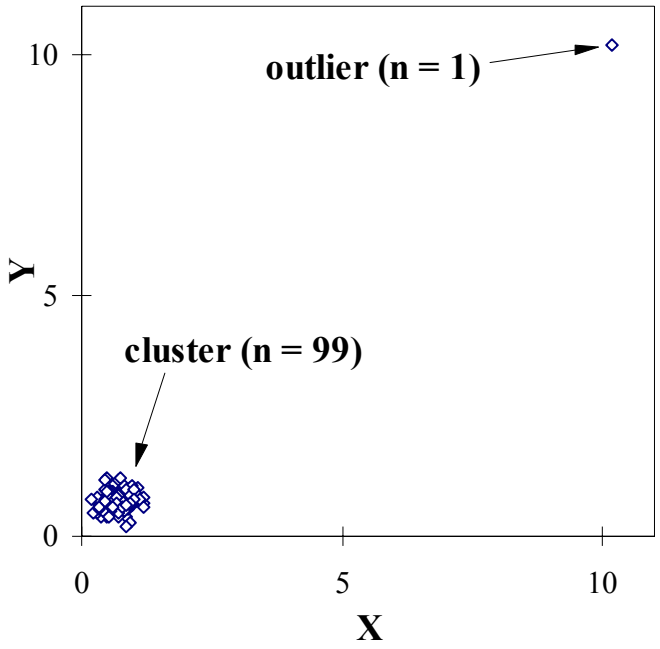


Figure D-7. Effect of an outlier on measured correlation: r=0.956 with outlier (*n*=100), whereas r=0.086 excluding outlier (*n*=99 clustered points).

If no soil samples were collected at a site (*n*=0), there is no information about the chemical concentrations in soil, and any guess may be considered an estimate. However, if the chemical concentration of a single sample (*n*=1) is measured, some information is obtained that partly restricts this estimate. As each additional independent sample is taken, more information is obtained, and the restriction on the estimate becomes more binding. If the same location is selected repeatedly for sampling, then the repeated measures, which may vary through time, will tend to be highly positively correlated; part of the information obtained from each sample is the same, and should not be counted more than once in estimating the site-wide soil concentration. Similarly, if immediately adjacent locations are sampled, the measures will often tend to be highly positively correlated (spatial autocorrelation). Once the first sample is taken, each additional sample provides only a fractional increment of new information about the site in its entirety.

D.4.2 EFFECTIVE SAMPLE SIZE (N*) AND DEGREES OF FREEDOM

Repeated measures can result in data clustering, which can be illustrated in a scatter diagram. Because two points determine a straight line, if (*n–1*) points cluster together on a scatter diagram while a single additional point occurs far away from this cluster (i.e., an outlier), then the resulting bivariate correlation will be very high (see Figure D-7). This situation alludes to the notion of effective sample size (*n**): the *n** is no longer equal to the number of observations (*n*), but rather is dramatically reduced by the presence of inter-observational correlation. For the example shown in Figure D-7, *n** is slightly greater than 2 rather than 100 (i.e., *n*).

Spatial autocorrelation plays an analogous role in georeferenced data. If a sampling network is arranged as a 25-by-25 square grid (one sample point per grid cell), and superimposed over a large site so that a very large distance separates nearby sample locations, then essentially zero spatial autocorrelation should be present in the geographic distribution of the concentrations of any given chemical. Concentrations will appear to be haphazard across the site, rendering the effective sample size as n^*=625. If the distance between nearby locations on the sampling mesh is decreased so that the spatial correlation is only r=0.050, then the effective sample size decreases to n^*=514. The effect of reducing the inter-sample distance on spatial autocorrelation and n^* for a 25-by-25 grid is shown in Exhibit D-4. If r increases to 1, then n^* reduces to 1. Therefore, obtaining a measure of latent spatial autocorrelation is essential to estimating n^*; this in turn is critical to determining confidence limits for estimates of mean concentrations, which are sensitive to sample size. The UCL for the mean will be biased *only* when very high levels of spatial autocorrelation are present; this is because the Student-t statistic used to estimate the UCL (assuming a normal distribution) changes very little as the degrees of freedom (related to sample size) increases above 10; part of the difference between n and n^* is offset by an inflation of the variance.

EXHIBIT D-4

EFFECT OF SPATIAL AUTOCORRELATION (r) ON EFFECTIVE SAMPLE SIZE (n^*)

r	n^*
0.000	625
0.050	514
0.539	64
0.957	3
1.000	1

The concept of effective degrees of freedom is important in exposure assessment because high positive spatial autocorrelation can bias the estimate of the UCL concentration if geospatial statistics are not considered. This should be of particular concern when specific locations at a site are intensively sampled (e.g., suspected "hot spots"), and other locations are relatively undersampled. Accordingly, the design of the sampling network itself can be evaluated from the perspective of geospatial statistics in order to ascertain the quality of sample information. The ideal sampling network should provide geographic representativeness, should be roughly uniformly distributed over a site, and is best implemented as a stratified random sampling design; that is, the site is partitioned into geographic stratum (e.g., EUs), and then a random sampling of points is selected within each strata. In practice, sample designs may need to focus on objectives that are in conflict with the above ideals. For example, intense sampling of suspected "hotspots" may be necessary at some sites, at the expense of a more representative spatial coverage of the site. In such cases, several statistical techniques are available for assessing the statistical benefit (in terms of reducing uncertainty) of additional sampling at undersampled locations.

D.4.3 ASSESSMENT OF ADDITIONAL SITE SAMPLING

Thiessen Polygons. In addition to calculating nearest neighbor statistics, the adequacy of a sampling network can be assessed by Voronoi (i.e., Thiessen polygon) surface partitioning, a popular approach used in mapping intra-site geographic distributions. This procedure divides a site into a mutually exclusive set of polygons, each polygon containing a single measured concentration. Each polygon has the unique property that any location within the polygon is closer to the polygon's sample location than to any other sample point (Clifford et al., 1995). The concentration measured at the sample point in the polygon is assigned to the entire area of the polygon. The intensity of sample points on a surface can be measured by Equation D-3 mean inverse polygon areas:

$$SI = \frac{1}{m} \sum_{i=1}^{m} A_i^{-1}$$ Equation D-3

where *SI* is a measure of the sampling intensity, A_i is the area of the i^{th} polygon, and *m* is the number of interior polygons (those not along the edge of the site); $m < n$. The variance of the sampling intensity can be expressed by Equation D-4:

$$SI_{Variance} = \frac{1}{m-1} \left[\sum_{i=1}^{m} A_i^{-2} - \frac{1}{m} \left(\sum_{i=1}^{m} A_i^{-1} \right)^2 \right]$$ Equation D-4

If the sampling network is uniform (i.e., polygon areas are equal), the variance will be essentially zero. The variance will increase as the network deviates from uniform. This measure can be used to assess whether or not additional samples will improve the spatial coverage.

☞ *Sampling locations that would yield a dramatic reduction in the variance should be given priority for future sampling efforts.*

Thiessen polygons can be used to develop area-weighted estimates of the arithmetic mean concentration ($C_{soil,w}$) according to the following general equation:

$$C_{soil,w} = \sum_{i=1}^{n} C_i \frac{A_i}{A_T}$$ Equation D-5

where C_i is the concentration in the i^{th} polygon, A_i is the area of the i^{th} polygon in the EU, and A_T is the total area of the EU. The weight for each measurement is essentially the ratio of the area of each polygon to the total area of the site. Clifford et al. (1995) applied this approach to an ecological risk assessment of the burrowing owl with the following simplifying assumptions: habitat range is circular, size of EU is constant (75 ha) although location may vary, and organisms spend equal time in all portions of their habitat. Given these assumptions, a nonparametric bootstrap method can be used to determine the approximate 95% UCL for the mean concentration (see Appendix C). Using Monte Carlo analysis, $C_{soil,w}$ can be estimated for different locations of the EU according to Equation D-5, and confidence limits can be generated from the multiple bootstrap estimates. Burmaster and Thompson (1997) demonstrate a similar approach in which the EU (with constant area but random rectangular dimensions) is overlayed on the Theissen polygon surface and 95% UCL for the mean is calculated from the bootstrap sample.

Linear Regression. Another diagnostic is found in the linear regression literature. The locational tag coordinates (e.g., longitude, latitude) can be converted to z-scores (say z_u and z_v) for the following calculation:

$$Y = \frac{1}{n} + \frac{z_u^2 + z_v^2 - 2r_{u,v}z_u z_v}{(n-1)(1-r_{u,v}^2)}$$ Equation D-6

where Y is a measure of the sampling network, r_{uv} is the correlation between the coordinate axes, and n is the number of samples. Any sampling location (z_u, z_v) in which $Y > 9/n$ may be considered too isolated in the sampling network. Additional sampling locations would be positioned closer to it to improve the overall coverage of the sampling network.

D.4.4 MAP GENERALIZATION

Another important application of geospatial statistics to risk assessment is that of map generalization, which draws on the subjects of geostatistics and spatial autoregression. Techniques developed for both topics exploit spatial autocorrelation in order to produce a map.

Kriging and Semivariograms. Geostatistics may employ kriging, which yields statistical guesses at values of a chemical at unsampled locations based on information obtained from sampled locations. Kriging assumes that the underlying geographic distribution is continuous, evaluates spatial autocorrelation in terms of distance separating sample points, and employs a scatter diagram similar to the Moran scatter plot to portray this relationship (i.e., the semivariogram plot: half the squared difference between measured concentrations for two sampled locations versus distance separating these two locations). The best-fit line to this scatter of points is described by one of about a dozen equations (semivariogram models).

Many different kriging approaches can be applied to quantify the spatial relationships among geographic attributes within an exposure unit. For example, site-specific chemical concentrations may be correlated with geologic information, such as glacial deposits, soil characteristics of core samples, and attributes that represent favorable habitats for ecological receptors. This information can be used to expand the available data and improve estimates of chemical concentrations at unsampled locations by employing a technique called co-kriging.

Thiessen Polygons and Spatial Autoregression. Spatial autoregression assumes a discretized surface, uses the Thiessen polygon surface partitioning to construct a Moran scatter plot, and can be used to estimate values at selected points with a regression-type equation. Theoretically, the exponential semivariogram model relates to the conditional autoregressive model, and the Bessel function semivariogram model relates to the simultaneous autoregressive model; in practice, though, the spherical semivariogram model often provides the best description of a semivariogram plot. Regardless of which approach is taken to map generalization, one relevant contribution of these two subjects is the following observation:

☞ *Including positive spatial autocorrelation results in more accurate variance estimates; this in turn yields more accurate estimates of the 95% UCL for the mean concentration.*

D.4.5 IMPLEMENTATION ISSUES RELATED TO GEOREFERENCED DATA

Estimation of parameters, for either geostatistical or spatial autoregressive models, cannot be achieved with ordinary least squares (OLS) techniques; nonlinear least squares must be used. While OLS provides unbiased regression coefficients, these estimates are not necessarily sufficient (i.e., they do not summarize all of the information in a sample pertaining to the population), efficient (i.e., the standard errors often are incorrect), and consistent (i.e., the asymptotic sampling distribution concentration will not be at the parameter value). In other words, OLS essentially uses the wrong degrees of freedom in its calculations, as described in Section D.4.2. Two additional complications of georeferenced data that do not appear in other types of data are (1) spatial autocorrelation might be directional (i.e., directional dependency); and (2) variance might be nonconstant over space as well as over the magnitude of the dependent variable, Y (e.g., chemical concentration). Several statistical approaches, which are beyond the scope of this guidance, are available for analyzing these potential sources of bias in the exposure concentration estimates (Isaaks and Srivastava, 1989; Cressie, 1991; Griffith, 1993; Ginevan and Splitstone, 1997).

D.5.0 EXPERT JUDGMENT AND BAYESIAN ANALYSIS

Up to this point in RAGS Volume 3: Part A, risk has been characterized as having a population probability distribution with parameters (e.g., mean, standard deviation) that can, theoretically, be estimated from observation. In theory, risk estimates could be derived by repeatedly measuring risk in subsets of the population of interest (e.g., repeated measurements of site-related cancer risk). The unstated expectation, or goal, is that the PRA model will accurately simulate this *real* risk distribution. This approach derives from a *classical* view of probability. The *classical* or *frequentist* view defines the probability of an event as the frequency with which it occurs in a long sequence of similar trials. From the *frequentist* perspective, the probability of having a flipped coin land *heads-up* is given by the frequency distribution of heads-up results derived from repeated similar trials of coin flips. For real-world decisions such as those informed by Superfund risk assessments, there is uncertainty that the sample data are representative of the population (see Chapter 1, Section 1.2.4).

Bayesian View of Probability. A Bayesian perspective on probability allows distributions to be constructed based on the judgment of an expert in the field. The subjectivist or Bayesian view is that the probability of an event occurring is the degree of belief a person has in the occurrence. Probabilities can be assessed by experts using scientific knowledge, judgment, data, past experience, and intuition. Different people may assign different probabilities to an event, and a single individual may assign different probabilities to the same event when considered at different times. The consequence is that probabilities become conditional and the conditions must be explicitly stated (Howson and Urbach, 1989; Morgan and Henrion, 1990; Ott, 1995; Sivia, 1996). These conditional probabilities can, of course, be updated with new information.

Using the coin flip analogy above, a Bayesian perspective might be that, based on experience with coins, assuming that most coins are *fair*, and that a fair coin would be expected to land heads-up half the time, the expected probability of the tossed coin landing heads-up is 0.5. If the outcome of repeated trials was different from the expected, the Bayesian approach would be to update the probability based on the new data. In the coin flip example, both the Bayesian and frequentist approaches will arrive at the same conclusions, because the outcome is amenable to rigorous experimentation. Where the two approaches can be expected to differ is in the assignment of probabilities to events that cannot be rigorously measured; for example, the probability of a site-related cancer risk, or the probability of a child ingesting a specific amount of soil.

The subjective judgment of experts is, therefore, an important tool in the Bayesian approach to risk assessment. For example, the input distributions for a PRA may be based upon the judgment of one or more experts who rely upon estimates from the literature, data from experimental studies, and any other information they consider relevant. Even when formal elicitations of expert opinion are not done, the final selection of the form and parameters of the input distributions usually involves some subjective judgment by the analyst. One of the challenges of incorporating judgments from experts or lay people is that there can be overconfidence bias (i.e., people tend to underestimate their uncertainty). There is a rich literature about the protocol for conducting expert elicitations and using the results to support decisions (Lichtenstein and Fischoff, 1977; Morgan and Henrion, 1990; Shlyakhter and Kammen, 1992). Elicitation of expert judgment has been used to obtain distributions for use in risk assessments (Morgan and Henrion, 1990; Hora, 1992; U.S. EPA, 1997;) and in developing air quality standards (U.S. EPA, 1982).

In addition to providing input distributions for PRAs, Bayesian analysis allows the current state of knowledge, expressed as a probability distribution, to be formally combined with new data to reach an updated information state. The distribution expressing the current knowledge is the *prior distribution* and may be the output of a PRA (Figure D-8). An appropriate *likelihood function* for the data must also be formulated. The likelihood function is based upon an understanding of the data gathering process and is used to determine the probability of observing a new set of data given that a particular risk estimate is true.

EXHIBIT D-5

COMPONENTS OF BAYES THEOREM IN PRA

- Input probability distributions for exposure (or toxicity) based on available data or expert judgment
- Prior probability distribution for risk based on input probability distributions (output from PRA)
- New data
- Likelihood function, expressing the probability of observing the new data conditional on prior risk estimates
- Posterior (updated) probability distribution for risk

Once the prior distribution is determined, the new data values are collected, and the likelihood function is assumed, Bayes theorem (Exhibit D-5) provides a systematic procedure for updating the probabilistic assessment of risk. The updated information state is called the *posterior distribution* and reflects the reduction in uncertainty arising from the new information.

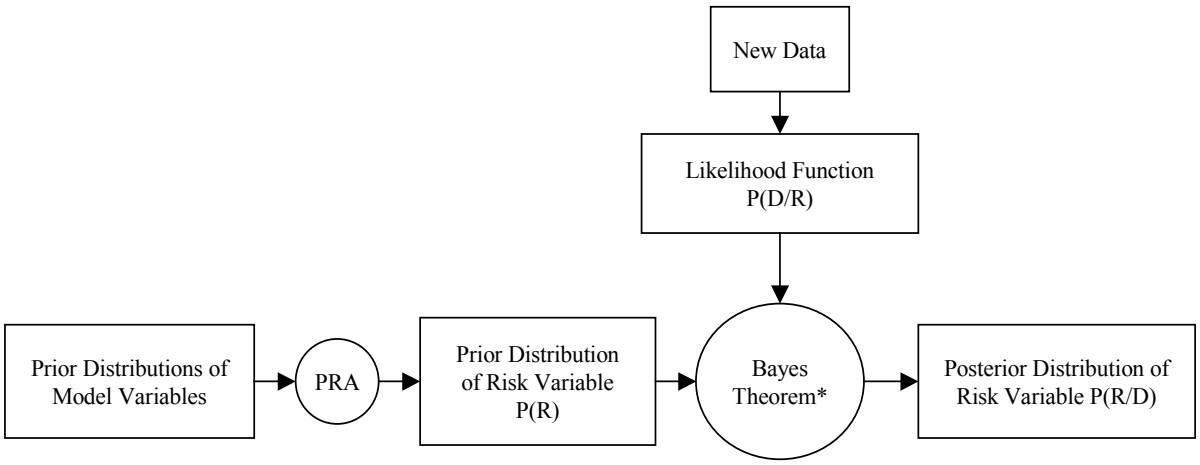


Figure D-8. Conceptual model of Bayesian Monte Carlo analysis. A PRA simulation yields a prior distribution of risk based on probability distributions for input variables. Given new data for an input variable, and a likelihood function for risk, Bayes Theorem (Eq. D-7) can be used to generate a posterior distribution of risk. The expression P(D/R) refers to a conditional probability, *"the probability of D, given R"*. Conditional probabilities can be thought of as relative frequencies, where R is the information given, and D is the event being computed when a particular value of R occurs.

$$\textit{Bayes Theorem}^{*}: \quad P(R_i/D) = \frac{P(D/R_i)\,P(R_i)}{\sum_{j=1}^{N} P(D/R_j)\,P(R_j)} \qquad \text{Equation D-7}$$

D = new data
R_i = i^{th} risk prediction associated with new data
Rj = j^{th} risk estimate simulated from PRA model
N = number of risk estimates from the PRA model

For example, suppose a model is available to relate soil tetrachlorodibenzodioxin (TCDD) concentrations at a site with serum concentrations of TCDD. A probability distribution of soil concentrations is created based upon expert judgment and a limited amount of site specific data. Using the model, the soil concentrations can be associated with a distribution of serum TCDD concentrations (P®), the prior distribution). New site-specific data (D) are subsequently collected on serum TCDD concentrations in order to reduce uncertainty in the risk estimate. Assume that it is known that serum TCDD concentrations generally follow a lognormal distribution and that the best estimate of the parameters of this distribution come from the prior distribution on serum TCDD. This creates the likelihood function (P(D|R)). Using Bayes Theorem, the new data are used to form a revised distribution of serum TCDD. This is the posterior distribution (P(R|D)).

Bayesian Monte Carlo analysis. In the past, the use of Bayesian analysis was limited by the degree of mathematical complexity involved. Using Monte Carlo analysis to carry out the PRA, rather than mathematical equations to describe the distributions, allows the calculations to be done much more easily. This variation on traditional Bayesian methods is called Bayesian Monte Carlo analysis (Patwardan and Small, 1992; Dakins et al., 1996). In the TCDD example discussed above and illustrated in Figure D-7, the required calculations are carried out for each of the N iterations of the Monte Carlo analysis (I and j go from 1 to N).

Bayesian Monte Carlo analysis is appropriate in several situations. If a model has been created and a distribution developed using PRA, new information may be incorporated without the need to repeat the entire analysis. This information could be on one of the uncertain parameters of the model or on the model output variable. Similarly, a generalized risk model with generic parameter distributions may be used for a Superfund risk assessment with the model predictions fine-tuned using data from a particular site of interest. Finally, after a distribution is developed, the amount of uncertainty that exists may be too large for the risk manager to make a decision. In this case, the risk manager might seek out new information that would refine the analysis and decrease the uncertainty.

Bayesian Monte Carlo analysis can also be combined with techniques from decision analysis to help determine the type and quantity of data that should be collected to reduce uncertainty. Decision analysis is a technique used to help organize and structure the decision maker's thought process and identify a best strategy for action. To determine the appropriate action, one defines the range of possible decisions, evaluates the expected value of the utility or loss function associated with each decision, and selects the decision that maximizes the expected utility or minimizes the expected loss.

> ☞ *Decision analysis provides a quantitative approach for evaluating the benefits of including an expanded assessment of uncertainty and the subsequent benefits of reducing this uncertainty.*

Value of Information. Value of information (VOI) analysis involves estimating the value that new information can have to a risk manager before that information is actually obtained (Clemen, 1996). It's a measure of the importance of uncertainty in terms of the expected improvement in a risk management decision that might come from better information. Examples of VOI quantities are the expected value of including uncertainty (EVIU), the expected value of sample information (EVSI), the expected value of perfect information (EVPI). Calculation of these quantities can be done using mathematical methods, numerical integration (Finkel and Evans, 1987), or Monte Carlo techniques (Dakins, 1999)

Value of information calculations require the specification of either a utility or a loss function. A loss function states the losses associated with making different types of decision errors including both direct monetary costs and losses associated with other consequences. Loss functions take various forms depending on the risk management situation (Morgan and Henrion, 1990).

Expected Value of Including Uncertainty. The expected value of including uncertainty, EVIU, is a measure of the value of carrying out a PRA. It's the difference between the expected loss of a decision based on a point estimate risk assessment and the expected loss of the decision that considers uncertainty (Figure D-9). If uncertainty in a risk assessment has been estimated using Monte Carlo techniques and a loss function has been specified, the EVIU can be easily calculated. First, the management decision from the point estimate assessment is determined. The loss from making this decision is calculated for each iteration of the Monte Carlo, each time assuming that the risk estimate from that iteration is true. The expected loss is the average of these individual losses. The expected loss for the PRA is determined by calculating the expected loss for a full range of management decisions and selecting the decision with the lowest expected loss. The EVIU is calculated by subtracting the loss associated with the PRA from that associated with the point estimate risk assessment.

Expected Value of Sample Information. The expected value of sample information is the difference between the expected loss of the decision based on the PRA and the expected loss of the decision from an improved information state. As such, the EVSI is a measure of the value that may result from the collection and use of new information (Figure D-9). Calculation of the EVSI involves a technique called preposterior analysis and is somewhat more complicated.

This type of analysis is termed "preposterior" because it involves the possible posterior distributions resulting from potential samples that have not yet been taken. For each replication from the Monte Carlo simulation, the predicted value from the model is used to randomly generate a set of K data points. Each set of data points is then used to calculate the posterior probabilities for the N Monte Carlo simulated values. These posterior probabilities are then used to obtain the optimal answer to the management question at this new level of uncertainty by selecting the decision that minimizes the expected loss over all possible management decisions.

This procedure is repeated for each of the N replications of the Monte Carlo analysis resulting in N posterior distributions, N management decisions, and N associated expected losses. Because each of these outcomes is equally weighted, the expected loss associated with the state of uncertainty expected to exist after the data collection program is carried out is simply the average of the N expected losses. The EVSI is the difference between the expected loss based on the results of the PRA and the expected loss from the updated information state.

Expected Value of Perfect Information. The EVPI is the difference between the expected loss of the decision based on the results of the PRA and the expected loss of the optimal management decision if all uncertainty were eliminated. In actual application, no research plan or data collection program can completely eliminate uncertainty, only reduce it. The EVPI is an upper bound for the expected value of efforts to reduce uncertainty and so provides the ultimate bound on what should be spent on research and data collection efforts.

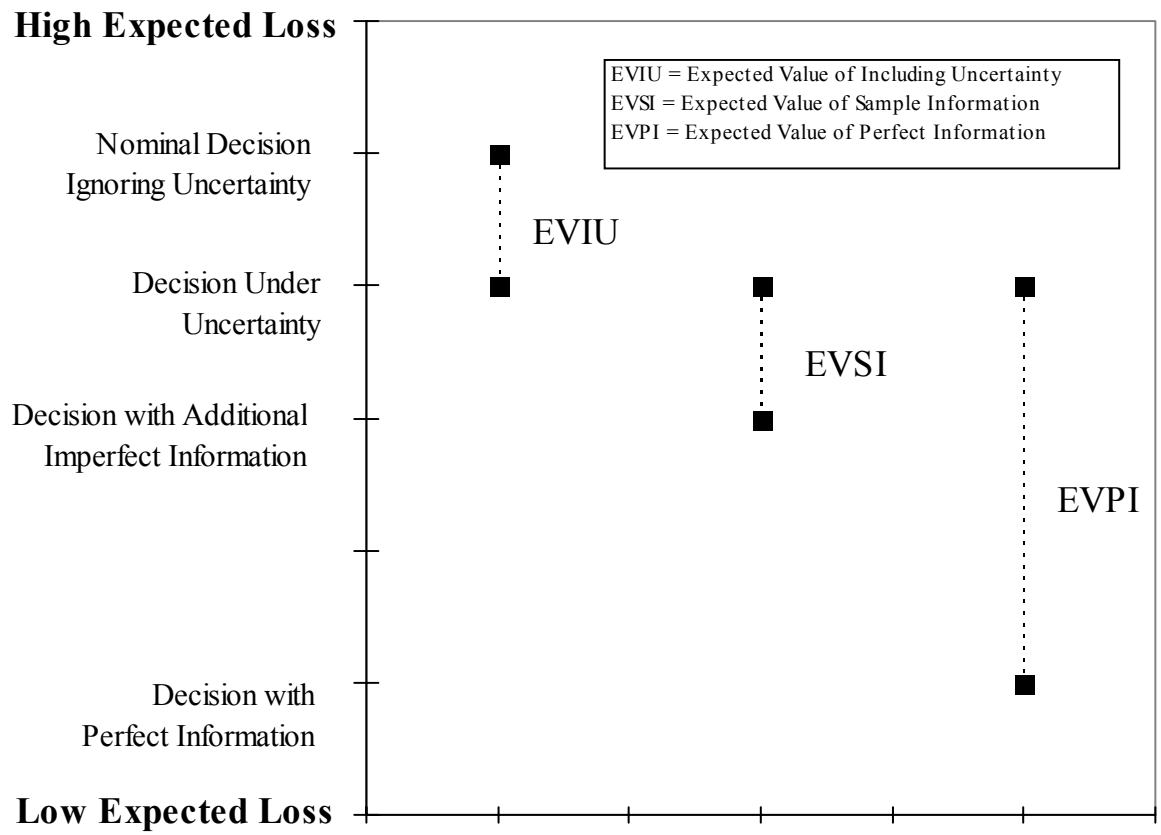


Figure D-9. Expected Loss associated with various types of information incorporated into a generic uncertainty analysis. The x-axis reflects different categories of value of information (VOI) quantities. The y-axis reflects the increasing Expected Loss with increasing uncertainty.

When a PRA has been carried out using Monte Carlo techniques, the expected loss associated with perfect information is calculated by determining the expected loss for each iteration of the Monte Carlo, assuming that the correct management decision, if that iteration were true, is made. As always, the expected loss is the average of these losses, and the EVPI is calculated by subtraction.

Uses of Value of Information in Risk Assessment. VOI analysis has many benefits for risk managers. First, VOI analysis makes the losses associated with decision errors explicit, balances competing probabilities and costs, and helps identify the decision alternative that minimizes the expected loss. VOI analysis can help a decision maker overcome a fear of uncertainty by developing a method to handle it. If the losses associated with making a poor decision are unclear, small uncertainties can take on major importance. Conversely, if the losses associated with different risk management decisions are similar, little additional effort need be expended to continue to consider the alternatives.

In addition, VOI analysis helps prioritize spending on research. It provides insights into how resources could be spent to achieve the most cost-effective reduction in uncertainty by identifying which sources of uncertainty should be reduced, what type of data should be obtained, and how much data is

needed. Finally, VOI analysis may help decision makers explain the rationale for their decisions to the public and help the public understand the multiple objectives considered in managing risks.

Expected Loss is usually greatest when uncertainty in risk estimates is ignored. For example, by quantifying uncertainty in risk (e.g., 2-D MCA, Bayesian Monte Carlo analysis) a risk manager may determine that the cleanup level associated with the 90th percentile of the risk distribution (rather than the 95th percentile) is adequately protective. Quantifying uncertainty may also result in lower expected loss when more soil remediation is required due to the losses associated with possible under-remediation, e.g., cost of additional sampling or lost revenue due to failure to meet land use requirements. The expected loss may be further reduced by collecting additional soil samples, which would presumably reduce uncertainty in estimates of mean exposure point concentrations. The expected loss may be minimized by obtaining "perfect" information (i.e., no uncertainty); however, as shown in Figure D-9, EVPI spans a wide range of expected loss because the value associated with reducing uncertainty may be tempered by costs associated with additional sampling and analysis. In practice, risk assessors consider this issue when deciding to obtain additional samples for site characterization.

The decision to obtain additional information in order to reduce uncertainty should be made on a site-specific basis, taking into account the potential impact that reducing uncertainty may have on the overall remedial decision. Important questions to consider include: (1) Are the risk estimates sufficiently sensitive to an exposure variable that collecting further data will reduce uncertainty? and (2) Are the confidence limits on the 95th percentile risk estimate sufficiently wide that reducing uncertainty may alter the cleanup goal? An example of decision framework applicable to PRA is presented in Figure D-10. The framework has three tiers. Tier 1 includes the point estimate approach and an assessment of the need for PRA. In Tier 2, the EVIU is calculated and, if warranted, a PRA is conducted. In Tier 3, the value of additional information is assessed and Bayes Theorem would be used to incorporate the new information and update probability distributions.

Limitations of These Techniques. Figure D-10 illustrates situations where Bayesian analysis and value of information quantities may not be helpful. For example, if point estimate risk assessment is selected as the appropriate method, these techniques do not apply. In addition, as site-specific data become available that are increasingly comprehensive and representative of the population of interest, Bayesian Monte Carlo analysis and the Monte Carlo analysis using the classical (frequentist) methods will approach the same result. This is because the site-specific data are incorporated into both approaches. To be representative and comprehensive, the data set must be sufficiently large, randomly selected, and represent the full range of variability that exists in the population (e.g., temporal, spatial, inter-individual). However, data sets are rarely perfect, often too small, suffer from relatively high sampling and/or measurement errors, or don't represent the entire population variability over time, space, age, gender, or other important variables. If the data cannot be assumed to describe the population distribution sufficiently well, then PRA will help to more fully develop the entire range of the population distribution and the Bayesian Monte Carlo analysis will act to refine the model estimates.

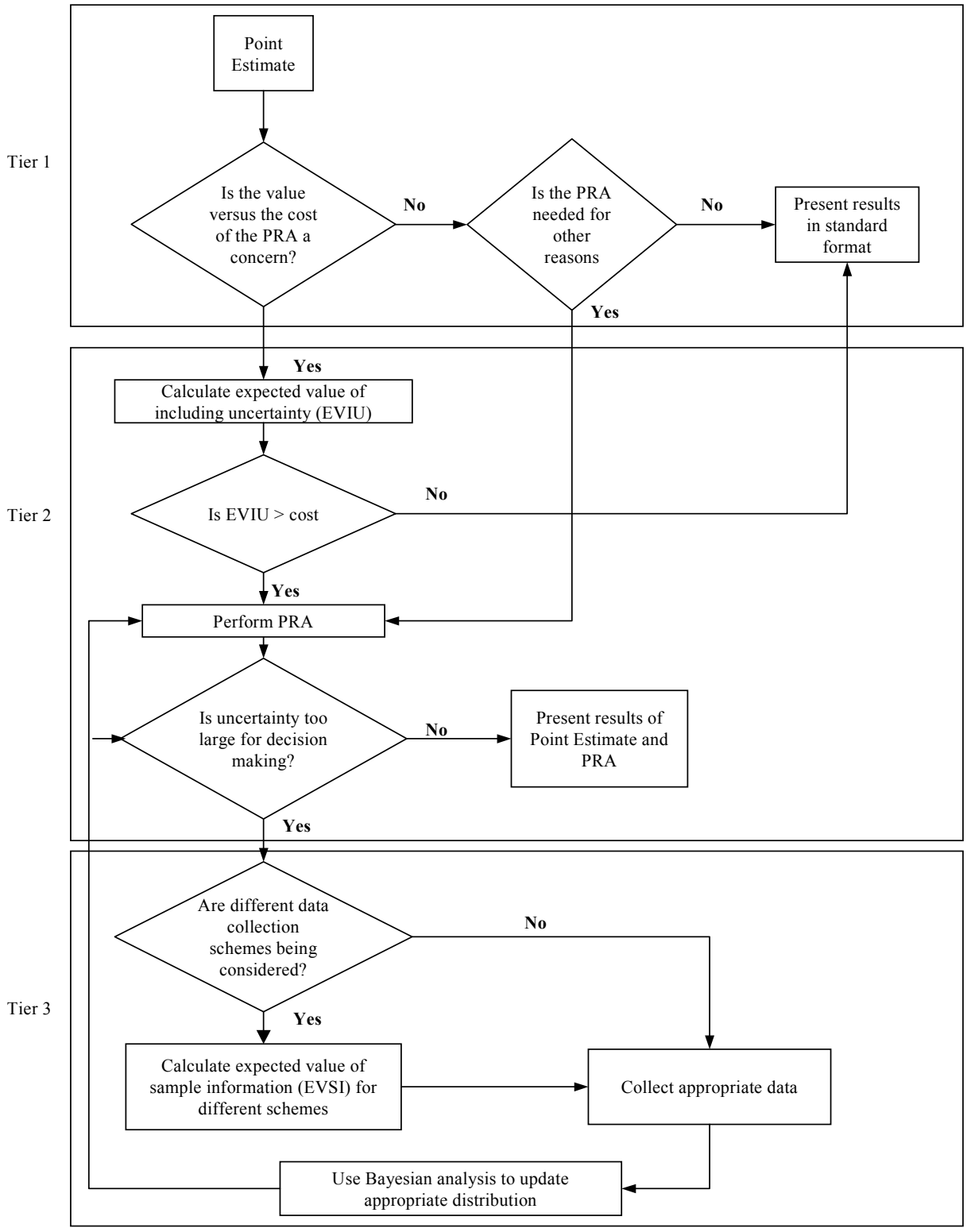


Figure D-10. Conceptual model for evaluating the expected value of including uncertainty in a Bayesian Monte Carlo anaylsis.

In order to carry out VOI calculations, a loss function must be assumed. Definition of the loss function may be complex due to multiple decision goals and/or multiple decision makers and may be difficult to capture in an equation. Finally, for Bayesian analysis and the calculation of the EVSI to be helpful, one or more sources of new data must exist. In addition, some information must be available about these data since a likelihood function describing its probability distribution must be assumed.

REFERENCES FOR APPENDIX D

Brattin, W.J., T.M. Barry, and N. Chiu. 1996. Monte Carlo Modeling with Uncertainty Probability Density Functions. *Hum. Eco. Risk Assess*. 2(4):820–840.

Buck, R.J., K.A. Hammerstrom, and P.B. Ryan. 1997. Bias in Population Estimates of Long-Term Exposure From Short-Term Measurements of Individual Exposure. *Risk Anal.* 17:455–466.

Burmaster, D.E. and K.M. Thompson. 1997. Estimating Exposure Point Concentrations for Surface Soils for Use in Deterministic and Probabilistic Risk Assessments. *Hum. Eco. Risk Assess.* 3(3):363–384.

Clemen, R.T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis*. Duxbury Press, Pacific Grove, CA.

Clifford, P.A., D.E. Barchers, D.F. Ludwig, R.L. Sielken, J.S. Klingensmith, R.V. Graham, and M.I. Banton, 1995. An Approach to Quantifying Spatial Components of Exposure for Ecological Risk Assessment. *Environ. Toxicol. Chem*. 14(5):895–906.

Cressie, N. 1991. *Statistics for Spatial Data.* Wiley, New York, NY.

Dakins, M.E., J.E. Toll, M.J. Small, and K.P. Brand. 1996. Risk-based Environmental Remediation: Bayesian Monte Carlo Analysis and the Expected Value of Sample Information. *Risk Anal.* 16:67–79.

Dakins, M.E. 1999. The Value of the Value of Information. *Human Eco. Risk Assess*. 5(2):281–289.

Finkel, A.M. and J.S. Evans. 1987. Evaluating the Benefits of Uncertainty Reduction in Environmental Health Risk Management. *J. Air Pollut. Control Assoc*. 37:1164–1171.

Ginevan, M.E. and D.E. Splitstone. 1997. Improving Remediation Decisions at Hazardous Waste Sites with Risk-Based Geostatistical Analysis. *Environ. Sci. Technol*. 31(2):92A–96A.

Gomez-Hernandez, J.J. 1996. Issues on Environmental Risk Assessment. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 1. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 15–26.

Goovaerts, P. 1996. Accounting for Local Uncertainty in Environmental Decision-Making Processes. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 2. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 929–940.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation.* Oxford University Press, NY.

Griffith, D.A. 1993. *Spatial Regression Analysis on the PC: Spatial Statistics Using SAS*. Association of American Geographers. Washington, DC.

Hope, B.K. 2000. Generating Probabilistic Spatially-Explicity Individual and Population Exposure Estimates for Ecological Risk Assessment. *Risk Anal.* 20(5):575–590.

Hope, B.K. 2001. A Case Study Comparing Static and Spatially Explicit Ecological Exposure Analysis Methods. *Risk Anal.* 21(6):1001–1010.

Hora, S.C. 1992. Acquisition of Expert Judgment: Examples From Risk Assessment. *J. Energy Eng*. 118:136–148.

Howson, C. and P. Urbach. 1989. *Scientific Reasoning: The Bayesian Approach*. Open Court, LaSalle, IL.

Isaaks, E. and R. Srivastava. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford.

Kriakidis, P.C. 1996. Selecting Panels for Remediation in Contaminated Soils via Stochastic Imaging. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 2. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 973–983.

Lichtenstein, S. and B. Fischoff, 1977. Do Those Who Know More Also Know More About How Much They Know? *Organizational Behavior and Human Performance* 20:159.

McKenna, S.A. 1998. Geostatistical Approach for Managing Uncertainty in Environmental Remediation of Contaminated Soils: Case Study. *Environ. Engin. Geosci.* 4(2), Summer, 175–184.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.

Ott, W.R. 1995. *Environmental Statistics and Data Analysis*. CRC Press. Boca Raton.

Patwardhan, A. and M.J. Small. 1992. Bayesian Methods for Model Ancertainty Analysis with Application to Future Sea Level Rise. *Risk Anal.* 12:513–523.

Price, P.S., C.L. Curry, P.E. Goodrum, M.N. Gray, J.I. McCrodden, N.H. Harrington, H. Carlson-Lynch, and R.E. Keenan. 1996. Monte Carlo Modeling of Time-Dependent Exposures Using a Microexposure Event Approach. *Risk Anal.* 16:339–348.

Price, P.S., J.Y. Young, C.F. Chaisson. 2000. *Assessing Aggregate and Cumulative Pesticide Risks Using LifeLine™* Version 1.0. A Report Submitted to U.S. EPA Science Advisory Panel. August 31.

Shlyakhter, A.I. and D.M. Kammen. 1992. Sea-level Rise or Fall? *Nature* 253:25.

Sivia, D.S. 1996. *Data Analysis: A Bayesian Tutorial*. Clarendon Press. Oxford.

Slob, W. 1996. A Comparison of Two Statistical Approaches to Estimate Long-Term Exposure Distributions from Short-Term Measurements. *Risk Anal.* 16: 195–200.

U.S. EPA. 1982. *Air Quality Criteria for Particulate Matter and Sulfur Oxides*. ECAO, Office of Research and Development. EPA/600/8-82/029.

U.S. EPA. 1997. *Exposure Factors Handbook. Update to Exposure Factors Handbook.* Office of Research and Development, NCEA. EPA/600/8-89/043, May 1989. August.

U.S. EPA. 2000. *Statistical Estimation and Visualization of Ground-water Contamination Data*. Office of Research and Development, Washington, DC. EPA/600/R-00/034.

Wallace, L.A., N, Duan, and R. Ziegenfus. 1994. Can Long-term Exposure Distributions be Predicted From Short-Term Measurements? *Risk Anal.* 14:75–85.

APPENDIX E

DEFINITIONS OF TERMS RELEVANT TO PRA AND REFERENCES FOR FURTHER READING

E.0 DEFINITIONS OF TERMS

Definitions for the specialized terms pertaining to probabilistic analysis are presented in this appendix. Some of the same terms are also defined at the beginning of each chapter, sometimes with additional examples that are relevant to concepts presented in the chapter. The definitions in this guidance are intended to be consistent with definitions used in the National Contingency Plan (NCP) and other Environmental Protection Agency (EPA) guidance, including the definitions of variability, uncertainty, and Monte Carlo simulation found in EPA's *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997a). Note that if a definition uses a term that is defined elsewhere in the appendix, it is highlighted in bold text.

Definitions of Terms Used in PRA	
50th percentile	The number in a distribution such that half the values in the distribution are greater than the number and half the values are less. The 50th **percentile** is equivalent to the **median**.
95th percentile	The number in a distribution such that 95% of the values in the distribution are less than or equal to the number and 5% of the values are greater than the number.
95% Upper Confidence Limit for a Mean	The 95 percent upper **confidence limit** (95% UCL) for a **mean** is defined as a value that, when repeatedly calculated for randomly drawn subsets of size *n*, equals or exceeds the true population **mean** 95% of the time. The 95% UCL provides a measure of u**ncertainty** in the **mean**; it is not a measure of **variability** and should not be confused with a 95th **percentile**. As sample size increases, the difference between the UCL for the **mean** and the true **mean** decreases, while the **95th percentile** of the distribution remains relatively unchanged, at the upper end of the distribution. EPA's Superfund program has traditionally used the 1-sided 95% UCL for the **mean** as the **concentration term** in **point estimates** of reasonable maximum exposure (**RME)** for human health **risk assessment** (U.S. EPA, 1992, 1997b).
Applicable or Relevant and Appropriate Requirements (ARARs)	Federal or state environmental standards; the NCP states that ARARs should be considered in determining **remediation goals**. ARARs may be selected as site-specific **cleanup levels**.
Arithmetic Mean (AM)	A number equal to the average value of a population or sample. Usually obtained by summing all the values in the sample and dividing by the number of values (i.e., sample size).
Assessment Endpoint	A term usually associated with **ecological risk assessment**; an explicit expression of an environmental value (ecological resource) that is to be protected, operationally defined by risk managers and risk assessors as valuable attributes of an ecological entity. Examples include 1) sustained aquatic community structure, including species composition and relative abundance and trophic structure; 2) reductions in populations of fish-eating birds; and 3) reductions in survival, reproduction or species diversity of indigenous benthic communities (U.S. EPA, 1997c, 1999a).

Definitions of Terms Used in PRA	
Backcalculation	A method of calculating a **preliminary remediation goal (PRG)** that involves algebraic rearrangement of the risk equation to solve for concentration as a function of risk, exposure, and toxicity.
Background Exposure	Exposures that are not related to the site. For example, exposure to chemicals at a different time or from locations other than the **exposure unit (EU)** of concern. Background sources may be either naturally occurring or anthropogenic (man-made).
Bayesian Analysis	Statistical analysis that describes the probability of an event as the degree of belief or confidence that a person has, given some state of knowledge, that the event will occur. Bayesian **Monte Carlo** combines a prior **probability distribution** and a **likelihood function** to yield a posterior distribution (see Appendix D for examples). Also called subjective view of probability, in contrast to the frequentist view of probability.
Bootstrap Methods	A method of sampling actual data at random, with replacement, to derive an estimate of a population **parameter** such as the **arithmetic mean** or the standard error of the **mean**. The sample size of each bootstrap sample is equal to the sample size of the original data set. Both parametric and nonparametric bootstrap methods have been developed.
Boxplot	Graphical representation showing the center and spread of a distribution, sometimes with a display of outliers (e.g., Figure 7-3). This guidance uses boxplots to represent the following **percentiles**: 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th}.
Cancer Slope Factor (CSF)	A plausible upper-bound estimate of the probability of a response per unit dose of a chemical over a lifetime. The CSF is used to estimate an upper-bound probability of an individual developing cancer as a result of a lifetime of exposure to a particular level of a potential carcinogen.
Central Limit Theorem	If random samples of size *n* are repeatedly drawn from a population of any distribution, the distribution of sample **means** converges to the normal distribution. The approximation improves as *n* increases.
Central Tendency Exposure (CTE)	A **risk descriptor** representing the average or typical individual in the population, usually considered to be the **arithmetic mean** or **median** of the risk distribution.
CTE Risk	The estimated risk corresponding to the **central tendency exposure.**
Cleanup Level	A chemical concentration chosen by the risk manager after considering both RGs and the nine selection-of-remedy criteria of the NCP (U.S. EPA, 1990; 40CFR 300.430(e)(9)(iii)). Also referred to as Final Remediation Levels (U.S. EPA, 1991), chemical-specific cleanup levels are documented in the Record of Decision (ROD). A cleanup level may differ from a **PRG** for several reasons, including various uncertainties in the risk estimate, the technical feasibility of achieving the **PRG**, and application of the nine criteria outlined in the NCP.
Coefficient of Variation	Ratio of the **standard deviation** (SD) to the **arithmetic mean** (AM) (CV=SD/AM). Dimensionless measure of the spread of a distribution, therefore, useful for comparing **probability density functions (PDFs)** for different **random variables**.

Definitions of Terms Used in PRA	
Community Advisory Group (CAG)	A group formed to provide a public forum for community members to present and discuss their needs and concerns related to the Superfund decision-making process. A CAG serves as the focal point for the exchange of information among the local community, EPA, State regulatory agency, and other pertinent Federal agencies involved in the cleanup of a Superfund site.
Community Involvement Coordinator (CIC)	As a member of the **CAG** and site team, the CIC coordinates communication plans (i.e., the **Communicty Involvement Plan (CIP)** and addresses site-specific **CAG** organizational issues.
Community Involvement Plan (CIP)	A plan that identifies community concerns and the preferences of the community for the communication of site-related issues.
Concentration Term	The concentration **variable** used in **exposure assessment**. Concentration terms are expressed in units applicable to the media of concern (e.g., mg/L for water, $\mu g/m^3$ for air; mg/kg for soil and dust.
Confidence Interval	A range of values that are likely to include a population **parameter**. **Confidence intervals** may describe a **parameter** of an input **variable** (e.g., **mean** ingestion rate) or output **variable** (e.g., **95th percentile** risk). When used to characterize **uncertainty** in a risk estimate, it is assumed that methods used to quantify **uncertainty** in the model inputs are based on statistical principles such as sampling distributions or **Bayesian** approaches. For example, given a randomly sampled data set, a 95% confidence interval for the **mean** can be estimated by deriving a sampling distribution from a Student's t distribution.
Confidence Limit	The upper or lower value of a **confidence interval**.
Continuous Variable	A **random variable** that can assume any value within an interval of real numbers (e.g., concentration).
Countably Infinite	Used to describe some discrete **random variables**, this term refers to a set of numbers that can be counted with integers (e.g., one, two, three) and that has no upper limit. Examples include the number of tosses required for a coin to show a head—we can count each toss, but it is possible that at least one more toss is needed. The number of dust particles in a volume of air is another example. Countably finite implies there is an upper limit (e.g., days of work per year).
Correlation	A quantitative relationship between two or more input **variables** of a model (e.g., body weight, inhalation rate, skin surface area). In analyses involving time-dependent **variables**, a change in one **variable** is accompanied by a change in another time-dependent, correlated **variable**. Ignoring correlations in **probabilistic risk assessment (PRA)** may lead to unrealistic combinations of values in a risk calculation. Correlations can also be defined as relationships between inputs and outputs.

Definitions of Terms Used in PRA	
Coverage	**Confidence intervals** are expected to enclose a true but unknown **parameter** according to a specified probability, such as 90% or 95%. This is the expected coverage of the **confidence interval**, given a specified significance level (alpha). The difference between the expected coverage and the actual coverage is one metric for evaluating statistical methods that yield different **confidence intervals**.
Credible Interval	A range of values that represent plausible bounds on a population **parameter**. Credible intervals may describe a **parameter** of an input **variable** (e.g., **mean** ingestion rate) or output **variable** (e.g., **95th percentile** risk). The term is introduced as an alternative to the term **confidence interval** when the methods used to quantify **uncertainty** are not based entirely on statistical principles such as sampling distributions or **Bayesian** approaches. For example, multiple estimates of an **arithmetic mean** may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the **arithmetic mean.**
Cumulative Distribution Function (CDF)	A graph that shows the cumulative probability of occurrence for a **random** independent **variable** (e.g., Fig. 6-1). The cumulative probability is typically given as the y-axis, ranging from 0 to 1.0. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*. Mathematically, the function that defines the CDF is obtained from the **PDF** by integration (in the case of a **continuous** random **variable**) or by summation (for **discrete random variables**).
Discrete Variable	A **random variable** that can assume any value within a finite set of values (e.g., number of rainfall events in one month) or at most a **countably infinite** set of values.
Empirical Distribution	A distribution obtained from actual data and possibly smoothed with interpolation techniques. Data are not fit to a particular **parametric distribution** (e.g., normal, lognormal), but are described by the **percentile** values.
Expected Value of Information (EVOI)	The expected increase in the value (or decrease in the loss) associated with obtaining more information about quantities relevant to the decision process. EVOI is a measure of the importance of **uncertainty** in risk and the potential for changing a **risk management** decision if **uncertainty** is reduced (see Appendix D).
Expert Judgment	An inferential opinion of a specialist or group of specialists within an area of their expertise. Expert judgment (alternatively referred to as professional judgment) may be based on an assessment of data, assumptions, criteria, models, and **parameters** in response to questions posed in the relevant area of expertise (see Appendix D).
Exposure Assessment	The qualitative or quantitative estimate (or measurement) of the magnitude, frequency, duration, and route of exposure. A process that integrates information on chemical fate and transport, environmental measurements, human behavior, and human physiology to estimate the average doses of chemicals received by individual receptors. For simplicity in this guidance, exposure encompasses concepts of absorbed dose (i.e., uptake and bioavailability).
Exposure Point Concentration (EPC)	The contaminant concentration within an **exposure unit** to which receptors are exposed. Estimates of the EPC represent the **concentration term** used in **exposure assessment**.

Definitions of Terms Used in PRA	
Exposure Unit (EU)	A geographic area where exposures occur to the receptor of concern during the time of interest. Receptors may be human or ecological (e.g., plants, birds, fish, mammals). For purposes of **PRA**, **probability distributions** for exposure and toxicity **variables** apply equally to all members of a population at a given exposure unit. Ecological exposure units often consider habitat and seasonality factors that enhance exposure in a spatial area usually related to home ranges.
Forward Calculations	A method of calculating a risk estimate that involves the standard arrangement of the risk equation to solve for risk as a function of concentration, exposure, and toxicity.
Frequency Distribution	A graph or plot that shows the number of observations that occur within a given interval; usually presented as a **histogram** showing the relative probabilities for each value. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.
Frequentist	A term referring to classical statistics in which the probability of an event occurring is defined as the frequency of occurrence measured in an observed series of repeated trials.
Geometric Mean (GM)	The n^{th} root of the product of n observations. For lognormal distributions, the GM is equal to the median and is less than the **arithmetic mean**. For normal distributions, all three measures of central tendency (GM, **AM**, **median**) are equal.
Geostatistics	Branch of statistics that focuses on data that have a spatial or geographic components. In risk assessment, geostatistics is a general term for a variety of techniques that are typically applied to chemical concentrations in soil or groundwater in which the sampling locations are considered in quantifying the **exposure point concentration**.
Goodness-of-Fit (GoF) Test	A method for examining how well (or poorly) a sample of data can be described by a hypothesized **probability distribution** for the population. Generally involves an **hypothesis test** in which the null **hypothesis** H_0 is that a **random variable** X follows a specific **probability distribution** F_0. That is, H_0: $F = F_0$ and H_a: $F \neq F_0$.
Hazard Index (HI)	The sum of more than one **hazard quotient** for multiple substances and/or multiple exposure pathways. The HI is calculated separately for chronic, subchronic, and shorter-duration exposures.
Hazard Quotient (HQ)	The ratio of estimated site-specific exposure to a single chemical from a site over a specified period to the estimated daily exposure level, at which no adverse health effects are likely to occur.
Hazardous Substance Research Centers (HSRC)	Research centers providing free technical assistance to communities with environmental contamination programs through two distinct outreach programs: **Technical Outreach Services for Communities (TOSC)** and Technical Assistance to Brownfields Community (TAB).
High-end Risk	A **risk descriptor** representing the high-end, or upper tail of the risk distribution, usually considered to be equal to or greater than the 90^{th} **percentile**.

Definitions of Terms Used in PRA	
Histogram	A graphing technique which groups the data into intervals and displays the count of the observations within each interval. It conveys the range of values and the relative frequency (or proportion of the sample) that was observed across that range.
Hypothesis Testing	Statistical test of an assumption about a characteristic of a population. The goal of the statistical inference is to decide which of two complementary hypotheses is likely to be true.
Image Analysis	A technique in **geostatistics** used to restore a degraded image or interpret images that have been contaminated by noise or possibly some nonlinear transformation.
Independence	Two events *A* and *B* are independent if knowing whether or not *A* occurs does not change the probability that *B* occurs. Two **random variables** *X* and *Y* are independent if the joint **probability distribution** of *X* and *Y* can be expressed as the product of the individual marginal **probability distributions**. That is, $f(X, Y) = f(X) \cdot f(Y)$. Independence of *X* and *Y* is *not* synonymous with zero **correlation** (i.e., $Cor(X, Y) = 0$). If *X* and *Y* are independent, then $Cor(X, Y) = 0$; however, the converse is not necessarily true because *X* and *Y* may be related in a nonlinear fashion but still maintain zero **correlation** (Law and Kelton, 1991).
Independent and Identically Distributed (IID)	**Random variables** that are independent and have the same **probability distribution** of occurrence.
Individual-Level Effect	An **assessment endpoint** that focuses on protecting a hypothetical or real individual in a population. Individual-based models may account for unique exposure and toxicological response to chemicals among individual receptors.
Iterative Reduction (IR)	A method of calculating a **PRG** that involves successively lowering the **concentration term** until the calculated risk is acceptable. This method can be applied to any medium.
Iterative Truncation	A method of calculating a **PRG** that involves developing an expression for the **concentration term** in which high-end values are "truncated" to reduce the maximum concentration, and calculating risks associated with the reduced concentration. The method may be repeated with consecutively lower **truncation** limits until risk is acceptable. Iterative **truncation** methods avoid difficulties associated with applying **Monte Carlo analysis** to a **backcalculation**.
Kriging	A statistical interpolation method that selects the best linear unbiased estimate of the **parameter** in question. Often used as a **geostatistical** method of spatial statistics for predicting values at unobserved locations based on data from the surrounding area. Information on fate and transport of chemicals within the area lacking data can be incorporated into kriged estimates.
Kurtosis	The measure of peakedness of a distribution. A uniform distribution has a lower kurtosis than a peaked distribution such as the normal and lognormal distribution. Kurtosis is referred to as the 4^{th} central **moment of a distribution**.

Definitions of Terms Used in PRA	
Land Method	The conventional method for calculating **uncertainty** in the **mean** concentration (e.g., **95% UCL**) when the sample data are obtained from a lognormal distribution (U.S. EPA, 1992).
Latin Hypercube Sampling (LHS)	A variant of the **Monte Carlo sampling** method that ensures selection of equal numbers of values from all segments of the distribution. LHS divides the distribution into regions of equal sampling **coverage**. Hence, the values obtained will be forced to cover the entire distribution. It is more efficient than simple random sampling, i.e., it requires fewer iterations to generate the distribution sufficiently.
Likelihood Function	A term from **Bayesian** statistics referring to a **probability distribution** that expresses the probability of observing new information given that a particular belief is true.
Local Sensitivity Analysis	Evaluation of the model sensitivity at some nominal points within the range of values of input variable(s).
Location Tag	The spatial coordinates of a sampling location (e.g., longitude, latitude).
Low-end Risk	A **risk descriptor** representing the low-end, or lower tail of the risk distribution, such as the 5^{th} or 25^{th} **percentile**.
Maximum Detected Concentration (MDC)	The maximum concentration detected in a sample.
Mean	**Arithmetic mean** or average; the sum of all observations divided by the number of observations. Referred to as the first central **moment of a distribution**.
Microexposure Event (MEE) Analysis	A method of assessing risk based on an aggregate sum of a receptor's contact with a contaminated medium. MEE analysis simulates lifetime exposure as the sum of many short-term, or "micro" exposures (see Appendix D). MEE approaches can be used to explore **uncertainty** associated with the model **time step** in **PRA** (e.g., use of a single value to represent a long-term average phenomenon, seasonal patterns in exposure, or intra-individual **variability**).
Mode	The most probable value of a **random variable**; a value with the largest probability or highest probability density (or mass for discrete **random variable**). The second **parameter** of a triangular distribution.
Moments of a Distribution	Similar to a **parameter**; constant that represents a mathematical description of a **random variable**. Central moments are defined with respect to the **mean**. **Mean**, **variance**, **skewness**, and **kurtosis** are the first, second, third, and fourth central moments of a **probability distribution**.

Definitions of Terms Used in PRA	
Monte Carlo Analysis (MCA) or Simulation	A technique for characterizing the **uncertainty** and **variability** in risk estimates by repeatedly sampling the **probability distributions** of the risk equation inputs and using these inputs to calculate a distribution of risk values. A set of iterations or calculations from Monte Carlo sampling is a simulation. For example, a single iteration for risk from ingestion of water may represent a hypothetical individual who drinks 2 L/day and weighs 65 kg; another iteration may represent a hypothetical individual who drinks 1 L/day and weighs 72 kg.
Monte Carlo Sampling	A method of simple random sampling used to obtain a distribution of values which may serve as an input to a **PRA**. The probability of obtaining any given sample is similar to the probability of a sample occurring within the distribution. Hence, for a given sample size, simple random sampling tends to produce values clustered around the **mean** of the distribution.
Multiple Regression Analysis	A statistical method that describes the extent, direction, and strength of the relationship between several (usually continuous) independent **variables** (e.g., exposure duration, ingestion rate) and a single continuous dependent **variable** (e.g., risk).
Nonparametric Method	A procedure for making statistical inferences without assuming that the population distribution has any specific form such as normal or lognormal. Sometimes referred to as *distribution-free* methods. Common examples are the sign test, **Spearman rank correlation**, and the **bootstrap**-t approach.
Numerical Stability	The property of a probabilistic simulation such that the a **parameter** value of the output distribution (e.g., **percentile**, **mean**, **variance**, etc.) remains sufficiently constant for a specified number of **Monte Carlo** iterations. Numerical **stability** is a measure of the precision of the output from a simulation; the tails of the distribution are typically less stable than the center. Sufficient precision is determined by professional judgment.
One-dimensional Monte Carlo Analysis (1-D MCA)	A method of simulating a distribution for an endpoint of concern as a function of **probability distributions** that characterize **variability** or **uncertainty**. In this guidance, distributions used to characterize variability may be abbreviated **PDF**v, whereas distributions used to characterize uncertainty may be abbreviated **PDF**u. It is good practice *not* to combine **PDF**s for **variability** and **uncertainty** in 1-D MCA.
Parameter	A value that characterizes the **probability distribution** of a **random variable**. For example, a normal **probability distribution** may be defined by two parameters (e.g., **AM** and SD). It is important to distinguish between this definition, and a second popular use of the term parameter when referring to an input variable in a mathematical equation or model. For this guidance, the term ***variable*** will be used to describe inputs to a model. For example, if body weight is a variable in the exposure assessment that we define with a probability distribution (e.g., normal) we would state that the variable is body weight and the parameters are the arithmetic **mean** and **standard deviation** values that characterize the normal distribution
Parametric Distribution	A theoretical distribution defined by one or more **parameters.** Examples are the normal distribution, the lognormal distribution, the triangular distribution, and the beta distribution.

Definitions of Terms Used in PRA	
Percentile	The *p^{th} percentile* of the distribution is the value such that *p* percent of the observations fall at or below it. Also called *quantiles* or *fractiles*; percentiles are expressed as a percent, ranging from 0 to 100, whereas quantiles or fractiles range from 0 to 1.
Point Estimate	A quantity calculated from values in a sample to represent an unknown population **parameter**. Point estimates typically represent central tendency or upper bound estimate of **variability**.
Point Estimate Risk Assessment	The familiar **risk assessment** methodology in which a single estimate of risk is calculated from a set of **point estimates**. The results provide **point estimates** of risk for the **CTE** and **RME** exposed individuals. **Variability** and **uncertainty** are discussed in a qualitative manner.
Point Pattern Analysis	A technique in **geostatistics** of restricting the analysis to location information, ignoring attribute information, addresses two location problems: (1) describing points according to spacing, and (2) describing points according to density.
Population-Level Effect	An ecological term for an **assessment endpoint** that focuses on protecting a group of individuals within a specified **exposure unit** and time that have similar exposures and toxicological responses to chemicals.
Posterior Distribution	A term from **Bayesian** statistics referring to a **probability distribution** that has been updated with new information.
Potentially Responsible Party (PRP)	Individuals, companies, or any other party that is potentially liable for Superfund cleanup costs.
Power	The probability that a test procedure detects a false null **hypothesis**; Power equals (1-β), where β is the probability of a **Type II error** (i.e., accepting H_0 when H_a is true). Power curves are a function of a fixed significance level (α), sample size, and **variability** (**SD**).
Preliminary Remediation Goal (PRG)	A chemical concentration in an environmental medium associated with a particular exposure scenario that is expected to be protective of human health and ecosystems. PRGs may be developed based on (**ARARs**), or exposure scenarios evaluated prior to a **risk assessment** (e.g., generic PRG) or as a result of the baseline **risk assessment** (site-specific PRG). Exhibit 5-1 provides further detail on generic and site-specific PRGs.
Prior Distribution	A **Bayesian** term referring to the hypothesized, expected, or calculated **probability distribution** for an event prior to the collection of new information.
Probabilistic Risk Assessment (PRA)	A **risk assessment** that uses probabilistic methods to derive a distribution of risk or hazard based on multiple sets of values sampled for **random variables**.

Definitions of Terms Used in PRA	
Probability Density Function (PDF)	A graph that shows the probability of occurrence of an unknown or **variable** quantity. A PDF is used to characterize a **continuous random variable;** the integral of all possible values is equal to 1.0 (i.e., the area under the curve). In PRA, PDFs can be used to display the shape of the distribution for an input variable (e.g., normal distribution for ingestion rate) as well as the output from a **Monte Carlo simulation** (e.g., risk distribution).
Probability Distribution	A function that associates probabilities with the values taken by a **random variable**. A probability distribution can be displayed in a graph (e.g., **PDF** or **CDF**), summarized in a table that gives the distribution name and parameters, or expressed as a mathematical equation. In PRA, the process of selecting or fitting a distribution that characterizes **variability** or **uncertainty** can also be referred to as applying a *probability model* to characterize **variability** or **uncertainty**. In this guidance, the probability model is considered to be one source of model uncertainty.
Probability Mass Function (PMF)	A **histogram** that shows the probability of occurrence of an unknown or **variable** quantity. A PMF is used to characterize a **discrete random variable**; similar to the **PDF**, the sum of all possible values of a PMF is equal to 1.0. The mass at a point refers to the probability that the **variable** will have a value at that point.
Random Variable	A **variable** that may assume any value from a set of values according to chance. Discrete random **variables** can assume only a finite or **countably infinite** number of values (e.g., number of rainfall events per year). A random value is continuous if its set of possible values is an entire interval of numbers (e.g., quantity of rain in a year) **variable** that may assume any of a set of values. The likelihood of each value is described by a **probability distribution**.
Range Sensitivity Analysis	Evaluation of the model sensitivity across the entire range of values of the input variable(s).
Rank	If a set of values is sorted in ascending order (smallest to largest), the rank corresponds to the relative position of a number in the sequence. For example, the set {7, 5, 9, 12} when sorted gives the following sequence {5, 7, 9, 12} with ranks ranging from 1 to 4 (i.e., rank of 5 is 1, rank of 7 is 2, rank of 9 is 3, and rank of 12 is 4).
Rank Correlation (Spearman Rank Order Correlation Coefficient)	A "distribution free" or nonparametric statistic *r* that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative **variables**.
Remedial Investigation/Feasibility Study (RI/FS)	Studies undertaken by EPA to delineate the nature and extent of contamination, to evaluate potential risk, and to develop alternatives for cleanup.
Reasonable Maximum Exposure (RME)	The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989, 1990). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.
RME Risk	The estimated risk corresponding to the **reasonable maximum exposure**.

Definitions of Terms Used in PRA	
Reference Dose (RfD)	An estimate of an exposure level for the human population, including sensitive subpopulations, that is likely to be without an appreciable risk of deleterious effects during a lifetime. Chronic RfDs are specifically developed to be protective for a long-term exposure to a chemical (e.g., >7 years) and account for **uncertainty** spanning perhaps an order of magnitude or greater.
Remediation Action Level (RAL)	Generally, a concentration such that remediation of all concentrations above this level in an **exposure unit** will result in the **95% UCL** being reduced to a level that does not pose an unacceptable risk to an individual experiencing random exposures. The RAL will depend on the **mean**, **variance**, and sample size of the concentrations within an **exposure unit** as well as considerations of acute toxicity of the chemicals of concern.
Remediation Goal	Generally, a health-based chemical concentration in an environmental medium chosen by the risk manager as appropriate for a likely land use scenario.
Risk Assessment	The use of available information to make inferences about the health effects associated with exposure of individuals or populations to hazardous materials or situations. Components of risk assessment include: hazard identification, dose-response assessment, **exposure assessment**, and **risk characterization** (NRC, 1983).
Risk Characterization	A component of **risk assessment** that describes the nature and magnitude of risk, including **uncertainty**. In assessments of Superfund sites, it includes the summary and interpretation of information gathered from previous steps in the site **risk assessment** (e.g., data evaluation, **exposure assessment**, toxicity assessment), including the results of a probabilistic analysis.
Risk Descriptor	A statistic (e.g., arithmetic **mean**, 95th **percentile**) that describes the risk to the **assessment endpoint**.
Risk Management	The process by which regulatory decisions are made using all available **risk assessment** information (including, but not limited to, the results of the **PRA**). The NCP provides nine criteria for remedial decisions (e.g., protection of human health, compliance with **ARARs**, etc.). Risk managers may include the Remedial Project Manager (RPM), section and branch chiefs, etc.
RME Range	The 90th to 99.9th **percentiles** of the risk distribution generated from a **PRA**, within which an **RME** risk value may be identified. The **95th percentile** is generally recommended as the starting point for specifying the **RME** risk in a Superfund **PRA**.
Scientific/Management Decision Point (SMDP)	A point during the **risk assessment** process when the risk assessor communicates results of the assessment at that stage to the risk manager. At this point, the risk manager determines whether the information is sufficient to arrive at a decision regarding **risk management** strategies and/or if additional information is needed to characterize risk.

Definitions of Terms Used in PRA	
Sensitivity Analysis	Process for identifying the important sources of variability and uncertainty in a model's output. Different techniques can be used in each of the 3 tiers of the tiered process for PRA (see Chapter 2). In Tier 1, sensitivity ratios are used to quantify the effects of changes in one or more model inputs on the model output. In Tiers 2 and 3, correlation analysis can be used to rank inputs based on their relative contribution to **variance** in risk. Local sensitivity refers to nominal changes in inputs within a plausible range, whereas range sensitivity refers to changes in inputs across the minimum and maximum values of the plausible range. Further explanations of the different methods for sensitivity analysis are given in Appendix A.
Sensitivity Ratio	Ratio of the change in model output per unit change in an input **variable**; also called *elasticity*.
Skewness	The measure of asymmetry of a distribution. Coefficients of skewness are zero for symmetric distributions (e.g., normal), positive for right-skewed distributions (e.g., lognormal), and negative for left-skewed distributions (e.g., specific forms of beta) . Referred to as the third central **moment of a distribution**.
Spatial Autocorrelation	The tendency of data from locations that are relatively close together to be geographically correlated.
Stakeholder	Any individual or group who has an interest in or may be affected by EPA's site decision-making process.
Stability	**Stochastic variability**, or "wobble" associated with random sampling, calculated as the average percent change in the model output after rerunning **Monte Carlo simulations** with the same set of input assumptions. Used as a metric for evaluating the adequacy of the number of iterations in a **MCA**.
Standard Deviation, Arithmetic and Geometric	Standard deviation (or arithmetic standard deviation, SD) is a common measure of the spread of a distribution. Calculated as the square root of the **variance.** The geometric standard deviation (GSD) is the anti-log of the standard deviation of the logarithms of each value. The GSD is a unitless quantity that gives a measure of the ratio of the variance to the mean, similar in concept to the **coefficient of variation**.
Step Function	A mathematical function that remains constant within each of a series of adjacent intervals but changes in value from one interval to the next. **Cumulative distribution functions** for **discrete random variables** are step functions.
Stochastic Dominance	Implies no intersection between the **CDFs**; distribution A stochastically dominates distribution B if, for every **percentile** of the **CDF**, A > B. This characteristic may not be apparent from the **PDFs** of the distributions, which may overlap.
Stochastic Process	A process involving **random variables**, and characterized by **variability** in space or time.
Target Population	The set of all receptors that are potentially at risk. Sometimes referred to as the "population of concern". A sample population is selected for statistical sampling in order to make inferences regarding the target population (see Appendix B, Section B.3.1, Concepts of Populations and Sampling).

Definitions of Terms Used in PRA	
Technical Assistance Grant (TAG)	A federal grant that is intended to provide a community with the opportunity to hire independent experts to help evaluate and explain the results of a **risk assessment**
Technical Outreach Services for Communities (TOSC)	A service of the **HSRC** with the aim to provide independent technical information and assistance to help communities with hazardous substance pollution problems.
Thiessen (Voronoi) Polygon Analysis	A method of spatial statistics in which an area is subdivided into subregions, or polygons, in order to predict values at unobserved locations.
Time Step	A **variable** in all exposure models that refers to the unit of time for which a **random value** is considered representative of intra-individual **variability** (e.g., average daily ingestion rates for an individual from one year to the next). A time step may be equal to an entire exposure duration (e.g., 30 years), or a fraction of the exposure duration during which changes in input **variables** may be expected (e.g., one year). Time steps need not be identical for all exposure **variables**, and should address the most rapidly changing **variable** in the risk equation. Time step can be an important consideration for **MEE analysis**.
Toxicity Reference Value (TRV)	A numerical expression of a chemical's dose-response relationship that is used in **ecological risk assessment**.
True Mean Concentration	The actual average concentration in an **exposure unit**. Even with extensive sampling, the true **mean** cannot be known. Only an estimate of the true **mean** is possible. A greater number of representative samples increases confidence that the estimate of the **mean** more closely represents the true **mean**.
Truncation	The process of setting lower and upper limits on the range of a distribution, in order to avoid unrealistic values for exposure **variables** (e.g., > 100% bioavailability). Most often used for continuous, unbounded **probability distributions** (e.g., normal).
Two-dimensional Monte Carlo Analysis (2-D MCA)	An advanced modeling technique that uses two stages of random sampling, also called nested loops, to distinguish between **variability** and **uncertainty** in exposure and toxicity **variables**. The first stage, often called the inner loop, involves a complete **1-D MCA** simulation of **variability** in risk. In the second stage, often called the outer loop, **parameters** of the **probability distributions** are redefined to reflect **uncertainty**. These loops are repeated many times resulting in multiple risk distributions, from which **confidence intervals** are calculated to represent **uncertainty** in the population distribution of risk.
Type I Errors	False positive; the error made when the null **hypothesis** is rejected in favor of the alternative, when in fact the null **hypothesis** is true.
Type II Errors	False negative; the error made when the null **hypothesis** is accepted when in fact the alternative **hypothesis** is true.
Uncertainty	Lack of knowledge about specific **variables**, **parameters**, models, or other factors. Examples include limited data regarding the concentration of a contaminant in an environmental medium and lack of information on local fish consumption practices. Uncertainty may be reduced through further study.

Definitions of Terms Used in PRA	
Variability	True heterogeneity or diversity in characteristics among members of a population (i.e., inter-individual variability) or for one individual over time (intra-individual variability). For example, body weights of a study population at one point in time will exhibit variability, and body weight will change as an individual ages. Further study (e.g., increasing sample size, *n*) will not reduce variability, but it can provide greater confidence in quantitative characterizations of variability.
Variable	A quantity that can assume many values.
Variance	Measure of the spread of a distribution, equal to the square of the **standard deviation** (SD). Calculated as the average of the squares of the deviations of the observations from their **mean**. Variance is referred to as the second central **moment of a distribution**.
Z-score	The value of a normally distributed **random variable** that has been standardized to have a **mean** of zero and a SD of one by the transformation $Z=(X-\mu)/\sigma$. Statistical tables typically give the area to the left of the z-score value. For example, the area to the left of z =1.645 is 0.95. Z-scores indicate the direction (+/-) and number of **standard deviations** away from the **mean** that a particular datum lies assuming X is normally distributed. Microsoft Excel's *NORMSDIST(z)* function gives the probability p such that $p=Pr(Z \leq z)$, while the *NORMSINV(p)* function gives the z-score z_p associated with probability p such that $p=Pr(Z \leq z_p)$.

E.1.0 ADDITIONAL INFORMATION

Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis (Morgan and Henrion, 1990) and *Probabilistic Techniques in Exposure Assessment* (Cullen and Frey, 1999) provide excellent philosophical and practical treatises on probabilistic risk assessment. These works are highly recommended to risk assessors who wish to know more about probabilistic risk assessment. The *Summary Report for the Workshop on Monte Carlo Analysis* (U.S. EPA, 1996) and the *Summary Report for the Workshop on Selecting Input Distributions for Probabilistic Assessments* (U.S. EPA, 1999b) are other sources of information to learn more about PRA. Other additional references for reading are listed in this Appendix.

REFERENCES FOR APPENDIX E

Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment.* A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press.

Law, A.M. and W.D. Kelton. 1991. Simulation Modeling and Analysis. McGraw-Hill, Inc., NY.

Morgan, G.M. and M. Henrion, 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.

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

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1991. *Risk Assessment Guidance for Superfund* (RAGS)*, Volume I: Human Health Evaluation Manual* (HHEM)*, Part B, Development of Risk-Based Preliminary Remediation Goals.* Office of Emergency and Remedial Response, Washington, DC. EPA/540/R-92/003. NTIS PB92-963333.

U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7-081.

U.S. EPA. 1997a. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May 15.

U.S. EPA. 1997b. *Lognormal Distribution in Environmental Applications*. Office of Research and Development, and Office of Solid Waste and Emergency Response, Washington, DC. EPA/600/R-97/006. December.

U.S. EPA. 1997c. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments.* Interim Final. U.S. Environmental Protection Agency, Environmental Response Team (Edison, NJ). June 5.

U.S. EPA. 1999a. Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. OSWER Directive 9285.7-28 P. Stephen D. Luftig for Larry D. Reed. October 7.

U.S. EPA. 1999b. *Summary Report for the Workshop on Selecting Input Distributions for Probabilistic Risk Assessment*. Risk Assessment Forum. EPA/630/R-98/004.

REFERENCES FOR FURTHER READING

Baird, BF. 1989. *Managerial Decisions Under Uncertainty*. John Wiley & Sons, Inc., NY.

Bevington, P.R. 1969. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, NY.

Bratley, P., B.L. Fox, and L.E. Schrage. 1987. *A Guide to Simulation*. Springer-Verlag, NY.

Burmaster, D.E. and P.D. Anderson. 1994. Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessment. *Risk Anal.* 14(4):477–481.

Clemen, R, 1990. *Making Hard Decisions*. Duxbury Press.

Conover, W.I. 1971. *Practical Nonparametric Statistics*. John Wiley & Sons, NY.

Cox, D.C. and P. Baybutt. 1981. *Methods for Uncertainty Analysis: A Comparative Survey*. *Risk Anal* 1(4):251–258.

Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment*. Plenum Press, NY.

D'Agostino, R. and M.A. Stephens, eds. 1986. *Goodness-of-Fit Techniques*. Marcel Dekker, Inc., NY.

Devroye, L. 1986. *Non-Uniform Random Deviate Generation*. Springer-Verlag, NY.

Evans, M., N. Hastings, and B. Peacock. 1993. *Statistical Distributions*. John Wiley & Sons, NY.

Finkel, A.M. 1990. *Confronting Uncertainty in Risk Management: A Guide for Decision-Makers*. Resources for the Future, Washington, DC.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, NY.

Hamby, D.M. 1994. A review of Techniques for Parameter Sensitivity Analysis of Environmental Models. *Environ. Monit. and Assess*. 32:135–154.

Hammersley, J.M. and D.C. Handscomb. 1964. *Monte Carlo Methods*. John Wiley & Sons, NY.

Hertz, D.B. and H. Thomas. 1983. *Risk Analysis and Its Applications*. John Wiley & Sons, NY.

Hertz, D.B. and H. Thomas. 1984. *Practical Risk Analysis - An Approach Through Case Studies*. John Wiley & Sons, NY.

Hoffman, F.O. and J.S. Hammonds. 1992. *An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment*. ES/ER/TM-35, Martin Marietta.

Hoffman, F.O. and J.S. Hammonds. 1994. *Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability*. *Risk Anal* 14(5):707–712.

Iman, R.L. and W.J. Conover. 1982. A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables. *Commun. Stat*, Part B 11:311–331.

Iman, R.L. and J.C. Helton. 1988. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Anal*. 8(1):71–90.

Iman, R.L., J.M. Davenport, and D.K. Zeigler. 1980. *Latin Hypercube Sampling (A Program Users Guide)*. Technical Report SAND 79:1473, Sandia Laboratories, Albuquerque.

Johnson, M.E. 1987. *Multivariate Statistical Simulation*. John Wiley & Sons, NY.

Johnson, N.L. and S. Kotz. 1970. *Continuous Univariate Distributions*. Vols. 1 & 2. John Wiley & Sons, NY.

Johnson, N.L., S. Kotz, and A.W. Kemp. 1992. *Univariate Discrete Distributions*. John Wiley & Sons, NY.

Kendall, M. and A. Stuart. 1979. *Advanced Theory of Statistics, Volume I - Distribution Theory, Volume II - Inference and Relationship*. MacMillan, Inc., NY.

Kennedy, W.J. and E. Gentle. 1980. *Statistical Computing*. Marcel Dekker, Inc., NY.

LePage, R. and L. Billard. 1992. *Exploring the Limits of Bootstrap*. John Wiley & Sons, NY.

Lipton, J., W.D. Shaw, J. Holmes, and A. Patterson. 1995. Short Communication: Selecting Input Distributions for use in Monte Carlo Analysis. *Regul. Toxicol. Pharmacol.* 21:192–198.

McKone, T.E. and K.T. Bogen. 1992. Uncertainties in Health Risk Assessment: An Integrated Case Based on Tetrachloroethylene in California Groundwater. *Regul. Toxicol. Pharmacol.* 15:86–103.

Megill, R.E., ed. 1985. *Evaluating and Managing Risk*. Penn Well Books, Tulsa, OK.

Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.

NCRP. 1996. Commentary No. 14. *A Guide for Uncertainty Analysis in Dose and Risk Assessments Related to Environmental Contamination*. National Committee on Radiation Programs, Scientific Committee 64-17. Washington, DC.

Palisade Corporation. 1994. *Risk Analysis and Simulation Add-In for Microsoft Excel or Lotus 1-2-3*. Windows Version Release 3.0 User's Guide, Palisade Corporation, Newfield, NY.

Press, W.H., B.P. Flannery, S.A. Teulolsky, and W.T. Vetterling. 1989. *Numerical Recipes in Pascal: the Art of Scientific Computing*. Cambridge University Press, NY.

Press, W.H., S.A. Teulolsky, W.T. Vetterling, and B.P. Flannery. 1992. *Numerical Recipes in FORTRAN: the Art of Scientific Computing*. Cambridge University Press, NY.

Press, W.H., S.A. Teulolsky, W.T. Vetterling, and B.P. Flannery. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, NY.

Read, T. and N. Cressie. 1988. *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer-Verlag, NY.

Rohatgi, V.K. 1984. *Statistical Inference*. John Wiley & Sons, NY.

Rubenstein, R.Y. 1981. *Simulation and the Monte Carlo Method*. John Wiley & Sons, NY.

Sachs, L. 1984. *Applied Statistics - A Handbook of Techniques*. Springer-Verlag, NY.

Saltelli, A and J. Marivort. 1990. Non-Parametric Statistics in Sensitivity Analysis for Model Output: A Comparison of Selected Techniques. *Reliab. Eng. Syst. Saf*. 28:299–253.

Schneider, H. 1986. *Truncated and Censored Distributions from Normal Populations*. Marcel Dekker, Inc., NY.

Seiler, F.A. 1987. Error Propagation for Large Errors. *Risk Anal*. 7(4):509–518.

Seiler, F.A. and J.L. Alvarez. 1996. On the Selection of Distributions for Stochastic Variables. *Risk Anal.* 16(1):5–18.

Slob, W. 1994. Uncertainty Analysis in Multiplicative Models. *Risk Anal*. 14(4):571–576.

Smith, A.E., P.B. Ryan, and J.S. Evans. 1992. The Effect of Neglecting Correlations when Propagating Uncertainty and Estimating the Population Distribution of Risk. *Risk Anal*. 12(4):467474.

Smith, R.L. 1994. Uses of Monte Carlo Simulation for Human Exposure Assessment at a Superfund Site. *Risk Anal* 14(4):433–439.

Sokal, R. and R. Rohlf. 1981. *Biometry: The Principles and Practice of Statistics in Biological Research*. Second Edition. W.H. Freeman & Co., NY.

U.S. EPA. 1978. *Source Assessment: Analysis of Uncertainty - Principles and Applications*. EPA/600/2-79-004.

U.S. EPA. 1992a. *Guidelines for Exposure Assessment*. *Federal Register*. 57(104):22888-22938. May 29.

U.S. EPA. 1992b. *Guidelines for Carcinogenic Risk Assessment*. *Federal Register*. 51(185):33992-34003. May 29.

U.S. EPA. 1996. *Summary Report for the Workshop on Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-96/010.

U.S. EPA. 1999. *Guidelines for Carcinogenic Risk Assessment*. Review Draft. Risk Assessment Forum. Washington, DC. NCEA-F-0644.

U.S. EPA. 2001. *Guidelines for Carcinogenic Risk Assessment*. *Federal Register*. 66(230):59593-59594. November 29.

Wilks, D.S. 1995. *Statistical Methods in the Atmospheric Sciences, An Introduction*. Academic Press, San Diego.

APPENDIX F

WORKPLAN AND CHECKLIST FOR PRA

F.0 INTRODUCTION

This appendix provides guidance on developing a workplan prior to the initiation of a probabilistic risk assessment (PRA), and using a checklist when reviewing a PRA. Like the quality assurance project plan (QAPP), the workplan for PRA generally should document the combined decisions or positions of the remedial project manager (RPM), risk assessor, and stakeholders involved in the risk assessment. Often there are many stakeholders in a risk assessment, and it is important to involve and engage all stakeholders early in the decision-making process. These are important steps that should save time and effort.

EXHIBIT F-1

EXAMPLES OF ELEMENTS OF THE WORKPLAN FOR PRA

1. Statement of the ecological assessment endpoints and/or human risk
2. Summary of the point estimate risk assessment
3. Potential value added for risk management by conducting a PRA and proceeding to the subsequent tiers (quantify variability, uncertainty, or both)
4. Discussion of adequacy of environmental sampling for PRA (e.g., data quality issues)
5. Description of the methods and models to be used (e.g., model and parameter selection criteria)
6. Proposal and basis for probability distributions and point estimates
7. Methods for deriving the concentration term
8. Proposal for probabilistic sensitivity analysis
9. Method for dealing with correlations
10. Bibliography of relevant literature
11. Software (i.e., date and version of product, random number generator)
12. Simulation approach (e.g., iterations, Monte Carlo or Latin Hypercube sampling, time step)
13. Proposed schedule and expertise needed

F.1.0 WORKPLAN

In general, PRAs may be developed by Environmental Protection Agency (EPA), EPA contractors, or a potentially responsible party (PRP) with appropriate EPA oversight. In each case, it is important to develop a workplan early in the risk assessment process. PRAs to be submitted by a contractor or PRP should generally be submitted for EPA review before commencing the analysis. The workplan should describe the software to be used, the exposure routes and models, and input probability distributions and their basis (e.g., relevance to the site-specific contamination and pathways), including appropriate literature references. Examples of the elements of a workplan are given in Exhibit F-1, as well as Exhibit 4-8 in Chapter 4 (Example Elements of a Workplan for Ecological PRA). It is important that the risk assessor and risk manager discuss the scope of the probabilistic analysis and the potential impact on the Remedial Investigation/Feasibility Study (RI/FS).

☞ *Given the time and effort that can be expected to be invested in conducting a PRA, it is important that a workplan undergo review and approval by EPA, prior to proceeding with the assessment.*

The EPA generally will not accept probabilistic analysis where a workplan for the analysis has not been initially submitted to the Agency and approved by the Regional risk assessor and RPM. Exceptions to this process may be considered on a case-by-case basis.

Conducting a PRA is an iterative process. In general, as new information becomes available, it should be used to evaluate the need to move to a higher tier. The decision to move an assessment to a higher tier of complexity should result in a revised workplan and consultation with the Agency. The previous PRA, and its sensitivity analysis, should be included in the revised workplan, along with a point estimate risk assessment based on any data collected as part of a lower tier. The assessment will often be restricted to the chemicals and pathways of concern that contribute the greatest risk.

Throughout the process of developing the PRA, the EPA risk assessor and the personnel involved in developing the assessment should have a continuing dialogue to discuss the many Agency decisions and their potential impact on the assessment. This dialogue, along with interim deliverables, will help to ensure that the risk assessment report will meet the needs of the Agency and that any problems are identified and corrected early in the process.

F.2.0 FOCAL POINTS FOR PRA REVIEW

In reviewing a PRA, it is recommended that a systematic approach be adopted to ensure that all key technical elements of the PRA are evaluated and potential weaknesses are identified. A review check list can facilitate this process and promote consistency in the reviews of PRAs. Such a list can be developed from EPA's guiding principles (U.S. EPA, 1997) and other reviews on the subject of PRA quality review (e.g., Burmaster and Anderson, 1994).

EXHIBIT F-2

KEY FOCAL POINTS FOR PRA REVIEW

1. Clarity of and conformation to objectives.
2. Scientific basis and documentation of input distributions and assumptions.
3. Model structure and computational mechanics.
4. Results, including, limitations, reasonableness, and clarity of documentation.

In general, the review of a PRA can be organized into four focal points listed in Exhibit F-2. PRAs can vary in complexity, from relatively simple to very complicated; thus, the review strategy may need to be customized for specific sites.

F.3.0 CHECKLIST FOR REVIEWERS

The exposure pathways and chemicals considered in a PRA should be clearly stated and related to the assessment endpoint. Often, the simplest way of doing this is to use the site conceptual model.

Table F-1 provides a list of major points that may be used to evaluate the quality of a probabilistic assessment. This is not an exhaustive list. The ultimate judgment of the acceptability of a PRA is the responsibility of the regional EPA personnel.

The issues that a reviewer should focus on may be different for each assessment. The workplan and the assessment should address each of the items on the checklist, but the workplan may include

additional items. The reviewer is responsible for ensuring that the workplan and the assessment are complete and of sufficient quality to help support a risk management decision under the National Contingency Plan (NCP).

The report should include a discussion of the results of assessment and how they relate to the point estimate of risk and hazard. A clear and concise description of what the results mean is an important part of each report.

F.4.0 INTERNAL AND EXTERNAL REVIEW

There are two levels of review that may be appropriate for a PRA. If an EPA reviewer feels the need for help with a review, other EPA personnel may be contacted formally or informally to provide additional review capabilities. The EPA personnel should also review the draft workplan for PRA to evaluate the appropriateness and consistency with Agency guidance. If EPA personnel are contacted early in the risk assessment process, the review can occur in a more productive and timely manner.

When the issues at a particular site are complex or contentious, EPA reviewers may also wish to obtain the services of outside experts for peer review (U.S. EPA, 2000). According to EPA's Peer-Review Policy Statement dated June 7, 1994 (U.S. EPA, 1994), "Major scientifically and technically based work products related to Agency decisions normally should be peer-reviewed." External peer review should be considered when allocating resources for a PRA. The EPA reviewers generally should select external peer reviewers who possess no bias or agenda regarding the process or methods of PRA.

Table F-1. Example of a Generic Checklist for Reviewers [2 pages]

Focal Point	✔	Evaluation Criterion
Objectives and Purpose		
Assessment Endpoints	✔	Are the human health and/or ecological assessment endpoints clearly stated and consistent with the workplan?
Benefits	✔	Are the rationales for, and benefits of, performing the PRA clearly stated and consistent with the workplan?
Site Conceptual Model	✔	Is there a description or graphic representation of the receptors and pathways considered in the assessment? Has the PRA addressed each of the pathways for completeness (e.g., sources, release mechanisms, transport media, route of entry, receptor)?
Separation of Variability and Uncertainty	✔	What is the modeling strategy for separating variability and uncertainty in the PRA? Is this strategy consistent with the assessment endpoints?
Model Structure and Computational Mechanics		
Flow Chart	✔	Is a diagram of the computational sequence provided so that the pathways of inputs and outputs and data capture can be understood and easily communicated?
1-D MCA / 2-D MCA	✔	Is a 1-D MCA or 2-D MCA being implemented in the PRA? What is represented by either or both dimensions?
Algorithms	✔	Are all algorithms used in the model documented in adequate detail to recreate the analysis?
Integration	✔	Are the algorithms used in numerical integration identified and documented?
Dimensional Analysis	✔	Has a unit analysis been conducted to ensure that all equations balance dimensionally?
Random Number Generation	✔	What random number generator is used in model computations? Is it robust enough? What reseeding approach is used to minimize repeated sequences?
Input Distributions and Assumptions		
Variability and Uncertainty	✔	Is there a clear distinction and segregation of distributions intended to represent variability from distributions intended to represent uncertainty?
Data sources	✔	Are the data or analysis sources used in developing or selecting the input distributions documented and appropriate for the site?
Distribution Forms	✔	Are the analyses used in selecting the form of the distribution adequately documented (i.e., understandable and repeatable by a third party?)
Distribution Parameters	✔	Are the analyses used to estimate the distribution parameters adequately documented?
Distribution Tails	✔	Do the estimation methods precisely depict the tails of the input distributions; how was this evaluated? Is there sufficient information to depict tails for empirical distributions? Are these estimated as exponential tails with bounding values?
Truncations	✔	Are any input distributions truncated? Do these truncations make sense? Should truncations be applied to any of the distributions?
Concentration Term	✔	Is the derivation of a point estimate or distribution for the concentration term adequately documented? Is sufficient information provided to enable the reviewer to recreate the concentration term?
Variable Correlations	✔	Have variable independence and correlations been addressed? Has the methodology for representing variable correlations in the model been documented and is it reasonable in terms of the variables, the site, and the statistical approach?

Focal Point	✔	Evaluation Criterion
Time Step	✔	Has the basis for the time step used in the model been documented? Is a single time step used, or do variables have different time steps? Are the time steps conceptually reasonable for the variables; for the site? Has the time step been evaluated in the sensitivity analysis?
Sensitivity Analysis	✔	Has a sensitivity analysis been conducted? Are the methods used in the analysis statistically valid? What did the analysis reveal about uncertainties in the assessment and the relative contributions of input variables to uncertainty?
Results of Modeling		
Completeness	✔	Are all the exposure routes identified in the site conceptual model and workplan addressed in the model results? Has the PRA fulfilled the objectives and satisfied the purpose stated in the workplan?
Point Estimate Calculation	✔	Has a point estimate calculation, using mean or median values of the input distributions, been performed? How do these results compare with the central tendencies calculated with the probabilistic model? How do the reasonable maximum exposure (RME) estimates compare? Have the similarities or differences between risk estimates from the point estimate and probabilistic approaches been adequately addressed?
Stability of Output Tails	✔	Has the stability of the high-end tail of the risk distribution been adequately evaluated? How stable are the estimated tails (in quantitative terms?) Is this level of stability adequate to support the risk management decisions that the model is intended to support?
Significant Figures	✔	Is the number of significant figures used in the output reasonable and consistent with model uncertainty?
Limitations	✔	Are the strengths and weaknesses of the PRA methodology and limitations of the results for decision making clearly presented?
Clarity	✔	Are the results and conclusions clearly presented and consistent with model output (e.g., central tendency exposure (CTE) and RME identified in the Executive Summary along with discussion of uncertainty)?
Graphics	✔	Are there graphics included that show both the risk distribution and PRA results (e.g., CTE and RME risk)?

REFERENCES FOR APPENDIX F

Burmaster, D.E. and P.D. Anderson. 1994. Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessment. *Risk Anal.* 14(4):477–481.

U.S. EPA. 1994. Memorandum from Deputy Administrator Carol Browner on *Peer Review and Peer Involvement at the U.S. Environmental Protection Agency*. June 7.

U.S. EPA. 1997. Memorandum from Deputy Administrator Fred Hansen on the *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment*, and *Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R–97/001. May 15.

U.S. EPA. 2000. *Peer Review Handbook: 2nd Edition.* Science Policy Council. Washington, DC. EPA/100/B-00/001. December.

APPENDIX G

FREQUENTLY ASKED QUESTIONS FOR PRA

INTRODUCTION

This section presents a few questions and answers relating to probabilistic risk assessment (PRA). The purpose of the frequently asked questions (FAQs) is to facilitate the understanding of PRA using a comparison with the traditional point estimate approach to risk assessment.

The FAQs presented here provide an overview of PRA with pointers to more detailed, and often more technical, discussions in other parts of the guidance.

(1) What is a risk assessment?

Risk assessment is a tool for organizing available information to make inferences about the potential human health or ecological effects associated with exposure to hazardous materials. The National Contingency Plan (NCP) addresses the use of a baseline risk assessment at Superfund sites to determine whether risks to human health and the environment are unacceptable. The NCP implements the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) of 1980.

Risk assessments traditionally provide single point descriptors of risk (e.g., a central tendency exposure (CTE) risk descriptor or a reasonable maximum exposure (RME) risk descriptor). As such, these types of risk assessments have been referred to as point estimate risk assessments.

In 1983, the National Research Council (NRC) described the following four steps for conducting human health risk assessments:

- ***Hazard identification***: the determination of whether a particular chemical is or is not causally linked to a particular health effect.
- ***Dose-response assessment***: the determination of the relation between the magnitude of exposure and the probability of occurrence of the health effects in question.
- ***Exposure assessment***: the determination of the extent of human exposure before or after application of regulatory controls.
- ***Risk characterization***: a description of the nature and often the magnitude of human risk, including attendant uncertainty (NRC, 1983).

Readers are referred to risk assessment guidance documents such as *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment)*(U.S. EPA, 1989a), *Risk Assessment Guidance for Superfund: Volume II. Environmental Evaluation Manual* (U.S. EPA, 1989b), and *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments* (U.S. EPA, 1997a) for more information about point estimate risk assessment methods and policies.

(2) What is a probabilistic risk assessment (PRA)?

Superfund risk assessments have traditionally provided single point estimates of risk. More recently, PRAs have been developed. A PRA is a risk assessment that provides a probability distribution, rather than a point estimate, of risk. A probability distribution conveys both a range of values and a likelihood of occurrence of each value. This may allow a risk assessor to make statements about the likelihood that risks will exceed a level of concern. The probability distribution for risk often represents variability in risk estimates for a potentially exposed population. This variability may be due to variability in exposure and/or toxicity. PRA may also be used to quantify uncertainty in risk estimates. This can be useful because it allows a risk assessor to make statements about the level of confidence in the likelihood that risks will exceed a level of concern.

Probabilistic methods often use computer simulations to combine multiple probabilistic distributions in a risk equation. Monte Carlo analysis (MCA) is perhaps the most widely used probabilistic method in PRA (see Question #7).

(3) How does PRA compare with the point estimate approach?

A single point estimate of risk does not explicitly characterize associated variability or uncertainty. However, multiple point estimates of risk (e.g., CTE or RME) can begin to characterize variability in risk as they use different points on each input distribution for exposure). A PRA can characterize variability in risk by using the full distribution of variability in exposure parameters in the risk equations. Advanced PRA techniques can also quantitatively characterize uncertainty. In appropriate circumstances, results of a PRA can lead to more informed risk management decisions.

A PRA can be more resource intensive than a point estimate risk assessment. Some PRAs can require greater effort than point estimate approaches to define model inputs (i.e., select and fit probability distributions), as well as additional steps in the planning, review, and communication of the risk assessment assumptions and results (see Chapter 6 and Appendix F). A PRA does not necessarily require more data than a point estimate approach, although it does provide a framework for incorporating more of the available information into the risk assessment. When information on important exposure variables is lacking, results from a point estimate approach and a probabilistic approach will be equally uncertain.

If a decision is made to conduct a PRA, this does not replace a point estimate risk assessment. Results of point estimate approaches should still be presented along with results of probabilistic approaches in Tier 2 or Tier 3.

(4) Why should I consider using PRA?

PRA can have several advantages over the traditional point estimate approach to risk assessment. PRA can often provide a more complete characterization of risk; a quantitative description of the uncertainties in the risk estimates; more informative sensitivity analysis; the ability to make probabilistic statements about risk; the ability to know where specific risk levels are on the potential distribution of risk; an increased understanding of risks; and opportunities for improved communication and risk management decision making.

(5) When should I consider using PRA?

A PRA may be considered as early as the planning stages of a point estimate risk assessment or as late as after the completion of a point estimate risk assessment. Ideally, PRA should be considered as early as possible in the planning of risk assessment activities at a site so that sampling plans and data collection efforts may be appropriately directed. A PRA may be used when the risk management decision is not apparent and when the results of a PRA may inform the risk management decision. Often a risk management decision is not apparent when the site-specific risk estimate is close to the regulatory level of concern. The NCP discusses a generally acceptable range for cumulative excess cancer risk of 1E-06 to 1E-04 for protecting human health (U.S. EPA, 1990). Noncancer risks to human health and ecological health are generally characterized by a ratio of exposure to toxicity, called a Hazard Quotient (HQ) or Hazard Index (HI) for multiple contaminants. The point of departure for evaluating noncancer risks may vary from site to site, but a HQ of 1 may be a good starting point for risk management decisions.

PRA may also be considered when the results of the point estimate risk assessment suggest that risks are clearly above a risk level of concern, and a preliminary remediation goal (PRG) is needed. Because PRA and point estimate risk assessments use different techniques for quantifying variability and uncertainty, they may support different PRGs. If the results are dramatically different, further steps may be warranted to reevaluate the choices for input variables - both the point estimates, and the probability distributions and parameters (including truncation limits) for the 1-D MCA.

PRA will not be needed in many cases. Point risk estimates often produce results which are sufficient for making remedial decisions (e.g., sites are usually either heavily contaminated or only marginally contaminated). A tiered approach to risk assessment has been developed by Environmental Protection Agency (EPA) and is recommended for use in deciding when to move from point estimate risk assessments to PRAs of varying complexities. A workplan should be developed and submitted for review before beginning a PRA at any stage in the tiered process. As a general rule, if the potential value added by a PRA outweighs the additional resource required to conduct it, PRA may be warranted (see Chapter 2).

(6) How is the risk distribution from PRA used for decision making?

The EPA's *RAGS Volume I* (U.S. EPA, 1989a) and the NCP Preamble (U.S. EPA, 1990) state that the RME will generally be the principal basis for evaluating potential human health risks at Superfund sites. Ecological assessments also often consider an RME endpoint. The point estimate Superfund risk assessments use a combination of average and high-end input values to arrive at the RME. In PRA, risks are described by a probability distribution instead of a point estimate. To use a risk distribution for decision making, one needs to identify a percentile value that corresponds to the RME. *EPA's Guidelines for Exposure Assessment* (U.S. EPA, 1992a) states that, "the high-end risk means risks above the 90^{th} percentile of the population distribution", and "the high-end estimator should not exceed the 99.9^{th} percentile" due to uncertainty in specifying the upper tail of the input distributions in a Monte Carlo analysis. Similarly, the 90^{th} to 99.9^{th} percentiles of the risk distribution are recommended in this guidance as the RME range for decision making in PRA. Selection of a single point within the RME range generally requires consideration of the level of uncertainty in the risk distribution. The EPA recommends that the 95^{th} percentile of the risk distribution be used as a starting point for risk management decisions in the absence of site-specific information.

(7) What is Monte Carlo Analysis (MCA)?

MCA is a numerical technique for PRA. MCA was developed in the 1940's during the beginnings of the nuclear power industry. MCA combines statistical analysis with modern computational techniques to calculate risk estimates, by randomly choosing different sets of input values each time. Each calculation is an iteration and a set of iterations is called a simulation. The output of a simulation used for risk assessment is a continuous probability distribution, which can be displayed in a graph in the form of either a probability density function (PDF) or corresponding cumulative distribution function (CDF). Both displays represent the same distribution, but are useful for conveying different information. For example, the PDF for risk is a good way for displaying relative probability using an area under the bell-shaped curve. The CDF for risk is generally S-shaped and can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., 95th percentile=1E-06). Other uses of PDFs and CDFs are presented in Chapter 1, Exhibit 1-3. In 1997, EPA published a policy accepting the use of MCA to perform human health and ecological risk assessments (U.S. EPA, 1997a). This guidance focuses on MCA as a method of quantifying variability and uncertainty.

(8) What is the policy on using PRA to characterize variability or uncertainty in toxicity or dose response?

In human health risk assessments, probability distributions for risk should reflect variability or uncertainty in exposure. In ecological risk assessments, risk distributions may reflect variability or uncertainty in exposure and/or toxicity (see Chapter 1, Sections 1.4 and 1.4.1, Item 3).

Approaches to characterizing variability and uncertainty in toxicological information should reflect both the latest developments in the science of hazard and dose-response evaluation and consistent application of EPA science policy. This statement is consistent with the *1997 EPA Policy Statement* presented in Section 1.4 above (U.S. EPA, 1997g). Probabilistic approaches to ecological dose-response assessment may be explored, as discussed and demonstrated in Chapter 4. This guidance does not develop or evaluate probabilistic approaches for dose-response in human health assessment and, further, *discourages undertaking such activities on a site-by-site basis*. Such activities require contaminant-specific national consensus development and national policy development. Parties wishing to undertake such activities should contact the OERR to explore ways in which they might contribute to a national process for the contaminant of interest to them.

(9) What is the policy on using PRA at EPA and in Superfund?

In the spring of 1997, EPA released the memorandum, *Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment* (U.S. EPA, 1997b). The policy states that probabilistic analysis techniques, "given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments." As such, a PRA, "will be evaluated and utilized in a manner that is consistent with other risk assessments submitted to the Agency." Together with this Policy Statement, the Agency released a set of guiding principles for use and review of probabilistic analyses. Hence, both RAGS and Agency-wide guidance emphasize the importance of review of the scientific and technical merit of a probabilistic analysis to determine whether or not the assessment is of sufficient quality to support a remedial decision. This guidance, *RAGS Volume 3: Part A*, provides risk assessors with comprehensive guidance on when and how to conduct PRAs using MCA within the Superfund program (see Preface and Chapter 1).

(10) What are the challenges of using PRA?

Although PRA may have several advantages over the traditional point estimate approach to risk assessment, the use of PRA tends to be more resource intensive and may introduce some additional challenges to risk communication efforts. Risk communication helps build trust with the stakeholders and disseminate the risk information. In general, EPA staff and stakeholders are accustomed to a point estimate of risk and are unfamiliar with PRA and the quantitative estimates of uncertainty that PRA can support. Although, quantitative risk estimates may be more informative, they also may be more difficult to communicate and may not be well received due to stakeholder desires for certainty (Slovic, et al. 1979). Early and frequent communication with stakeholders is key in implementing PRA successfully. Often PRA requires additional data collection efforts as well as more time and resources to select and fit probability distributions.

REFERENCES FOR APPENDIX G

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

Slovic, P., B. Fischoff, and S. Lichtenstein. 1979. Rating the Risks. *Environ.* 21(3):14–20 and 36–39.

U.S. EPA. 1989a. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1989b. *Risk Assessment Guidance for Superfund.* (RAGS)*: Volume II. Environmental Evaluation Manual*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/001.

U.S. EPA. 1990. National Oil and Hazardous Substances Pollution Contingency Plan. Final Rule. 40 CFR 300: *55 Federal Register*, 8666-8865, March 8.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. *57 Federal Register*, 22888-22938, May 29.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments*. Interim Final. Environmental Response Team, Edison, NJ. EPA/540/R-97/006, OSWER Directive No. 9285.7-25, June.

U.S. EPA. 1997b. *Memorandum from Deputy Administrator Fred Hansen on the Use of ProbabilisticTechniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

APPENDIX H

INDEX