Health Canada Santé Canada

Your health and safety… our priority. *Votre santé et votre sécurité… notre priorité.*

Part V:

Guidance on Human Health Detailed Quantitative Risk Assessment for Chemicals ($DQRA_{Chem}$)

Federal Contaminated Site Risk Assessment in Canada

Canada

Health Canada is the federal department responsible for helping the people of Canada maintain and improve their health. *We assess the safety of drugs and many consumer products, help improve the safety of food, and provide information to Canadians to help them make healthy decisions. We provide health services to First Nations people and to Inuit communities. We work with the provinces to ensure our health care system serves the needs of Canadians.*

Published by authority of the Minister of Health.

Federal Contaminated Site Risk Assessment in Canada, Part V: Guidance on Human Health Detailed Quantitative Risk Assessment for Chemicals ($DQRA_{Chem}$).
is available on Internet at the following address:

www.healthcanada.gc.ca

Également disponible en français sous le titre :
L'évaluation des risques pour les sites contaminés fédéraux au Canada, Partie V : L'évaluation quantitative détaillée des risques pour la santé humaine associés aux substances chimiques ($ÉQDR_{chim}$)

This publication can be made available on request in a variety of alternative formats.

For further information or to obtain additional copies, please contact:

Publications
Health Canada
Ottawa , Ontario K1A 0K9
Tel.: 613-954-5995
Fax: 613-941-5366
E-mail: info@hc-sc.gc.ca

Cat.: H128-1/11-639E-PDF
ISBN: 978-1-100-17926-1

FEDERAL CONTAMINATED SITE RISK ASSESSMENT IN CANADA

PART V: GUIDANCE ON HUMAN HEALTH DETAILED QUANTITATIVE RISK ASSESSMENT FOR CHEMICALS ($DQRA_{CHEM}$)

September 2010

Prepared by:
Contaminated Sites Division
Safe Environments Directorate

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EXAMPLES

PREFACE

The Federal Contaminated Sites Action Plan (FCSAP) is a program of the Government of Canada designed to ensure improved and continuing federal environmental stewardship as it relates to contaminated sites located on federally owned or operated properties. Guidance documents on human health risk assessment (HHRA), prepared by the Contaminated Sites Division of Health Canada in support of the FCSAP, are available on our website, and may also be obtained by contacting the Contaminated Sites Division at cs-sc@hc-sc.gc.ca.

This guidance document (*Federal Contaminated Site Risk Assessment in Canada, Part V: Guidance on Human Health Detailed Quantitative Risk Assessment for Chemicals (DQRA$_{Chem}$)*) was prepared to provide guidance for custodial departments. It addresses complex contaminated sites and those sites for which a preliminary quantitative risk assessment (PQRA) is considered too conservative or not adequate to support a risk management plan. As is common with any national guidance, this document will not satisfy all of the requirements presented by contaminated sites, custodial departments, or risk assessors in every case.

As the practice of HHRA advances and the FCSAP proceeds, new and updated information on soil quality guidelines, drinking water guidelines, toxicological reference values (TRVs), contaminant bioavailability, human characteristics and exposure factors, and other aspects of HHRA will be published. As a result, it is anticipated that revisions to this document will be necessary from time to time to reflect this new information. Health Canada should be consulted at the address below to confirm that the version of the document in your possession is the most recent edition, and that the most recent assumptions, parameters, etc., are being used.

In addition, Health Canada requests that any questions, comments, criticisms, suggested additions, or revisions to this document be directed to: Contaminated Sites Division, Safe Environments Directorate, Health Canada, 99 Metcalfe Street, 11th Floor, Address Locator: 4111A, Ottawa, ON K1A 0K9. E-mail: cs-sc@hc-sc.gc.ca.

See also: http://www.hc-sc.gc.ca/ewh-semt/contamsite/index-eng.php.

ABBREVIATIONS AND ACRONYMS

ADI	acceptable daily intake
ATSDR	Agency for Toxic Substances and Disease Registry (United States)
BMC	benchmark concentration
BMD	benchmark dose
BMR	benchmark response
CCME	Canadian Council of Ministers of the Environment
CDF	cumulative distribution function
COPCs	contaminants of potential concern
CSM	conceptual site model
DL	detection limit
DQRA	detailed quantitative risk assessment
EDI	estimated daily intake
ESA	environmental site assessment
ET	exposure term
FCSAP	Federal Contaminated Sites Action Plan
HHRA	human health risk assessment
HQ	hazard quotient
ILCR	incremental lifetime cancer risk
LOAEL	lowest observed adverse effect level
NOAEL	no observed adverse effect level
NOEC	no observed effect concentration
PAHs	polycyclic aromatic hydrocarbons
PCBs	polychlorinated biphenyls
PCDDs	polychlorinated dibenzo-*p*-dioxins
PCDFs	polychlorinated dibenzofurans
PDF	probability density function
PHCs	petroleum hydrocarbons
PQRA	preliminary quantitative risk assessment
PRA	probabilistic risk assessment
QA	quality assurance
RfC	reference concentration
RfD	reference dose
SF	slope factor
TC	tolerable concentration
TC_{05}	tumorigenic concentration found to induce a 5% increase in the incidence of, or deaths due to, tumors considered to be associated with exposure
TD_{05}	tumorigenic dose found to induce a 5% increase in the incidence of, or deaths due to, tumors considered to be associated with exposure
TDI	tolerable daily intake
TLV	threshold limit value
TRV	toxicological reference value
UR	unit risk
U.S. EPA	United States Environmental Protection Agency
WHO	World Health Organization

1.0 BACKGROUND AND CONTEXT

1.1 Health Canada's Role in the Federal Contaminated Site Action Plan

A new contaminated sites initiative has emerged within the federal government. The Federal Contaminated Sites Action Plan (FCSAP) has been established to assist in identifying, assessing, and managing the risks at contaminated properties under the custodial care of Canadian federal departments.

Under the FCSAP, Health Canada is designated to provide expert support to federal departments with respect to providing guidance, training, and advice on human health risk assessment (HHRA) and public outreach at federal contaminated sites. Environment Canada, Fisheries and Oceans Canada, and Public Works and Government Services Canada are also designated as expert support departments in their areas of expertise.

Available HHRA guidance, contractor reports, and other information from Health Canada, in support of FCSAP, can be found at http://www.hc-sc.gc.ca/ewh-semt/contamsite/index-eng.php.

1.2 Human Health Risk Assessment for Federal Contaminated Sites

Whether at the screening level, preliminary quantitative level, or at a more complex level, HHRA is not an exact science. International, national, and provincial health and environmental agencies offer a wide variety of advice and direction regarding the conduct of HHRA. Different risk assessors access and rely differently on the available regulatory advice and direction, resulting in potentially wide variability in the estimates of chemical exposure and health risk. For example, in 1997, the Canada Mortgage and Housing Corporation (CMHC, 1997) commissioned a study in which nine consulting firms were commissioned to estimate the human health risks posed by a contaminated residential property. The resulting estimates of exposure and risk produced by the different firms varied over 9 orders of magnitude for non-cancer endpoints and over 10 orders of magnitude for cancer, despite being given the same site data set. The large variability related primarily to the differing human receptors and exposure scenarios assumed by the different firms. Variability was also introduced by the selection of different toxicological reference values (TRVs) for risk characterization.

Provincial regulatory agencies across Canada offer differing guidance on many aspects of HHRA. For example, definitions of acceptable cancer risk vary; British Columbia, Alberta, and the Atlantic provinces accept an incremental lifetime cancer risk (ILCR) of 1 in 100,000 (10^{-5}), whereas Ontario and Quebec target 1 in 1 million (10^{-6}). When characterizing the risks posed by exposure to non-carcinogenic substances, British Columbia accepts a hazard quotient (HQ) of 1.0 (exposure ≤ TRV), whereas Alberta and Ontario target 0.2 (exposure ≤ 1/5 TRV). Quebec accepts a HQ of 1.0; however, this value requires that background exposure be included in the calculation (MSSS, 2002).

Provinces also differ in the statistics preferred for exposure calculations, varying in the prescription of the maximum contaminant concentration, the 95% upper confidence limit on the mean concentration, or the 90th percentile or 95th percentile of the concentration data distribution. A comparison of preliminary quantitative risk assessments (PQRAs) conducted for a single hypothetical site following guidance from variety of provincial agencies (Dillon, 2004; Loney et al., 2007) demonstrated variable approaches, assumptions, and risk-related conclusions.

Based on the above observations, it was apparent that specific guidance on the conduct of detailed quantitative risk assessment (DQRA) was required at the federal level. Guidance for HHRA of simple federal contaminated sites is available: *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA), Version 2.0* (HC, 2010a). PQRAs generally prescribe methods and assumptions that ensure that exposures and risks are not underestimated and are primarily used to rank federal contaminated sites for FCSAP funding purposes. The purpose of this DQRA guidance is to prescribe, to the degree possible, a standard approach to quantitatively assess the potential chemical exposures and risks at complex contaminated sites under federal jurisdiction and those for which risk management is based on the HHRA.

The DQRA guidance presented herein is intended specifically for the assessment of sites that are to remain the property of a federal agency. For properties being divested to a private party or to provincial or territorial governments, or for assessments that address risks from off-site migration of contamination (to an adjacent provincial water body or neighbouring private property, for example), HHRAs may have to be completed in accordance with local provincial/territorial regulatory requirements. Local regulatory requirements may differ from the methods presented herein. When the methods being employed in such a case differ significantly from those presented herein, risk assessors should identify those assumptions, methods, and interpretations required by provincial agencies that differ from this method, and discuss the implications for the federal custodial department.

At first glance this guidance may seem overly complex. However, the length of this document stems predominantly from the inclusion of explanatory or "educational" text to ensure that the guidance and HHRA concepts are understandable, and that the rationales for various

requirements are apparent. In other words, an attempt has been made to describe **why** the methods are suggested, not just to delineate those methods.

The guidance presented herein is designed for application at sites where simple screening level or PQRA methods are inadequate. Because of the anticipated complex nature of such sites, this guidance necessarily lacks the prescriptive direction common to simpler PQRA guidance. It is impossible to anticipate or predict all possible land uses, receptor characteristics (particularly time–activity patterns, and the frequency and duration of site occupation), exposure scenarios, and other factors that will influence the HHRA. Although the guidance presented herein is not prescriptive in nature, it attempts to present a variety of issues and aspects that may have to be addressed and how to approach those issues.

This guidance is not designed or intended as a substitute for the sound professional judgment of a qualified and experienced practitioner of HHRA. It is recognized that each site to be subjected to a DQRA presents unique situations that cannot be predicted or anticipated. Risk assessors are encouraged to ensure that their HHRAs are complete and address all relevant human health risks. Such unique circumstances require that all necessary data, assumptions, equations, and approaches be sufficiently and unambiguously documented and described, both to enable peer review and to ensure the defensibility and validity of the resulting health risk estimates and subsequent risk management plans.

1.3 Petroleum Hydrocarbons and Radiological Contaminants

The guidance presented herein focuses on typically encountered inorganic and organic chemical contaminants. Although generally applicable to petroleum hydrocarbons (PHCs), additional guidance is available from the Canadian Council of Ministers of the Environment (CCME), specifically the Canada-Wide Standards for PHCs in soil (CCME, 2006, 2008a, 2008b, 2008c). Additional approaches are available from Atlantic Partners in Risk-Based Corrective Action Implementation (Atlantic PIRI, 2003) as well as contractor reports and spreadsheet models available from Health Canada via our website at http://www.hc-sc.gc.ca/ewh-semt/contamsite/index-eng.php.

With respect to sites contaminated with radiological substances, the following Health Canada guidance should also be consulted: *Federal Contaminated Site Risk Assessment in Canada, Part VI: Guidance on Human Health Detailed Quantitative Radiological Risk Assessment ($DQRA_{Rad}$)* (HC, 2010b).

1.4 References

Atlantic Partners in Risk-Based Corrective Action Implementation (Atlantic PIRI). 2003. *Atlantic RBCA (Risk-Based Corrective Action) Version 2.0 for Petroleum Impacted Sites in Atlantic Canada: User Guidance*.

Canadian Council of Ministers of the Environment (CCME). 2006. *A Protocol for the Derivation of Environmental and Human Health Soil Quality Guidelines [Revised].* CCME, Winnipeg. Report CCME PN 1332.

CCME. 2008a. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil*. CCME, Winnipeg.

CCME. 2008b. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil: Scientific Rationale.* Supporting Technical Document. CCME, Winnipeg.

CCME. 2008c. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil: Technical Supplement*. CCME, Winnipeg.

Canada Mortgage and Housing Corporation (CMHC). 1997. *Evaluation of Site-Specific Risk Assessment for Contaminated Lands.* Contractor report submitted by Golder Associates Ltd., to CMHC.

Dillon Consulting Ltd. (Dillon). 2004. *Comparison of Provincial Human Health Risk Assessment Methods for Contaminated Sites – Final Report*. Contractor report submitted to Environmental Health Assessment Services, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa.

Health Canada (HC). 2010a *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA)*, Version 2. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

HC. 2010b *Federal Contaminated Site Risk Assessment in Canada, Part VI: Guidance on Human Health Detailed Quantitative Radiological Risk Assessment ($DQRA_{Rad}$)*. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

Loney, C.B., G.M. Richardson, B.D. Leece, R.M. Wilson, U. Klee, and B.J. MacLean. 2007. Comparison of Contaminated Site Human Health Risk Assessment Approaches in Canada: Application of Provincial Methods to a Hypothetical Site. *Human Ecol. Risk Assess.* 13(6): 1228–1254.

2.0 OVERVIEW OF HUMAN HEALTH RISK ASSESSMENT FRAMEWORK

2.1 Introduction

This section presents an overview of the framework for HHRA that is the basis for the guidance provided in the subsequent sections of this document. The HHRA framework (Figure 2.1) is a component of an integrated risk management process (Figure 2.2). The framework includes linkages to other components of this integrated process (such as site investigation and remediation), and emphasizes the need for communication between parties (stakeholders) involved in the process, including the risk assessor, risk manager (federal custodial department), the regulatory agencies, potentially affected communities, and the public in general.

Figure 2.1 Risk Assessment Framework

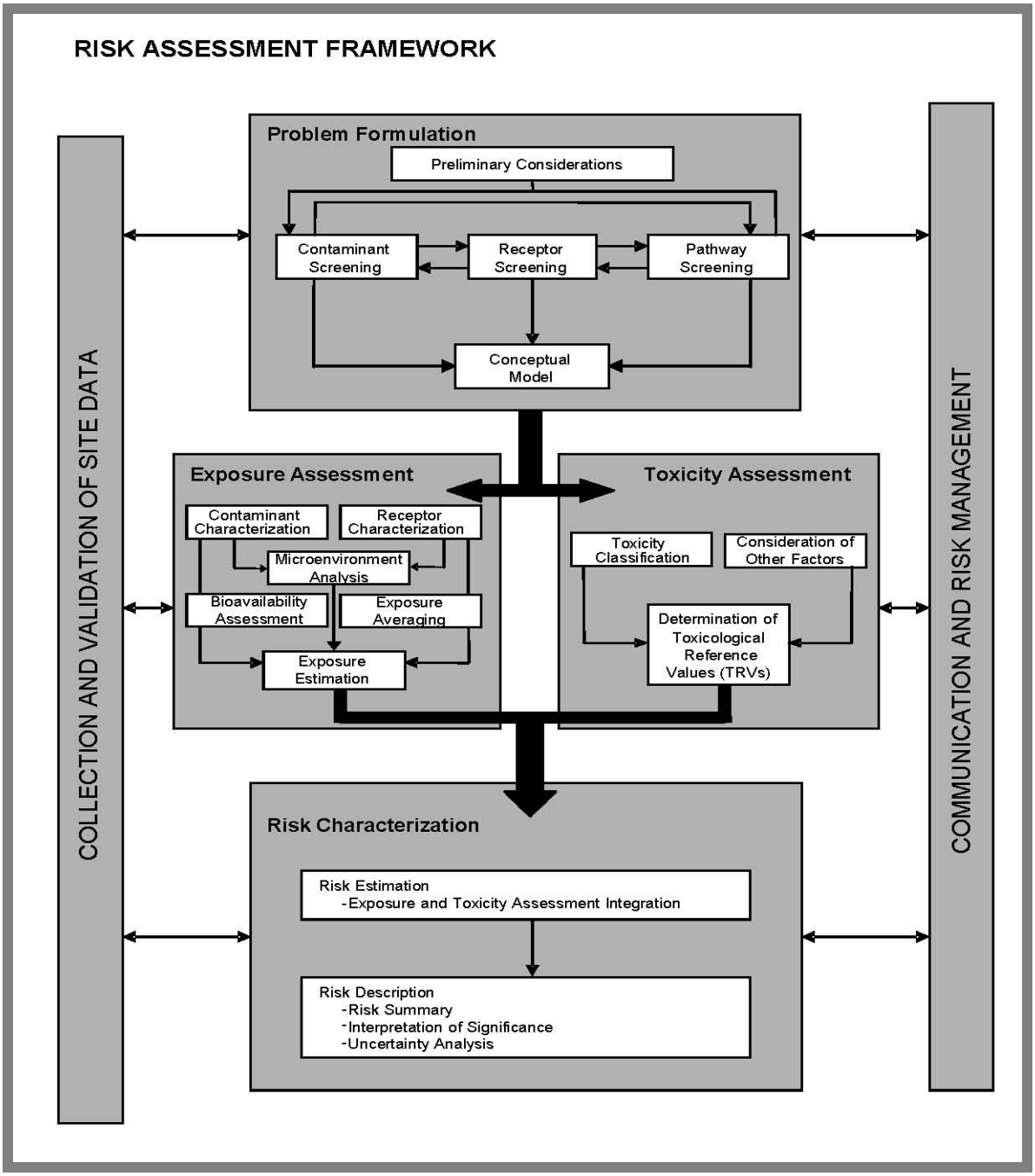
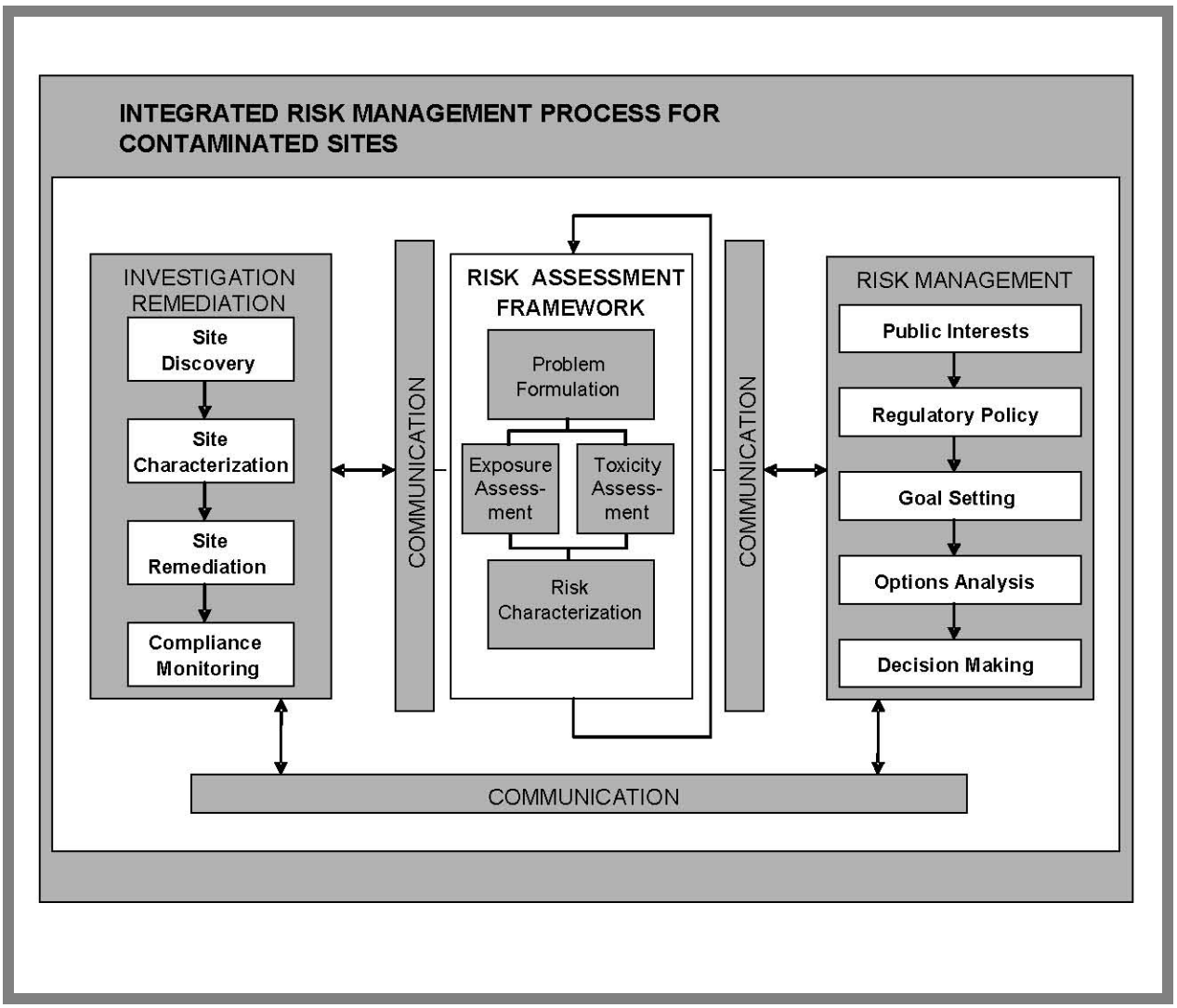


Figure 2.2 Integrated View of the Risk Management Process



2.1.1 Concepts, framework, and linkages to management of contaminated sites

The fundamental concept of HHRA of contaminated sites is based on three essential components: contaminants, exposure pathways, and receptors. A contaminant is a chemical or inorganic element present at a contaminated site at concentrations exceeding the natural or background concentration or the soil quality guidelines for a specific site use. Exposure pathways refer to the routes by which a contaminant may come into contact with a receptor from the contaminant source (soil, groundwater, etc.) to the point of exposure. A receptor is an individual, group, or population of humans (or animals, plants, or other organisms for ecological risk assessment) that may be exposed to the contaminant.

Each of these three components must be present in order for potential human health risk to exist; the absence of one or more of these components would eliminate the risk. The absence of one or more of these components is not necessarily unconditional; changes in site conditions, land use, and/or frequency or duration of site occupancy may introduce the missing component and create a risk where none previously existed. A risk can be eliminated unconditionally only by the removal of the contaminant source.

The framework for HHRA described in this document is consistent with approaches that have been widely used and accepted since the 1980s. Modelled after the framework established by the United States Environmental Protection Agency (U.S. EPA) for HHRA and ecological risk assessment of Superfund sites (NAS, 1983; U.S. EPA, 1986, 1989, 1998), this framework was originally described by BC Environment (1993)The framework, illustrated in Figure 2.1, comprises four main stages: problem formulation, exposure assessment, toxicity assessment, and risk characterization. Different terminology for these components may be found in other guidance; for example, toxicity assessment is also often referred to as hazard assessment. However, the overall process of HHRA for contaminated sites is generally consistent among jurisdictions.

Although the four stages of risk assessment comprise a complete and stand-alone undertaking, an HHRA is commonly conducted as part of the overall process of assessment and management of contaminated sites; this may include site data collection and validation, risk assessment, risk management, remediation, and consultation/communication. The linkages to other components of site management are illustrated in figures 2.1 and 2.2.

2.1.2 Applications and goals of human health risk assessment

In the context of the management of contaminated sites, a number of situations can arise in which a HHRA may be appropriate. Although an HHRA can theoretically be conducted in any situation where contaminants are present at a site, the majority of contaminated sites are typically managed through other means, such as the application of published numerical risk-based environmental quality guidelines or standards, whereby health and environmental risks are assessed and managed implicitly. Owing to data and resource requirements, detailed HHRAs are conducted in a relatively small proportion of cases and where the explicit evaluation of risk is necessary to support a risk management decision and related communications. PQRAs may be carried out in a greater number of instances to assess the need for, and feasibility of a DQRA and the associated data requirements.

Examples of applications in which HHRAs may be undertaken include the following:

- assessment of baseline human health risks (i.e. the risks posed by a site or facility in its present condition without remediation or other form of management);
- determination of site-specific risk-based remediation or risk management objectives where generic numerical standards or guidelines are inappropriate;
- assessment and management of sites affected by contaminants for which no standards or guidelines are available;
- assessment of human health risks associated with residual contamination, either following remediation or during the course of long-term remediation;
- assessment of human health risks arising from actual remediation activities (addressed where necessary for federal sites under the *Canadian Environmental Assessment Act*); and
- assessment of human health risks due to “background” exposure (not associated with a particular site or point source) to contaminants of potential concern (COPCs).

Although HHRAs are clearly applicable to the above situations, the decision to undertake a complex HHRA or DQRA, given the relatively extensive data and resource requirements, would also take into account its likely value in terms of the overall costs and benefits of the available risk management options, as well as the information and communication needs of stakeholders and regulators. Ultimately, the objective of any risk management decision is to focus the technical and management resources of the federal custodial departments, the relevant federal regulatory agencies, and the FCSAP to optimize the cost-effectiveness

of the solution while ensuring the protection of human health and the environment.

2.1.3 How a human health risk assessment is conducted

In concept, a DQRA is a linear step-wise process, originating with the problem formulation and culminating in the estimation of risk and/or the determination of risk-based remediation objectives. In reality, it is an iterative process in which data gaps are identified and addressed, and the key stages in the HHRA are successively refined. The scope of the HHRA (i.e. the level of effort) and the number of iterations will depend on the complexity of the site and the outcome of each stage or iteration in the process, as well as the overall goals of the assessment. Regulators and, in some cases, other stakeholders will need to be consulted at key stages for the identification and resolution of issues (see section 2.6).

HHRA activities can be divided into the respective stages of the process, and would be documented within the corresponding sections of the HHRA report. However, each stage is not necessarily independent of the others, although the first stage, problem formulation, is often a stand-alone task that forms the basis for regulatory and stakeholder consultation. The key risk assessment stages are described briefly in the following text and in greater detail in the respective sections of this guidance document.

2.2 Problem Formulation

Problem formulation is the first stage of HHRA and involves screening of the three main components of human health risk: chemicals, exposure pathways, and receptors (see Figure 2.1). The screening is based on preliminary considerations of the site, including likely site activities and land use scenarios. The objective of the problem formulation stage is to develop a conceptual site model (CSM) that will assist in determining how much additional data may be required to complete the HHRA, and which of the chemicals, pathways, and receptors are significant and most relevant to the site in question. The goal of this stage is to focus the quantitative HHRA on the contaminants, pathways, and receptors that have the greatest potential to contribute to potential human health risk.

Figure 2.1 shows the framework on which the problem formulation stage should be based. The arrows linking the three screening processes illustrate the first level of iterations that may occur at this stage. The problem formulation stage is broad reaching and has linkages to many aspects of the HHRA framework and to the management of site remediation as a whole. Early use of problem formulation principles can ensure that the data to be generated in the site investigation meet the needs of the subsequent HHRA stages. Further refinement of the hypotheses, which occurs as the site becomes better characterized, provides the basis and focus (i.e. the CSM) for the more detailed exposure and toxicity assessments and risk characterization to follow.

2.3 Exposure Assessment

The exposure assessment is conducted for the COPCs, exposure pathways, and receptors described by the CSM developed during problem formulation. The exposure assessment involves the estimation of the intake of COPCs by human receptors. The total intake for a given contaminant is the sum of the intakes estimated for each operable pathway identified in the problem formulation. Estimation of the intake rate, or dose, involves the determination (by direct measurement or predictive modelling) of the contaminant concentration in each relevant exposure medium in combination with the intake rate of the respective medium by the receptor(s). Although the exposure assessment is identified as the second stage of the HHRA, both the exposure assessment and toxicity assessment are typically conducted concurrently at this point because both follow the problem formulation and precede the risk characterization (see Figure 2.1), and may entail some interdependence.

2.4 Toxicity Assessment

The third stage of the HHRA is the toxicity assessment (sometimes termed hazard assessment), typically conducted in parallel with the exposure assessment as noted above and illustrated in Figure 2.1. Toxicity assessment is conducted for all COPCs. It involves identification of the potential toxic effects of these chemicals and the determination of either (i) a maximum dose or concentration of each chemical to which a receptor can be exposed without an appreciable risk of adverse health effect (threshold dose or concentration), or (ii) the relationship between dose and incidence or severity of adverse effect (dose-response relationship); this is particularly relevant to substances deemed to be carcinogenic. The toxicity assessment may involve the selection of published TRVs recommended by an appropriate regulatory agency or, in cases where appropriate published regulatory values are not available, the development of de novo values based on the critical analysis of toxicity studies. It is not uncommon for de novo TRVs to require subsequent regulatory acceptance or approval prior to their use in an HHRA.

2.5 Risk Characterization

Risk characterization involves quantifying the potential risks to receptors resulting from the estimated exposure to COPCs, and describing these estimated risks. Risk characterization is conducted for all chemicals and exposure pathway/receptor combinations identified in the problem formulation stage. Risk characterization involves:

- integrating the results of the exposure and toxicity assessments to determine whether a human health risk may be expected;

- analyzing, quantifying (where appropriate), and discussing uncertainty in the overall HHRA process, thus providing some indication of the validity or confidence of risk estimates; and
- describing the risks in terms of magnitude, type, and uncertainty involved.

2.6 Integration of Risk Assessment with Environmental Site Assessment, Remediation, Risk Management, and Communication

Although HHRA is the focus of this guidance manual, references are made throughout the document to the other key elements of the overall risk management process for contaminated sites, as illustrated in Figure 2.2. It is important that these elements be integrated with the risk assessment components because they will ultimately affect the acceptance of the risk assessment and subsequent adoption of appropriate measures to ensure that chemicals from the site do not adversely affect human health or the environment. A brief description of each element follows.

Environmental site assessment (ESA) involves the collection of environmental site data to characterize the sources, amounts, and concentrations of chemicals present, and to evaluate the potential pathways of contaminant transport and exposure.

Remediation is the evaluation, selection, and implementation of measures to reduce the concentrations of contaminants present at the site or in the relevant environmental media.

Risk management is the process of satisfying public concern and regulatory policy by implementing measures to manage health and environmental risks associated with chemical hazards present at a site. Risk management often describes measures aimed at reducing or eliminating exposure to contaminants as opposed to source reduction (remediation). Remediation is a form of risk management.

Communication is the exchange of information regarding site conditions, risks, and goals and objectives among the risk assessor, the site manager (federal custodial department), regulators, the affected community, and the public in general. Effective communication is the product of effective public involvement.

To further increase transparency in the manner in which a contaminated site is managed, Health Canada encourages custodial departments to implement public involvement strategies during all phases of contaminated site management (from the moment a site has been identified, and through the site investigation, risk assessment, risk management, and remediation phases). To help custodial departments undertake public involvement activities, Health Canada has developed the following guidance materials:

Guidance documents

- *Addressing Psychosocial Factors Through Capacity Building: A Guide for Managers of Contaminated Sites* (HC, 2005)
- *Improving Stakeholder Relationships, Public Involvement and the Federal Contaminated Sites Action Plan: A Guide for Site Managers* (HC, 2006)
- *A Guide to Involving Aboriginal Peoples in Contaminated Sites Management*

Training courses

- Improving Stakeholder Relationships: Public Involvement and Contaminated Sites
- Communicating Health Risk Information to Stakeholders

Fact sheets

- Benefits of Public Involvement for Custodial Departments
- Public Involvement Program
- Risk Assessment and Public Involvement at Contaminated Sites
- Health Impacts of Site Remediation
- Understanding Risk Assessment

Health Canada also has a team of public outreach specialists who can be contacted for more information and assistance at cs-sc@hc-sc.gc.ca. Fact sheets relating to public involvement and outreach may be found at http://www.federalcontaminatedsites.gc.ca/fcsap_pascf/public-eng.aspx.

2.7 Levels of Human Health Risk Assessment

HHRAs may be conducted to varying levels of detail and complexity, depending on the goals for the HHRA, the extent of available data, and the results or outcomes of the initial steps. In many cases, a preliminary assessment (i.e. PQRA) may be followed by a more detailed evaluation as part of an iterative process to assessing risk; in other situations, a PQRA may provide sufficient information to enable a risk management decision to be made.

A PQRA and a more complex DQRA are not independent, but represent opposite ends of a continuum of complexity. The general characteristics of a DQRA versus a PQRA are outlined in Table 2.1.

Table 2.1 Specific Characteristics of a Preliminary Quantitative Risk Assessment (PQRA) Versus a Detailed Quantitative Risk Assessment (DQRA)

	PQRA	DQRA
Environmental media sampled	Generally soil only; occasionally groundwater, if a concern	Generally includes soil, groundwater, vegetation, indoor air, outdoor air (volatiles and/or particulate), indoor dust, surface water, sediment, other environmental media as required
Quantity of data	Limited: generally restricted to data collected during Phase II/III ESA for confirmation of contamination and very limited delineation of hot spots	Extensive: generally includes a sampling plan designed to provide reliable and representative quantification of the contaminant(s) in each environmental medium/pathway
Statistic used to represent contaminant of potential concern (COPC) level(s)	Generally, the maximum measured concentration	Generally, the arithmetic average or the upper 95% confidence limit on the arithmetic average
Use of modelling	Extensive because COPC concentrations in all media, except soil (and perhaps groundwater), are usually estimated with the use of models	Limited: generally, direct data collected for all environmental media that are expected to be contaminated and/or to contribute significantly to exposure
Characterization of site	Limited to measurement of COPCs in soil (and perhaps groundwater)	Extensive; physical (soil texture, depth to groundwater, etc.) and chemical (pH, organic carbon content, buffering capacity, etc.) characterization of on-site soils and groundwater; precise measurement of distance from on-site structures (house, etc.) or proposed structures to contamination sources (hot spots); delineation of contamination in various media; other characteristics as required
Characterization of receptors	Limited to standard conservative assumptions available from published sources	May be site-specific, particularly with respect to the nature and extent of land use as well as time–activity patterns (when and how the land is used by receptors); quantification of receptor characteristics tends toward greater precision and less conservatism
Risk characterization	For non-carcinogens, based on 20% of the tolerable daily intake (TDI) because exposure from other media and background sources (unrelated to the site) may not be quantified For carcinogens, based on 100% of the acceptable risk value of 1×10^{-5} because the incremental lifetime cancer risk (ILCR) is independent of background sources	For non-carcinogens, can be based on 100% of the TDI because exposure from other media and background sources is quantified For carcinogens, based on 100% of the acceptable risk value of 1×10^{-5} because the ILCR is independent of background sources

A DQRA may be particularly appropriate in situations where there is a large amount of variation at the site in terms of land use, contaminant types and concentrations, soil and other site characteristics, and receptors and their interaction with the site. The increased detail and complexity of a DQRA will generally decrease the degree of uncertainty associated with risk estimates (relative to those from a PQRA, for example). This will result in the more accurate, precise, realistic, reliable, and defensible quantification of risks, as well as serving as a critical tool in the identification of complex remedial and risk management alternatives.

The required level of effort, detail, and scope of the risk assessment are usually determined at the problem formulation stage; this is discussed in greater detail in section 3.0. The terms **screening-level risk assessment**, **preliminary quantitative risk assessment**, and **detailed quantitative risk assessment** are commonly used to describe various levels of risk assessment. These terms are explained below; however, it should be noted that the level of detail, complexity, and degree of "realism" may vary among risk assessments conducted at a given level.

2.7.1 Screening-level risk assessment

The term screening-level risk assessment is generally applied to an assessment that does not result in a quantitative determination of human health risk (for example, see CCME, 2008), although the term does not uniquely describe the type or method of assessment. A screening-level risk assessment may be qualitative in that the problem formulation process is used to identify whether sources, complete exposure pathways, and receptors (and hence potential risks) may exist. Alternatively, it may be semi-quantitative, where concentrations are compared with relevant numerical criteria to assess whether potential risks might be anticipated to be significant, or to establish a relative risk ranking among the contaminants, sites, or receptors. In some cases, screening-level risk assessments may involve preliminary risk or exposure calculations. Screening-level risk assessments may be sufficient to establish that there is likely no risk, or they may form a preliminary step in a detailed risk assessment.

2.7.2 Preliminary quantitative risk assessment (PQRA)

A PQRA involves a quantitative determination of risk, generally using conservative contaminant concentrations and exposure assumptions. A PQRA is commonly carried out on the basis of relatively limited site information to provide an approximate, but conservative, estimate of potential human health risk. This may be used for preliminary site-ranking purposes, such as by Health Canada under the FCSAP. A PQRA may also be conducted to determine the need for a more detailed or site-specific quantitative risk assessment, including additional data collection requirements. Health Canada has defined the requirements of a PQRA in the specific context of Canada's FCSAP, and has published guidance for conducting PQRAs. A PQRA may involve fate and transport modelling, although simplified models are often used, and the assumptions used in the modelling are, again, generally conservative.

2.7.3 Detailed quantitative risk assessment (DQRA)

The purpose of a site-specific or DQRA is to produce a more accurate (realistic), defensible, and representative estimate of risks than that generated by a PQRA. Although the level of detail of such an HHRA can vary considerably, depending on the objectives of the assessment, a DQRA typically uses more comprehensive site characterization data, more representative or site-specific exposure information and, in many cases, a higher level of sophistication in the fate, transport, and exposure modelling. A common feature of a DQRA is the use of a statistical representation of the average source concentration of a contaminant rather than a conservative worst-case value. The need for a greater level of detail is usually assessed in the context of the benefits of reduced uncertainty in the risk estimate as compared to the costs and resources to collect the additional data and conduct a more detailed assessment, or as compared to the costs and resources required to conduct remediation/risk management based on the conservative PQRA.

2.7.4 Deterministic and probabilistic risk assessments

PQRAs and many DQRAs are deterministic. In a deterministic assessment, single "point estimate" values are selected for the assumptions and the parameters used in the risk calculations, leading to a single value or point estimate of risk. Depending on the objectives and availability of data, DQRAs may also be conducted using probabilistic methods. In a probabilistic assessment, probability distributions are used to represent uncertainty and variability in key assumptions and parameters. Equally, probabilistic risk assessment provides the range and frequency of risk estimates across an entire (exposed) population, rather than for a single hypothetical receptor. The resulting probability distribution of estimated risk provides more information and insight for risk management decision making; however, site data requirements are generally more onerous. Probabilistic risk assessment is discussed further in section 7.0 and Appendix A.

2.8 Desired Attributes of a Risk Assessment and Preferred Report Organization

The end-users of an HHRA, aside from the risk assessor, include the federal custodial department (the site risk manager), the regulatory agencies, the affected community, and the public in general. Each of these stakeholders has a

different level of understanding of scientific and policy issues and an interest in different aspects of the risk assessment itself. The overriding requirements of an HHRA are that the scope and objectives should be clearly and explicitly stated, that the scope and objectives satisfy all stakeholders, and that the presentation should be comprehensive and understandable.

Guidance regarding the deliverables associated with each of the major stages of an HHRA is provided in the respective sections of this guidance document. However, the following are considered desirable attributes of all stages of the assessment, including preparation of the report.

- The presentation should be complete, credible, and fully defensible, and it should include a description of any integral review process that was employed.
- The basis for selection of all critical assumptions should be documented and referenced scientific assumptions should be clearly distinguished from policy decisions.
- The influence of uncertainties in key assumptions on the risk estimate should be discussed, with sensitivity analyses conducted where appropriate.
- Conclusions should be drawn from the report results, and they should be relevant to risk management decisions.
- The report should contain a clear and concise summary in "layperson" terms that summarizes the objectives and conclusions and that includes a balanced treatment of relevant contentious issues.

A suggested outline for a DQRA report is presented in Table 2.2.

Table 2.2 Suggested Outline of a Detailed Quantitative Risk Assessment Report

Background and Objectives
Problem Formulation
Exposure Assessment
Toxicity Assessment
Risk Characterization
Recommendations

1.0 Introduction

1.1 Background and Objectives
- Background to problem
- Objectives of risk assessment

1.2 Site Description
- Physical setting and surrounding land use
- Site history and operations
- Summary of data collection activities
- Site geology and hydrogeology
- Contaminants present

1.3 Scope of Risk Assessment
- Complexity of risk assessment and rationale
- Overview of scope and methodology

2.0 Problem Formulation

2.1 Site Characterization
- Contaminant concentrations in soil
- Contaminant concentrations in groundwater
- Contaminant concentrations in other media
- Background concentrations

2.2 Hazard Identification
- Summary of contaminant concentrations
- Contaminant screening
- Identification of contaminants of potential concern

2.3 Receptor Identification
- Present and future human receptors
- Sensitive receptors
- Receptor screening

2.4 Exposure Pathway Identification
- Summary of contaminant sources and release mechanisms
- Identification of transport and exposure media (by contaminant class)
- Exposure routes (by receptor)
- Exposure pathway screening

2.5 Site conceptual model

3.0 Exposure Assessment

3.1 Characterization of Contaminant Concentrations
- Source concentrations
- Concentrations in exposure media (if available)
- Background concentrations

3.2 Characterization of Parameters Governing Contaminant Fate and Transport
- Contaminant-specific parameters
- Site parameters (geological, hydrogeological, etc.)

3.3 Fate and Transport Modelling
- Rationale for model selection
- Summary of model assumptions and inputs
- Estimation of exposure concentrations

3.4 Receptor Characterization
- Identification and characterization of critical receptors

3.5 Bioavailability Assessment (if applicable)

3.6 Exposure Estimation
- Exposure averaging/amortization
- Estimation of contaminant intake by contaminant, pathway, and exposure route
- Estimation of background exposure

3.7 Summary of Exposure Assessment

4.0 Toxicity Assessment

4.1 Threshold Contaminants (by contaminant)
- Summary of toxicological reference values (TRVs) and rationale
- Bioavailability assessment

4.2 Non-threshold Contaminants (by contaminant)
- Summary of toxicological reference values (TRVs) and rationale
- Bioavailability assessment

4.3 Evaluation of Potential Toxic Interactions

5.0 Risk Characterization

5.1 Threshold Contaminants (by contaminant)
- Estimation of pathway hazard quotients
- Additivity of pathways (if applicable)
- Additivity of contaminants (if applicable)

5.2 Non-Threshold Contaminants (by contaminant)
- Estimation of incremental lifetime cancer risks by pathway
- Additivity of pathways (if applicable)
- Additivity of contaminants (if applicable)

5.3 Uncertainties
- Uncertainties in site characterization
- Uncertainties in exposure assessment
- Uncertainties in toxicity assessment
- Potential for synergistic/antagonistic effects
- Uncertainties in risk estimates
- Identification of major factors influencing uncertainty

5.4 Discussion
- Comparison with background exposures
- Comparison with regulatory protection goals
- Comparison with site-specific health studies (if any)
- Identification of major factors driving risks
- Confidence in risk estimate

5.5 Summary of Risk Characterization

6.0 Recommendations

6.1 Additional Data Collection Requirements

6.2 Recommended Risk Management Measures

APPENDICES

A. Site-Specific Data
B. Models and Equations for Exposure Assessment and Sample Calculations
C. Toxicity Review (by contaminant) with Rationale for Selection of TRV(s)

Source: Adapted from U.S. EPA, 1989.

2.9 References

BC Environment, 1993. *Quantitative Human Health Risk Assessment: Phase I – Review of Methods and Framework Recommendations*. Prepared by Golder Associates under contract to BC Ministry of Environment, Lands and Parks.

Canadian Council of Ministers of the Environment (CCME). 2008. *National Classification System for Contaminated Sites: Guidance Document*. CCME, Winnipeg. PN 1403. ISBN 978-1-896997-80-3 PDF.

HC. 2005. *Addressing Psychosocial Factors Through Capacity Building: A Guide for Managers of Contaminated Sites*. Contaminated Sites Division, Safe Environments Directorate, Ottawa. H46-2/05-430E.

HC. 2006. *Improving Stakeholder Relationships, Public Involvement and the Federal Contaminated Sites Action Plan: A Guide for Site Managers*. Contaminated Sites Division, Safe Environments Directorate, Ottawa. Cat. H128-1/05-441E.

HC, 2011. *A Guide to Involving Aboriginal Peoples in Contaminated Sites Management.* Contaminated Sites Division, Safe Environments Directorate, Ottawa.

National Academy of Sciences (NAS). 1983. *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.

United States Environmental Protection Agency (U.S. EPA). 1986. *Superfund Public Health Evaluation Manual. Office of Emergency and Remedial Response*. U.S. EPA, Washington, DC. EPA/540/1-86/060.

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

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment*. Office of Solid Waste and Emergency Response. Washington, DC. EPA/630/R-95/002F.

3.0 PROBLEM FORMULATION

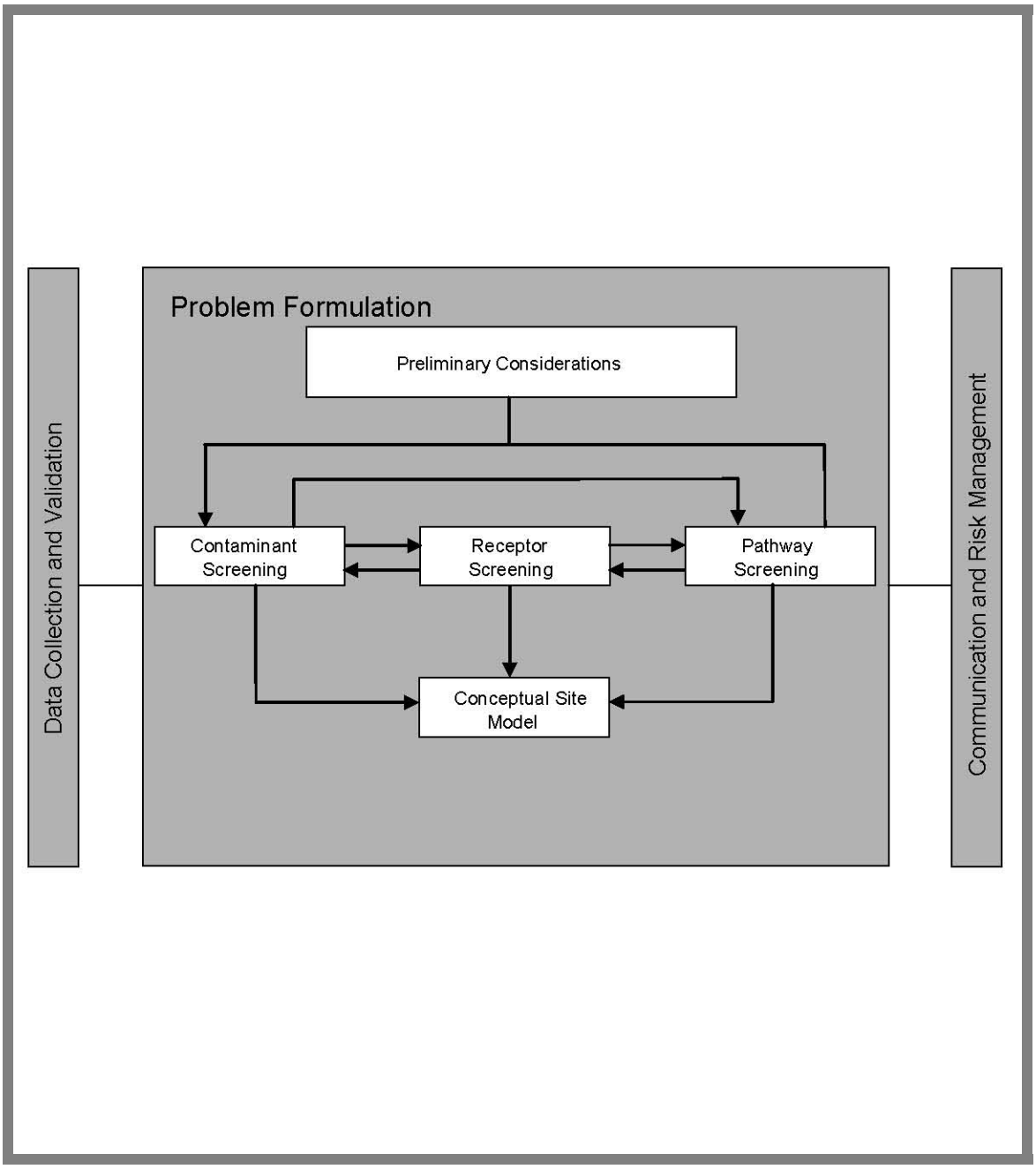
3.1 Introduction and Linkages to Other Risk Assessment Tasks

Problem formulation is the first stage of contaminated site risk assessment. During this stage, the risk assessor develops a focused understanding of the contaminated site, defines the goals of the risk assessment, and conceptually defines the working hypotheses of how receptors, contaminants, exposure, and toxicity interact to result in potential health risks. Collectively, these hypotheses form a CSM that underpins the subsequent risk assessment process.

The problem formulation stage of risk assessment is sometimes referred to as hazard identification, and many risk assessors use these terms interchangeably. Although the essential steps of problem formulation and hazard identification are the same, and are aimed at identifying chemical hazards, receptors, and exposure pathways, problem formulation also encompasses the scoping of the risk assessment and the process of deciding how to proceed beyond this initial stage.

As the critical initial stage of the risk assessment process, problem formulation involves rigorous preliminary assessments; screening of contaminants, receptors, and pathways; and effective planning. This helps to reduce the total number of overall iterations, and prevents delays caused by unexpected data gaps or misunderstandings with stakeholders in the process. Figure 3.1 outlines the essential components upon which the problem formulation stage should be based. The arrows linking the three screening processes illustrate the potential for iterations that may occur during this phase. For example, if contaminant screening has been conducted against environmental guidelines and background levels but the subsequent screening of exposure pathways suggests that a chemical need no longer be considered, a further refinement of the contaminant screening may be justified. In the case where new site information is acquired, or a change in goals occurs during later phases of the remediation process, additional iterations involving problem formulation may be required, as illustrated in Figure 3.1.

Figure 3.1 Problem Formulation



The problem formulation stage is broad reaching, and it has linkages to many aspects of the risk assessment framework and to the management of site remediation as a whole. Most fundamentally, the principles used and the tasks conducted during problem formulation will help to crystallize why, or even whether, the risk assessment is needed with respect to site remediation. For example, initial consideration of chemical concentrations, receptors, and exposure pathways may indicate implicitly that the health risks are highly unacceptable and that remediation is required without further assessment of baseline conditions. Conversely, problem formulation may reveal that a contaminated site has no operable exposure pathways, and therefore no associated toxicological health risks under the present conditions. In such a case, a baseline risk assessment might be unnecessary, unless changes in land use or divestiture (sale) of the property is planned.

Problem formulation is also linked to site investigation, and can provide guidance on how further efforts and resources should be expended to characterize the site. In addition, early use of problem formulation principles assures that the data to be generated during the site investigation will meet the needs of subsequent risk assessment steps. Further refinement of the problem formulation process, as the site becomes better characterized, provides the basis and focus (i.e. the CSM) for the more detailed assessments of chemical exposure, toxicity, and risk characterization.

Linkages also exist among problem formulation and other considerations outside the risk assessment process, including regulatory issues, consultation, public outreach and communication, and broader societal and risk management issues. Because regulatory and societal issues often drive risk management, it is important that these issues are considered when conducting problem formulation so the risk assessment is relevant, as well as scientifically defensible. The risk assessor must define the goals of the risk assessment incorporating these issues, as well as the scope of the project according to these goals. In addition, the risk assessor must communicate effectively with the regulators and other stakeholders to ensure that the issues receive adequate consideration during problem formulation and the development of the CSM.

3.2 Objectives of Problem Formulation

The objectives of the problem formulation stage are to:

- address regulatory and societal issues to define goals of the risk assessment
- establish scope and complexity of the risk assessment
- review existing data (or collect initial data, where applicable)
- identify and screen COPCs
- identify and screen receptors
- identify and screen exposure pathways
- develop a CSM
- identify data gaps and additional data collection requirements (if necessary)

3.3 Scoping the Risk Assessment

The problem formulation checklist described in *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA)*, *Version 2.0* (HC, 2010) provides an effective tool to summarize land use, receptors, and pathways that are critical to the site in question.

3.3.1 Defining goals

Strategic communication with the client, stakeholders, and regulators should give the risk assessor a clear understanding of the goal(s) of the intended risk assessment. These should be stated clearly and unambiguously in the report chapter produced at the conclusion of this stage (see also section 3.9 on recommended deliverables).

It is not sufficient to simply state that the primary goal is to determine the human health risks of a contaminated site; this is implicit. Rather, the statement of goals should also include **why** the risk assessment is being conducted (e.g. to establish site-specific remediation objectives, to establish whether unacceptable health risks may exist as a screening step in identifying the need for further action, to distinguish the most critical pathways contributing to risk for subsequent risk management considerations). This is where the regulatory and stakeholder perspectives come to bear. If the risk assessor cannot articulate why the risk assessment is being undertaken, it will be difficult to scope the complexity of the project and to rationalize "how" to conduct the risk assessment.

For example, if an investigation of a large industrial site has identified many areas of concern that require remediation, and the goal of the risk assessment is to prioritize the areas for remediation, the approach taken may not require sophisticated modelling or uncertainty analysis. On the other hand, if the site is an entire town and involves public and/or residential exposure, concerned citizens will likely want to know about detailed baseline health risks, and sophisticated uncertainty analysis may be necessary.

3.3.2 Determining complexity and level of effort

The complexity of the risk assessment and required level of effort will be dictated to some extent by the defined goals, as illustrated by the above examples. At the outset, the assessor should be able to determine whether a screening-level risk assessment, a PQRA, or a DQRA is warranted. However, in addition to the identified goals, a number of other factors also

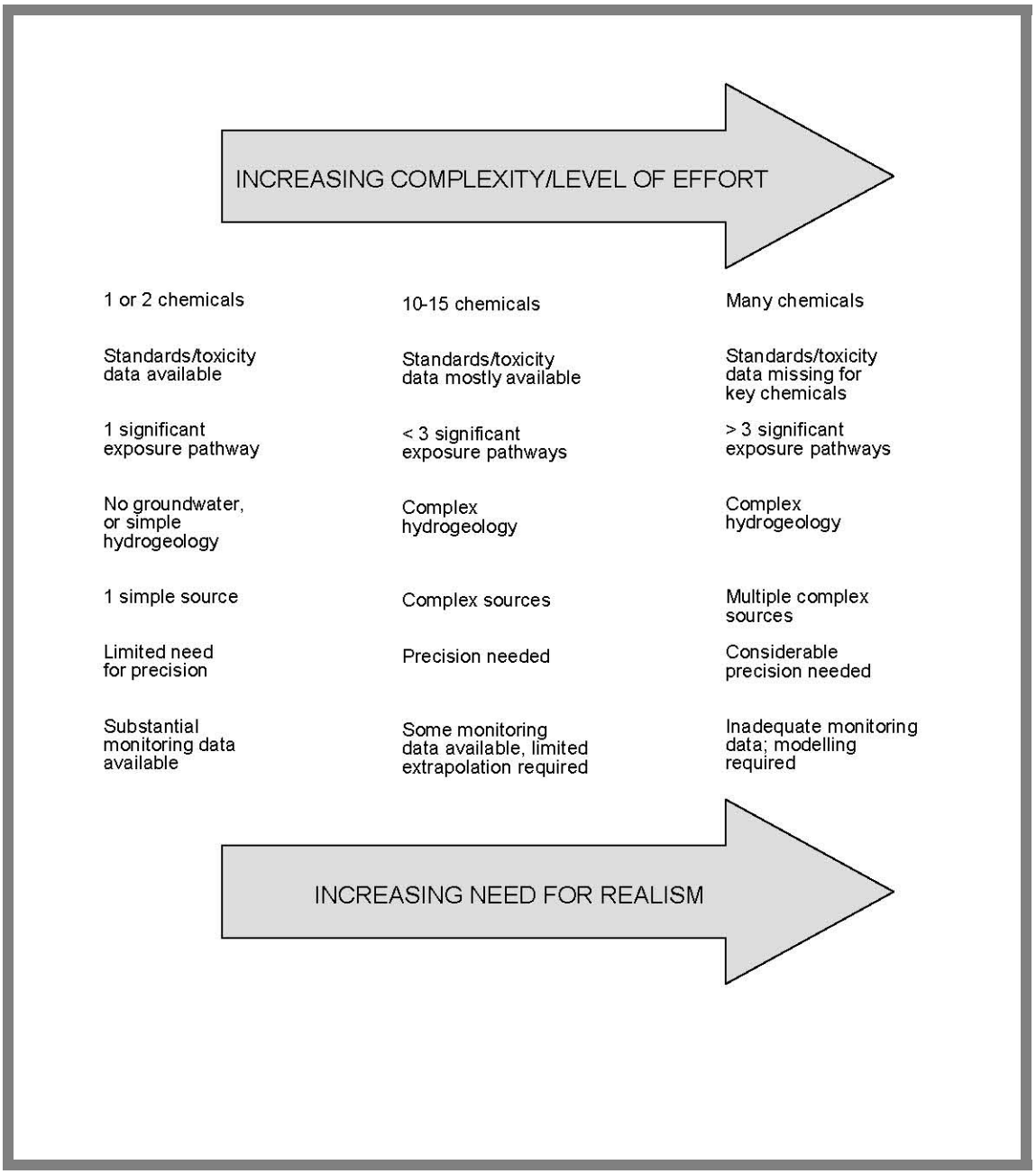
affect the complexity and level of effort, including the extent and quality of available data, complexity of site conditions, and financial and time constraints.

A useful method of scoping the problem is to scale the level of effort and resources required to successfully complete the risk assessment to the complexity of the site. Site-specific factors affecting the level of effort include (Zamuda, 1989):

- number and types of chemicals present
- availability of applicable criteria and toxicity data
- number and complexity of exposure pathways
- quality of the site investigation
- necessity for precision in the results

Figure 3.2 illustrates the continuum of analytical complexity and the relative level of resource requirements as determined by the complexity of the site.

Figure 3.2 Continuum of Analytical Complexity for Risk in the Contaminated Site Remediation Process



Sites having a low level of complexity (i.e. those to the left of the continuum) will require the lowest level of effort, whereas complex sites (i.e. those to the right of the continuum) are likely to require a high level of effort. For moderately complex sites, a combination of simple and complex analytical techniques may be used, depending on the variables in question. Therefore, in some situations, it may not be necessary to implement sophisticated techniques, such as fate and exposure models and uncertainty analysis, whereas in others it may be essential.

In many cases, the risk assessor may not become involved in a risk assessment of the site until after considerable effort has been expended on site investigation. When such a case arises, the assessor must ascertain whether the existing data are sufficient in quantity and quality to achieve the assessment objectives established by the site manager (custodial department) in consultation with stakeholders. Whether or not the opportunity exists to obtain further necessary or useful data may significantly influence the scope and complexity of the assessment because certain issues or questions may not be able to be addressed in a valid and confident manner when the necessary data are lacking. Ideally, a risk assessment will be contemplated and the problem formulation phase initiated prior to the collection of detailed site data. The level of detail of a risk assessment conducted at that point would necessarily be lower; a qualitative or semi-quantitative screening-level assessment may be appropriate, and may be used to guide the collection of additional data. The key point regarding available data is that the complexity of the risk assessment must be commensurate with the extent and quality of existing information. A detailed risk assessment involving sophisticated modelling techniques or probabilistic methods, conducted on the basis of inadequate data, may be misleading, the level of detail may obscure the underlying uncertainties, and the overall (lack of) confidence in the results may not meet assessment goals or expectations of the client and stakeholders.

Although not usually conducted at the problem formulation stage, sensitivity analyses permit the required level of detail or complexity to be assessed, based on the relative influence of parameters considered in a PQRA, particularly with respect to key or "driving" contaminant pathways and receptors. By focusing data collection, characterization, and modelling efforts on those parameters and variables that are significant drivers for the outcome of the risk assessment, the scope of the risk assessment can be bounded prior to embarking on more detailed data collection and modelling. An iterative approach would therefore be appropriate in some cases, whereby a PQRA would be conducted at the outset to refine the scope and essential data requirements of the detailed assessment.

3.3.3 Determining need for deterministic or probabilistic risk assessment

After defining the goals and scoping the complexity of the risk assessment problem, and conceptualizing the factors that govern potential risks, the risk assessor should determine whether a deterministic or probabilistic assessment should be conducted, or whether a combination of these analytical approaches should be used. As discussed in more detail in section 7.0, a deterministic analysis involves the use of point estimate values to characterize variables in the risk model; the result is a point estimate of risk with no quantitative information on the underlying variability or uncertainty. A probabilistic analysis, on the other hand, uses probability distributions to define the full range and frequency of values for variables in the risk model to account for uncertainty and/or variability in some or all of the variables; a probability distribution of estimated risk is produced (hypothetically) representing the full range and frequency of risks across the entire population of receptors. This additional information may be desirable in various risk management decisions, particularly when the risk estimates are of a magnitude that is close to the level of concern.

It is important to determine, as soon as possible, whether a deterministic or probabilistic approach will be used because the outcome of the decision may require new data or different techniques to be employed in subsequent phases of the risk assessment. A deterministic analysis will almost always be conducted as an initial step even if a probabilistic analysis is planned. The deterministic assessment, together with an associated sensitivity analysis, would be used to establish the scope and data requirements of the subsequent probabilistic analysis, particularly with respect to identifying and characterizing those variables that are to be represented by probability distributions.

Given the differences in interpretation between the two methods of analysis, it is important to have an understanding of when it is most appropriate to employ a deterministic method, a probabilistic method, or a combination of both. Table 3.1 lists considerations for selecting one or both methods.

Table 3.1 Considerations for Selecting a Deterministic Versus a Probabilistic Method of Analysis

Deterministic	Probabilistic
Relatively low degree of site-specific uncertainty	Relatively high degree of site-specific uncertainty
Deterministic result significantly higher or lower than level of concern	Deterministic result close to level of concern
Small-scale project (scope, budget, schedule)	Large-scale project (scope, budget, schedule)
Routine application	Non-routine application
Large number of potential contaminants and/or pathways (screening tool)	Small number of potential contaminants and/or pathways (or screening has been conducted)
Initial model development	Model refinement
Qualitative uncertainty analysis	Quantitative uncertainty analysis
	Value of information analysis (additional data collection)

Deterministic assessment can be especially effective as a first-cut screening tool and in spatially small-scale risk assessments with limited budget or constraints in a reporting schedule; this is because deterministic assessment is generally conservative (i.e. risk estimates will be biased on the side of safety following the precautionary principle). Upper-bound (worst-case, reasonable maximum exposure, etc.) estimates are used to characterize input variables in order to increase confidence that risks have not been underestimated. However, just how conservative a deterministic assessment is (e.g. what proportion of the exposed population will be "safe") cannot be ascertained. This method is also less costly, and requires less time to perform than a probabilistic assessment. In addition, combinations of a low degree of public concern or public interest (where detailed communication of the uncertainties in risk estimates is not essential), low site-specific uncertainty (e.g. small sites with relatively uniform contaminant levels and well-defined land use), and routine applications where remediation is not expected to be problematic (e.g. low-cost remediation, such as excavation, with little likelihood of residual contamination) would also be suitable circumstances for the use of a deterministic approach. For such cases, uncertainty need not be explained in quantitative terms, and conducting a probabilistic risk assessment would produce results that would likely not influence the ultimate risk management decision.

Deterministic PQRAs should follow the guidance provided by Health Canada for federal contaminated sites.

Certain conditions exist where the use of probabilistic methods provides distinct advantages:

- where there is a high degree of site-specific uncertainty or variability, because probabilistic methods are able to quantify uncertainties and variability and can provide added insight by describing the risk in terms of a given percentile of exposure, or a specified degree of conservatism;
- where there is a high degree of public interest, such as controversy over the relevance or meaning of a risk assessment to public health (i.e. where qualitative estimates of uncertainty will not be credible or acceptable);

- where the deterministic result is close to, or perhaps exceeds, the action level (i.e. where risk managers could significantly overestimate or underestimate actual risk);
- where initial risk calculations are conducted on the basis of professional judgment, and where such judgment can differ among professionals, uncertainty analysis can assist in refinement of the model assumptions, thereby leading to more reliable and more broadly accepted risk estimates; or
- where a value-of-information analysis, which is facilitated by probabilistic modelling, could save significant time, money, and resources by focusing on key parameters when collecting data for evaluating remediation alternatives.

From the above, it is clear that the decision to use a deterministic, a probabilistic, or a combination method of analysis is not always readily apparent. In many instances, it will be appropriate to use deterministic analysis for initial CSM development during problem formulation and probabilistic analysis for the remainder of the risk assessment. The exact choice, however, is a project management decision that the risk assessor must make in consultation with the client and stakeholders after considering the regulatory, societal, and risk management issues, as well as data limitations. Considerations in selecting between probabilistic and deterministic methods are discussed further in section 7.0 and elsewhere (e.g. Burmaster, 1996; Richardson, 1996; Cullen and Frey, 1998; U.S. EPA, 2001).

3.3.4 Determining need for specific models

Most risk assessments require some level of environmental migration and fate modelling to aid in the characterization of exposure, and occasionally in the characterization of toxic response. Modelling is often required to estimate (particularly into the future) exposure concentrations where measured contaminant concentrations in exposure media were not explicitly collected, as well as to predict future contaminant concentrations in time and space when future land use scenarios are being considered and where the long-term safety of a site is in question. Models are also used to estimate human intake or uptake of contaminants, as well as physiological and metabolic fates of contaminants within the human body.

Models employed in risk assessment typically address some aspect of the exposure pathway(s). The need for contaminant fate and transport modelling will be a function of the source of contamination, the transport and exposure media, and the form of interaction between the human receptor(s) and the exposure media. Identification of the requirement for modelling is facilitated by the CSM, but modelling requirements also depend upon the availability (or lack thereof) of data with respect to contaminant concentrations in all relevant media.

Models for estimating contaminant fate, transport, and exposure are discussed in further detail in section 4.0. The selection of an appropriate model for use in a risk assessment depends upon various factors, including the goals and objectives of the assessment. The choice of model may also significantly affect the scope and complexity of the assessment. The level of sophistication of a model should be consistent with the available data. The use of a complex or sophisticated model is generally unjustified or indefensible if site data are limited for application in a simple risk assessment. An exceedingly complex model will be inappropriate where site characterization failed to collect data on all or a significant number of the variables required to confidently apply the model to the site in question. On the other hand, the use of simpler models may introduce greater uncertainty into the risk assessment process, often with implicit conservativeness that may be inconsistent with the desired level of confidence or accuracy expressed by stakeholders. It is important to **ensure** that the level of uncertainty is appropriate given the intended application of the results of the assessment. Selection of an appropriate model should also reflect its suitability for application to a deterministic or probabilistic analysis, and selection should be discussed with the client and stakeholders prior to its application to the site in question.

3.4 Hazard Identification

The overall objectives of hazard identification are to identify (screen) and characterize the COPCs that will be considered in the ensuing risk assessment. Factors considered in the identification of COPCs include concentration and amount (mass or volume) present, toxicity, mobility, persistence, and potential to bioaccumulate or biomagnify. In practice, selection of the COPCs typically involves a screening process whereby measured contaminant concentrations are compared with reference criteria and with naturally occurring levels in the case of inorganic elements. The screening process requires adequate site characterization data as well as appropriate reference criteria and background concentrations for comparison purposes. Minimum data requirements for site characterization are discussed in this document and in the Health Canada environmental site assessment (ESA) guidance (HC, unpublished).

3.4.1 Site characterization requirements

Site characterization data are generally obtained by means of an ESA. ESAs are commonly conducted in phases. A Phase I ESA makes use of historical information to identify contaminants likely to be associated with past site activities and to guide the subsequent collection and analysis of soil, groundwater, and other samples, but typically does not involve intrusive investigation or sampling. A Phase II and maybe a Phase III ESA is therefore required to obtain quantitative information regarding concentrations of chemicals in environmental media and their spatial and vertical distribution. For a more detailed review of this topic, refer to

the *Manual for Environmental Site Characterization in Support of Human Health Risk Assessment* (HC, unpublished) or other available resources, including the CCME or provincial guidance. Sampling and analysis should be conducted by qualified individuals and follow standard reference analytical methods, and protocols should be documented. Chemical analyses should be performed by accredited laboratories, and an appropriate number of samples should be analyzed for quality assurance/quality control purposes.

The actual requirements for site characterization data cannot be prescribed herein; they will depend upon the intended application, scope, and level of detail of the risk assessment, but should also be consistent with the size and complexity of the site. Health Canada should be consulted about the availability or recommendation of guidance on data quantity, data quality, and designing site sampling plans for collection of data that will be employed to estimate potential risks posed by a contaminated site. Professional judgment will be applied in determining the number of samples and range of analytes to ensure confidence in the final risk assessment results. As a general rule, the data should be sufficient to determine statistically representative, spatially representative, and/or temporally representative contaminant concentrations, including the spatial extent and distribution of the contamination, as well as local background concentrations.

Sampling during ESA Phase II often targets areas or strata of known or suspected contamination. Such data have a reasonable certainty of measuring the maximum or near-maximum concentration. However, such data may not be "representative" as intended or required to meet the objectives of the risk assessment. Data should be obtained for all contaminants that are known or suspected to be present at the site, including toxic degradation products of contaminants known to degrade. It should be noted that the extent of information required for chemical screening at the problem formulation stage may be different from that required to undertake the site-specific risk assessment, particularly if a probabilistic analysis is to be conducted. The collection of further data may therefore be required at a subsequent stage. For example, chemicals are typically screened on the basis of maximum measured concentrations. Although this requires sufficient analyses to ensure that the areas of likely highest concentrations have been sampled, a higher number of samples would generally be needed to establish a concentration (or probability distribution thereof) that is representative of the site as a whole or of the areas of the site frequented by receptors for the quantitative estimation of risk.

3.4.1.1 Soil particle size distribution relevant to human health risk assessment

Risk assessors should ensure that the range of soil particle sizes collected and analyzed should be relevant to on-site exposure to soil-borne contaminants. Neither contaminant concentration (see Bright et al., 2006) nor bioavailability (see Richardson et al., 2006) is uniform across soil size fractions. Both concentration and bioavailability may increase as particle size decreases, but this pattern is by no means consistent, universal, or predictable.

Soil samples sieved to ≤ 250 µm (or simply bulk soil) are not necessarily representative of the soil size fraction of greatest relevance to HHRA. The U.S. EPA (2000) concluded that 250 µm represents a reasonable upper bound for the size range of ingested soil particles. However, the studies on which that conclusion was based were limited in the particle size ranges examined; focus was generally on the ≤ 250 µm fraction. There is evidence that the critical size fraction is still finer than this. In a review of available literature conducted for Health Canada (Globaltox, 2005), it was concluded that the < 150 µm fraction of soil may better represent the particle sizes resulting in exposure. However, that literature provides considerable uncertainty about what particle size fraction(s) may be most significant. Data published by Duggan and Inskip (1985) demonstrated preferential adherence of particles < 53 µm to the digits (fingers). Later work by this group (Duggan et al., 1986) demonstrated that 90% to 98% of the soil that adhered to children's hands, and was thereby potentially available for transmission to the mouth for ingestion, was < 10 µm in diameter. More recent investigations by Sheppard and Evenden (1994) and Kissel et al. (1996) also demonstrated the preferential adherence to the hands (for subsequent potential transfer to the mouth for ingestion) of the finer particle size fractions, relative to the distribution of size fractions in the soils investigated. Sheppard and Evenden (1994) reported particle size fractions in the 0.5–25 µm range as preferentially adhering to hands, whereas Kissel et al. (1996) demonstrated this same phenomenon for the < 65 µm fraction (smallest size fraction investigated). Edwards and Lioy (1999), investigating house dust, determined that > 95% of the total number of particles adhering to hands (or a sampling device designed to mimic the hand surface) were ≤ 2.5 µm in size.

3.4.1.2 Other site characteristics required for risk assessment

Aside from the characterization of contaminant concentrations, additional data required for risk assessment includes information on land use, and the frequency and duration of visits or occupation by human receptors on site, as well as the physical, chemical, and geological conditions that govern the fate and transport of chemicals in the environment. Such data could include, but would not be limited to, the following:

- soil type, texture, stratigraphy, porosity
- soil properties such as moisture content, organic carbon content, bulk density, pH
- depth to groundwater

- direction of groundwater flow, hydraulic conductivity, hydraulic gradient
- direction and distance to surface-water bodies
- land and water use, distance to human receptors
- presence of structures or other features that may influence human activity patterns or areas of activity on site

Some of the above information is required at the problem formulation stage in order to facilitate screening of receptors and exposure pathways (see sections 3.5 and 3.6). However, detailed data that may be used in the modelling of contaminant fate and transport may be collected at a later stage once the CSM of the site has been established.

3.4.2 Chemical screening

The purpose of chemical screening at the problem formulation stage is to identify chemicals that pose, may pose, or have the potential to pose risks or hazards to human health. These chemicals are carried forward to the subsequent stages of the risk assessment and are referred to as chemicals of potential concern (COPCs).

The approach to chemical screening involves a number of steps:

- identifying appropriate screening criteria, such as CCME soil and surface-water quality guidelines
- comparing chemical concentrations with screening criteria
- comparing chemical concentrations with background conditions
- rationalizing/excluding innocuous substances
- selecting COPCs

3.4.2.1 Identifying appropriate screening criteria

To be considered **appropriate** for the purpose of screening COPCs, criteria should be risk based, scientifically defensible, up-to-date, and acceptable to the governing regulatory agency. For federal contaminated sites in Canada, the CCME *Canadian Environmental Quality Guidelines* (CCME, 1999, with more recent updates), and the CCME Canada-Wide Standards for PHCs in soil (CCME, 2008a, 2008b, 2008c) should be used. For substances lacking guidelines from CCME, similar provincial guidelines and standards should be used. Where no Canadian jurisdiction has established a human health-based environmental quality guideline for a particular contaminant, criteria derived by other jurisdictions, such as the most recent U.S. EPA preliminary remediation goals or soil-screening levels (U.S. EPA, 1996, 2002), may be used with appropriate adjustments, as described later in the text. Where a criterion other than those developed by CCME is used for screening, a detailed rationale for the use of the criterion should be provided. The rationale should include the basis for the criterion and any adjustments that were made to the criterion.

Most published risk-based environmental quality guidelines and standards have been developed using standardized (or "generic") assumptions regarding exposure conditions. In some circumstances, where site-specific conditions differ from those assumed in the generic case such that the potential degree of human exposure is greater, the existing guidelines or standards may not be sufficiently protective. In all cases, values derived for the protection of human health under the most conservative applicable land use category should be used.

Where the criteria adopted for screening purposes are obtained from sources other than CCME or Health Canada, they should be adjusted as necessary to be consistent with the health protection endpoints prescribed by Health Canada and CCME. For example, if the health-based guidelines for carcinogens (non-threshold substances) are derived based on a target incremental cancer risk of 1×10^{-6} (1 in 1 million), the criteria can be adjusted to a target incremental risk of 1×10^{-5} in accordance with Health Canada's essentially negligible risk level. For non-carcinogens (threshold substances), guidelines from other jurisdictions, such as the U.S. EPA, are commonly based on 100% of the tolerable daily intake (TDI) or reference dose (RfD). These guidelines should be divided by 5 to make them approximately equivalent to CCME guidelines that are based on only 20% of the TDI or RfD.

Note that some sources may identify criteria or standards (e.g. for drinking water) that are based on an aesthetic (e.g. odour or taste) threshold because, in some cases, this may be a more sensitive endpoint for regulatory purposes. Such criteria should be acknowledged because they may affect site remediation, but they are inappropriate for purposes of health risk assessment and should not be used in this context without consultation with the relevant regulatory agency.

Should an appropriate comparative criterion not be identified for a particular chemical, the chemical should be retained as a COPC and carried forward to the detailed risk assessment stage.

3.4.2.2 Comparing chemical concentrations with screening criteria

For screening site chemical concentrations against criteria, the maximum observed concentration is used in a risk assessment. The purpose of this screening step is to compile a conservative list of chemicals for further evaluation. **If the maximum chemical concentration does not exceed the appropriate criterion, the chemical is considered to be acceptable for that land use and is normally excluded from further detailed risk assessment.**

It is important to note that the criteria developed by provincial and federal regulatory bodies may not account for potential interactive effects of similarly acting chemicals. For example, it is possible that the soil at a site contains numerous types of carcinogenic polycyclic aromatic hydrocarbons (PAHs) that have interactive effects. Although none of the individual chemical concentrations may exceed the identified criteria, it is possible that collectively they could pose significant risks if, for example, the sum of their potency equivalents for benzo[*a*]pyrene exceeds the criterion for benzo[*a*]pyrene. It is essential that the interactive effects of similarly acting chemicals be considered, or the whole group could be inappropriately excluded on the basis of data on individual chemicals. Currently, this concern applies to carcinogenic PAHs, polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCDDs/PCDFs), and dioxin-like polychlorinated biphenyls (PCBs).

3.4.2.3 Comparing chemical concentrations with background concentrations

The list of COPCs derived following screening with applicable regulatory guidelines may then be screened, based on local geochemistry. Information on background concentrations for a limited number of inorganic elements is available from the Geological Survey of Canada and some provincial sources. This issue was also reviewed on behalf of Health Canada by Water & Earth Science Associates (WESA, 2005). Because some chemical concentrations may be elevated naturally, the next step involves consideration of site history, land use, and background data (local and regional) to determine if the detected concentrations are a natural anomaly. Although the risk assessment principles of this document could be applied to a site with natural contamination in excess of CCME guidelines, the intended focus is on anthropogenic sources because the latter are the focus of risk management plans. In many cases this step can be completed rapidly and intuitively if the particular chemical itself is anthropogenic (e.g. most persistent organic pollutants), and appears to be logically associated with historical land use of the site. This latter information should be readily available as a result of the site investigation. However, for naturally occurring substances (e.g. arsenic, mercury, selenium, PAHs, PCDDs/PCDFs) and for other widely distributed chemicals (e.g. pesticides, radionuclides, lead historically deposited by automobile exhaust, and PCBs), comparisons should be made with available local or regional surveys of background concentrations, or measurements from a control site, to ascertain whether or not the chemicals in question stem from local anthropogenic sources. The control or reference site must be carefully chosen because the chemicals of interest must be attributable solely to natural (i.e. non-anthropogenic) sources. The reference site must be shown to be free of any possible anthropogenic point source influence with regard to the chemicals of interest.

The reference site should closely match the contaminated site of interest in "geographic area and scope." This encompasses location, topography, size/area, physical and chemical characteristics of soil geology, and hydrology. Within sites, preference should be given to vacant land, naturally wooded areas, parks, or large residential lots. Sites with obvious vegetation damage should be avoided. The history of the reference site and adjacent land, including current and past activities, must be considered and documented.

If concentrations of a contaminant are considered to be naturally elevated, or if measured concentrations are within the range of local or regional background conditions, the contaminant should be excluded from further consideration as a COPC unless specifically directed to retain it as a COPC as part of the project scope (because of community/public interest, for example).

The concept of **background concentrations** of chemicals varies among jurisdictions (see WESA, 2005). This is an important area for communication with the affected community or public in general. Potential background levels can be derived from data taken from control sites located close to, but outside the influence of, the contaminated site, or they may be based on a more regional view (e.g. some percentile of province-wide values) as defined by the appropriate province in which the site is located. Whichever concept is employed, it may be prudent at this stage to consult with the custodial department, Health Canada, Geological Survey of Canada, and provincial agencies to ascertain the availability and appropriateness of data on background levels, and to determine the need for sampling of background levels at a reference site.

3.4.2.4 Excluding innocuous substances

Several naturally occurring substances, such as calcium, iron, magnesium, potassium, and sodium, are included in routine analytical chemical analyses. Government agencies often do not develop regulatory criteria for these and other innocuous substances. Unless these substances are knowingly associated with on-site activities, they should be excluded from the risk assessment.

The rationale for exclusion of these chemicals must be recorded so that the decision process is understood, transparent, easily retraced, and verifiable. If exclusion cannot be rationalized during this step, the chemical is regarded as a COPC and retained for detailed HHRA.

3.4.2.5 Selecting contaminants of potential concern

All chemicals identified and/or investigated at a site should be summarized, along with their maximum measured concentrations (as well as other basic statistical information such as number of samples, mean concentration, and range). Screening criteria should also be tabulated, with appropriate references to source documentation.

Any contaminant whose maximum concentration exceeds the appropriate screening criterion, and that was not excluded owing to background concentrations or other considerations, would be classified as a COPC and included in the detailed risk assessment. Any contaminant for which appropriate screening concentrations are not available should also be retained as a COPC. Chemicals for which insufficient data have been collected to permit their definitive exclusion from further consideration should also be considered as COPCs and/or subject to further investigation or data collection. The basis for any decision to include or exclude a chemical as a COPC at the problem formulation stage should be clearly documented.

3.5 Receptor Identification

The objective of receptor identification is to identify and screen those human receptors likely to experience significant exposure, and hence risk, to the COPCs. Although the receptor screening process will focus on those receptors likely to receive the greatest degree of exposure, care must be taken to ensure that receptors who may experience less exposure, but who are subject to potentially greater risk (e.g. as a result of higher sensitivity), are not excluded.

Key receptors will be primarily determined by two factors:

1. land use (e.g. adults are key for industrial lands, whereas toddlers are key for assessment of non-carcinogens at residential properties)

2. chemical-specific toxicity where a particular receptor group is considered the key receptor for the substance in question (e.g. the foetus with respect to methylmercury toxicity)

The receptor identification and screening step should also be conducted in conjunction with the identification and screening of exposure pathways, as the steps are closely linked.

3.5.1 Potential human receptors

At the outset, human receptors are identified in a generic sense in accordance with the land use of the site and surrounding area. The term land use, as used in this section, also encompasses air and water use. The land use generally defines which age groups or other subsets of the overall human population are expected to be present and hence potentially exposed. Receptor groups may include the general public, on-site residents, employees, or members of specific subpopulations, such as Aboriginal communities, or other categories with distinct and identifiable site access, exposure characteristics, and/or behavioural patterns relevant to the site and contaminants in question. Receptor groups are typically divided into age classes; in Canada, this includes infants, toddlers, children, teens, and adults. In some cases, these may be further divided to include categories such as seniors or separate categories for males and females.

Four generic land use classifications have been defined by CCME (2006): agricultural, residential/parkland, commercial, and industrial. Each of these is commonly associated with one or more receptor group(s). For example, the general public would be expected to be potentially present, or have access to a site, under agricultural, residential, or commercial land use; access under industrial land use may be limited to adult employees. Members of Aboriginal communities may be present under any of the land uses, depending upon site location. These groups may be particularly relevant to northern and remote sites that may be frequented as camps or locations for hunting and fishing. On a site-specific basis, the preceding generic land use classifications may not be representative with regard to the exposed population or their potential degree of exposure, in which case additional land use definitions or characterization may be required. Receptors and exposure characteristics should be identified on a site-specific basis if they are likely to result in exposure of different receptor groups. For example, remote sites with limited access may be visited exclusively by maintenance personnel on an infrequent basis. On the other hand, some remote locations may be subject to more intensive use by First Nations and Inuit populations engaged in harvesting of traditional foods.

Depending on the receptor group, a **critical receptor** may be defined on the basis of intake rate, body weight, and other exposure factors, such as frequency and duration of exposure. This is part of the receptor screening process and is discussed in further detail below.

When selecting receptors based on land use, consideration should be given to the most sensitive current land use, or likely future land use if the scope of the risk assessment is to address site development and land use change. Care should be exercised to ensure that the most sensitive usage is reflected in the land use classification. For example, daycare centres are commonly co-located with commercially zoned areas; a daycare facility would represent a more sensitive land use than a typical commercial enterprise. It must also be recognized that any change in land use not considered within the risk assessment may invalidate that risk assessment or lead to land use restrictions or other administrative controls. Regardless of site-specific land use, adjacent land use should also be considered if it is, or may in future become, more sensitive (such as a residential development constructed

adjacent to an industrial site), or if on-site contamination is migrating off-site and potentially impacting adjoining properties.

3.5.2 Sensitive and unique receptors

In identifying potential receptors, consideration should be given to potentially sensitive and/or unique receptors who may be exposed to increased levels of risk. Examples include seniors, pregnant or nursing mothers and infants (particularly where contaminants exhibiting potential neurotoxic or fetotoxic effects are present), and consumers of higher quantities of local foods (particularly in the presence of bioaccumulating substances).

Workers involved in excavation, construction, or remediation activities associated with maintenance, remediation, or redevelopment of a contaminated site may be exposed to higher levels of risk owing to more intimate contact with contaminated environmental media. However, such individuals may be exposed for a relatively short duration, and may also use protective clothing, equipment, and procedures that would serve to minimize exposure. Health and safety plans can include mitigation measures for construction workers if the risk assessment indicates this is required.

3.5.3 Receptor screening

It is becoming standard practice to assess exposure for all relevant age groups for identified receptor populations. Such information is particularly valuable in public consultations and communications. However, it is generally not necessary to conduct a complete detailed risk assessment for all identified receptors. In most cases, a critical receptor can be identified for each class of contaminants, each land use, and potentially for each exposure pathway. The critical receptor is normally the member of the applicable receptor group that is expected to receive the highest exposure to a chemical, expressed as average daily intake on a per unit body weight basis. For example, for threshold chemicals at a site where all age classes are present, the toddler would normally be considered the critical receptor. However, if a particular member of the population is more sensitive to a given level of exposure because of a specific vulnerability or life-stage dependent toxic response, this may influence the selection of critical receptors for specific chemicals. Selection of a critical receptor requires consideration of human exposure factors for different populations and age classes; these are discussed in section 4.5. As noted above, the selection of a critical receptor is also dependent on exposure pathway and is therefore closely related to the identification and screening of exposure pathways, which are discussed in the following text.

3.6 Exposure Pathway Identification

The objective of exposure pathway identification at the problem formulation stage is to identify and screen pathways of potential concern that involve contaminants and receptors identified in the previous tasks. For risk assessment purposes, an exposure pathway consists of a contaminant source, a release mechanism, a transport mechanism within the relevant environmental medium (or media), a point of exposure (receptor), and an exposure route. The exposure route refers to the route by which a chemical physically contacts and is absorbed into the human body.

A chemical that is not screened out (i.e. one that is retained as a COPC) is a contributor to risk only if a pathway to a receptor exists or is likely to exist, and if the pathway leads to exposures above a concentration or dose of concern. For an exposure pathway to exist, all components of the exposure pathway previously listed must be present; if one or more is absent or inactive, the exposure pathway is incomplete and therefore inoperative. Owing to the interrelationship of exposure pathways and receptors, it is important that the screening of exposure pathways be conducted in conjunction with the screening of receptors.

Identifying exposure pathways involves the systematic identification and documentation of sources, release mechanisms, transport mechanisms, receptors, and exposure routes. Exposure pathway identification and screening at the problem formulation stage involves a qualitative assessment of site conditions, and is not a quantitative assessment of chemical release, transport, and exposure. Quantitative assessment forms part of the subsequent exposure assessment stage.

3.6.1 Sources and release mechanisms

A contaminant source may be a leaking tank, vessel, or other container, or a volume of waste material whose contents are being released (or have been released) to the soil, groundwater, air, or other receiving medium. Where the contaminant has already been “released” to the environment, the release mechanism is the process of mobilization of the contaminant into a transport medium by volatilization, dissolution, leaching, erosion, bioaccumulation, bioconcentration, etc.

Table 3.2 lists some common contaminant sources, release mechanisms, and receiving or transport media for contaminated sites.

Table 3.2 Common Contaminant Sources, Release Mechanisms, and Receiving Media

Receiving Medium	Release Mechanism	Contaminant Source
Air	Volatilization	Surface wastes – lagoons, ponds, pits, spills Contaminated surface water or groundwater Contaminated surface soil Contaminated wetlands Leaking drums
	Fugitive dust generation	Contaminated surface soil Waste piles
Surface water	Surface runoff	Contaminated surface soil
	Episodic overland flow	Lagoon overflow Spills, leaking containers
	Groundwater seepage	Contaminated groundwater
Ground water	Leaching	Surface or buried wastes Contaminated soil
Soil	Leaching	Surface or buried wastes
	Surface runoff	Contaminated surface soil
	Episodic overland flow	Lagoon overflow Spills, leaking containers
	Fugitive dust generation/deposition	Contaminated surface soil Waste piles
	Tracking	Contaminated surface soil
Sediment	Surface runoff, episodic overland flow	Surface wastes – lagoons, ponds, pits, spills Contaminated surface soil
	Groundwater seepage	Contaminated groundwater
	Leaching	Surface or buried wastes Contaminated soil

Source: Adapted from U.S. EPA, 1989.

3.6.2 Environmental media and fate and transport mechanisms

The transport component of an exposure pathway involves the movement of a contaminant within or between environmental media from the receiving medium to the point of exposure. One or more of the receiving, transport, and exposure media may be the same; for example, air receiving a volatile chemical can move via wind then be inhaled. However, exposure pathways frequently involve the transfer of a contaminant from one medium to another. Environmental media include soil (and other geological materials such as bedrock and sediment), groundwater, surface water, air, and biota. Food may also be considered an environmental medium where food chain exposure pathways are applicable.

A number of mechanisms govern contaminant fate and transport within and between environmental media (U.S. EPA, 1989):

- physical transport with wind, water, etc. (e.g. advection, diffusion, particulate transport)
- physical transformation (e.g. volatilization, dissolution, precipitation)
- chemical transformation (e.g. photolysis, hydrolysis, oxidation, reduction)
- biological transformation (e.g. biodegradation)
- accumulation (e.g. bioaccumulation, bioconcentration, biomagnification)

The identification of fate and transport mechanisms and the relevant environmental media involves the consideration of physical and chemical properties of the COPCs, as well as soil and hydrogeological conditions. Relevant contaminant-specific physical, chemical, and environmental fate parameters that should be considered in evaluating contaminant transport are summarized in Table 3.3. Site-specific factors affecting contaminant fate and transport are summarized in Table 3.4. Available measurements of contaminant concentrations in transport and/or exposure media should also be used to assist in the identification of fate and transport processes and hence exposure pathways.

Table 3.3 Contaminant-Specific Physical and Chemical Factors Governing Contaminant Fate and Transport

Bioconcentration factor (BCF) provides a measure of the extent of chemical partitioning at equilibrium between a biological medium (e.g. fish tissue, plant tissue) and an external medium (e.g. water). The higher the BCF, the greater the accumulation in living tissue.

Diffusivity describes the movement of a molecule in a liquid or gas medium as a result of differences in concentration. It is used to calculate the dispersive component of chemical transport. The higher the diffusivity, the more likely a chemical is to move in response to concentration gradients.

Henry's Law constant provides a measure of the extent of chemical partitioning between air and water at equilibrium. The higher the Henry's Law constant, the more likely a chemical is to volatilize than to remain in water.

***K_{oc}* organic carbon–water partition coefficient** provides a measure of the extent of chemical partitioning between organic carbon and water at equilibrium. The higher the K_{oc}, the more likely a chemical is to bind to soil or sediment than to remain in water.

***K_d* soil/sediment–water partition coefficient** provides a soil or sediment-specific measure of the extent of the chemical partitioning between soil or sediment and water, unadjusted for dependence upon organic carbon. To adjust for the fraction of organic carbon present in soil or sediment (f_{oc}), use $K_d = K_{oc} \times f_{oc}$. The higher the K_d, the more likely a chemical is to bind to soil or sediment than to remain in water.

***K_{ow}* octanol–water partition coefficient** provides a measure of the extent of chemical partitioning between water and octanol at equilibrium. The greater the K_{ow}, the more likely a chemical is to partition to octanol than to remain in water. Octanol is used as a surrogate for lipids (fats), and K_{ow} can be used to predict bioconcentration in aquatic organisms.

Media-specific half-life provides a relative measure of the persistence of a chemical in a given medium. Actual values can vary greatly, depending on site-specific conditions. The greater the half-life, the greater the persistence of the chemical.

Solubility is an upper limit of a chemical's dissolved concentration in water at a specified temperature. Aqueous concentrations in excess of solubility may indicate sorption onto sediments, the presence of solubilizing chemicals such as solvents, or the presence of a non-aqueous phase liquid.

Vapour pressure is the pressure exerted by a chemical vapour in equilibrium with its solid or liquid form at any given temperature. It is used to calculate the rate of volatilization of a pure substance from a surface, or to estimate a Henry's Law constant for chemicals with low water solubility. The higher the vapour pressure, the more likely a chemical is to exist in a gaseous state.

Source: U.S. EPA, 1989.

Table 3.4 Site-Specific Factors Governing Contaminant Fate and Transport

Stratigraphy and soil conditions describe subsurface conditions in terms of soil type and texture, bedrock conditions, depositional features (e.g. layers), structural features (e.g. fractures or fissures), and other effects, such as weathering. Stratigraphic and soil conditions affect the physical movement of contaminants in the saturated and unsaturated (vadose) zones of the subsurface. Of particular significance are layers and structural features that provide a preferential path for the migration of fluids (groundwater, non-aqueous-phase liquids and vapours).

Porosity is related to soil type (texture) and density, and is a measure of the relative volume of pore spaces or voids in the soil. A higher porosity soil of a given texture (e.g. coarse-grained) is generally more transmissive to fluids than a lower porosity soil. However fine-grained soils may exhibit a higher porosity than coarse-grained soils because of their colloidal structure.

Air-filled porosity and **moisture content** describe the proportion of the pore spaces of an unsaturated soil that are filled with air or water. These properties affect air permeability, and hence vapour-phase contaminant transport in the unsaturated zone.

Hydraulic conductivity is a measure of the ability of a soil or bedrock unit to transmit groundwater. It is related to soil type (texture) and density or porosity, but is also affected by the presence of structural features.

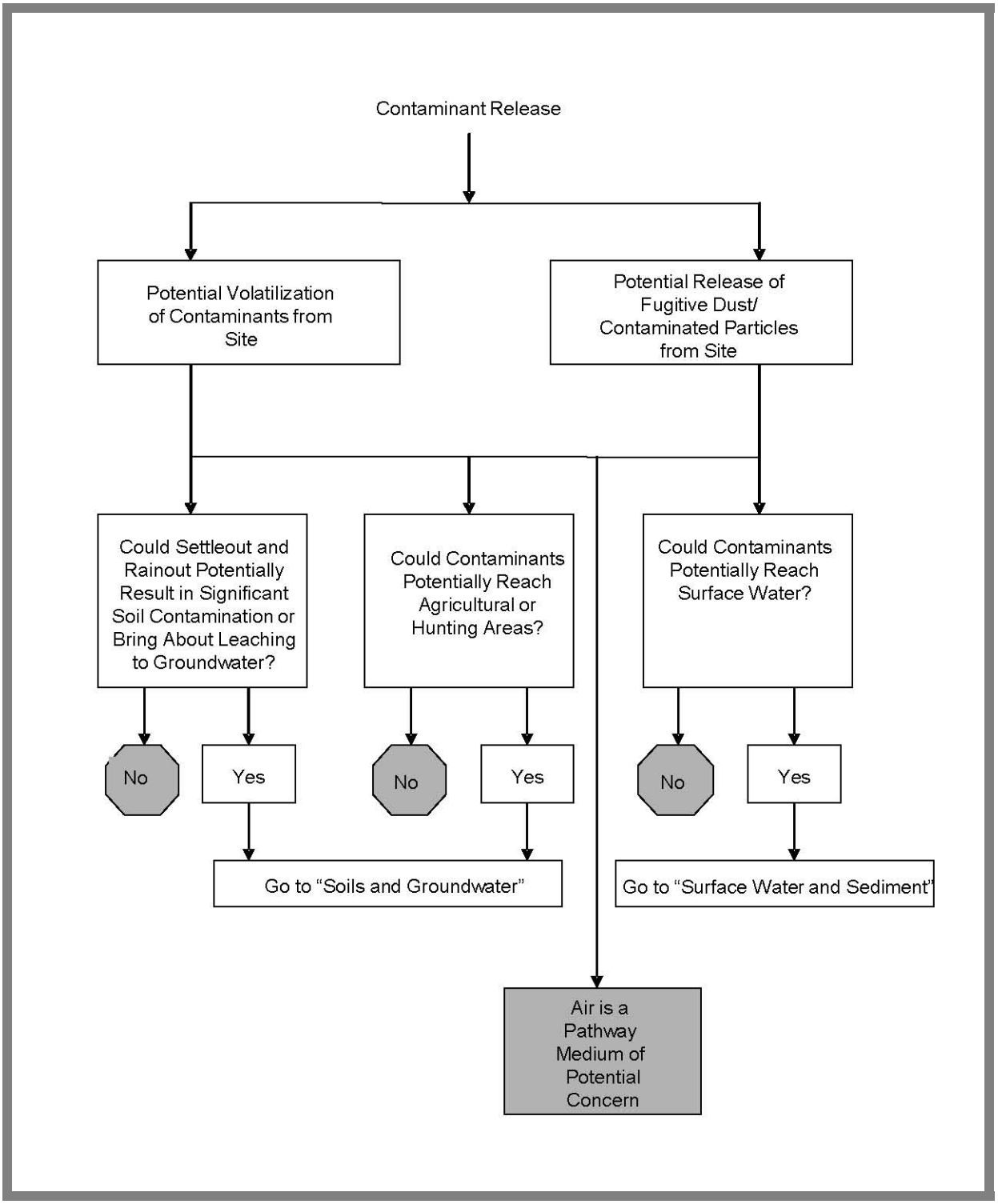
Fraction of organic carbon is the gravimetric ratio of organic carbon to mineral soil. Because organic chemicals are adsorbed by organic carbon in the soil in accordance with the chemical's organic carbon–water partition coefficient, the fraction of organic carbon affects partitioning of a contaminant among soil, groundwater, and vapour phases. A higher fraction of organic carbon also retards the migration of an organic chemical through the soil.

Hydraulic gradient, in combination with hydraulic conductivity, determines the flow rate of groundwater through a soil unit, and hence the rate of migration of chemicals dissolved in the groundwater.

Other factors that influence the fate and transport of contaminants in the subsurface include pH value, soil chemistry (e.g. presence of oxidizing/reducing conditions), microbial environment, temperature, and the presence of other chemicals.

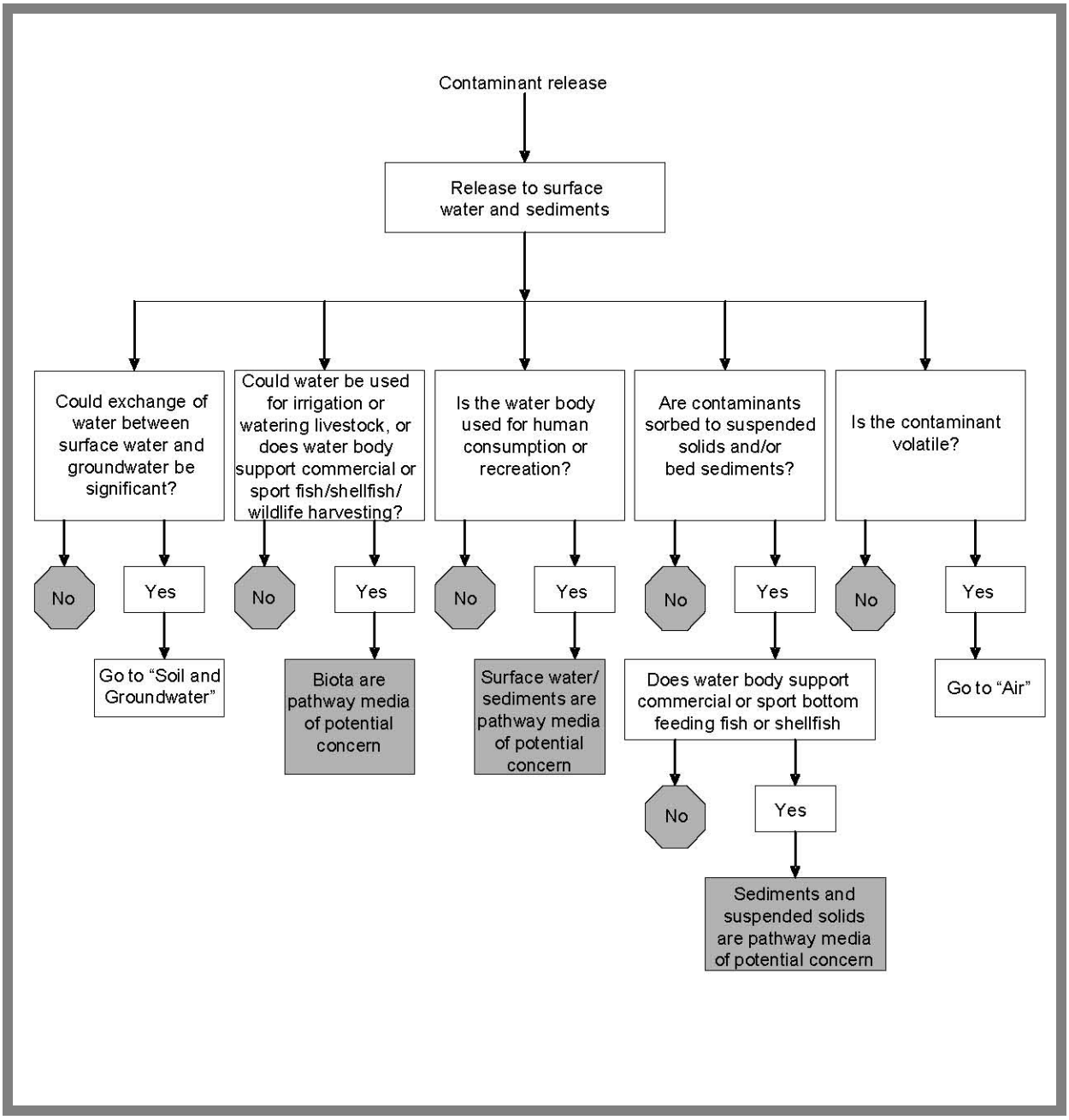
Figures 3.3, 3.4, and 3.5 present processes for assessing the potential for chemical transport in air, surface water/sediment, and soil/groundwater, respectively. These processes are integrated. For example, the release of a chemical into the air (Figure 3.3) could result in pathways to soil, groundwater, surface water, and sediment. Similarly, releases to surface water and sediment (Figure 3.4) could result in pathways to soil, groundwater, and air. Releases to soil (Figure 3.5) could lead to pathways involving surface water, sediment, air, and groundwater. A number of these pathways can also result in uptake and accumulation in biota, including those used as food.

Figure 3.3 Pathway Screening Process – Air



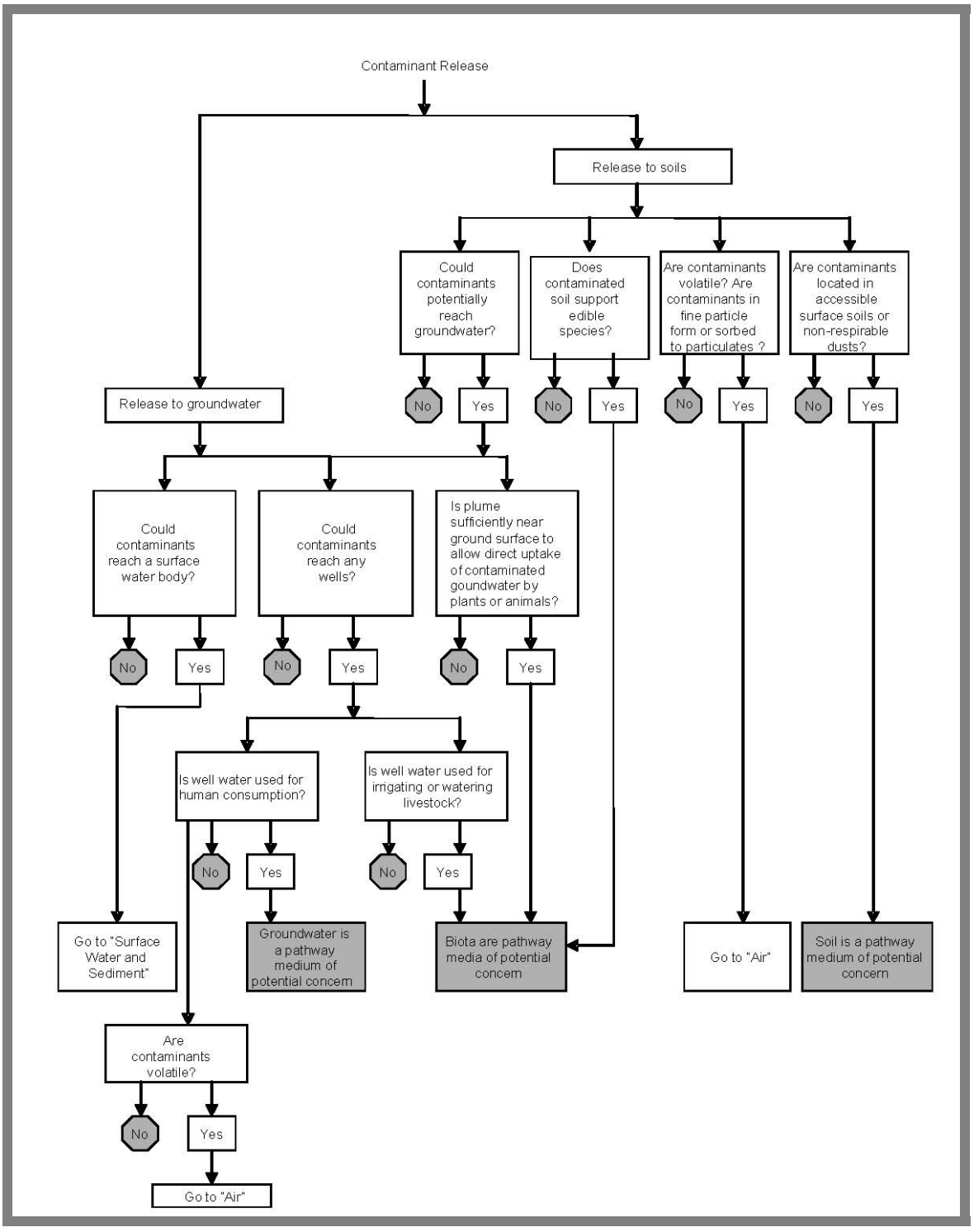
Source: Adapted from U.S. EPA, 1989.

Figure 3.4 Pathway Screening Process – Surface Water and Sediment



Source: Adapted from U.S. EPA, 1989.

Figure 3.5 Pathway Screening Process – Soil and Groundwater



Source: Adapted from U.S. EPA, 1989.

3.6.3 Exposure routes

Exposure routes are the mechanisms of intake of a chemical to the human body. These include the following:

- ingestion (soil, water, produce, fish and shellfish, game)
- inhalation (vapour, suspended particulate matter)
- dermal absorption (soil, water, air)

Identification of potential exposure routes involves consideration of the exposure media and the behavioural characteristics of the identified receptors at the point(s) of exposure.

3.6.4 Exposure pathway screening

Upon initial identification of all potential exposure pathways, some may be excluded immediately on the basis of the absence of one or more of the required components of an exposure pathway, described in the preceding text. Following this initial screening, all complete or operative pathways should be listed together with the applicable receptor groups. Not all operative pathway/receptor combinations will be necessarily evaluated in detail in the subsequent risk assessment. Further screening may be conducted to exclude pathways for which the probability of exposure is very low, or the potential magnitude of exposure is negligible or much lower than another pathway involving the same medium and exposure point. Pathways may also be excluded on the basis of monitoring data showing that the pathway is not currently active, or on the basis of the implementation of risk measures that effectively prevent exposure; these situations may change with time, and may require ongoing management or monitoring. Certain receptors may be excluded from further analysis by focusing the assessment of a given pathway on a critical or sensitive receptor. Sound justification should be provided for the exclusion of any complete exposure pathway and receptor from further consideration in the risk assessment.

Some exposure pathways may not be subjected to a full quantitative exposure assessment for other reasons. For example, a lack of data on which to base estimates of contaminant release and transport, uncertainties in the modelling methods, and the absence of monitoring data to validate exposure point estimates may preclude quantitative assessment if the resulting uncertainties are sufficiently great. In these situations, the exposure pathways would not be excluded at this stage, but would be carried forward in order to permit a qualitative assessment of the associated risks, a detailed uncertainty analysis, or to support the need to collect further site data to fill the data gap.

3.7 Conceptual Site Model Development

A key output of the problem formulation stage of a risk assessment is the **conceptual site model**. The CSM provides a complete description, usually in schematic form, of the COPCs, their sources and release mechanisms, transport pathways, and exposure routes to identified critical receptors. Health Canada PQRA guidance provides a tabular problem formulation checklist that is intended to summarize critical land uses, receptor groups, and exposure pathways within PQRAs. However, an illustrated CSM facilitates a clearer common understanding, by the risk assessor and other stakeholders, of the issues associated with the site. The CSM, which is qualitative in nature, provides the basis and guidance for the subsequent quantitative risk assessment. It also serves to focus attention on the critical aspects of the problem, and can be used to guide further data collection, as well as consultation and risk communication. Figures 3.6, 3.7, 3.8, and 3.9 present examples of CSMs in different formats.

Figure 3.6 Example of Conceptual Site Model in Checklist Format

Land Uses Check (✓) as appropriate		Receptor groups(s) Check (✓) as appropriate		Critical Receptors Check (✓) as appropriate		Exposure Pathways Check (✓) as appropriate	
	Agriculture		General Public		Infant		Soil ingestion
	Residential/urban parkland		Employees		Toddler		Soil dermal absorption
	Commercial with daycare		Construction workers		Child		Particulate inhalation
	Commercial without daycare		Canadian native communities		Teen		Vapour inhalation
	Industrial		Other (specify)		Adult		Groundwater ingestion
	Other (specify				Other (specify)		Water dermal absorption
							Produce ingestion
							Fish ingestion
							Wild game ingestion
							Other (specify)

Source: Adapted from Health Canada, 2010.

Figure 3.7 Example of Conceptual Site Model in Tabular Format

Primary source	Secondary source	Hazard	Transport mechanism	Pathway	Medium of exposure	Receptor
Fuel tank	Contaminated soils	Dizziness, Central nervous system depression	Vapour transport through unsaturated zone	Inhalation of vapours	Air	Humans (recreational users)
Fuel tank	Contaminated soils	Skin irritation, contact dermatitis in extreme	Direct contact with contaminated soil	Dermal contact at surface	Soil	Humans (recreational users)
Fuel tank	Contaminated soils	Flammability	Vapour transport through unsaturated zone	Vapour buildup in basement void	Air	Humans (residential)
Fuel tank	Contaminated soils	Central nervous system depression, asphyxiation	Vapour transport through unsaturated zone	Vapour buildup in basement void	Air	Humans (residential)
Spills from customer activity	Contaminated soils	Dizziness, Central nervous system depression	Vapour transport through unsaturated zone	Inhalation of vapours	Air	Humans (forecourt users)

Source: Adapted from DEFRA, 2000.

Figure 3.8 Example of Conceptual Site Model in Flow Chart Format

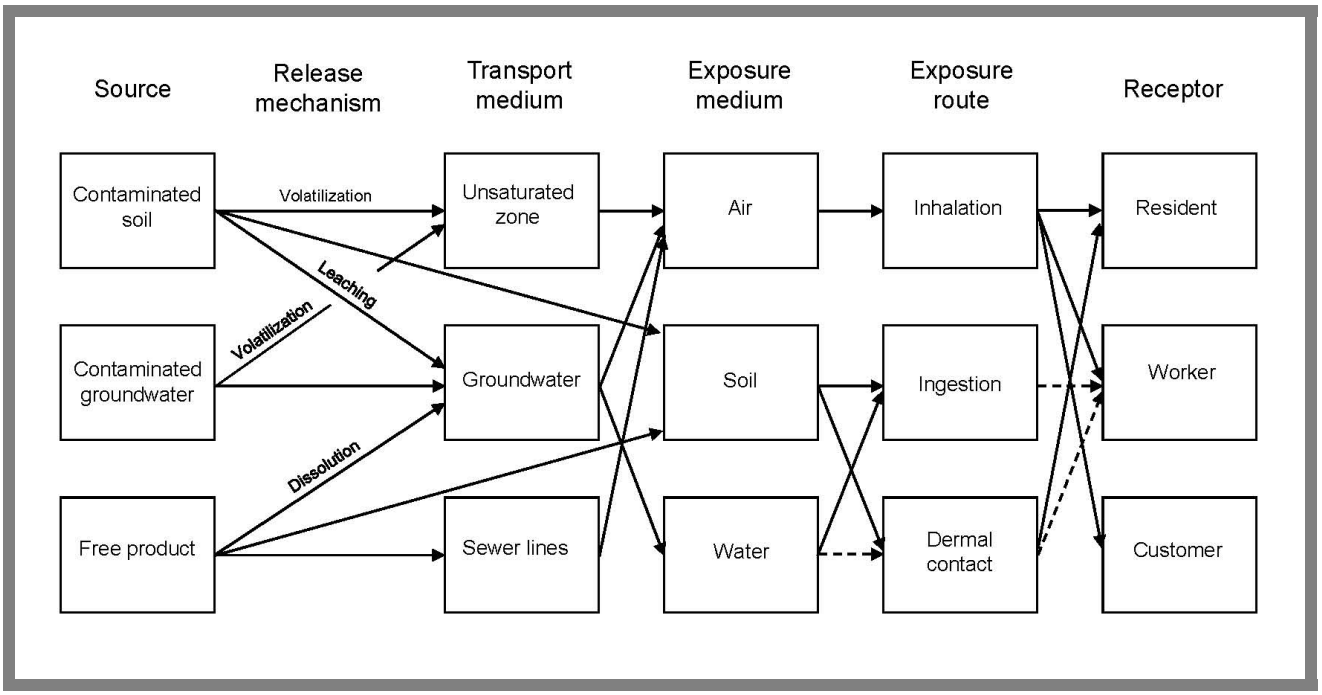
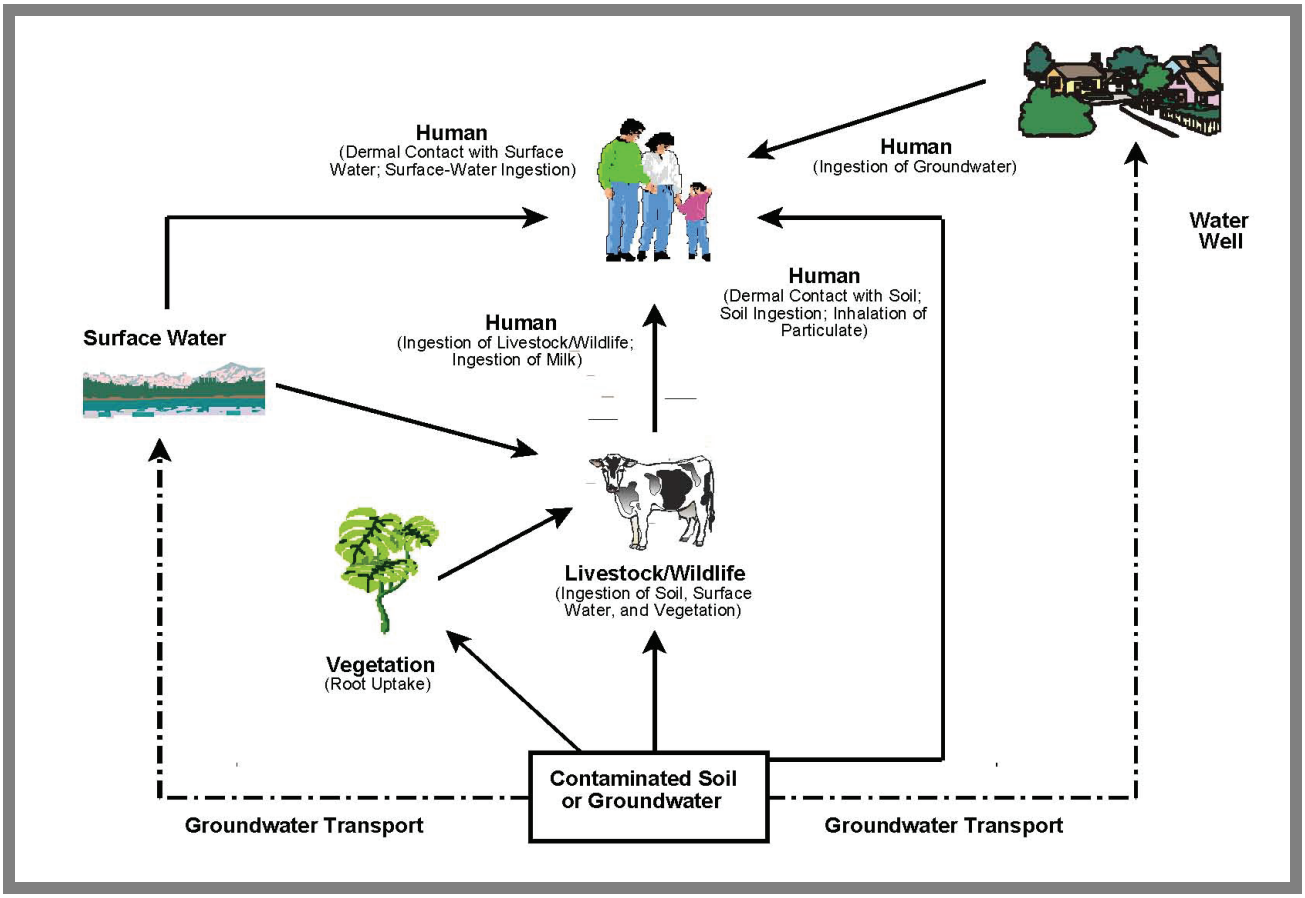


Figure 3.9 Example of Conceptual Site Model in Pictorial Format



3.8 Risk Assessment Decision

As noted previously, one of the purposes of the problem formulation stage is to determine the scope and level of detail of the ensuing risk assessment. Although the objectives of the risk assessment are normally established at the outset, the process required to meet those objectives may not be apparent until completion of the problem formulation and the development of the CSM. For example, the CSM may indicate that no complete and operative exposure pathways exist for the identified COPCs, or that the likelihood of exposure is minimal. In this situation, more detailed evaluation may not be required, other than a discussion of uncertainties. At the other extreme, the contaminant screening may indicate that contaminant source concentrations are sufficiently elevated with respect to screening criteria so that the completion of a more detailed and realistic risk assessment using site-specific data would not result in a significantly different assessment of risk. In this case, a more detailed risk assessment may focus on the risks associated with different remedial or risk management alternatives.

If the contaminant concentrations are sufficiently close to screening criteria so that no clear conclusions can be drawn with respect to acceptability or unacceptability of the potential health risks, further risk assessment would likely be required, particularly where exposure pathways involve contaminant transport by mechanisms that have not been implicitly considered or incorporated into the screening criteria. In this case, increased realism through the use of site-specific data likely would provide a more confident and defensible estimate of risk. The risk assessor may choose an iterative approach, based on the use of progressively more realistic assumptions and/or the collection of additional site-specific data.

3.9 Recommended Deliverables

At the conclusion of the problem formulation stage, an interim technical report should normally be prepared. This report may ultimately form a section or chapter of the detailed risk assessment report, but in its interim form it should serve as a stand-alone report documenting the methods, rationale, and results of the screening tasks, and presenting the CSM. It is noted that the report describing problem formulation often provides guidance and focus not only to the risk assessment itself but also to the entire site investigation, remediation, and risk management process. It is also a key tool in

communicating with stakeholders, affected communities, and the public as a whole.

The report should summarize site conditions, including the history, physical conditions, contaminant information, land use, and exposure scenarios. It should also describe contaminants of concern, receptors, and exposure pathways. Results of the contaminant, receptor, and pathway screening should be fully documented, as described in the relevant preceding sections. A complete rationale for the contaminants and receptor/pathway combinations ultimately selected for (and those excluded from) detailed analysis should be included.

The report should clearly state the objectives of the risk assessment in the context of the proponent's goals, regulatory and stakeholder requirements, as well as project and technical constraints and limitations. Because the problem formulation stage is, among other things, a scoping exercise for the detailed risk assessment, the approach to be taken in the subsequent stages of the assessment should be discussed and substantiated. Data limitations and requirements for further investigation should also be identified.

In many cases, a critical review stage occurs following problem formulation, where regulatory agencies, communities, and other stakeholders may be asked to provide input prior to the risk assessment process. The problem formulation report should be prepared with this in mind. It should be clear and concise, use non-technical vocabulary where possible, and yet be detailed enough to support the CSM developed and the proposed risk assessment methodology.

3.10 References

Bright, D.A., G.M. Richardson, and M. Dodd. 2006. Do current standards of practice measure what is relevant to human exposure at contaminated sites? I: A discussion of soil particle size and contaminant partitioning in soil. *Human Ecol. Risk Assess*. 12(3): 591–605.

Burmaster, D.E. 1996. Benefits and costs of using probabilistic techniques in human health risk assessments – with an emphasis on site-specific risk assessments. *Human Ecol. Risk Assess*. 2(1): 35–43.

Canadian Council of Ministers of the Environment (CCME). 1999. *Canadian Environmental Quality Guidelines*. CCME, Winnipeg. Updated online. Accessed December 10, 2009, from http://ceqg-rcqe.ccme.ca/.

CCME. 2006. *A Protocol for the Derivation of Environmental and Human Health Soil Quality Guidelines [Revised].* CCME, Winnipeg. Report CCME PN 1332. ISBN 13-978-1-896997-45-2.

CCME. 2008a. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil*. CCME, Winnipeg.

CCME. 2008b. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil: Scientific Rationale.* Supporting Technical Document. CCME, Winnipeg.

CCME. 2008c. *Canada-Wide Standards for Petroleum Hydrocarbons (PHCs) in Soil: Technical Supplement*. CCME, Winnipeg.

Cullen, A.C., and H.C. Frey. 1998. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Society for Risk Analysis. Plenum Press, New York.

DEFRA (Department for Environment Food and Rural Affairs). 2000. *Guidelines for Environmental Risk Assessment and Management.* Accessed online September 4, 2010: http://www.defra.gov.uk/environment/quality/risk/eramguide/

Duggan, M.J., and M.J. Inskip. 1985. Childhood exposure to lead in surface dust and soil: A community health problem. *Public Health Rev.* 13(1-2): 1–54.

Duggan, M.J., M.J. Inskip, S.A. Rundle, and J.S. Moorcroft. 1986. Lead in playground dust and on the hands of schoolchildren. *Sci. Tot. Environ*. 44(1): 65–79.

Edwards, R.D., and P.J. Lioy. 1999. The EL sampler: A press sampler for the quantitative estimation of dermal exposure to pesticides in housedust. *J. Exp. Anal. Environ. Epidemiol*. 9: 521–529.

Globaltox. 2005. *Does Size Matter? The Role of Particle Size in Exposure Estimates for Risk Assessment. A Review of Associated Factors Including Respiratory Tract Penetration and Skin Adherence*. Contractor report to Environmental Health Assessment Services, Health Canada, Ottawa.

Health Canada (HC). unpublished. Final Draft Guidance Manual for Environmental Site Characterization in Support of Human Health Risk Assessment. Contaminated Sites Division, Safe Environments Directorate, Ottawa. Available upon request by contacting the Contaminated Sites Division at cs-sc@hc-sc.gc.ca.

HC. 2010 *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA)*, *Version 2.0*. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

Kissel, J.C., K.Y. Richter, and R.A. Fenske. 1996. Factors affecting soil adherence to skin in hand-press trials. *Bull. Environ. Contam. Toxicol*. 56: 722–728.

Ontario Ministry of Environment and Energy (OMEE). 1996. *Rationale for the Development and Application of Generic Soil, Groundwater and Sediment Criteria for Use at Contaminated Sites in Ontario*.

Richardson. G.M. 1996. Deterministic versus probabilistic risk assessment: Strengths and weaknesses in a regulatory context. *Human Ecol. Risk Assess*. 2(1): 44–54.

Richardson, G.M., D.A. Bright, and M. Dodd. 2006. Do current standards of practice measure what is relevant to human exposure at contaminated sites? II: Oral bioaccessibility of contaminants in soil. *Human Ecol. Risk Assess*. 12(3): 606–616.

Sheppard, S.C., and W.G. Evenden. 1994. Contaminant enrichment and properties of soil adhering to skin. *J. Environ. Qual*. 23: 604–613.

United States Environmental Protection Agency (U.S. EPA). 1989. *Risk Assessment Guidance for Superfund (RAGS): Volume I – Human Health Evaluation Manual (Part A)*. Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002.

U.S. EPA. 1996. *Soil Screening Guidance, Technical Background Document*. U.S. EPA, Washington, DC. EPA/540/R-95/128.

U.S. EPA. 2000. *Short Sheet: TRW Recommendations for Sampling and Analysis of Soil at Lead (Pb) Sites*. Office of Solid Waste and Emergency Response, Washington, DC. EPA/540-F-00-010.

U.S. EPA, 2001. *Risk Assessment Guidance for Superfund (RAGS): Volume III – Part A, Process for Conducting Probabilistic Risk Assessment*. Office of Emergency and Remedial Response, Washington, DC. EPA 540-R-02-002.

U.S. EPA. 2002. Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites. Office of Solid Wasate and Emergency Response, City. OSWER 9355.4-24.

Water & Earth Science Associates Inc. (WESA). 2005. *Incorporating Background Soil Quality Into Risk Assessment*. Contract report to Environmental Health Assessment Services, Health Canada, Ottawa.

Zamuda, C. 1989. Superfund risk assessments: The process and past experience at uncontrolled hazardous waste sites. In: *The Risk Assessment of Environmental Hazards*. D.J. Paustenbach (ed.). J Wiley & Sons, New York.

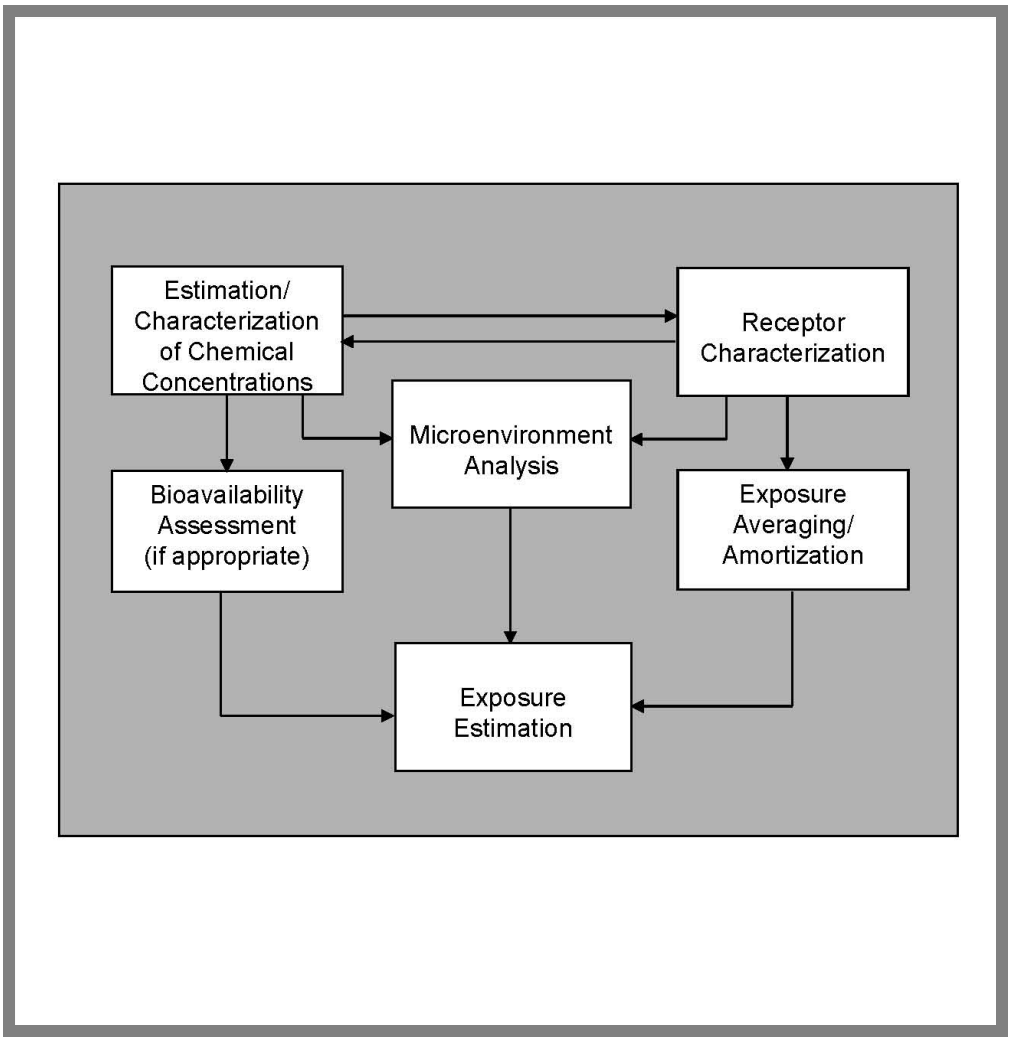
4.0 EXPOSURE ASSESSMENT

All possible receptor groups, exposure scenarios, ameliorating or influencing factors, and possible exposure pathways cannot be predicted for every possible contaminated site. Therefore, the guidance in this section is necessarily generic in nature. As site-specific scenarios and exposure models evolve, it is essential that all variables and factors be clearly and transparently explained and rationalized.

4.1 Introduction and Linkages to Other Risk Assessment Tasks

In the context of HHRA, the exposure assessment stage involves the estimation of the quantity of each chemical received by human receptors per unit time (e.g. daily intake or dose). The rates of exposure to chemicals from the various environmental media are usually expressed in units of mass of chemical intake per unit of body mass per unit of time (e.g. mg/kg bw/d); sometimes exposure is also evaluated using exposure concentrations (e.g. $\mu g/m^3$ in air), particularly if concentration-based TRVs are used to express toxicity. For chemicals eliciting toxic response at the site of bodily contact (locally acting chemicals: e.g. irritants), the exposure is often expressed as a concentration, an exposure duration, and (if applicable) an exposure frequency. The basic components of exposure assessment are illustrated in Figure 4.1. Exposure assessment is conducted for all chemicals, human receptors, and exposure pathways identified as being of concern during the problem formulation stage of the DQRA.

Figure 4.1 Components of Exposure Assessment



The exposure assessment is normally conducted concurrently with the toxicity assessment (section 5.0), and both must be completed before risk characterization (section 6.0) can be undertaken. The exposure assessment uses the CSM developed during problem formulation (section 3.0), including exposure pathways, receptors, and COPCs.

The exposure assessment and toxicity assessment affect each other, and therefore cannot be done in isolation. The toxicological characteristics of COPCs affect the exposure periods that must be considered; for example, a chemical with a potent toxic effect following short-term exposure (e.g. a caustic agent) may lead to a requirement for the estimation of short-term exposure doses (over minutes or hours). Likewise, the risks of chronic, relatively low-dose exposure cannot be adequately characterized from toxicological data (or TRVs) that address only short-term high-dose exposure. Also, the toxicity assessment may identify particularly sensitive receptors that must be considered in the exposure assessment.

4.2 Objectives of Exposure Assessment

The exposure assessment is conducted in order to estimate the amount of each COPC that comes into contact with and/or is taken into the body of the affected human receptors over a specified time period. The exposure rates (dose rates), when combined with the toxic potencies of the chemicals determined during the toxicity assessment, allow the estimation of risk from exposure to chemical contaminants at a site.

The exposure assessment should critically evaluate all potential exposure routes through which contact with a chemical could occur. Concentrations in environmental media must be determined either by measurement or modelling, and the duration and frequency of receptor exposure to the respective media must be described mathematically. Absorption (bioavailability) of the chemicals into the body following entry or contact may also require quantification for inclusion as a variable in exposure equations. Bioavailability is discussed in section 4.7.

4.3 Steps in Exposure Assessment

The main steps in exposure assessment, for the purposes of this document, follow.

Characterization/estimation of chemical concentrations: The concentrations of all COPCs must be determined in each environmental medium that receptors may be exposed to. This is accomplished through direct measurement of chemical concentrations and/or environmental modelling.

Receptor characterization: Physical and behavioural characteristics of human receptors that influence the frequency, duration and intensity of exposure (e.g. body weight; time–activity patterns; food, air, and water intake rates; dermal contact rates with environmental media.) must be quantified.

Exposure averaging and amortization: The amount of time for which a receptor is exposed (frequency of exposure events combined with duration of each event) over the total possible exposure period is determined so that the exposure dose can be adjusted, where appropriate, to reflect an average exposure rate.

Bioavailability assessment: If the TRV determined during the toxicity assessment is based on an absorbed dose or on a different exposure route, then the amount of the substance absorbed by the body for each relevant exposure route is determined.

Microenvironment analysis: If certain parts of the site are used more than others, or if there are distinctly different chemical concentrations in different areas of the site, then the site may be broken down into microenvironments in order to more realistically estimate exposure.

Exposure estimation: The exposure dose is calculated for each receptor, chemical, and exposure route.

Each of these steps is further detailed in the following sections. The steps do not necessarily have to be performed in order, aside from exposure estimation, which uses the results of the other steps and is therefore completed last.

4.4 Characterization and Estimation of Chemical Concentrations

Adequate characterization and/or estimation of chemical concentrations in various environmental exposure media is a critical component of the exposure assessment. For each COPC, an estimated concentration is required for each environmental medium of interest at the site (e.g. soil, groundwater, surface water, sediment, indoor air, outdoor air, dust/particulate, food). In many cases, available data from previous site investigations is useful for the determination of chemical concentrations, although a detailed evaluation of these data is necessary and further investigation is frequently needed to meet the data needs of a DQRA.

Chemical concentrations can be estimated using two general approaches:

1. direct measurement (i.e. sampling and chemical analysis of environmental media at the site)
2. environmental modelling (i.e. using mathematical models to predict chemical concentrations in exposure media)

A combination of the two approaches is often used for DQRAs. Considerations for deciding whether to use measured or modelled concentrations include:

- accessibility of media for sampling
- necessity of predicting future concentrations
- cost of adequate sampling versus modelling
- reliability and appropriateness of available models
- technical difficulties with sampling and/or analysis (e.g. sampling highly volatile chemicals or determining metal speciation)

Direct measurement generally provides the most realistic estimate of current chemical concentrations in the environment. If measured concentrations are used, the quantity of data required to adequately characterize a given medium for a specific site must be determined. The level of sampling effort required will be governed by the goals and scope of the risk assessment. The size of the site and the distribution of contamination (spatially, and with depth) will be a major determinant of required sampling effort. Spatial variability in chemical concentrations also affects the number of samples that are required during a sampling event. Consideration also needs to be given to the sampling strategy (e.g. sampling locations chosen to identify and delineate contamination sources or plumes, random sampling, or systematic grid sampling). Samples can be either discrete point samples or composites from several sampling locations. Although composites may allow for the estimation of average concentrations with fewer laboratory analyses than discrete samples, they cannot identify maximum chemical concentrations or spatial distributions. Also, mixing samples may result in the loss of volatile compounds. More detailed technical considerations for environmental sampling and analysis have been published by various agencies, including CCME (1993), ASTM (2002), and many provincial environment ministries. The U.S. EPA has also published numerous documents on various aspects of environmental sampling and analysis (e.g. U.S. EPA, 1991, 1997, 1999). In addition, useful information on environmental sampling and analysis has been published by various scientific authors (e.g. Gilbert, 1987; Bodger, 2003).

Environmental modelling is particularly useful in situations where certain environmental media are not accessible or there are significant technical difficulties associated with obtaining reliable measured concentrations. Modelling approaches are also very useful if future concentrations need to be determined (e.g. future chemical concentrations at a location that has not yet been reached by a mobile contaminant plume, or concentrations after a period of natural attenuation). Models range in complexity from "screening-level" models with simple equations and few input parameters to advanced complex numerical simulations requiring many input parameters and detailed three-dimensional site characterization. Models chosen must be scientifically defensible and appropriate for the scenario being modelled, and should be available for review by the appropriate regulatory agency. The complexity of the model should be consistent with the needs and scope of the risk assessment and the level of site detail available. Criteria for selecting an appropriate model and examples of commonly used models are outside the scope of this document.

Regardless of whether direct measurement or environmental modelling is used, both spatial and temporal variability need to be characterized. Spatial definition of the site is particularly important for the application of any microenvironment analysis. Temporal definition of the site is needed to address changes in chemical concentrations over time (e.g. seasonal variations, effects of natural attenuation, future concentrations in areas not yet reached by a contaminant plume) so that long-term and future risks at the site can be characterized, if required.

Future concentrations of chemicals at the site may be affected by environmental fate processes. The concentrations of organic chemicals at a site will often decrease over time if more of the chemical is not being added owing to processes such as dispersion, photodegradation, volatilization, and biotransformation. If concentrations decrease over time, future exposure of human receptors should be reduced based on degradation half-lives; this can have a significant effect on chronic exposure to chemicals that degrade rapidly. It should be noted that degradation rates for a given chemical can vary considerably depending on site conditions, so evaluation of degradation should be undertaken with considerable caution and an appropriate level of conservatism. Additionally, degradation products of the original chemicals present at the site may also be toxic, in some cases more toxic than the original chemicals; as a result, the estimation of the future concentrations of degradation products may also be essential for the risk assessment.

Potential environmental transport of chemicals should also be considered. This is particularly the case for groundwater, where transport may occur over a period of years; for example, a water well located some distance away from the contamination may initially have no measurable levels of a chemical, but several years later the chemical may be detected. Future risks cannot be assessed without first evaluating the environmental transport; this is normally accomplished with the use of mathematical fate and transport models.

For deterministic exposure assessments, chemical concentrations are represented by point estimates. These point estimates may be based on the arithmetic mean, upper 95% confidence interval of the mean, 95th percentile of the data distribution, or some other statistic depending on the quality and quantity of data available. Adequate data permitting, Health Canada prefers use of the mean or upper 95% confidence interval of the mean. However, for PQRAs where data are more limited, the 95th percentile of the data distribution or the maximum measured concentration will more likely be employed. For probabilistic assessments, chemical

concentrations will be represented by the full distribution of all measured values at the site (see section 7.0).

4.4.1 Concentrations in soil

Concentrations of COPCs in soil are relevant for sites where direct human exposure to soil is possible (generally any site with uncovered soil). Concentrations in soil are also frequently used as a basis for estimating concentrations in several other media, including groundwater (for soluble compounds) and indoor air (for volatiles), when environmental modelling is used.

Chemical concentrations in soil are most often determined by direct measurement. Direct exposure to soil (ingestion, dermal contact, inhalation of suspended soil particles) is predominantly a function of surface soil contamination. For these pathways, a reasonably representative sampling of surface soil is required. Consideration also needs to be given to selecting appropriate sampling depths, based on the location, and vertical distribution of contamination and potential indirect exposure pathways being evaluated in the risk assessment (e.g. groundwater transport, volatilization, and migration to enclosed spaces). In some cases, field-screening methods are used to identify worst-case locations prior to selecting samples for laboratory analyses.

Background soil concentrations are generally required for DQRA, particularly when dealing with naturally occurring substances. In some cases, background concentrations can be determined on a site-specific basis by sampling soils from nearby locations that have not been affected by contamination. If this approach is used, care should be taken to select reference sites that have a similar geography and geology to the contaminated site. Also, reference soils should be of a similar type, grain size, and origin when compared with the contaminated soils. Alternatively, values reported in the literature or data and values available from provincial agencies or federal agencies (GSC, 2006) may be appropriate.

Modelling may be used to determine soil concentrations as well, particularly if future chemical concentrations in soil are being estimated, or if concentrations need to be estimated for an inaccessible area (e.g. off-site properties if access is not granted, or beneath structures). Scenarios modelled may include hypothetical chemical spills, deposition of airborne contaminants on soil, or transport of soil by surface runoff or wind erosion. Environmental modelling may also be needed to account for environmental fate processes, such as degradation, when considering future exposure, particularly to toxic degradation products (such as vinyl chloride associated with trichloroethylene [TCE] and perchloroethylene [PCE] degradation).

Estimated soil concentrations of individual chemicals are designated as C_s in equations presented in this document.

4.4.2 Concentrations in groundwater

Concentrations of chemicals in groundwater are relevant for sites where groundwater may be ingested (e.g. as a drinking water source) or, in some situations, used for showering where subsequent inhalation of vapours and aerosols might occur (in the case of TCE and PCE, for example). Groundwater concentrations can also be used to estimate concentrations in other media, such as modelling the indoor infiltration of volatile organic contaminants or concentrations in food sources when groundwater is used for livestock watering or irrigation.

Chemical concentrations in groundwater can be directly measured, typically using groundwater- monitoring wells that are screened across an appropriate depth interval. Because concentrations can fluctuate seasonally as well as change over time owing to environmental fate processes, multiple samples through direct measurement are often needed over an extended time period to adequately characterize chemical concentrations in groundwater.

Concentrations in groundwater can also be estimated from soil concentrations using partitioning relationships. Environmental fate and transport models are frequently used to estimate future groundwater concentrations or to predict the movement of groundwater contamination.

Background groundwater concentrations can often be measured in the same aquifer by collecting samples hydraulically upgradient of a contaminated area, provided that there are no additional sources of contamination potentially affecting the upgradient groundwater. Alternatively, literature values or background concentrations published by jurisdictions may be appropriate for comparison and screening purposes.

Estimated groundwater concentrations of individual chemicals are designated as C_{gw} in this document.

4.4.3 Concentrations in soil gas

In most cases, human receptors are not normally exposed directly to soil gas. However, soil gas concentrations of volatile contaminants (e.g. benzene, toluene, ethylbenzene, xylenes, TCE, PCE, vinyl chloride) are often used to estimate potential chemical concentrations in indoor air. The use of soil gas concentrations has the advantage that, unlike indoor air concentrations, the source of the chemical can generally be attributed to soil or groundwater contamination and not other potential sources, such as consumer products stored inside the building. Humans may be exposed directly to soil gas when working in excavations.

Obtaining accurate soil gas concentrations can be difficult. Generally, specially designed soil gas wells are preferred for collection of the samples, and sampling rates must be

appropriate for the soil type to ensure that there is no short-circuiting that can lead to dilution of the sample by ambient outdoor air. Guidance on the installation of soil gas monitoring wells is provided by Health Canada (see *Manual for Environmental Site Characterization in Support of Human Health Risk Assessment* (HC, unpublished). Soil gas concentrations may be higher beneath buildings or impermeable surface features than beneath bare ground. Chemical concentrations in soil gas can vary over time both owing to seasonal fluctuations and the time required for a vapour plume to reach a particular point (which can be several years). Therefore, in most cases, multiple samples collected over an extended period of time, and sometimes at multiple depths, are needed to accurately characterize soil gas concentrations. In addition to Health Canada's *Manual for Environmental Site Characterization in Support of Human Health Risk Assessment*, protocols and guidelines are available for developing soil gas sampling programs (e.g. ASTM, 1992; API, 1998; U.S. EPA, 1997; API, 2005).

Chemical concentrations in soil gas can also be estimated using theoretical partitioning relationships (see *Federal Contaminated Site Risk Assessment in Canada, Part VII: Guidance for Soil Vapour Intrusion Assessment at Contaminated Sites* [HC, 2010a]), although these approaches generally assume equilibrium conditions. Non-equilibrium methods also exist for the estimation of soil gas concentrations (e.g. fugacity calculations); however, these methods may not have the same level of regulatory acceptance in many jurisdictions.

Estimated soil gas concentrations of individual chemicals are designated as C_{sg} in this document.

4.4.4 Concentrations in indoor air

Indoor air concentrations are relevant at any site affected by volatile chemicals and having (or may have) occupied buildings on site or nearby. Chemical concentrations in indoor air can be measured directly using a variety of methods such as evacuated canisters, air sampling pumps, and exposure badges. Concentrations are generally measured over a discrete time interval (the averaging time), typically ranging from a few minutes to a day or more, that must be reported. The method used depends on factors such as the chemical being measured and the required analytical detection limit. An important consideration for indoor air sampling is the potential contribution of multiple background sources to indoor air concentrations; in particular, many consumer products and cigarette smoke can contribute significant concentrations of chemicals to indoor air, thereby potentially confounding source attribution for chemicals of interest in a contaminated site risk assessment. As a result, it is important to identify potential sources of the COPCs within the building, and interpret results within the context of typical background concentrations measured in similar buildings away from the contaminated site. It should also be noted that indoor air concentrations of chemicals vary over time, and may demonstrate diurnal or seasonal fluctuations owing to changes in building use, climatic factors, and air circulation.

Modelling techniques are frequently used to estimate indoor air concentrations, either to identify the contribution of subsurface contamination to indoor air concentrations or if the building is not accessible for sampling. In particular, models relating indoor air concentrations to soil gas concentrations for volatile chemicals have widespread use; models for the volatilization of chemicals from water being used domestically (e.g. showering, washing) may also be appropriate if contaminated water is used indoors. Models are also sometimes used to estimate indoor air concentrations from outdoor air concentrations. Guidance and methods on the modelling (estimation) of indoor air concentrations from groundwater, soil, and soil gas measurements are presented in *Federal Contaminated Site Risk Assessment in Canada, Part VII: Guidance for Soil Vapour Intrusion Assessment at Contaminated Sites* (HC, 2010a).

Estimated indoor air concentrations of individual chemicals are designated as C_{ia} this document.

4.4.5 Concentrations in outdoor air

Chemical concentrations in outdoor air are relevant at virtually all sites where human exposure is possible. Outdoor air concentrations can be measured using similar methods to those used for indoor air concentrations. Averaging time is particularly important for outdoor air concentrations because chemical concentrations can change rapidly with changing meteorological conditions. Concentrations can vary by orders of magnitude over both relatively short distances and relatively short times, making it difficult to obtain representative samples without a well-planned sampling program. Much like for indoor air, at times it can be difficult to determine the source of a chemical measured in outdoor air.

Outdoor air concentrations can also be estimated using modelling methods that estimate the emission rate and evaluate atmospheric transport. Emission rates can be estimated either by direct measurement (e.g. mercury vapour: Rasmussen et al., 1998; Richardson et al., 2003), or calculated using emission factors (e.g. U.S. EPA, 1995). A wide variety of mathematical models are available for evaluating the atmospheric transport of chemicals; the appropriate model depends on the chemicals being evaluated, release conditions, terrain, availability of meteorological data, and regulatory requirements. Models are also available for predicting outdoor air concentrations of chemicals originating in soil or soil gas. However, in many cases, outdoor air concentrations resulting from subsoil contamination are not evaluated because dispersion and dilution is rapid, and indoor air concentrations are generally much higher, thereby making outdoor air relatively insignificant in the context of total exposure.

Estimated outdoor air concentrations of individual chemicals are designated as C_{oa} in this document.

4.4.6 Concentrations in dust/airborne particulates

For many chemicals, particularly non-volatile chemicals, concentrations in air are primarily associated with a particulate phase. In some cases, these chemicals may be evaluated using measured concentrations in air directly, but often they are evaluated by determining the concentration of the chemical on the airborne particulates, particularly if environmental modelling approaches are being used.

Concentrations of a chemical in dust and airborne particulate are often assumed to be equivalent to concentrations in the soil from which they are derived. In some cases, metals may be at a higher (or lower) concentration in airborne particulates if the metals are associated with soil particles of a particular size (e.g. if the metal contamination in soil originates from deposition of airborne industrial emissions). In these situations, use of an enrichment factor based on available data or literature may be appropriate or the direct measurement of metal concentrations in airborne particulate (e.g. collecting particulate for analysis using a high-volume air sampler) and dust (e.g. collecting dust swabs from residences). Accurate determination of chemical concentrations in dust or airborne particulate becomes especially critical if particulate inhalation may be a significant (or the only) exposure pathway, such as for receptors located downwind of a tailings pile.

In addition to the chemical concentration in airborne particulate, the amount of respirable airborne particulate (i.e. PM_{10}, particulate matter with an aerodynamic diameter $\leq$10 µm) present in the air must also be known to evaluate exposure through this medium, either by direct measurement or by estimation. In the absence of site-specific data, a typical airborne particulate matter concentration of 0.76 µg/m^3 may be assumed for most sites (HC, 2010b: based on U.S. EPA, 1992a) or a concentration of 250 µg/m^3 for sites with vehicle traffic on contaminated unpaved surfaces (HC, 2010b: based on Claiborn et al., 1995).

Estimated dust and airborne particulate concentrations of individual chemicals are designated as C_d in this document.

4.4.7 Concentrations in surface water

Chemical concentrations in surface water are relevant at sites where humans may be exposed to surface water, either as a domestic water source or through recreational use (e.g. swimming). Surface water concentrations may also be useful for estimating concentrations of chemicals in food sources, such as fish present in the water or crops irrigated with the water.

Chemical concentrations can be directly measured in surface water. However, chemical concentrations in surface water can be highly variable both spatially (laterally and vertically because of stratification) and temporally; sampling programs need to address this variability in order to adequately characterize chemical concentrations. Background concentrations can be determined from literature or upstream measurements, or in some cases from measurements in nearby similar water bodies.

Environmental modelling can be used to predict future concentrations of chemicals in water or transport of chemicals within surface-water bodies. In some cases, jurisdictions may allow dilution factors to be used when estimating concentrations in surface water, based on discharge from groundwater or surface runoff, or direct chemical discharge into surface water. In all cases, the modeller should ensure that all potential sources of the chemical in the surface water are evaluated.

Estimated surface water concentrations of individual chemicals are designated as C_{sw} in this document.

4.4.8 Concentrations in sediments

Chemical concentrations in sediments are relevant for sites where humans may be exposed to sediments directly (e.g. through recreational activities), or where food sources may accumulate chemicals from the sediments (e.g. contaminants entering aquatic food webs via bottom-feeding fish, aquatic plants, or invertebrates).

Concentrations in sediments can be measured directly. Much like for soils, the heterogeneity of sediments must be considered when designing a sampling program. Environmental modelling can also be used to estimate sediment concentrations, particularly when future concentrations are needed.

Estimated sediment concentrations of individual chemicals are designated as C_{sd} in this document.

4.4.9 Concentrations in food

The accumulation of chemicals in food (agricultural produce, backyard urban gardens and allotments, "country foods") can result in significant exposure to human receptors; as a result, accurate measurement or estimation of chemical concentrations in food is critical for risk assessments where this pathway is active.

Concentrations of chemicals can be measured directly in plants and animals used as food sources. Because chemical concentrations vary in different tissues and organs, the samples analyzed should reflect the tissues normally consumed by humans. It is also relevant in some cases to consider concentrations in foods as they are prepared for consumption (e.g. cooked versus raw, washed versus unwashed, peeled versus unpeeled).

Modelling approaches should consider all mechanisms of chemical uptake, including uptake from air, water, and soil; uptake from food sources should also be considered when evaluating chemical concentrations in animals. Chemicals are also removed from plants and animals by excretion or metabolization (potentially into toxic metabolites), factors that may be considered when modelling tissue biotransfer and accumulation with time.

Chemical concentrations in food are often estimated using bioconcentration or biotransfer factors (typically relating soil or water concentrations to concentrations in the plant or animal, as well as from fodder or prey in the case of animals) reported in the literature. This approach is subject to considerable uncertainty because bioconcentration factors have been shown to vary by several orders of magnitude for the same chemical for different species, soil conditions, and chemical concentrations in the source medium Therefore, care must be taken in selecting appropriate and conservative bioconcentration or biotransfer factors.

4.5 Receptor Characterization

4.5.1 Physical and behavioural characteristics

Exposure to chemicals from environmental media is highly dependent on the physical and behavioural characteristics of the exposed human receptors. The rate at which receptors contact environmental media (e.g. soil ingestion rate, air inhalation rate) affects the amount of exposure, as does the frequency of contact and the duration of each contact event during which receptors are exposed. Because exposure doses are frequently expressed on a per unit body weight basis, the receptor body weight also affects the exposure dose.

Receptor characteristics vary among individuals and age groups, and from region to region. The characteristics used for the risk assessment must be appropriate for the exposed population. In addition, some characteristics may be positively or negatively correlated with each other (e.g. body weight and skin surface area); values chosen need to be consistent with each other.

Health Canada has identified five age groups into which the physical characteristics of the human population should be classified for most risk assessments:

- infant (0 to 6 months inclusively)
- toddler (7 months to 4 years inclusively)
- child (5 years to 11 years inclusively)
- teen (12 years to 19 years inclusively)
- adult (20 years to 80 years inclusively)

Not all of these age groups may need to be considered in every risk assessment. For example, if exposure is being evaluated for an operating industrial facility with restricted access, it may be appropriate to consider only adults. If access to the site is not restricted, then in most cases all age groups should be considered. Other receptor groups may be necessary for unique or peculiar sites and situations.

The specific physical and behavioural characteristics needed for the risk assessment will depend on the exposure pathways being evaluated. Characteristics that are typically needed include (but are not necessarily limited to):

- body weight
- soil ingestion rate
- air inhalation rate
- water ingestion rate
- skin surface area
- soil loading to exposed skin
- food ingestion rates
    - root vegetables
    - other vegetables
    - fish
    - wild game
    - dairy products
- frequency and duration of site visits or exposure events

The *Compendium of Canadian Human Exposure Factors for Risk Assessment* (Richardson, 1997) summarizes Canadian data for many of the receptor characteristics. Where Canadian data or guidance on a required characteristic are lacking, the

U.S. EPA exposure factors handbooks (U.S. EPA, 2008, 2009) should be consulted. However, because characteristics can be different for local regions or specific cultural groups, the risk assessor should ensure that the values used are appropriate for the exposed population.

Many of the relevant exposure characteristics are lifestyle dependent, or may be a function of job-specific or facility-specific work requirements if the site is an industrial, commercial, or other occupational site. It is therefore important that the local conditions, habits, and site use patterns be considered. Lifestyle habits, such as smoking, can also affect background exposure or chemical sensitivity, as can occupational exposure. Potential methods for evaluating the local or site-specific characteristics include:

- direct consultation with local residents (e.g. open houses, conversations with residents)
- surveys of time–activity patterns and frequency and duration of site access/use
- sites visits and observations
- review of literature on local population (e.g. public consultants' reports, demographic information)
- review of scientific literature on similar populations (i.e. similar geographic and climatic area, cultural background, etc.)

For sites used for commercial, industrial, or other occupationally related activities, interviews and questionnaires to staff and/or management can be particularly valuable in identifying unique situations relevant to assessing risks on site.

4.5.2 Receptor characterization for deterministic risk assessment

Natural variability is present in both physical and behavioural characteristics of human receptors. This variability can result from factors such as age, gender, genetic factors, physical condition, cultural background, diet, individual history, or occupation. Because of this variability, each human receptor exposed at a site will experience a different exposure dose.

When conducting a deterministic risk assessment, receptor characteristics are represented by single point estimates. Because a single value obviously cannot reflect the full range of variability present in the population, the point estimate values must be chosen with care.

For federal sites in Canada, receptor characteristics will generally be represented by the "typical" or arithmetic average value for the relevant age group. However, an upper percentile (such as 95th percentile), or a "reasonable" worst-case value may also be appropriate in some cases, particularly if the characteristic(s) in question is not well understood or quantified. It is important for the risk assessor to understand the uncertainty and variability in the receptor characteristics and the effects of the chosen point estimates on the risk assessment results. If all receptor characteristics are represented by typical or average values, then the risk assessment may not be protective of all members of the population unless conservatism is present in other aspects of the risk assessment (such as in the setting of regulatory TRVs in the estimation of media concentrations, etc.). Using 95th percentile or worst-case values for all parameters, however, may lead to significant overestimation of exposures and risks.

4.5.3 Receptor characterization for probabilistic risk assessment

Receptor characteristics used in a probabilistic risk assessment may be characterized by probability distributions that reflect the full range and variability across a population or receptor group. Receptor characterization for probabilistic risk assessment is discussed further in section 7.0.

4.5.4 Sensitive receptor groups

Some human receptors may be more sensitive to the COPCs than the general population, either because of susceptibility to contaminants or because of unusually high exposure. The former group is evaluated during the toxicity assessment (section 5.0); however, in some cases this may also require the definition of additional receptor groups (e.g. if women are more susceptible to a chemical than men, their exposure should be evaluated separately). Receptors that may have unusually high exposure should be specifically identified and addressed during the exposure assessment.

Typical examples of sensitive receptor groups include people with a high rate of fish consumption (for methylmercury and other bioaccumulative contaminants), gardeners (increased dermal exposure to soil-borne contaminants), people consuming produce growing on a contaminated site, or some types of occupational receptors. These receptors can be evaluated in two ways:

1. defining one or more receptor groups consisting of these sensitive receptors
2. ensuring that characteristics of the sensitive receptors are encompassed in the distributions used for receptor characteristics for a probabilistic exposure assessment

Additionally, in some cases, receptors may be exposed to unusually high doses of chemicals over a relatively short period of time. An example of this is acute ingestion exposure of children during pica events (Calabrese et al., 1997). In these situations, the acute exposure should be evaluated separately, in addition to the chronic exposure at the site (see section 4.6.1).

4.6 Exposure Averaging and Amortization

Exposure amortization is the process used to derive the average dose (typically per day) of a chemical over a given exposure period (i.e. average dose rate, often expressed as mg/kg bw/d) by taking into account the overall duration of exposure and the pattern(s) of exposure in the scenarios selected for the site. Exposures to chemicals from a site are seldom continuous (i.e. exposures are usually not 24 hours per day, 7 days per week, 52 weeks per year for a lifetime). Instead, almost all exposure patterns will involve interruptions in events leading to exposure. Some exposure patterns will occur on a regular basis, whereas others will be intermittent, sporadic, or may involve only a single brief exposure at the site. Consequently, exposure estimates for selected receptors should consider the likely duration of and possible interruptions in exposure. Mathematically, exposure amortization involves the establishment of the exposure term (ET), which is used as a parameter in many equations to estimate exposure. However, to establish the ET, the risk assessor should consider the basic principles of exposure amortization presented in the sections that follow. Where acute or subchronic exposures are assumed, a scientific rationale must be provided that is consistent with the toxicity data for each COPC.

4.6.1 Classification of exposure duration

For the purposes of exposure assessment, three different lengths of exposure are used in the classification of exposure duration for human receptors:

- acute (here assumed to be less than 14 days, but often involving a single high-intensity exposure)
- subchronic (here assumed to be greater than 14 days and less than 90 days)
- chronic (greater than 90 days)

The above designations are consistent with Health Canada's general approach, but are not identical to the classification scheme recommended by the U.S. EPA (1989), the Agency for Toxic Substances and Disease Registry (ATSDR) (ATSDR, 2005a), or some Canadian provincial regulatory agencies. In addition, there are several key factors to consider in the classification of exposure duration. These factors are discussed in the following text.

4.6.1.1 Short-duration or acute exposures

Short-duration exposures of human receptors are termed acute exposures. These are exposure scenarios that occur over a period not longer than 14 days. However, in some situations, acute exposure scenarios may consist of a single exposure event, or they may occur as a result of accidents or events that do not occur on a regular basis. Acute exposure scenarios include events such as a child gaining access for varying time intervals to a contaminated site or a site under remediation, or the interaction of the child with the site during such exposures is a significant concern.

4.6.1.2 Subchronic exposures

Human exposures that occur over a period ranging from 14 to 90 days are termed subchronic exposures. An example would be exposure received by remediation workers who may be at a site for 5 days per week for 6 weeks. However, if the same workers subsequently move on to similar work at other contaminated sites, the above example may actually contribute to a chronic exposure scenario as well.

4.6.1.3 Chronic exposures

Human exposures that occur over a period of longer than 90 days are considered chronic exposures. These are continuous or repeated exposures that occur on a regular basis for a period of 90 days to 80 years or more.

These categories of exposure duration roughly correspond to those used to describe the durations in animal toxicity studies. For example, animal studies are typically categorized as acute (less than 14 days), subchronic (30 to 90 days), or chronic (major part of lifetime). Although a human exposure that lasts more than 90 days but less than the major part of the lifetime is not truly chronic (i.e. does not represent a major part of the lifetime), it is greater than one of subchronic duration. Consequently, in order to be conservative, it is recommended in most cases that human exposures occurring over a period greater than 90 days be considered chronic (see section 5.0). If in doubt about the appropriate exposure duration to consider at a federal site in Canada, Health Canada should be consulted to clarify the preferred method for classifying the exposure duration.

Another critical factor is the approach taken for the evaluation of acute or subchronic exposures to chemicals that have only chronic TRVs, such as tolerable daily intakes (TDIs), slope factors (SFs), and unit risks (URs). Generally, it is preferred that chronic TRVs be used in these circumstances, recognizing that the resultant risks will likely be overestimated, independent of any amortization of exposure; this is discussed further in section 5.0.

4.6.1.4 Determination of the exposure term

Once an exposure scenario has been classified as acute, subchronic, or chronic in duration, the ET must be determined. The ET is based on the total amount of time spent exposed to a chemical and the total duration over which exposure is amortized. It must be noted that TRVs selected for risk characterization must be consistent with this ET; TRVs are discussed in section 5.0.

4.6.1.5 Exposure term for chronic and subchronic exposure

Exposure is normally amortized over the total exposure period. For subchronic and chronic exposures, the total exposure period normally has a duration of days or years, respectively, and reflects the period of time starting the moment the receptor is first exposed to a chemical and ending when the receptor is last exposed to a chemical. This period may be the period of time spent working in a particular occupational setting, or the number of years spent living at a residence near a contaminated site. For example, if a chronic exposure scenario involves a person being exposed 5 days a week for 15 years, the total exposure period would be 15 years.

Because it is necessary to estimate the typical exposure for a given time period (i.e. usually the daily rate of exposure), exposure amortization may be used for intermittent or non-continuous exposure estimates. For example, in a chronic exposure scenario a person may be exposed to a systemically acting chemical (e.g. arsenic) 8 hours per day, 5 days per week, 50 weeks per year for 5 years. In these circumstances, it is appropriate to estimate the daily rate of exposure as the average daily rate of exposure of the person that occurs for every day of the 5 years. In other words, the risk assessor may estimate the total dose that occurs from exposure at the site for 5 years (i.e. the sum of all daily exposures occurring from the site). The total dose is then divided by the number of days in 5 years (i.e. 5 x 365 days = 1,825 days) to arrive at the average daily dose rate that occurred from exposure at the site over the 5 year period.

The ET used in many exposure calculations is the time exposed divided by the total exposure period; in the above example, the ET would be 5 days/week x 50 weeks/year x 5 years divided by 1,825 days = 0.685. In some cases, the ET would be further modified by the hours per day exposed (i.e. 8/24), if exposure to the environmental medium in question would be expected to occur throughout the day (e.g. inhalation of air). If the daily exposure to the environmental medium may occur per exposure event irrespective of the duration of that event, then the number of hours per day exposed would not normally be considered in the ET. This is often assumed to be the case for ingestion of soil or dermal contact with soil, which is expected to be dominated by discrete exposure events, or with ingestion of water because a receptor's daily water supply may come from a single source.

Exposure amortization may not be appropriate for some exposure scenarios, such as repeated acute or subchronic exposure—for example, a gardener who is exposed to a chemical in the soil 1 month every spring for 20 years. Although exposure is occurring for 20 years, this scenario does not represent a true chronic exposure scenario that can be amortized over 20 years. In these circumstances, it would be more conservative to estimate the typical daily dose rate that occurs during the month of greatest exposure each year. This exposure should then be compared to both a TRV based on subchronic toxic effects and a TRV based on chronic toxic effects. In all cases, where less-than-lifetime exposures are being evaluated, the risk assessor should ensure that the exposure estimate is not being compared to a TRV that was developed for protection of a shorter duration of exposure.

Consideration should also be given to chemical exposures that may manifest effects based on the pattern or timing of exposure rather than on the total exposure that occurs over a given time interval. As an example, a chemical that may cause developmental toxicity may produce effects only if exposure occurs during a particular stage of development (e.g. during the period of organogenesis). The possibility of such effects will influence the appropriateness of exposure amortization procedures.

4.6.1.6 Exposure term for acute exposure

The preceding approach, normally used for most subchronic and chronic exposure scenarios, also may be applicable to some acute exposure scenarios. In such cases, the exposure estimate should represent the average daily rate of exposure that has occurred for the acute time frame. However, for acute exposure scenarios of very short duration, the exposure estimate should represent exposures that occur for durations measured in time units that are compatible with the acute TRV being used to characterize risks in the assessment (in some cases hours or minutes, especially for certain respiratory irritants). Because the effects from some acute exposures are very dependent on the maximum chemical concentration a person is exposed to, irrespective of the duration of exposure, amortization of an exposure may not always be appropriate. As a result, for acute exposure scenarios, the total exposure period must not be greater than the exposure time frame over which the toxic effects have been observed in the toxicity assessment. For example, if the toxicity assessment indicated that exposures to an acutely acting chemical must not exceed 200 mg/m^3 for more than 2 hours, it would not be appropriate to amortize an exposure that lasts for 3 hours over a period of days.

Another situation in which amortization of exposures may not be appropriate is for scenarios involving chemicals that manifest toxicity at their site of contact with the body (i.e. local toxicity). Adverse effects of chemicals that cause local toxicity are a function of the chemical concentration and exposure duration. As an example, the risk of acute adverse effects from skin irritants or corrosive agents is clearly more effectively assessed through consideration of their concentrations rather than through amortization of the exposure period. In other situations, the duration of exposure may also play a role in the manifestation of adverse effects, and care should be taken before conducting exposure amortization. For example, a person may be exposed to a

locally acting chemical at an air concentration of 100 mg/m^3 every day for 20 minutes for 5 years. Experimental evidence may show that repeated short-duration exposures to 75 mg/m^3 can cause adverse effects. Therefore, if the 20 minute exposure were amortized to a daily exposure, the potential for adverse health effects may be underestimated. However, there are situations where some locally acting chemicals may cause adverse health effects based on typical daily exposures. It is therefore critical that any exposure amortization be consistent with the toxicity assessment.

4.6.2 Key issues pertaining to exposure amortization

From the preceding discussion, it is apparent that the decision to amortize exposures is based on considerations of the specific site and the specific chemical (i.e. on a case-by-case basis). Assignment of the total exposure period, and whether to proceed with amortization, should be based on the exposure conditions, the expected mechanism, and site of toxicity (e.g. local effects versus systemic effects; see section 5.0). Care should be taken to ensure that potential short-duration exposure effects are not overlooked. As exposure amortization is a complicated issue in some circumstances, risk assessors are advised to consult Health Canada if there is doubt concerning the acceptability of the proposed approach for establishing an averaging time.

Exposure amortization calculations are outlined in Example 4.1.

Example 4.1 Exposure Amortization

Chronic Exposure

Problem

A lead-contaminated site is being used for commercial purposes. It is determined that a worker would spend 8 hours per day, 5 days per week, 48 weeks per year at the site. Exposure is expected to occur through incidental ingestion of soil, dermal contact with soil, and inhalation of soil particulate. What is the exposure term?

Solution

Although exposure by soil ingestion and dermal contact is not continuous throughout the day, Health Canada recommends assuming that the full daily inhalation exposure could occur at the site. Therefore, the exposure term would not consider the hours per day spent at the site, and would be 5 days/7 days x 48 weeks/52 weeks ≈ 0.659. Because inhalation occurs throughout the day, it is appropriate to consider the hours per day at the site for inhalation of soil particulate; therefore, the exposure term for this pathway would be 8 hours/24 hours x 5 days/7 days x 48 weeks/52 weeks ≈ 0.220.

Subchronic Exposure

Problem

During the remediation of a contaminated site, a backhoe operator is expected to be exposed to benzene in ambient air inside an excavation. It is anticipated that the operator will be exposed 10 hours per day, 5 days per week, for a period of 6 weeks. What is the exposure term?

Solution

The exposure term would be 10 hours/24 hours x 5 days/7 days ≈ 0.298. Note that it is not appropriate to amortize this exposure over the number of weeks per year exposed because this is a subchronic exposure with a total exposure duration of 6 weeks.

Acute Exposure

Problem

Tetrachloroethylene contamination from a former dry-cleaning facility has been identified in a utility corridor. During site operations to remove contamination from the utility corridor and install a barrier to protect the underground utilities, it is anticipated that a worker could be exposed to a tetrachloroethylene concentration in air of 200 µg/m^3 for 6 hours. During the toxicity assessment, an occupational exposure limit of 170 µg/m^3 as an 8 hour time-weighted average is identified. What is the amortized exposure concentration?

Solution

The most conservative approach is to apply the exposure concentration (200 µg/m^3) directly without amortization. If the toxicity reference value allows the concentration to be adjusted to an 8 hour time-weighted average, then the calculation would be 200 µg/m^3 x 6 hours/8 hours = 150 µg/m^3. In that case, the exposure concentration should also be evaluated against shorter-term exposure limits if available (e.g. 15 minute average exposure limit) to ensure that these are not exceeded.

4.7 Bioavailability Assessment

The toxicity of systemically acting chemicals is related to the amount of the chemical absorbed and retained in the body (the bioavailable fraction). Therefore, it may be appropriate to adjust exposure to reflect bioavailability, particularly if the environmental exposure route/scenario is different from the exposure route/scenario considered in the toxicity assessment (e.g. using an oral TDI or RfD developed for ingestion to evaluate dermal contact exposure). It is critical that any assessment of bioavailability is integrated with the toxicity assessment; if the toxicity benchmarks do not consider bioavailability, then the exposure dose calculations should also omit this factor. A bioavailability assessment is not typically undertaken for locally acting chemicals (e.g. irritants) because toxicity is generally based on the contact dose for these substances.

A bioavailability assessment may be important under any of the following conditions.

- The form of the on-site chemical (e.g. metal species) is different from that used in experiments to describe the toxicity of the chemical, leading to a difference in bioavailability, and this difference can be measured and reliably quantified.
- The route of exposure relevant to the site is different from that used in the key study to determine the TRV of the

chemical, leading to a difference in bioavailability contributing to total system load, and this difference can be reliably quantified.

- The TRV identified in the toxicity assessment has been adjusted for bioavailability (i.e. represents or considers absorbed dose rather than delivered dose).
- Multiple routes of chemical exposure are considered, but the TRV for the chemical is developed from studies that used a single route of exposure.

When PQRAs are performed, a value of 1 (i.e. 100%) is normally used to represent the proportion of the chemical absorbed into the body following oral (and inhalation) exposure. However, in a complex DQRA, it may be pertinent to conduct in vivo or in vitro assays to more accurately characterize oral bioavailability of soil-borne contaminants on site. If quantified for the site in question, the arithmetic average of all measurements may be used to represent site-specific bioavailability, provided enough measurements (data quantity) and consistent results (data quality) are available. However, if only a few measurements were made and/or there was high variability in those results, then maximum measured bioavailability should be used for site exposure pathways. See section 4.7.3 for further discussion.

For probabilistic risk assessments, where there are sufficient data, a distribution representing the variability of bioavailability/bioaccessibility determined on a site-specific basis may be used.

It is important that, if bioavailability is assessed, it is assessed for both the environmental exposure and the TRV. Toxicity studies and TRVs are generally quantified in terms of the delivered dose rather than the absorbed dose. Absorption in toxicity studies is seldom 100% (see AMEC, 2005). It is inappropriate to assume 100% bioavailability for the TRV while using a lower measure of absolute bioavailability for the site-specific soil being assessed. All bioavailability adjustments must be made relative to the known or likely bioavailability in the relevant toxicity study. It is also possible, albeit rare, that bioavailability could be higher from the on-site exposure scenario than from the vehicle and exposure route used to evaluate toxicity, particularly if using a TRV derived from an oral toxicity study to evaluate inhalation exposure, or a TRV derived from a food exposure study to evaluate exposure through drinking water.

Chemical absorption by the body can be highly variable. It is affected by chemical species, properties of the environmental medium (e.g. soil particle size and organic carbon content, pH, etc.), and properties and characteristics of the receptor (may vary by age, fasting versus non-fasting). The bioavailability assessment should consider this variability, and ensure that uncertainties are addressed in a conservative manner.

Example 4.2 outlines sample bioavailability calculations.

Example 4.2 Bioavailability

Problem

Dermal exposure to arsenic-contaminated soil and water (e.g. swimming or bathing) is being evaluated. The *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, *Version 2.0* (HC, 2010) guidance has specified a toxicity reference value based on oral exposure (ingestion of water). How should bioavailability be incorporated?

Solution

Based on a study by Wester et al. (1993), the risk assessor determines that dermal absorption of inorganic arsenic from water is approximately 6.4% and dermal absorption from soil is approximately 4.5%. Absorption from ingestion of water is estimated at 74.4% based on a study by Roberts et al. (2002).

Two approaches can be taken to adjust the dermal exposure doses for bioavailability.

1. The risk assessor can multiply the exposure dose by 0.045 for dermal exposure from soil and by 0.064 for dermal exposure from water. However, the toxicity reference value must also be adjusted during the toxicity assessment to reflect the estimated oral absorption from water of 0.744.
2. Relative absorption factors could be calculated for dermal exposure from soil (0.045/0.744 = 0.060) and water (0.064/0.744 = 0.086); the dermal exposure doses would then be multiplied by the relative absorption factors and compared directly to the published toxicity reference value.

Note that the absorption values presented in this example are for illustrative purposes only and may not apply to all types of soil.

4.7.1 Dermal bioavailability

Bioavailability adjustment is most commonly considered for dermal exposure pathways because dermal toxicity studies for systemic toxicants are rare. In most cases, an oral toxicity benchmark or potency factor is used for the evaluation of dermal exposure, and exposure should therefore be adjusted to reflect dermal bioavailability relative to the oral exposure route. However, site-specific assays of dermal bioavailability are seldom, if ever, conducted. Should assays of dermal bioavailability be planned, Health Canada should be consulted as to preferred in vivo and in vitro methods. The *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, *Version 2.0* (HC, 2010c) gives relative dermal absorption factors for several chemicals that may be applied for estimating exposures via the dermal route. Alternatively, data and information from recognized agencies such as the U.S. EPA (e.g. U.S. EPA, 1992b, 2004) or other sources can be employed as appropriate, as long as the source is clearly identified and the use of those particular values adequately supported with a scientific rationale.

An exception to this is when using exposure equations with permeability factors, such as the equation for dermal contact with water published by the U.S. EPA (1992b). The permeability factors in these equations already consider the absorption of the chemical, although it may be appropriate to evaluate the oral bioavailability for the toxicity benchmark or potency factor in these cases. The use of permeability factors and lag times may allow for a more realistic estimate of dermal bioavailability than the application of relative absorption factors if sufficient data are available.

4.7.2 Inhalation bioavailability

If the TRV used for the evaluation of inhalation exposure is derived from an inhalation-based toxicity study, then bioavailability is not generally considered. However, if inhalation toxicity is evaluated using a TRV derived for the oral route, then a bioavailability assessment should be performed. In some cases, inhalation absorption may be greater than ingestion absorption. If this is the case, the relative bioavailability adjustment factor will be > 1.0 (>100%).

Bioavailability can also influence the risk assessment if a chemical is associated with airborne particulate matter, particularly if the particle sizes can be characterized; this is because absorption is likely to be greater from smaller particles that penetrate further into the respiratory system (see Bright et al., 2006). One in vitro means of estimating inhalation bioavailability of particle-borne contaminants is to measure their solubility in simulated lung fluid. Such assays are commonly applied in health physics when assessing tissue distribution and organ-specific doses of inhaled radioactive particles (see Bright et al., 2006). However, the relevance and utility of such assays for DQRA is questionable (see Bright et al., 2006). It is not recommended at this time that assays of lung fluid solubility be used within risk assessments for particle-borne chemicals at Canadian federal contaminated sites.

4.7.3 Oral bioavailability

In most cases, TRVs used to evaluate oral exposure are based on toxicity studies with oral exposure. Therefore, bioavailability adjustments are not normally done within PQRAs for oral exposure. However, in DQRAs in which more accurate and realistic estimates of exposure and risk are the objective, such bioavailability adjustments may be desirable and will be considered on a case-by-case basis by Health Canada. However, Health Canada will only consider such adjustments where site-specific data are collected for the relevant environmental medium (i.e. must be soil or other terrestrial material to which receptors can conceivably be exposed via ingestion; for example, media such as submerged sediments will not be accepted as representative of terrestrial soils).

4.7.3.1 In vivo assays of site-specific oral bioavailability from soil

No specific or standard protocol for in vivo bioavailability studies is defined by Health Canada at this time. Any plans for such studies should be discussed with Health Canada well in advance of study initiation. It should be noted, however, that laboratory rat species appear to be inappropriate for in vivo investigations of oral bioavailability from soil. A review of published studies pertaining to arsenic (As), cadmium (Cd), and lead (Pb) concluded that bioavailability of these soil-borne contaminants in rats was consistently low relative to other mammalian species, including humans (JWEL, 2005).

4.7.3.2 In vitro assays of site-specific oral bioaccessibility from soil

No specific or standard protocol for the in vitro assay of soil bioaccessibility is defined by Health Canada at this time. Any plans for such studies should be discussed with Health Canada well in advance of study initiation. However, some factors to be considered if assaying bioaccessibility of inorganic contaminants are discussed in the following text. These are reviewed and discussed in greater detail by Richardson et al. (2006).

Bioaccessibility assays are becoming an increasingly common component of human exposure assessments for contaminated sites, despite the lack of in vivo validation (beyond Pb and As). Prohibitive costs will likely prevent the widespread application of in vivo assays on a site-specific basis for all but the largest and most complex (and expensive to remediate) sites. Prohibitive costs will also likely prevent the in vivo validation of bioaccessibility results for more than a select few contaminants. However, in vitro bioaccessibility

assays may still be valid and applicable at specific sites, provided that some basic data requirements are fulfilled to permit Health Canada to ascertain the validity and defensibility of bioaccessibility results for application in federal contaminated site risk assessments. As a minimum, the following data are required.

1. A range of soil particle size fractions (e.g. ≤45 μm, ≤125 μm, ≤250 μm) should be assayed to determine if soil particle size significantly influences measured bioaccessibility at the site in question. If it does not, then soil characteristics based on the < 250 μm size fraction, including bioaccessibility, soil-borne contaminant concentration (and other factors), can be employed in estimating exposures and risks. However, in situations where soil particle size is influential, data for the smallest size fraction should be used in the risk assessment.
2. For a subset of assayed soil samples, a range of ratios of simulated gastric fluid (mL) to soil mass (g) must be employed, ranging from 100:1 to perhaps 5,000:1, and possibly up to 10,000:1, if analytical detection limits are not limiting. If measured bioaccessibility is not significantly influenced by this ratio, then further assays for that site can proceed with the standard 100:1 (or other) methodology. However, in situations where measured bioaccessibility is influenced by these ratios, and particularly where bioaccessibility increases as the ratio of simulated gastric fluid volume to soil mass increases, then bioaccessibility adjustments within the risk assessment should be based on assays employing the maximum ratio possible, up to 10,000:1 where feasible.
3. Further to requirement (2), statistical analysis of the bioaccessibility data should confirm that contaminant solubility is not confounding the measure of bioaccessibility. Except in rare instances where a ratio of 10,000 mL of gastric fluid to 1 g of soil can be employed, the final data used to determine site-specific absolute bioaccessibility should demonstrate the independence of measured bioaccessibility and the concentration or mass of contaminant in soil being assayed.
4. For application as a bioavailability adjustment factor in risk assessment, the absolute bioaccessibility (as directly measured in a bioaccessibility assay) must be adjusted **relative** to the (likely) bioavailability of the contaminant in the key toxicological or epidemiological study upon which the TRV (e.g. TDI or cancer SF) was derived. To facilitate the determination of relative bioaccessibility, Health Canada is currently compiling oral bioavailabilities in key toxicological studies upon which Canadian federal TRVs have been established.

4.7.3.3 Environmentally relevant soil fraction

As discussed above in section 3.4, and as reviewed by Bright et al. (2006) and Richardson et al. (2006), soil samples sieved to ≤ 250 μm may not be representative of the soil size fraction of greatest relevance to HHRA. Various studies and authors suggest that smaller size fractions, possibly as small as ≤ 2.5 μm (Edwards and Lioy, 1999), may be the most relevant fraction resulting in exposure to soil-borne contaminants. Therefore, risk assessors should be confident that the soil particle size fractions tested for chemical concentration, bioaccessibility, and/or other characteristics is the size fraction of greatest relevance at the site in question. Health Canada should be consulted where uncertainties exist; it may be necessary to characterize a range of soil fractions to determine if fraction-related increases or decreases in those characteristics are evident.

4.7.3.4 Simulated gastric fluid volume: soil mass

Various researchers have noted that, at least for some contaminants, bioaccessibility increases with an increasing ratio of acidic leachate to soil mass (reviewed by JWEL, 2005; Richardson et al., 2006). A variety of ratios of simulated gastric fluid to soil have been employed for bioaccessibility assays, ranging from 5:1 to 5,000:1 (reviewed by JWEL, 2005). The most common assay designs for measuring the bioaccessibility of soil-borne contaminants employ a ratio of leachate volume to soil mass of 100 mL:1 g, but none approach the ratios likely to exist in the toddler or adult gastrointestinal tract—a ratio on the order of up to 10,000 mL:1 g soil or greater. Although the 100:1 ratio used in in vitro studies does not approach the ratios likely to exist in the toddler or adult gastrointestinal tract, in vivo data correlates reasonably well, at least for As and Pb, suggesting it is a reasonable surrogate for these element in most cases. Unfortunately, at present, it is impossible to predict when or how the use of a 100:1 ratio will **not** produce representative results for other elements.

Interpretation of Bioaccessibility Results

Measuring contaminant saturation rather than bioaccessibility

The relatively low ratio of simulated gastric fluid volume to soil mass commonly used in bioaccessibility assays can cause problems associated with potentially exceeding the saturation point for the contaminant being investigated, particularly when that substance is sparingly soluble. For valid bioaccessibility results, it is essential that the measured bioaccessibility be independent of solubility limitations. This is particularly true when the bioassay design employs a gastric fluid to soil mass ratio substantially lower than what occurs in the human gastrointestinal tract. The validity of bioaccessibility results can be tested where multiple soil samples representing a

range of contaminant concentration have been assayed. If bioaccessibility is independent of solubility limitations, then (i) the curve associating bioaccessibility with soil-borne concentration should not be negative, and (ii) the curve associating the mass of extracted substance and soil-borne concentration should have a statistically significant positive slope. Both of these relationships will indicate that the proportion of the substance extracted from the soil is independent of soil-borne concentration and is, thereby, independent of assay design.

Absolute versus relative bioavailability/bioaccessibility

Absolute bioavailability is defined herein as the proportion of the mass of a soil-borne (or sediment-borne) contaminant reaching the gastrointestinal tract (or lungs or skin) that is systemically absorbed (in vivo assay), or that is deemed to be available for systemic absorption (in vitro bioaccessibility assay). Relative bioavailability is defined herein as the absolute bioavailability (or bioaccessibility) from the site-specific soil samples divided by the absolute bioavailability of the same substance under the conditions used to derive the TRV (Paustenbach et al., 1997).

The gastrointestinal absorption of any element or contaminant is seldom complete. However, when adjusting exposure assessment calculations for the bioavailability of a contaminant in soil, this must be done **relative** to the gastrointestinal absorption of the same substance in the toxicological or epidemiological study upon which the TRV is based. Failing to do so will result in a bioavailability adjustment that may significantly underestimate bioavailability of the toxic moiety and, as a result, on-site risks will be underestimated. For example, the regulatory TDI often employed to characterize the risk of Ni exposure is that for soluble Ni salts published by Health Canada (1996). That TDI was based on a study of Ni sulphate toxicity in rats dosed via their diet (Ambrose et al., 1976). When mixed with food, the absorption of Ni species is significantly < 100%, likely in the range from < 10% to 40% (ATSDR, 2005b). If the absolute bioaccessibility for Ni in soil samples has been measured to be 20%, the **relative** oral bioavailability of soil-borne Ni would appropriately be established as 50% (20% ÷ 40% x 100) to as much as 100%, noting that absorption from food can be < 20%.

Health Canada is in the process of compiling the bioavailabilities of chemicals in key toxicological studies, so that site-specific relative oral bioavailability adjustment factors can be computed.

4.8 Microenvironment Analysis

When conducting an exposure assessment, it can be important to consider particular microenvironments (Lioy et al., 1992; Schwab et al., 1992; Whitmyre et al., 1992).

Microenvironments are defined as smaller regions of the site that are characterized by specific ranges of environmental concentrations of chemicals and by physical features that would affect exposures of receptors to chemicals, depending on the manner in which the site is used. Analysis of microenvironments can identify areas where unacceptable exposures could occur that would be missed entirely using data-averaging techniques to describe the site as a whole.

4.8.1 When to use microenvironment analysis

Use of microenvironment analysis may be important under the following conditions.

- The concentrations of the chemicals are not uniformly distributed throughout the entire site (i.e. distinct concentration ranges of the COPCs are present in different areas of the site).
- The site would not be used by receptors in a uniform manner, and the patterns of use can be reliably estimated or quantified.

One such example would be a children's playground within a larger site. The playground will preferentially attract children; therefore, the frequency and duration spent there and contaminant data specific to this area will be very important to accurately assessing risks to children.

As a general rule, if both of the above conditions are true, microenvironment analysis will assist in improving the accuracy of the exposure assessment and consequently the overall risk assessment; in these cases, a microenvironment analysis should always be applied. If only one of the conditions is true, the microenvironment analysis may still improve the accuracy of the exposure assessment, and consideration should be given to applying a microenvironment analysis if possible. If neither condition is applicable, microenvironment analysis will not improve the overall risk assessment and may be omitted from the exposure assessment.

4.8.2 Considering microenvironments

The consideration of microenvironments improves the realism in exposure assessments and helps to ensure that rates of exposure to chemicals from the site are not underestimated for specific types of site use.

It is not possible to provide a definitive set of rules or methods to identify the existence of microenvironments that should be considered. Instead, based on site investigational data and proposed and current uses of the site, the risk assessor should examine whether or not there is the possibility of microenvironments existing that should be specifically accounted for in the risk assessment.

Microenvironments can be identified based on both the patterns of use by the receptor and on the chemical distribution patterns at the site (U.S. EPA, 1989). For example, if there are localized "hot spots" (i.e. areas of greater than average chemical concentrations) at a site, these could be treated as microenvironments. Alternatively, if hypothetical human receptors spend disproportionate amounts of time at certain locations within the site (at a playground, for example), these areas frequented by receptors may also be considered microenvironments.

Microenvironments may also be receptor age-specific. Although, young children may be more likely to spend time near playground equipment at an urban park, adolescents may spend more time in open areas of the park (e.g. playing sports). These use patterns would be particularly critical if, for example, the risk assessment was dealing with arsenic contamination originating from the playground equipment, or if herbicide application to maintain the sports field was an issue. If use patterns such as these are identified, the increased or decreased usage of microenvironments should be incorporated into the exposure assessment.

Once site microenvironments have been identified, the risk assessor should estimate the exposure that occurs for each microenvironment (see section 4.9). Scenarios should be developed describing the on-site behaviour of receptors and the amount of time spent in each microenvironment estimated from these scenarios. In addition, concentrations of the COPCs in each microenvironment should be adequately characterized. The total exposure from the site is estimated as the sum of the exposures from each microenvironment. For the preceding urban park example, the risk assessor might assume that toddlers and children spend all of their time at the park exposed to the arsenic concentrations measured in the vicinity of the playground equipment.

In addition to estimating exposure (and subsequent risks) under site conditions prior to remediation, microenvironment analysis may also benefit remediation and risk management decisions. Microenvironment analysis allows risk assessors to evaluate how remediation of certain areas of the site will affect the total estimated risks associated with specific uses of the site. As an example, the risk assessor can estimate the reduction in risk as a result of remediation of a specified "hot-spot" at the site, thereby enabling more realistic estimates of the degree of site remediation necessary to achieve a specified risk reduction. Alternatively, risks could be estimated based on frequently used parts of the site being remediated while less-used parts could safely remain contaminated.

It should be stressed that microenvironment analysis should be used only to increase the realism of the risk assessment. As such, it should rely on observed rather than speculative data or behaviour. There are two exceptions to this: (i) if the intent is to construct an extreme worst-case scenario wherein intense use is combined with high chemical concentrations, and (ii) site redevelopment plans are being prepared and locations within the site that could be unsuitable to specific uses (as a playground, for example) are being identified.

4.9 Exposure Estimation – Deterministic

Exposure is estimated for each chemical and for each exposure pathway, based on the considerations discussed earlier in this section. Depending on the circumstances, pathway, and chemical-specific exposures may be summed to determine a total exposure. Background exposures may also be assessed in some circumstances.

Several general equations that can be used in exposure assessments are presented in this section. These equations are presented as typical examples only to illustrate the principles and essential elements in exposure estimation. Exposure assessments reviewed by Health Canada are evaluated individually on their own merits, and any exposure assessment that is completed using scientifically sound principles will likely be acceptable. Other potential sources for exposure equations include various documents published by the U.S. EPA (e.g. U.S. EPA, 1989, 1992b) or provincial jurisdictions.

4.9.1 Estimation by pathway and exposure route

All significant exposure pathways identified during the problem formulation must be evaluated. The objective of the exposure assessment is to estimate an exposure dose. The exposure dose is often expressed as a **potential** dose because the scenarios are hypothetical in nature; they are seldom confirmed through direct observation, biomonitoring, or other means of direct exposure assessment. The exposure dose represents the rate at which a chemical comes into contact with the body relative to body weight. Typical units are mg/kg bw/d or µg/kg bw/d. If bioavailability is evaluated for a systemically acting chemical (see section 4.7), then the exposure dose may be expressed as an internal dose rate, representing the rate at which a chemical is absorbed into the body relative to body weight. For example, if a 70 kg person is exposed to 7 mg of a chemical each day but absorbs only 50%, the potential dose rate would be 0.1 mg/kg bw/d whereas the internal dose rate would be 0.05 mg/kg bw/d. Care must be taken to ensure that exposure doses and TRVs are in the same units and both represent either delivered or absorbed doses.

Exposure to locally acting chemicals (e.g. irritants) is often more appropriately expressed as a concentration of the chemical in the specific environmental medium that is contacting affected tissues, and the duration and frequency of exposure. The specific form in which exposure is reported must be consistent with the exposure endpoint determined during the toxicity assessment (section 5.0).

Exposure equations generally comprise the product of the chemical concentration in the environmental medium and the contact rate with that medium, divided by the body weight and adjusted for the exposure frequency/duration. Examples of general exposure equations for commonly evaluated exposure pathways are presented in Table 4.1. Exposure is calculated for each chemical, for each exposure pathway, and for each receptor group being evaluated. Exposures are normally summed for all pathways (e.g. soil ingestion, water ingestion, food ingestion—all pathways that are attributable to the oral route) relating to each major exposure route (oral, dermal, and inhalation), where the toxic effect is independent of the route of exposure (i.e. where all routes contribute to a total systemic load). However, where route-specific TRVs exist, it may not be appropriate to sum exposures from different routes. In cases where mixtures of similar compounds are evaluated using toxic equivalency factors (TEFs) (such as PCBs, carcinogenic PAHs, dioxins and furans; see section 5.6.4 for further discussion of chemical mixtures), exposures should be adjusted and summed to the reference chemical-equivalent concentration or dose. Otherwise, exposures to individual chemicals (with unique modes of action and target tissues and effects) should not be summed. Section 6.0 discusses how to characterize risks posed by different chemicals with the same endpoints or modes of action.

Table 4.1 Typical Exposure Equations for Contaminated Site Exposure Assessment

Inadvertent Ingestion of Contaminated Soil:

$$\text{Dose (mg/kg/d)} = \frac{C_s \times IR_s \times RAF_{GIT} \times ET}{BW}$$

Inhalation of Contaminated Soil Particles:

$$\text{Dose (mg/kg/d)} = \frac{C_s \times P_{Air} \times IR_A \times RAF_{Inh} \times ET}{BW}$$

Indoor Inhalation of Contaminant Vapours:

$$\text{Dose (mg/kg/d)} = \frac{C_{ia} \times IR_A \times RAF_{Inh} \times ET}{BW}$$

Ingestion of Contaminated Groundwater:

$$\text{Dose (mg/kg/d)} = \frac{C_{gw} \times IR_W \times RAF_{GIT} \times ET}{BW}$$

Dermal Contact with Contaminated Soil:

$$\text{Dose (mg/kg/d)} = \frac{C_S \times \Sigma\,(SA_i \times SL_i) \times RAF_{Skin} \times EF \times ET}{BW}$$

Ingestion of Contaminated Food:

$$\text{Dose (mg/kg/d)} = \frac{\Sigma\,(C_{Foodi} \times IR_{Foodi} \times RAF_{GITi} \times ET_i)}{BW}$$

Dermal Contact with Surface Water:
Determined in accordance with methods recommended by U.S. EPA, 1992b.

BW = body weight (kg)
CF_{Foodi} = concentration of contaminant in food type "*i*" (mg/kg)
C_{gw} = concentration of contaminant in groundwater (mg/L)
C_{ia} = concentration of contaminant in indoor air (mg/m^3)
C_S = concentration of contaminant in soil (mg/kg)
EF = exposure frequency (events/d)
ET = exposure term (unitless)
IR_A = air inhalation rate (m^3/d)
IR_{Foodi} = ingestion rate of food type "*i*" (kg/d)
IR_S = soil ingestion rate (kg/d)
IR_w = water ingestion rate (L/d)
RAF_{GIT} = relative absorption factor from the gastrointestinal tract (unitless)
RAF_{Inh} = relative absorption factor for inhalation (unitless)
RAF_{Skin} = relative absorption factor for skin (unitless)
SA_i = exposed skin surface area for body part "*i*" (cm^2)
SL_i = soil loading to skin for body part "*i*" (kg/cm^2/event)

Exposures to locally acting agents (such as irritants or caustic chemicals) are not normally summed across exposure routes and pathways, since these chemicals do not normally act on the same target tissues/organs via multiple exposure routes.

Sample calculations are shown in Example 4.3.

Example 4.3 Exposure Estimation – Deterministic

Problem

A site contaminated with tetrachloroethylene is being used for residential purposes. Drinking water for the residence is obtained from an on-site water well.

Concentrations of tetrachloroethylene were determined in various media by a combination of direct measurement and modelling:

C_s = 3 mg/kg in soil
C_{gw} = 2 mg/L in groundwater
C_{ia} = 0.1 mg/m^3 in indoor air

The problem formulation indicated that the main potential exposure pathways include incidental ingestion of soil, dermal contact with soil, ingestion of groundwater, and inhalation of indoor air.

The toxicity assessment did not identify a toxicity reference value for dermal exposure; therefore, a bioavailability assessment was undertaken for this pathway, and a relative dermal absorption factor of 0.1 was derived.

What is the exposure dose for each pathway for a toddler (7 months to 4 years inclusively) living at the site?

Solution

Receptor characteristics are based on "typical" values specified by Health Canada:

Body weight (BW) = 16.5 kg
Soil ingestion rate (IR_S) = 0.08 g/d
Water ingestion rate (IR_W) = 0.6 L/d
Air inhalation rate (IR_A) = 8.3 m^3/d
Exposed skin surface area – hands (SA_H) = 430 cm^2
Exposed skin surface area – arms (SA_A) = 890 cm^2
Exposed skin surface area – legs (SA_L) = 1690 cm^2
Soil loading to skin – hands (SL_H) = 1 x 10^{-4} g/cm^2/event
Soil loading to skin – arms and legs (SL_A and SL_L) = 1 10^{-5} g/cm^2/event

In this case, it is conservatively assumed that a toddler could be at the site full time; therefore, exposures were not amortized for the exposure duration.

Exposure doses are calculated below.

Incidental soil ingestion

$$\text{Dose (mg/kg/d)} = \frac{C_s \times IR_s \times RAF_{GIT} \times ET}{BW}$$

$$\text{Dose (mg/kg/d)} = \frac{3\ mg/kg \times 0.08\ g/d \times 0.001\ kg/g \times 1 \times 1}{16.5\ kg}$$

Dose = 1.5 x 10^{-5} mg/kg/d from soil ingestion

Dermal contact with soil

$$\text{Dose (mg/kg/d)} = \frac{C_S \times \Sigma (SA_i \times SL_i) \times RAF_{Skin} \times EF \times ET}{BW}$$

$$\text{Dose (mg/kg/d)} = \frac{3\ mg/kg \times (430\ cm^2 \times 10^{-4}\ g/cm^2/event + 2580\ cm^2 \times 10^{-5}\ g/cm^2/event) \times 0.001\ kg/g \times 0.1 \times 1 \times 1}{16.5\ kg}$$

Dose = 1.3 x 10^{-6} mg/kg/d from dermal contact

Ingestion of groundwater

$$\text{Dose (mg/kg/d)} = \frac{C_{gw} \times IR_W \times RAF_{GIT} \times ET}{BW}$$

$$\text{Dose (mg/kg/d)} = \frac{2\ mg/L \times 0.6\ L/d \times 1 \times 1}{16.5\ kg}$$

Dose = 0.073 mg/kg/d from groundwater ingestion

Inhalation of indoor air

$$\text{Dose (mg/kg/d)} = \frac{C_{ia} \times IR_A \times RAF_{Inh} \times ET}{BW}$$

$$\text{Dose (mg/kg/d)} = \frac{0.1\ mg/m^3 \times 8.3\ m^3/d \times 1 \times 1}{16.5\ kg}$$

Dose = 0.05 mg/kg/d from inhalation of indoor air

Total ingestion exposure dose = 1.5 x 10^{-5} + 1.3 x 10^{-6} + 0.073 = 0.073 mg/kg/d
Total inhalation exposure dose = 0.05 mg/kg/d
Total exposure dose = 1.5 x 10^{-5} + 1.3 x 10^{-6} + 0.073 + 0.05 = 0.12 mg/kg/d

It is apparent from these results that exposure occurs mainly from two pathways: ingestion of contaminated groundwater and inhalation of indoor contaminant vapours.

4.9.2 Determination of background exposure

Many of the chemicals evaluated during a risk assessment occur naturally. For example, most metals are ubiquitous in soils, water, air, and food, and PAHs can be generated naturally by forest fires. Other chemicals, such as many persistent organic pollutants, may be present throughout the environment because of human activities. Therefore, human receptors are often exposed to COPCs independent of the site being evaluated. This background exposure should be quantified during the risk assessment.

Background exposure generally occurs through all of the same pathways as exposure at the site, and frequently through additional pathways. In particular, exposure through commercial foods and consumer products can be major sources of background exposure to chemicals. Ambient concentrations of COPCs in local air, water, soil, and food items should be estimated from off-site monitoring data or published literature, where available. In situations where local data are unavailable or inadequate, background concentrations in various media should be estimated from other similar areas in Canada, where possible. National exposure estimates have been performed for some chemicals, e.g. through the *Canadian Environmental Protection Act* Priority Substances List I and List II (HC,

1994, 1996]). These assessments may be particularly useful for evaluating exposure through pathways such as supermarket food that would not be expected to vary regionally as much as exposure through soil or air.

Two different types of background exposure assessment may be performed, depending on the requirements of the jurisdiction and the needs of the risk assessment. The first is to estimate exposure occurring off site and combine it with the on-site exposure to determine the total exposure (site + background) of a human receptor to the chemical. The second is to estimate background exposure only (i.e. the exposure a human receptor would experience if the site did not exist). Performing both of these methods allows for the comparison of exposures between those affected by the site and those not affected by the site in order to determine the incremental increase in exposure resulting from the site. An example of both background exposure estimates is presented below as Example 4.4.

Example 4.4 Background Exposure

As part of a risk assessment for an aluminum-contaminated site, background exposure to aluminum is evaluated for adults with an assumed body weight of 70 kg.

First, background exposure is evaluated for a person not exposed to the contaminated site, using the following data.

Exposure Source	Aluminum Concentration	Source Intake Rate	Exposure Dose (mg/kg bw/d)
Drinking Water	0.05 mg/L	1.5 L/d	0.0011
Food			0.324
dairy products	0.78 mg/kg	0.283 kg/d	
meat, poultry, fish, eggs	3.7 mg/kg	0.183 kg/d	
cereal products	82 mg/kg	0.247 kg/d	
fruit and fruit products	2.1 mg/kg	0.186 kg/d	
vegetables	0.66 mg/kg	0.250 kg/d	
fats	3.7 mg/kg	0.025 kg/d	
nuts and dried legumes	2.9 mg/kg	0.012 kg/d	
food, primarily sugar	2.8 mg/kg	0.057 kg/d	
mixed dishes and soup	0.72 mg/kg	0.100 kg/d	
soft drinks and alcohol	2.3 mg/kg	0.255 kg/d	
Outdoor Air*	0.0005 mg/m^3	2 m^3/d	0.001
Indoor Air*	0.0013 mg/m^3	13.8 m^3/d	0.018
Soil	61,000 mg/kg	0.00002 kg/d	1.22
Total Intake			**1.56**

* The dose is based on an assumption of 3 hours/day spent outdoors and 21 hours/day spent indoors.

Following the determination of background exposure for the general Canadian population, background exposure can be evaluated for people who are exposed to aluminum at the contaminated site. If a worker spends 8 hours/day, 5 days/week, 48 weeks/year at the site (site exposure term = 8/24 x 5/7 x 48/52 = 0.22), it would be appropriate to adjust background exposure estimates for certain pathways by a factor of 0.78 (1–0.22). Unless food and water are obtained from the site, background exposure from these pathways would not be adjusted; only those pathways for which exposure is being evaluated at the site should be adjusted. In this case, indoor air (0.018 mg/kg/d x 0.78 = 0.014 mg/kg/d) and soil exposure (1.22 mg/kg/d x 0.78 = 0.95 mg/kg/d) might be adjusted, resulting in an estimated exposure of 1.29 mg/kg bw/d, plus the estimated exposure occurring at the site.

Sources: Data adapted from Environment Canada/Health Canada, 2000; Health Canada, 1994.

For chemicals causing adverse health effects through a non-threshold mechanism (e.g. genotoxic carcinogens), the use of a background exposure assessment allows for a more complete estimate of the incremental risks associated with the site. For chemicals causing adverse health effects through a threshold mechanism, the use of a background exposure assessment allows the determination of whether the exposure from the site, when combined with the background exposure, results in exceeding the threshold.

This information can be very useful to risk managers, particularly in cases where background exposure concentrations themselves may be sufficient to be associated with adverse health effects, and where on-site exposure might result in an increased exposure that might be deemed biologically insignificant given the range of background exposure across Canada or within the local area.

4.9.3 Determination of lifetime and life-stage exposure

Using the approach outlined in the previous sections, exposure may be estimated for all relevant age groups of concern (e.g. infant, toddler, child, teen, and adult). However, for threshold contaminants, exposure may be estimated only for the most sensitive receptors. For threshold-response chemicals, it may be appropriate simply to express the exposure estimates as the individual life-stage dose rates. These will then be evaluated against the toxicity estimates during the risk characterization (section 6.0). However, for non-threshold-response chemicals, it may be appropriate to provide lifetime estimates of risks based on lifetime exposure rates. The most straightforward approach is to calculate exposure for each life stage, then determine a weighted-average exposure based on the duration of each life stage.

4.10 Probabilistic Exposure Estimation

The preceding section describes the estimation of exposure in the context of a deterministic risk assessment. Exposure estimation may also be conducted probabilistically, as part of a probabilistic risk assessment. Key assumptions in the exposure equations would be assigned probability distributions representing the range and frequency of all possible values of equation variables (representing either inherent variability or uncertainty), and a probability distribution of estimated exposures would be generated. Further description of probabilistic assessment is presented in section 7.0.

4.11 Alternate Methods of Estimating Exposure

Traditional exposure estimation, based on concentrations of chemicals in environmental media and receptor physical/behavioural characteristics, is not the only method of evaluating exposure to chemicals. Two alternate methods, personal exposure monitoring and biological monitoring, are presented in the following text. These methods are appropriate only for evaluating existing exposures under certain conditions and are not commonly used in risk assessment of contaminated sites.

4.11.1 Personal exposure monitoring

Personal exposure monitoring is a method whereby exposure is measured directly, rather than estimated based on chemical concentrations in the environment and an assumed exposure scenario.

Exposure is measured using devices collectively referred to as personal exposure monitors. Examples include various dosimeters designed to monitor workplace exposure to hazardous chemicals (typically airborne exposure to volatile organic chemicals or gases such as carbon monoxide).

Personal exposure monitoring can also be conducted for food or water, although this generally involves an individual actively testing water or food prior to ingestion, and measuring all amounts (mass or volume) of such media that are ingested.

The primary advantage of personal exposure monitoring is that exposure to an individual can be directly measured, reducing the uncertainty in the exposure estimate. However, this approach also has several limitations.

- Exposure can be determined only for the time period during which the monitoring takes place; future and historical exposure cannot be readily predicted.
- Personal exposure monitors are available only for certain chemicals and only for certain exposure routes (typically air inhalation, and sometimes food and water ingestion).
- Personal exposure monitoring requires active cooperation from the receptors.
- Most personal exposure monitors cannot distinguish the source of the chemical, and many are subject to interference from similar chemicals.
- Results apply only to the individual(s) being evaluated; large numbers of individuals may need to be monitored in order to extrapolate the results to a population.
- Costs can be high, especially when evaluating a large exposed population.

Despite these limitations, when used appropriately, personal exposure monitoring can be a valuable tool for conducting exposure assessments, possibly in combination with a more traditional exposure assessment based on environmental concentrations.

4.11.2 Biological monitoring

Biological monitoring involves measuring indicators of exposure on members of the exposed or potentially exposed population. This includes measurement of a chemical or its metabolites in blood, urine, breast milk, body fat, exhaled air, hair, or fingernails/toenails; in some cases, the amount of a chemical or metabolite bound to target molecules is measured instead. Biological monitoring also includes measurement of the effects of exposure on receptors.

Biological monitoring can demonstrate that exposure has taken place and that absorption of the chemical has occurred. When used appropriately, it can also be used to estimate the quantity of the chemical absorbed. However, this method has its own limitations.

- It must be possible to establish a relationship between exposure or risk and the parameter being measured;

currently, this can be reliably done only for a small number of substances.

- Reference results (e.g. biological monitoring results for a similar population not affected by a contaminated sites) are generally necessary.
- Active cooperation of human receptors (and potentially a control population) is needed.
- Costs may be high, especially when evaluating a large exposed population.
- Established and accurate biological test methods must be available for the chemical being evaluated.
- In-depth knowledge of the chemical fate in the body is required (i.e. where the chemical accumulates, body residence time, metabolites, whether toxicity is caused by the chemical or a metabolite, etc.).
- Generally no information about the source of exposure is obtained.
- Timing of the monitoring is critical, especially for substances with short body residence times.

When measuring chemical or metabolite concentrations in tissues/fluids, it is important to ensure that the appropriate tissue or fluid is selected. The following examples are adapted from the enHealth Council (2002).

- **Blood**, depending on the biological residence time of the chemical involved, may reflect either recent exposure (within a few hours) or longer-term exposure over several years.
- **Urine** generally reflects integrated exposure over recent hours or days.
- **Hair, fingernails, and/or toenails** provide integrated exposure estimate over an extended period of time; however, external contamination of hair in particular cannot generally be distinguished from contamination accumulating within hair.
- **Breast milk** provides an integrated exposure assessment for a time period related to the body residence time of the chemical; often used for persistent organic pollutants.
- **Exhaled air** can be used to evaluate recent exposure to volatile chemicals such as ethanol and some solvents.

Biological monitoring, in combination with monitoring for effects of exposure (e.g. health monitoring) is more commonly performed for epidemiological studies or occupational monitoring programs than traditional risk assessments, because in most cases a high level of exposure is required to produce a distinguishable effect.

4.11.3 Pharmacokinetic modelling

Pharmacokinetic modelling refers to the use of mathematical models to predict the concentration of a chemical in different body compartments (e.g. blood, kidney), often on a time-dependent basis. The modelling accounts for physiological processes occurring within the body, as well as the characteristics of the chemical. The models typically divide the body into several compartments for modelling purposes.

Many of the pharmacokinetic models used in risk assessment, such as the U.S. EPA *Integrated Exposure and Uptake Biokinetic Model (IEUBK) for Lead in Children* (U.S. EPA, 1994), evaluate both the exposure to the chemical at the site and its fate within the body.

The main advantage of these models is the ability to predict concentrations of the chemicals in body tissues; this may be more meaningful than an administered dose estimate and more reliable than a dose adjusted by a relative absorption factor. Pharmacokinetic models include the following limitations.

- Relatively large amounts of input data may be needed.
- Human pharmacokinetics is often estimated from data for other mammalian species.
- The uncertainties associated with such modelling are largely unquantifiable and/or ignored.
- Detailed understanding of the chemical's fate in the body is necessary.
- Extensive data validation is required to ensure that the model performs reliably.
- It must be possible to relate the predicted tissue concentrations of a chemical to a toxicological response.

4.12 Recommended Deliverables

By using scientifically valid methods for estimating exposure (such as those discussed in this section), the exposure assessment provides estimates of the rates of exposure of human receptors (categorized by age) to the COPCs at a specific site. Exposure estimates may be presented as point estimates (in the case of deterministic analysis techniques) or as distributions (in the case of probabilistic analysis techniques). Depending on how a particular chemical acts on biological systems, exposure estimates should be expressed in the form of a dose rate (for systemically acting chemicals) or as exposure concentration/frequency/duration estimates (for locally acting chemicals). Depending on the requirements of the regulatory authority, exposure estimates may be required for a number of different scenarios.

In communicating the results of the exposure assessment, the risk assessor should present in a suitable format a list of the dose rates for each receptor, for each chemical, for each exposure route, and for each scenario. In addition, all bioavailability factors, if used, should be listed and their derivations documented. All formulae used in the calculation of exposures should be described, and referenced for explanations of their derivation provided. Point values and/or distributions for all parameters that are used as input variables in the exposure formulae should be listed, with the rationale provided, and their sources documented.

4.13 References

Ambrose A.M., D.S. Larson, J.R. Borzelleca, and G.R. Hennigar Jr. 1976. Long-term toxicologic assessment of nickel in rats and dogs. *J. Food Sci. Technol*. 13: 181–187.

AMEC. 2005. *Summary of Absorption Efficiency Values for TRV Compounds Reported by Health Canada*. Contract report prepared for Health Canada.

American Petroleum Institute (API). 1998. *Assessing the Significance of Subsurface Contaminant Vapor Migration to Enclosed Spaces: Site-Specific Alternatives to Generic Estimates*. API, Washington, DC. API Publication 4674.

API 2005. *Collecting and Interpreting Soil Gas Samples from the Vadose Zone: A Practical Strategy for Assessing the Subsurface Vapor-to-Indoor-Air Migration Pathway at Petroleum Hydrocarbon Sites*. API, Washington, DC. API Publication 4741.

American Society for Testing of Materials (ASTM). 1992. *Standard Guide for Soil Gas Monitoring in the Vadose Zone*. ASTM, West Conshohocken, PA . ASTM D 5314-92.

ASTM. 2002. *ASTM Standards on Environmental Site Characterization*, 2nd Edition. ASTM, West Conshohocken, PA.

Agency for Toxic Substances and Disease Registry (ATSDR). 2005a. *ATSDR Public Health Assessment Guidance Manual*. US Department of Health and Human Services, Public Health Service, Atlanta. Last revised January 2005.

ATSDR. 2005b. *Toxicological Profile for Nickel*. U.S. Department of Health and Human Services, Public Health Service, Atlanta.

Bodger, K. 2003. *Fundamentals of Environmental Sampling*. Government Institutes, Blue Ridge Summit, PA.

Bright, D.A., G.M. Richardson, and M. Dodd. 2006. Do current standards of practice measure what is relevant to human exposure at contaminated sites? I: A discussion of soil particle size and contaminant partitioning in soil. *Human Ecol. Risk Assess*. 12(3): 591–605.

Calabrese, E.J., E.J. Stanek, R.C. James, and S.M. Roberts. 1997. Soil ingestion: A concern for acute toxicity in children. *Environ. Health Perspect*. 105: 1354–1358.

Canadian Council of Ministers of the Environment (CCME). 1993. *Guidance Manual on Sampling, Analysis, and Data Management Guidelines for Contaminated Sites. Vol. I: Main Report* (CCME EPC-NCS62E), and *Vol. II: Analytical Method Summaries.* CCME, Winnipeg. CCME EPC-NCS66E.

Claiborn, C., A. Mitra, G. Adams, L. Bamesberger, G. Allwine, R. Kantamaneni, R. Lamb, and H. Westberg. 1995. Evaluation of PM10 emission rates from paved and unpaved roads using tracer techniques. *Atmos. Environ*. 29(10): 1075–1089.

Edwards R.D., and P.J. Lioy. 1999. The EL sampler: A press sampler for the quantitative estimation of dermal exposure to pesticides in housedust. *J. Exp. Anal. Environ. Epidemiol.* 9: 521–529.

enHealth Council. 2002. *Environmental Health Risk Assessment: Guidelines for Assessing Human Health Risks from Environmental Hazards*. Australia National Public Health Partnership, Canberra.

Environment Canada/Health Canada (EC/HC). 2000. *Canadian Environmental Protection Act, 1999. Priority Substances List, State of the Science Report for Aluminum Chloride, Aluminum Nitrate and Aluminum Sulfate*. Environment Canada/Health Canada, Ottawa.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Geological Survey of Canada (GSC). 2006. *Geochemical Background in Soil and Till*. Prepared by A.N. Renz, R.G. Garrett, S.W. Adcock, and G.F. Bonham-Carter. GSC, Ottawa.

Health Canada (HC). 1994. *Canadian Environmental Protection Act. Human Health Risk Assessment for Priority Substances*. Minister of Supply and Services Canada, Ottawa. Cat. No. En40-215/41E; ISBN 0-662-22126-5.

HC. 1996. *Canadian Environmental Protection Act – Priority Substances List Supporting Documentation: Health-Based Tolerable Daily Intakes/Concentrations and Tumorigenic Doses/Concentrations for Priority Substances*. Environmental Health Directorate, Health Protection Branch, Ottawa. Cat. No. H46-2/96-194E.

HC. 2010a. *Federal Contaminated Site Risk Assessment in Canada, Part VII: Guidance for Soil Vapour Intrusion*

Assessment at Contaminated Sites. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

HC. 2010b. *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA)*, Version 2.0. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

HC. 2010c. *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, Version 2.0. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

Jacques Whitford Environmental Ltd. (JWEL). 2005. *Ingestion Bioavailability of Arsenic, Lead and Cadmium in Human Health Risk Assessments: Database, Critical Review, and Recommendations*. Contract report to the Contaminated Sites Division, Health Canada, Ottawa.

Lioy, P.J., N.C.G. Freeman, T. Wainman, A.H. Stern, R. Boesch, T. Howell, and S.I. Shupack. 1992. Microenvironment analysis of residential exposure to chromium-laden wastes in and around New Jersey homes. *Risk Anal.* 12: 287–299.

Paustenbach D.J., G.M. Bruce, and P. Chrostowski. 1997. Current views on the oral bioavailability of inorganic mercury in soil: Implications for health risk assessment. *Risk Anal*. 17(5): 533–544.

Rasmussen, P.E., G.C. Edwards, R.J. Kemp, D.R. Fitzgerald-Hubble, and W.H. Schroeder. 1998. Towards an improved natural sources inventory for mercury. In: *Proceedings of Metals and the Environment: an International Symposium*. J. Skeaff (ed.). CIM Annual General Meeting, May 1998.

Richardson, G.M. 1997. *Compendium of Canadian Human Exposure Factors for Risk Assessment*. O'Connor Associates Environmental Inc., Ottawa.

Richardson, G.M., D.A. Bright, and M. Dodd. 2006. Do current standards of practice measure what is relevant to human exposure at contaminated sites? II: Oral bioaccessibility of contaminants in soil. *Human Ecol. Risk Assess*. 12(3): 606–616.

Richardson, G.M., I.A. Mitchell, M. Mah-Paulson, T. Hackbarth, and R.G. Garrett. 2003. Natural emissions of mercury to the atmosphere in Canada. *Environ. Rev*. 11: 17–36.

Roberts, S.M., W.R. Weimar, J.R.T. Vinson, J.W. Munson, and R.J. Bergeron. June 2002. Measurement of arsenic bioavailability in soil using a primate model. *Toxicol. Sci.* 67: 303–310.

Schwab, M., A. McDermott, and J.D. Spengler. 1992. Using longitudinal data to understand children's activity patterns in an exposure context: Data from Kanawha County Health Study. *Environ Inter*. 173–189.

United States Environmental Protection Agency (U.S. EPA). 1989. *Risk Assessment Guidance for Superfund (RAGS): Volume I – Human Health Evaluation Manual (Part A). Interim Final*. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002.

U.S. EPA. 1991. Site *Characterization for Subsurface Remediation*. EPA seminar publication. Office of Research and Development. Washington, DC. EPA/625/4-91/026.

U.S. EPA. 1992a. *Risk Assessment Guidance for Superfund: Volume I – Human Health Evaluation Manual (Part B, Development of Risk-based Preliminary Remediation Goals).* Washington, DC. EPA/54-/R-92/003.

U.S. EPA. 1992b. *Dermal Exposure Assessment: Principles and Applications*. Exposure Assessment Group, Office of Health and Environmental Assessment, Washington, DC. EPA/600/8-91/011B.

U.S. EPA. 1994. *Integrated Exposure and Uptake Biokinetic Model (IEUBK) for Lead in Children.* Accessed December 10, 2009, from http://www.epa.gov/superfund/health/contaminants/lead/products.htm.

U.S. EPA. 1995. *Compilation of Air Pollutant Emission Factors.* Office of Transportation and Air Quality, Washington, DC. AP-42, Fifth Edition.

U.S. EPA. 1997. *Expedited Site Assessment Tools for Underground Storage Tank Sites: A Guide for Regulators*. U.S. EPA, Office of Underground Storage Tanks, Office of Solid Waste Emergency Response, Washington, DC. EPA/510/B/97/001.

U.S. EPA. 1999. *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air – Second Edition*. Center for Environmental Research Information, Cincinnati, OH. EPA/625/R-96/010b.

U.S. EPA. 2004. *Risk Assessment Guidance for Superfund. Volume I: Human Health Evaluation Manual – Part E, Supplemental Guidance for Dermal Risk Assessment.* Office of Superfund Remediation and Technology Innovation, Washington, DC. EPA/540/R/99/005.

U.S. EPA. 2008. *Child-Specific Exposure Factors Handbook.* Final report. National Center for Environmental Assessment, Washington, DC. EPA/600/R-06/096F.

U.S. EPA. 2009. *Exposure Factors Handbook: 2009 Update.* External review draft. U.S. EPA, Washington, DC. EPA/600/R-09/052A.

Wester, R.C., H.I. Maibach, L. Sedik, J. Melendres, and M. Wader. 1993. In vivo and in vitro percutaneous absorption and skin decontamination of arsenic from water and soil. *Fund. Appl. Toxicol.* 20: 336–340.

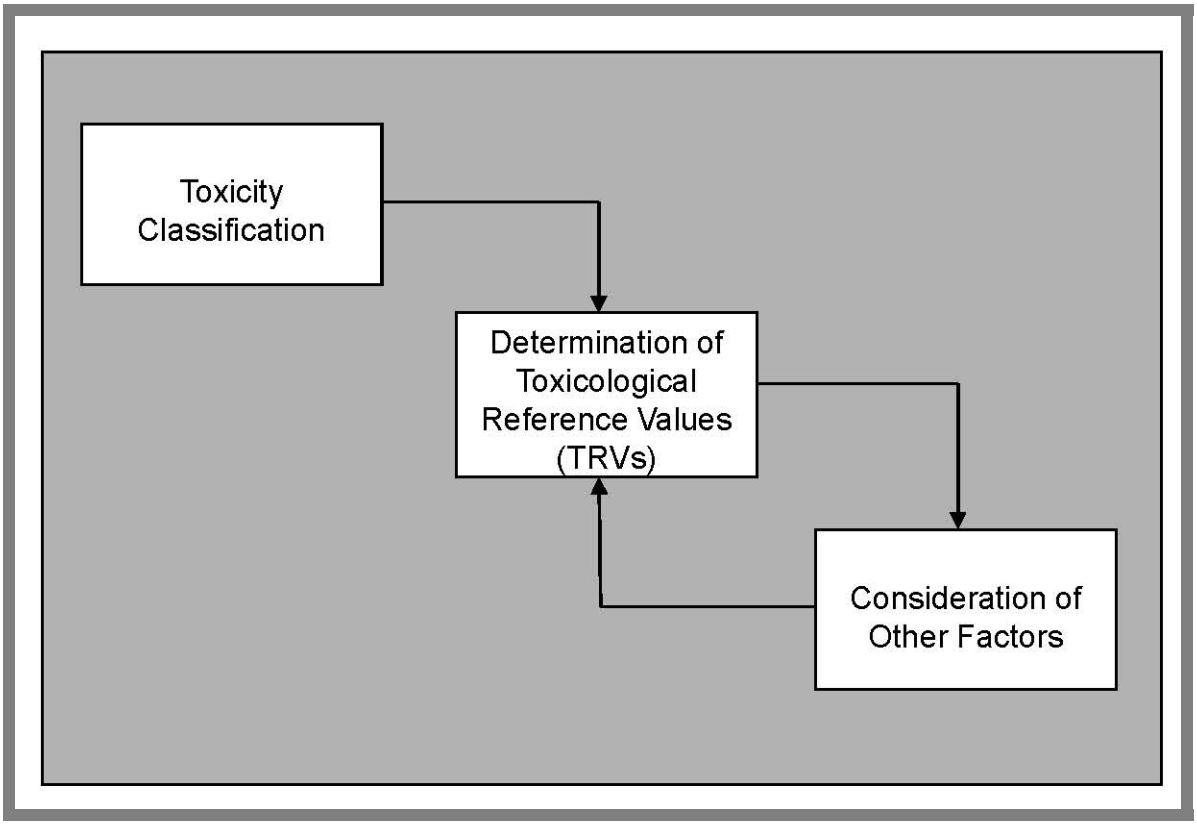
Whitmyre, G.K., J.H. Driver, M.E. Geneva, R.G. Tardiff, and S.R. Baker. 1992. Human exposure assessment II: Quantifying and reducing the uncertainties. *Toxicol. Indus. Health* 8: 321–342.

5.0 TOXICITY ASSESSMENT

5.1 Introduction and Linkages to Other Risk Assessment Tasks

In the context of HHRA, the toxicity assessment stage involves identifying the potential toxic effects of COPCs and establishing TRVs with which to characterize potential risks. It is normally conducted for all identified COPCs and considers all possible toxic effects associated with different routes of exposure for the identified receptor groups, as well as sensitive receptors. Depending on the mechanism of toxicity, the toxicity assessment provides either an estimate of how much exposure to a chemical can occur without any anticipated adverse health effects (threshold effect chemicals), or establishes a relationship between the exposure dose of a chemical and the probability of developing an adverse health effect such as cancer (non-threshold effect chemicals). The basic components of toxicity assessment are presented in Figure 5.1.

Figure 5.1 Components of Toxicity Assessment



The toxicity assessment is conducted after the problem formulation stage (section 3.0), and often concurrently with the exposure assessment (section 4.0). The toxicity assessment considers the CSM of the site (exposure pathways, receptors, and chemicals of concern) developed during the problem formulation because TRVs are often

exposure-route specific and occasionally specific to certain sensitive receptors. The predicted exposures to chemicals from the site are compared to the TRVs in order to estimate and characterize the potential health risks to the identified human receptors (section 6.0).

Although it is a separate step, the toxicity assessment should be conducted in conjunction with the exposure assessment. Information determined during the exposure assessment, such as exposure duration (short term versus long term), can affect the toxicity assessment; in turn, the mechanisms of toxic action (e.g. local versus systemic) can affect how the exposure assessment is performed. The TRVs and exposure doses must be compatible with each other (i.e. if the exposure is expressed as a daily dose per unit body weight, the TRV should be in the same form).

5.2 Objectives of Toxicity Assessment

The purpose of the toxicity assessment is to define the toxic potential of the identified COPCs; specifically, there are two major objectives.

1. Identify the potential toxicological effects associated with the COPCs.
2. Select or develop TRVs that can be used in combination with the calculated exposures or doses for risk characterization.

The TRVs can take the form of (i) a tolerable exposure dose such as a TDI, (ii) a tolerable or reference air (and sometimes water) concentration, (iii) a risk-specific dose (RSD), or (iv) a toxic potency factor such as a cancer slope factor (SF). Types of TRVs are discussed further in section 5.5.1.

5.3 Steps in Toxicity Assessment

There are two main steps in the toxicity assessment:

- classification of chemicals based on toxicological action – chemicals are classified based on whether or not there is an exposure threshold below which adverse health effects are not expected; and
- determination of TRVs – reference values are established to quantify the potential risk of toxic effect that could result from exposure to the COPCs.

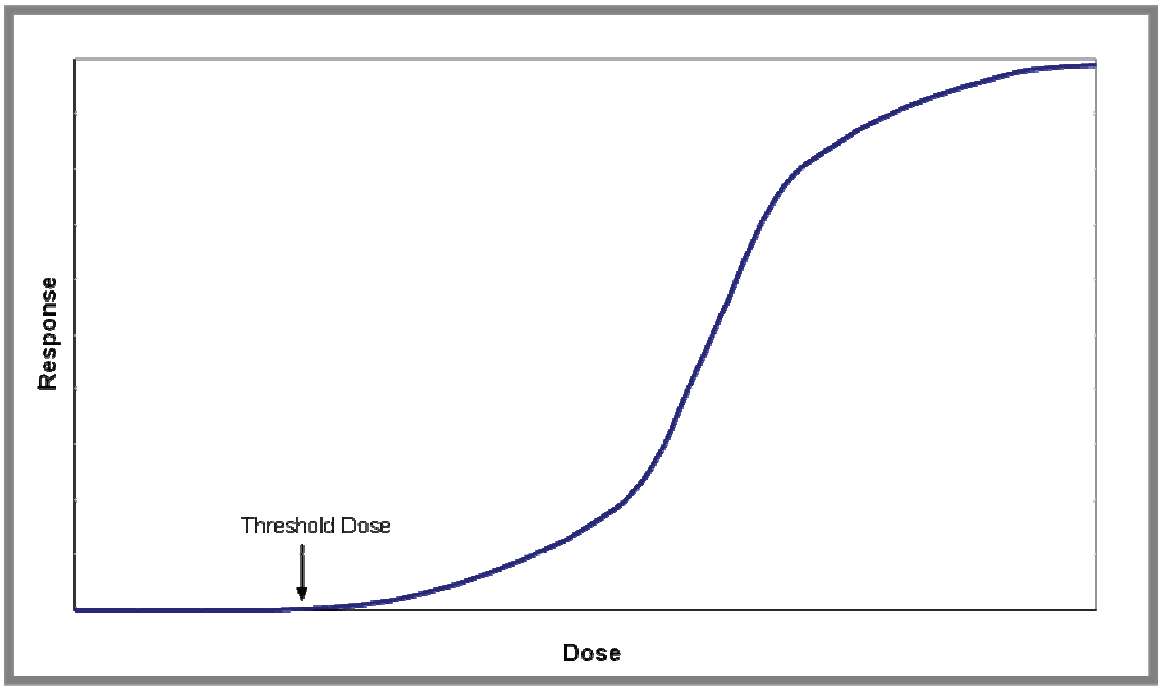
The toxicity assessment is performed for each COPC and each exposure route. As outlined in section 5.6, various factors that can affect toxicity must also be considered.

5.4 Classification of Chemicals Based on Toxicological Action

There are several ways in which toxic chemicals can be classified, such as by effect (e.g. mutagen), by target organ (e.g. hepatotoxin), or by mechanism of action (e.g. cholinesterase inhibitor). Although these classifications provide valuable information, for purposes of the toxicity assessment the key classification is based on the type of dose-response relationship, specifically whether the chemical is considered to be a threshold or non-threshold chemical. A threshold (or threshold-response) chemical is one that is considered to show adverse health effects only once a certain dose (the threshold) is exceeded. A non-threshold chemical is considered to have some potential to cause adverse health effects at any dose. Health Canada normally applies this latter classification to genotoxic carcinogens and genotoxic teratogens (see Health Canada, 1996). The type of dose-response relationship (threshold or non-threshold) determines the method and assumptions used for deriving TRVs.

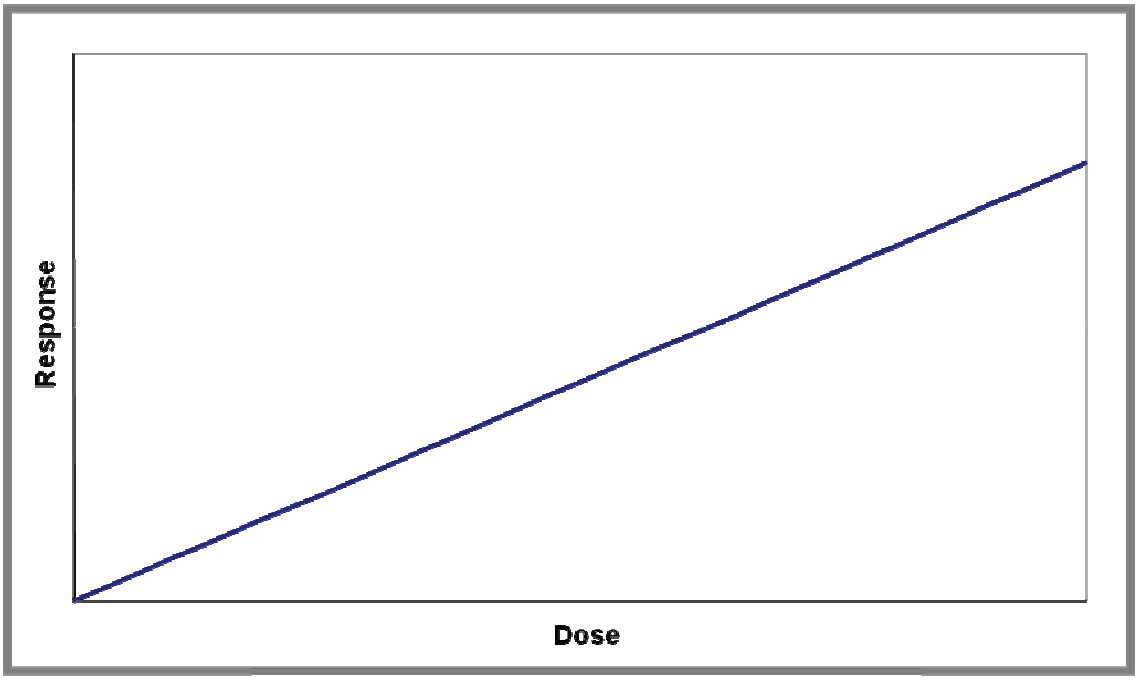
Threshold chemicals exhibit a non-linear dose-response curve (Figure 5.2) with a clear threshold dose below which no toxic effects are observed. In practice, the threshold dose cannot be determined exactly, but is represented by a "no observed adverse effect level" (NOAEL) established during toxicity testing.

Figure 5.2 Dose-Response Curve for a Threshold Chemical



Non-threshold chemicals are assumed to have dose-response curves that do not exhibit a clear threshold dose (Figure 5.3); any dose other than zero is believed to have some potential for producing a toxic effect. Non-threshold chemicals generally cause toxicity through mechanisms that result in self-propagated lesions (i.e. lesions to genetic material that once they occur can continue to progress to a disease endpoint, such as cancer or certain birth defects, even if exposure to the chemical should cease). Toxicological data for non-threshold chemicals routinely do, in fact, define a NOAEL. Toxicological bioassays and human epidemiological studies are never statistically powerful enough (primarily because of limits on the number of doses tested and number of animals or humans in those studies) to detect or observe cancers, for example, at low environmentally relevant exposure levels. Where it is determined that a chemical acts through a genotoxic mechanism, a non-threshold dose-response relationship is merely **assumed**.

Figure 5.3 Dose-Response Curve for a Non-Threshold Chemical



If published TRVs are being used (see section 5.5.2), then the classification as a threshold or non-threshold chemical will normally have been undertaken by the jurisdiction publishing the TRVs (e.g. Health Canada, U.S. EPA, ATSDR, WHO). If new TRVs are being developed for the risk assessment (see section 5.5.3), where no recognized jurisdiction has so far classified the chemical, then a weight-of-evidence assessment must be undertaken as the first step in developing a TRV.

It is important to note that some chemicals can exhibit both threshold and non-threshold effects. For some chemicals, different dose-response classifications will apply for different exposure routes; for example, cadmium is currently classified by Health Canada as a threshold chemical for oral exposure and a non-threshold chemical for inhalation exposure. Furthermore, genotoxic carcinogens may also have threshold effects associated with them. Although the non-threshold effect is the most critical response for chronic exposure in most cases, a threshold effect may be more critical for shorter exposure durations. The risk assessor must ensure that the classification used is appropriate for the exposure route, conditions, and exposure duration. If there is any doubt as to the most critical effect, then risks should be evaluated based on both the threshold and non-threshold effects.

5.5 Determination of Toxicological Reference Values (TRVs)

For most HHRAs, TRVs are obtained from published sources; for Canadian federal contaminated sites, TRVs published by Health Canada will be used where available. However, the supporting documentation for these TRVs should be reviewed to ensure that they are current and appropriate for the exposure scenario occurring at the site. De novo determination of TRVs should only be undertaken by individuals qualified and experienced in toxicology and in consultation with Health Canada.

5.5.1 Types of TRVs

TRVs are commonly reported in several different formats, and it is possible to express the same degree of toxicity in different ways. Although different formats may represent equivalent toxicity, they are applied differently during the risk characterization. Common formats for the representation of chronic TRVs follow.

5.5.1.1 Tolerable daily intake

The TDI represents the maximum dose of a threshold substance to which an individual could be exposed daily over a lifetime without any expected deleterious effects (Health Canada, 1996). It is expressed as the amount of substance per unit body weight per unit time (i.e. mg/kg bw/d). Some regulatory agencies may refer to this type of TRV as a reference dose (RfD) or acceptable daily intake (ADI).

5.5.1.2 Tolerable concentration

The tolerable concentration (TC) represents the maximum concentration (usually airborne) of a threshold substance to which a person may be continually exposed over a lifetime without any expected deleterious effects (Health Canada, 1996). It is expressed as a concentration (e.g. $\mu g/m^{3)}$. Some regulatory agencies may refer to this type of TRV as a reference concentration (RfC).

5.5.1.3 Slope factor

A slope factor (SF) relates the exposure dose of a non-threshold substance to the expected probability of developing cancer. It is expressed as the inverse of a dose – e.g. $(mg/kg\ bw/d)^{-1}$ – and quantifies the number of predicted cancers per unit dose. The exposure dose multiplied by the SF is the expected cancer risk. The SF is referred to by some agencies as a cancer potency factor, and denoted as q_1^*.

5.5.1.4 Unit risk

A unit risk (UR) represents the amount of risk per unit concentration of a non-threshold substance to which an individual is continually exposed. It is expressed as the inverse of a concentration – e.g. $(mg/L)^{-1}$ – and is commonly used for both air and water concentrations. The amortized chemical concentration multiplied by the UR is the expected cancer risk.

5.5.1.5 Risk-specific dose

A risk-specific dose (RSD) is the dose of a non-threshold substance that is expected to lead to a specified cancer risk. It is expressed in the same units as the exposure dose (i.e. mg/kg bw/d); it is also essential that the risk associated with the RSD is specified (e.g. RSD for a cancer risk of 1×10^{-5}). The tumorigenic dose 05 (TD_{05}) values published by Health Canada (1996) for some substances are RSDs for a 5% cancer risk.

5.5.1.6 Risk-specific concentration

A risk-specific concentration (RSC) is the concentration of a non-threshold substance that is expected to lead to a specified cancer risk from continual exposure. It is expressed as a concentration (e.g. $\mu g/m^3$), and generally relates to concentrations in air, although it can be applied for concentrations in water or other media. As for the RSD, it is essential that the risk associated with the RSC is also specified (e.g. RSC for a cancer risk of 1×10^{-5}). The tumorigenic concentration 05 (TC_{05}) values published by Health Canada (1996) for some substances are RSCs for a 5% cancer risk.

Subchronic and acute TRVs are discussed in section 5.5.4.

5.5.2 TRVs from published sources

The most straightforward approach to determining TRVs is to use values published by an appropriate regulatory body. For sites under Canadian federal jurisdiction, TRVs published in *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors,* Version 2.0. (HC, 2010) should be employed where available. In the absence of a Health Canada TRV, values can be obtained from other regulatory agencies. TRVs published by the following agencies are generally accepted by Health Canada, in the following order of preference:

1. Other Health Canada TRVs
2. U.S. EPA Integrated Risk Information System (IRIS):

   http://www.epa.gov/iris/
3. World Health Organization (WHO) – various sources, including:

   http://www.inchem.org/

   http://jecfa.ilsi.org/index.htm
4. Netherlands National Institute of Public Health and the Environment (RIVM)

   http://www.rivm.nl/bibliotheek/rapporten/711701025.pdf
5. Agency for Toxic Substances and Disease Registry (ATSDR)

   http://www.atsdr.cdc.gov/toxprofiles/index.asp
6. California Environmental Protection Agency (Cal EPA)

   http://www.oehha.ca.gov/risk/ChemicalDB/index.asp

TRVs developed by other jurisdictions or in peer-reviewed scientific literature may also be acceptable with sufficient justification and supporting data. However, because different jurisdictions use different methods for the development of TRVs, any TRVs adopted from sources other than Health Canada must be evaluated for consistency with the level of protection adopted by Health Canada.

If a federal site has off-site impacts on properties under the regulatory authority of another jurisdiction (e.g. province or territory), that other jurisdiction may have its own TRVs or preferences for sources; the recommendations of any other regulatory authorities with interest in the site should be determined before finalizing the TRVs. In these cases, it may prove necessary and appropriate to select the most conservative TRVs from conflicting sources so as to satisfy both jurisdictions that risks are not underestimated.

It is essential that the supporting information for the TRVs be reviewed to ensure they are appropriate prior to adoption. In

particular, TRVs must be appropriate for the exposure route(s) and exposure duration as determined during the exposure assessment (section 4.0). If no appropriate TRVs are identified, then published TRVs may require modification, or new TRVs may be needed. It is also important that the TRVs selected be as up-to-date as possible.

5.5.3 De novo determination of TRVs

If no published TRVs are available, or if there is compelling evidence that the published TRVs are inappropriate (e.g. outdated or based on a different exposure route or chemical form), then new TRVs may be required. De novo development of TRVs should only be undertaken by individuals qualified and experienced in toxicology, and only after Health Canada has been consulted.

Procedures for de novo determination of TRVs are detailed in Appendix B.

5.5.4 Acute and subchronic TRVs

In many ways, the approach for addressing acute and subchronic TRVs is similar to that described for chronic TRVs. A similar hierarchical approach is recommended whereby values provided by Health Canada are considered first, followed by the U.S. EPA, WHO, RIVM, ATSDR, and Cal EPA. In addition, other jurisdictions and the peer-reviewed literature can also be considered. Finally, if no TRV is available, it is possible to determine a de novo TRV using the procedures detailed in Appendix B.

Notwithstanding the preceding, there are some important considerations in identifying and applying TRVs for less-than-lifetime exposures. First, the establishment of acute and subchronic TRVs has not been a priority for most chemicals evaluated by health agencies. Although acute and subchronic TRVs are often available for irritant chemicals in air and certain systemically acting chemicals in the environment, most efforts of health agencies have been directed toward the estimation of acceptable exposures for lifetime exposures (e.g. chronic TDIs, RfDs, and ADIs). Of the health agencies listed in the preceding paragraph, only the ATSDR routinely provides TRVs for acute and subchronic duration (but even the ATSDR does not provide TRVs for all COPCs). Consequently, it is not uncommon to find situations where an acute or subchronic TRV is not available for a chemical. Nevertheless, this does not mean that acute and subchronic health risks should be ignored; there are alternatives that can be used when appropriate TRVs are not available.

Second, it is important that acute and subchronic TRVs should be established for a time period that is at least as long as the exposure period for human receptors being considered in the risk assessment. In the estimation of risks from chronic exposures to chemicals, it is often not possible to ensure that environmental exposures occur for durations less than the toxicity studies used to establish the chronic TRVs (i.e. laboratory exposures in mice and rats typically do not occur for more than 72 and 104 weeks, respectively). Although this approach is considered to be acceptable for chronic risk assessment, Health Canada recommends that the duration of environmental exposures being evaluated in the risk assessment should be equal to or less than the duration of toxicity studies used to establish the acute or subchronic TRVs. In other words, if a person is being exposed to a chemical for 2 months, it is not appropriate to use a 3 week toxicity study to evaluate health risks. Owing to the nature of possible mechanisms of action, it is possible that a given dose rate of a chemical may only manifest toxicity after a certain duration of exposure because:

- some chemicals only cause toxicity after sufficient time for repair mechanisms to be overcome has occurred (e.g. a 3 week exposure may allow for sufficient repair that cannot occur after 2 months of exposure); and
- some chemicals may cause damage after a single exposure event, but only after sufficient events will the toxicity be readily observed (e.g. damage caused by a 3 week exposure may not be readily detectable, whereas the damage may be easily detectable after 2 months of exposure).

It is for the above reasons (and others) that acute and subchronic exposures involving human receptors should not be compared to TRVs developed for shorter durations. Instead, the acute and subchronic TRV should match as closely as possible the duration of exposure that humans will receive from the site of concern.

In cases where TRVs are not available for sufficient durations, it is usually prudent to use a TRV from the next highest duration. For example, if a subchronic TRV was not available for addressing 1 month exposures to mercury, the default assumption would be to use the chronic TRV for assessment of human health risks. Although qualitatively it should be clear that a subchronic exposure is usually less severe than a chronic duration exposure, the risks from such exposures should only be quantified with precise toxicological data and/or TRVs that are developed for the specific duration of interest. Thus, in lieu of not having acute or subchronic TRVs available, the risk assessor should complete the risk assessment by comparing acute or subchronic exposures to chronic TRVs.

Somewhat related to the above, it is recommended that exposures not be spread out over too long a duration that would result in the TRVs being used out of context. Once again, the TRV should match as closely as possible the duration of exposure that humans will receive from the site. It is important that amortization of exposures does not underestimate potentials for exceeding threshold effects. For

example, if a person spends 1 month per year at a summer camp that is impacted by lead, it would not be appropriate to spread or amortize the daily dose that person receives during the 1 month period over the entire year (i.e. from a toxicological perspective, 12 μg lead/kg bw/d for 1 month is not equivalent to 1 μg/kg bw/d for an entire year). In this case, the risk assessor would have to either identify a TRV that addresses repeat subchronic exposures to lead (those that occur 1 month per year on an ongoing lifetime basis) or else compare the exposures that occurred during the 1 month period to the chronic TDI for lead. In the latter scenario, the risk assessor could discuss qualitatively how the risk from the 1 month per year exposure is likely overestimated by the use of the chronic TDI that was established to assess daily exposures over the entire year. However, without precise toxicological information addressing the time duration of concern, it is not recommended that amortization of dose rates beyond the exposure period be used to develop quantitative risk estimates.

In cases where acute and subchronic TRVs are identified, it should be clear that such values are usually not meant for the assessment of repeated exposures. Many acute and subchronic TRVs have been established in laboratory animals on exposures that occurred for periods of less than 3 months once per lifetime, and repeated exposures were not considered (e.g. animals were not exposed for 3 months, allowed to recover for 12 months, and then re-exposed for 3 months). Thus, unless otherwise specified by the health agency responsible for derivation of the acute or subchronic TRV, it should be assumed that these are appropriate for assessment of one exposure period per lifetime. For example, for a child who spends 1 month per year at a summer day camp, it would not be appropriate to use a subchronic TRV to estimate risks from the child returning year after year. In such a scenario, the chronic TDI would likely be the more prudent choice of a toxicological value, with a qualitative discussion on the conservative nature of this assumption.

A review of the typical ratio of acute and subchronic TRVs to chronic TRVs has indicated that there is no single conversion factor that can be uniformly applied to estimate short-term TRVs from chronic data. However, if a chemical is not known to cause portal of entry or otherwise irritant effects at the dose rate/concentrations of concern, a minimum conversion of seven times may already be in use in some instances where specific information suggesting otherwise is not available (i.e. acute/subchronic TRV = chronic TRV x 7). A 7-fold factor does not exceed the commonly cited additional 10-fold uncertainty factor that regulatory agencies commonly use for estimation of TDIs from subchronic toxicity studies. The U.S. EPA (2002) provides a thorough review for development of acute and subchronic TRVs. In this approach, various uncertainty factors are recommended, depending upon the database uncertainties. The U.S. EPA (2002) approach may be more appropriate than adopting the default "7-fold" factor previously discussed, and may serve as reasonable guidance. In any event, a decision to use a conversion factor should be completed only on a chemical-specific basis, should be attempted only by a qualified scientist, and should be properly documented.

Overall, Health Canada encourages a scientifically defensible approach be used in the evaluation of less-than-lifetime exposures at contaminated sites. Identification and application of acute and subchronic TRVs can be an appreciable challenge to the risk assessor and communication with Health Canada is encouraged where the risk assessor has doubt regarding the process.

5.6 Consideration of Other Factors Affecting Toxicity

5.6.1 Metal speciation and organic isomers

Many chemicals can be present in several different forms in the environment. Metals often have several different valence states; organic chemicals with the same chemical formula can occur as isomers (i.e. with different structures).

The toxicity assessment must consider the metal species or organic isomer present at the site being evaluated because toxicity can vary significantly among different forms of a chemical. For example, chromium is normally present in soils mainly as Cr^{3+}, but under highly oxidizing conditions may be present as Cr^{6+} (ATSDR, 2000); hexavalent chromium (Cr^{6+}) is far more toxic than trivalent chromium (Cr^{3+}) (EC/HC, 1994). Likewise, the toxicity differs between 2,4,5-trichlorophenol and 2,4,6-trichlorophenol (Health and Welfare Canada, 1987). Therefore, incorrectly identifying the chemical form present can have significant effects on the risk assessment.

In some cases, it is difficult to determine the form of the chemical present. In most cases, tests of metal speciation will not be performed; in this case, it will be assumed that the most toxic form of the metal is present. Also, many metals may readily transform from one valence state to another, and the laboratory analyses may therefore not reflect the situation in situ at the site, even if metal speciation is performed. In these cases, it is appropriate to assume that the most toxic form is present, or to make conservative assumptions (with detailed rationale) about the relative proportions of different forms based on a review of literature.

5.6.2 Bioavailability

Generally, bioavailability is not considered in the toxicity assessment or when setting TRVs. Dose-response

relationships are virtually always based on delivered dose, not absorbed dose. The bioavailability of the toxicant in the key study upon which the TRV is based is important to establishing relative bioavailability adjustment factors used in exposure assessment (see section 4.7). TRV bioavailability values are in the process of being compiled and will be published by Health Canada elsewhere in the near future.

When characterizing risks of dermal exposures or inhalation exposures relative to an oral TRV, Health Canada prefers that the dermal and inhalation exposure be adjusted relative to oral bioavailability (see section 4.7.1); a dermal or inhalation TRV should not be estimated from the oral TRV. This would give the false impression that dermal or inhalation route toxicity studies had been conducted, which would not have been the case.

Further details on evaluating bioavailability are presented in section 4.7.

5.6.3 Multiple exposure routes

The exposure route can significantly influence the type, location (e.g. local versus systemic toxicity), and magnitude of adverse effects caused by some chemicals. Often, exposure through different routes may result in effects on different target organs or effects through different mechanisms. Therefore, a single TRV may not be adequate or appropriate for all exposure routes, and it is crucial that route-specific toxicity be accounted for in the toxicity assessment when possible. Normally, the exposure route for TRVs published by regulatory agencies is clearly identified and discussed in the documentation. The risk assessor should ensure that the TRV is appropriate for the exposure route being evaluated.

Some chemicals may produce both local and systemic effects. Depending on the exposure scenario, a TRV that was developed based on the systemic effects may not be protective of local effects or vice versa. For these chemicals, both types of effects should be evaluated where appropriate.

For some substances, published TRVs may not be available for one or more applicable exposure routes. In particular, TRVs for chronic dermal exposure are rare. It is inappropriate to extrapolate a TRV for one exposure route to another. Instead, as long as it is reasonable to assume that the toxic mode of action and target organ(s) would be similar for the different exposure routes, the exposures from multiple exposure routes should be combined for comparison to the oral TRV. This extrapolation requires an assessment of the relative bioavailability from the exposure routes in question (see section 5.6.2).

As part of the toxicity assessment, the risk assessor should determine whether risks/hazards determined for different exposure routes during the risk characterization (section 6.0) should be combined. In general, if the substance has systemic effects with a similar mode of action and the same target organ(s) across different exposure routes, the risks/hazards should be combined for these routes.

5.6.4 Toxicity of mixtures

Many contaminated sites involve more than one toxic chemical. For Canadian federal contaminated sites, Franz (2005) has summarized the frequency of occurrence of mixtures. Often, the toxic effects of individual chemicals are influenced by interactions with other chemicals in the mixture. It is therefore critical that the mixture of chemicals at the site be evaluated to identify any chemicals that may have toxic interactions.

Most toxicity databases do not provide information on interactions among chemicals because of the excessive number of possible chemical combinations that can occur and the relatively limited database of adequate studies quantifying the interactions. Therefore, the risk assessor generally must rely on knowledge of the mechanisms of toxicity and potential interactions, along with any available advice from regulatory agencies or the toxicological literature. Recently, some agencies have begun to publish evaluations of toxicological interactions in mixtures, such as the ATSDR Interaction Profiles, although these are still available for only a small number of chemical combinations.

The main forms of toxicological interactions follow.

5.6.4.1 Additive

Additive effects occur when the combined toxic effect of chemicals equals the sum of the effects of each individual chemical. Additivity is believed to be the most common type of chemical interaction, particularly at environmentally relevant (relatively low) concentrations and doses. Additivity often occurs when chemicals act via similar toxicological mechanisms or have similar metabolites that cause a toxic effect. For example, various PCDDs/PCDFs have similar chemical structures and have been shown to produce similar toxic effects although at different potencies (HC, 1994). Consequently, the toxicity of these compounds is believed to be additive in mixtures, once adjusted for relative toxic potency.

5.6.4.2 Synergistic/potentiative (greater than additive)

Synergistic effects occur when the combined effects of two or more toxic chemicals are significantly greater than the sum of the effects of the individual chemicals; when a relatively non-toxic chemical enhances the toxicity of a second chemical, then the relationship is considered to be potentiative. Both of these relationships are considered to be greater than additive effects. These interactions can occur when one chemical reduces the detoxification of another, or when their

interactions affect their fate within the body. For example, simultaneous exposure to manganese and lead has been shown to result in higher lead concentrations in the brain than exposure to lead alone at the same exposure dose, resulting in greater toxicological effects (ATSDR, 2004).

5.6.4.3 Antagonistic (less than additive)

Antagonistic effects occur when the combined toxic effect of chemicals is less than the sum of the toxic effects of the individual chemicals (i.e. less than an additive effect). This may occur, for example, if toxic chemicals compete for the same chemical binding site. An example of antagonism is excess zinc exposure reducing the absorption of copper by the body, in turn reducing the toxic effects of copper exposure (ATSDR, 2004).

Mixture-specific TRVs can be developed where toxicity studies of the same or similar mixtures exist. However, because of the general lack of data on the effects of mixtures, Health Canada recommends that, in most cases, the risks should be summed for chemicals with similar mode of action and/or same target organ or tissue (see section 6.4.5). Otherwise, toxicity and risk should be assessed on a chemical-by-chemical basis.

5.6.4.4 Toxic equivalents

- There are certain families of chemicals for which a mixture approach, using isomeric or congener toxicity equivalent factors (TEFs), has been devised. These include PCDDs/PCDFs, dioxin-like PCBs, and carcinogenic PAHs. For PCDDs,/PCDFs and dioxin-like PCBs, this involves TEFs relative to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (see WHO, 1998). For carcinogenic PAHs, this involves carcinogenic potency factors relative to benzo[*a*]pyrene (see *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors, Version 2.0*. (HC,2010).

Additional information on methods for mixtures has been published by the U.S. EPA (2000). The selection of an appropriate method is based on the availability of data, professional judgment, and in consultation with Health Canada as necessary and appropriate.

5.6.5 Sensitive populations and life stages

For a variety of reasons related to exposure rates and biological responses to chemicals, children, pregnant women, seniors, persons in poor health, and some ethnically distinct populations may be more sensitive to certain chemical exposures than the general population. For some chemicals, a child's defence mechanisms may not be fully developed, rendering the child unable to tolerate chemical concentrations that would not normally cause adverse effects in adults; although for other chemicals, children may be less sensitive than adults. Increased sensitivity may occur in seniors because of factors such as compromised health, poor nutrition, compromised immune system, or reduced ability of the body's other defence mechanisms. Pregnant women may also be of special concern because some chemicals may be transferred via the placenta and affect the developing foetus (developmental toxicants). Asthmatics may be particularly sensitive to irritant gases (e.g. sulphur dioxide, nitrogen dioxide). As well, people of certain ethnic backgrounds may have reduced abilities to detoxify certain chemicals because of physiological and/or genetic differences. All of these population subgroups should be considered in the toxicity assessment when possible.

In most cases, potentially sensitive populations and life stages are incorporated into TRVs published by regulatory agencies through the application of an uncertainty factor (typically a factor of 10) to account for these groups. However, the risk assessor should review the supporting documentation for the TRV to ensure that sensitive populations are adequately protected.

Health Canada is currently reviewing the policy on dose averaging with respect to short-term exposure in cancer risk assessment. Until formal policy is provided, any amortization of short-term exposures should provide scientific rationale on a chemical-specific basis. For some chemicals, there may be adequate information available to quantify the toxicity to sensitive populations or life stages; in these cases, there may be separate TRVs available for different groups (e.g. separate TRVs for adults and children). The risk assessor must ensure the appropriate TRVs are used for each group that may be exposed.

If the potentially exposed population as determined during the problem formulation differs significantly from the general population (e.g. Native populations, ethnic communities, seniors' communities), the risk assessor should ensure that TRVs appropriate to the local population are applied.

5.7 Recommended Deliverables

The results of the toxicity assessment should be summarized in a table listing the TRV(s) used for each COPC. If bioavailability is applied, then bioavailability and the bioavailability adjusted TRVs should also be tabulated. The table should be logically organized, and the presentation and units of measure should be consistent with those used in the presentation of exposure assessment and risk characterization results. Chemicals should be classified as either threshold or non-threshold chemicals; in most cases, chemicals that interact (e.g. chemicals with additive toxicity) should be grouped together in the table.

The selection or development of the TRVs should be documented. For TRVs adopted from published regulatory values, the source should be noted, along with the rationale for the adoption of the TRV. For modified or de novo TRVs, the methodology used must be described, along with information sources used, in sufficient detail for the TRV derivation to be evaluated by Health Canada. De novo TRV development is discussed in greater detail in Appendix B.

5.8 References

Agency for Toxic Substances and Disease Registry (ATSDR). 2000. *Toxicological Profile for Chromium*. US Department of Health and Human Services. Public Health Service, Washington, DC.

ATSDR. 2004. *Interaction Profile for: Lead, Manganese, Zinc and Copper*. US Department of Health and Human Services, Public Health Service, Washington, DC.

Environment Canada/Health Canada (EC/HC). 1994. *Canadian Environmental Protection Act Priority Substances List Assessment Report: Chromium and Its Compounds*. EC/HC, Ottawa.

Franz, 2005. *Contaminant Mixture Study: Federal Contaminated Sites*. Contract report to Contaminated Sites Division, Health Canada, Ottawa.

Health Canada (HC). 1994. "Canadian Health-Based Soil Quality Criteria for Contaminated Sites: Polychlorinated Dibenzodioxins and Polychlorinated Dibenzofurans – Supporting Documentation." Environmental Health Directorate, Air and Waste Section, Ottawa. Unpublished report.

HC. 1996. *Health-Based Tolerable Daily Intakes/Concentrations and Tumorigenic Doses/Concentrations for Priority Substances*. Environmental Health Directorate, Health Protection Branch, Ottawa.

HC. 2010. *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, Version 2.0. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

Health and Welfare Canada. 1987. *Guidelines for Canadian Drinking Water Quality – Supporting Document: Chlorophenols*. Health and Welfare Canada, Water Quality and Health Bureau, Safe Environments Programme, Ottawa.

United States Environmental Protection Agency (U.S. EPA). 2000. *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures*. Risk Assessment Forum Technical Panel, Washington, DC. EPA/630/R-00/002.

U.S. EPA. 2002. *A Review of the Reference Dose and Reference Concentration Processes. Risk Assessment Forum Technical Panel*. Washington, DC. EPA/630/P-02/002F.

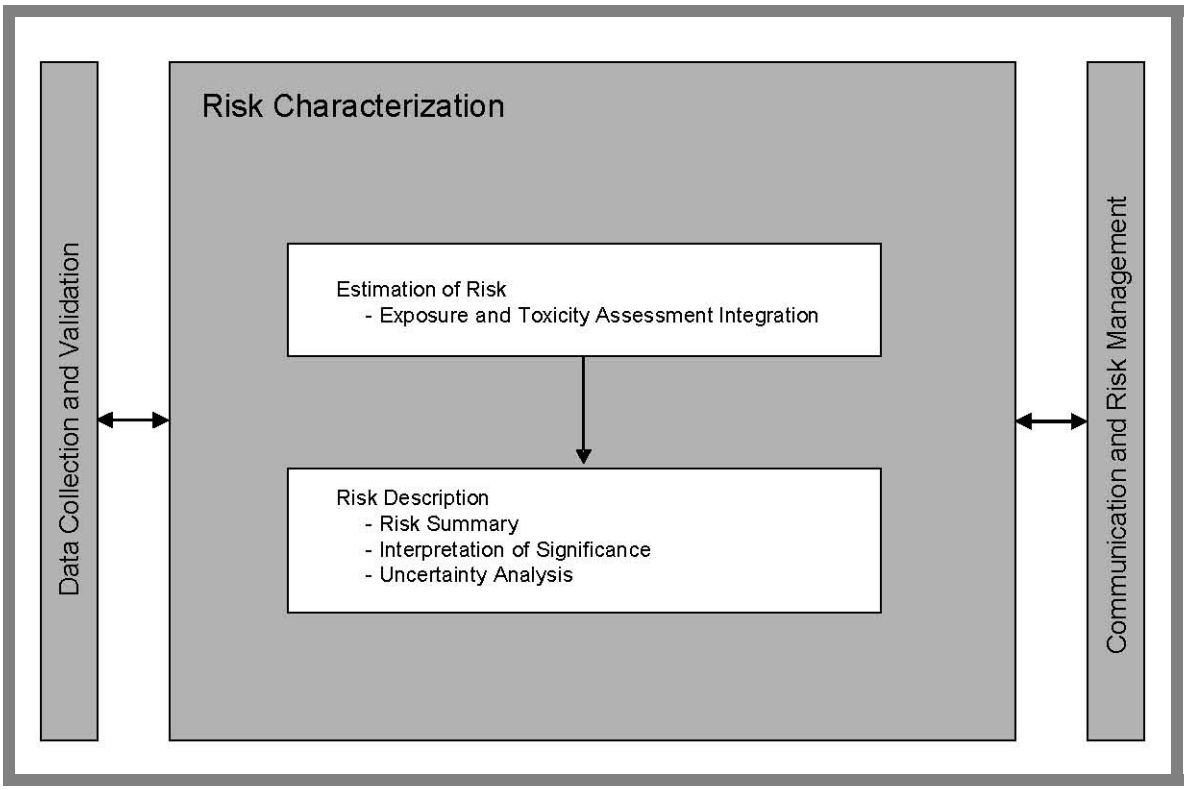
World Health Organization (WHO). 1998. *Assessment of the Health Risks of Dioxins: Re-evaluation of the Tolerable Daily Intake (TDI).* WHO European Centre for Environment and Health, International Programme on Chemical Safety, WHO, Geneva.

6.0 RISK CHARACTERIZATION

6.1 Introduction and Linkages to Other Risk Assessment Tasks

Risk characterization (Figure 6.1) is the quantification and evaluation of the estimated risks and hazards resulting from exposure to chemicals from a contaminated site. Risks/hazards are quantified by comparing the estimated exposure from the site (section 4.0) with the TRV (section 5.0). The evaluation includes a determination of whether or not the predicted risks/hazards are acceptable, tolerable, or essentially negligible, as well as a quantitative or qualitative evaluation of the uncertainties associated with the predictions.

Figure 6.1 Risk Characterization



Risk characterization is the last of the four major stages of a risk assessment to be performed because it uses the results of the exposure assessment (section 4.0) and toxicity assessment (section 5.0). However, results of the risk characterization phase may lead to the necessary revision of earlier stages of the risk assessment to address data gaps or to further refine assumptions, and may also lead into a subsequent risk management stage.

6.2 Objectives of Risk Characterization

The purpose of the risk characterization phase is to provide an estimate and explanation of the potential risks associated with exposures to chemicals from the site. Risk characterization also puts the estimated rates of exposure into perspective through comparisons to potential risks that may be associated with the ambient background environment independent of the site.

The results of risk characterization, along with various engineering, economic, and societal considerations, are used to make risk management decisions for the specific site under consideration.

6.3 Estimation of Risk

6.3.1 Non-carcinogens – single substance exposures

For substances presenting risks other than cancer, a hazard quotient (HQ) (analogous terms include exposure ratio and hazard ratio) is derived as the ratio of the estimated exposure (for each critical receptor) to the TDI or TC, as per (6.1) below.

(6.1)

$$\text{Hazard Quotient} = \frac{\textit{Estimated (on-site + background) Exposure } (\mu g/kg/d)}{\textit{Tolerable Daily Intake } (\mu g/kg/d)}$$

OR, in the case of airborne contaminants with a tolerable air concentration ($\mu g/m^3$):

$$\text{Hazard Quotient} = \frac{\textit{Air Concentration } (\mu g/m^3) * \textit{Fraction of Time Exposed}}{\text{Tolerable Air Concentration } (\mu g/m^3)}$$

HQs for individual exposure routes (dermal, ingestion, inhalation) should be presented where there are pathway-specific TRVs. Where exposures by multiple routes are being summed for comparison to a single TRV (e.g. where only an oral TDI exists), it is necessary only to display the HQ for the summed exposure.

For purposes of a DQRA, where both on-site and background exposures are combined (i.e. total exposure is being assessed), a HQ $\leq$ 1.0 will be deemed to represent an acceptable or negligible risk.

It is critical that the estimated exposure and TRV have the same units and reflect the same time frame (acute, subchronic, chronic). For example, a chronic exposure with units of mg/kg bw/d should be divided by a chronic TDI also with units of mg/kg bw/d, or an amortized subchronic exposure concentration with units of mg/m^3 could be divided by a subchronic tolerable concentration (TC) also with units of mg/m^3. It is also critical that if the exposure estimate was adjusted for bioavailability this adjustment was relative to that observed in the key study upon which the TDI was based (see section 4.7).

The HQ indicates whether or not the estimated exposure exceeds the TRV. A HQ less than 1 indicates that the total exposure (on-site + background) is less than the TRV, and a HQ greater than 1 indicates that the exposure exceeds the TRV. It is important to note that the magnitude of the HQ does not necessarily correspond to the magnitude of expected health effects. A TDI or RfD does not distinguish between health and disease. The TDI represents a conservative estimate of human dose that will be free of health effects in the vast majority of the population. The extent by which a TDI must be exceeded before health effects could occur is not known.

Calculation of the HQ is illustrated in Example 6.1.

Example 6.1 Calculation of Risk Estimates

Problem

The risk assessor discovers that there are five chemicals of concern for which risks must be characterized. The exposure assessment has provided the following exposure estimates for the chemicals and receptors of concern.

Chemical	Estimated Average Lifetime Dose Rate (mg/kg bw/day)
Benzo[*a*]pyrene	1.33×10^{-5}
Benzo[*a*]anthracene	7.54×10^{-4}
Dibenzo[*a*,*h*]anthracene	7.12×10^{-4}
Naphthalene	5.33×10^{-4}
Phenanthrene	9.78×10^{-5}

The toxicity assessment has provided the following toxicity estimates for the chemicals of concern.

Toxicity Estimate (unadjusted for bioavailability)		
CHEMICAL	**Toxicity Reference Value (TRV)**	**Potency Equivalence Factor (PEF)**
Benzo[*a*]pyrene*	2.3 $(mg/kg\ bw/d)^{-1}$	1
Benzo[*a*]anthracene*	0.23 $(mg/kg\ bw/d)^{-1}$	0.1
Dibenzo[*a*,*h*]anthracene*	2.3 $(mg/kg\ bw/d)^{-1}$	1
Naphthalene†	0.02 (mg/kg bw/d)	Not Applicable
Phenanthrene*	0.0023 $(mg/kg\ bw/d)^{-1}$	0.001

* TRVs are expressed as slope factors and adjusted for PEFs, referenced from *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA), Version 2.0* (HC, 2010).

† TRV for this chemical is expressed as a tolerable daily intake (TDI) because it is a threshold chemical.

What are the risks associated with the exposures from the site?

Solution

For this problem, a deterministic analysis approach is taken. However, a probabilistic approach would employ the same techniques, except that a computer program would be used to provide risk estimates in the form of probability distributions.

The following solutions do not take into account general considerations such as the evaluation of less-than-lifetime exposures, amortizing risk estimates across age groups, and multiple chemical exposures. Approaches to address these general considerations are discussed in later sections.

1. Threshold chemicals

The hazard quotient (HQ) can be estimated for threshold-response chemicals. As noted previously, HQs should be calculated for all threshold-response chemicals (in this case, naphthalene), receptors, and exposure scenarios of concern. Using the equation presented earlier (i.e. HQ = exposure divided by TRV), a HQ value for naphthalene was calculated and is presented in the following table.

Chemical	Hazard Quotient
Naphthalene	0.027

An important consideration in the preceding calculation is that this chemical did not have its TRV and exposure estimate adjusted for bioavailability, but when this is done, it is always done for both the TRV and the exposure estimate.

2. Non-threshold chemicals

Numerical cancer risks are estimated only for non-threshold-response chemicals. As noted previously, incremental lifetime cancer risks (ILCR) should be calculated for all non-threshold-response chemicals (i.e. benzo[*a*]pyrene, benzo[*a*]anthracene, dibenzo[*a*,*h*]anthracene, and phenanthrene). Using the equation presented earlier (i.e. risk = exposure multiplied by slope factor), cancer risks were calculated and are presented in the following table.

Chemical	Estimated ILCR
Benzo[*a*]pyrene	3.06×10^{-5}
Benzo[*a*]anthracene	1.73×10^{-4}
Dibenzo[*a*,*h*]anthracene	1.64×10^{-3}
Phenanthrene	2.25×10^{-7}

Two important considerations to note for the above calculations: (i) cancer risk estimates have not been amortized throughout all age groups over a lifetime (methods for this are discussed later in this section); and (ii) not all chemicals have had their cancer potency estimates and exposure estimates adjusted for bioavailability, but when this has been done, it is always done for both the cancer potency estimate and the exposure estimate.

For estimation of cancer risks, it is important to calculate the ILCR, which is an estimate of the cancer risk associated with the site in question, and does not include background exposures.

For threshold-response chemicals, adverse effects could potentially occur when the threshold rate of exposure is exceeded for duration that may be significantly less than a lifetime. Risk estimates for threshold-response chemicals should therefore not be averaged or weighted between the various stages of life or durations of exposure, and should not be averaged over a lifetime. A worked example of the estimation of total lifetime risks from chronic exposure to threshold-response chemicals is provided as Example 6.2.

Example 6.2 Estimate of Total Lifetime Risks from Continuous Lifetime Exposure to Threshold-Response Chemicals

Problem

A female receptor lives at a site for an entire lifetime. The site contains measurable levels of naphthalene (a threshold-response chemical). Using the approach outlined earlier, the risk assessor calculates hazard quotients for naphthalene for each of the age groups listed in the following table.

Receptor Age Groups	Hazard Quotient
Infants (0 to 6 months incl.)	0.780
Toddlers (7 months to 4 years incl.)	9.91
Children (5 years to 11 years incl.)	0.782
Teens (12 years to 19 years incl.)	0.223
Adults (20 years to 80 years incl.)	0.0911

What are the lifetime risks?

Solution

Because the risk value estimates are calculated based on comparisons of exposure estimates and tolerable daily intakes, it is **not** appropriate to weight the risk value estimates with respect to the fraction of lifetime they represent in order to obtain an average lifetime hazard quotient. This is because the exceedance of the toxicological reference value in early life may be enough to cause toxicity or at least warrant concern for potential toxicity, irrespective of the exposures later in life. Consequently, the hazard quotients are expressed individually for each age group. This concept is especially important in this example because it demonstrates that the tolerable daily intake for naphthalene may be exceeded during the 7 months to 4 years (inclusive) age group (i.e. hazard quotient is greater than one for the toddler). It is for this reason that estimates for threshold-response chemicals are typically completed for the toddler age group where a toddler may be present at a site.

6.3.2 Carcinogens – single substance exposures

For substances deemed to be carcinogenic, only on-site exposures are considered. The estimated exposure (amortized as appropriate) is multiplied by the appropriate SF or unit risk (UR) to derive a conservative estimate of the potential incremental lifetime cancer risk (ILCR) associated with that exposure. The ILCR is derived as:

ILCR = *Exposure* (μg/kg/d) × Cancer Slope Factor $(\mu g/kg/d)^{-1}$

OR, in the case of airborne contaminants with a unit risk value in $(\mu g/m^3)^{-1}$:

ILCR = *Air Concentration* ($\mu g/m^3$) × *Fraction of Time* Exposed × Cancer Unit Risk $(\mu g/m^3)^{-1}$

Where pathway-specific SFs or URs exist, the risks via inhalation and the risks via oral + dermal exposure should be estimated separately. In other cases, the cancer risks posed by simultaneous inhalation/dermal/oral exposure will be estimated.

Cancer risks will be deemed to be "essentially negligible" (de minimus) where the estimated ILCR is ≤ 1 in 100,000 ($\leq 1 \times 10^{-5}$). The rationale for this essentially negligible risk level is presented elsewhere (see Appendix 2 of *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA)*, *Version 2.0* [HC, 2010]).

The units for the cancer potency factor are the inverse of the units for the estimated exposure; for example, if the exposure is expressed in mg/kg bw/d, the potency factor must be expressed in $(mg/kg\ bw/d)^{-1}$. Cancer risks are normally related to chronic exposures.

Calculation of the cancer risk is illustrated in Example 6.1.

In contrast to threshold chemicals, risk estimates for non-threshold-response chemicals may be averaged or weighted across various life stages; this is because cancer potency estimates for these chemicals are based on an assumption of lifetime exposure. Two approaches are typically used to estimate risks.

1. Assume that the life stage with the greatest cancer risk persists for a lifetime.

2. Apply straight arithmetic weighting to average risks over a lifetime.

Straight arithmetic weighting may bias risk estimates because the actual risks associated with high rates of exposure to a non-threshold chemical over a short period of time (several months to a few years) may not necessarily result in the same risk of adverse effects as when the same total amount of exposure to the same chemical is received over a long period of time (a lifetime). On the other hand, assuming that the life stage with the greatest cancer risk persists for a lifetime may overestimate risks. Rationale should be provided for short-term exposures with amortization. An example of how to estimate total lifetime risks from exposure to non-threshold chemicals is provided in Example 6.3.

Example 6.3 Estimate of Total Lifetime Risks from Continuous Lifetime Exposure to Non-Threshold-Response Chemicals

Problem

A person lives at a site for an entire lifetime. The site contains measurable levels of benzo[*a*]pyrene (a non-threshold-response chemical). Using the approach outlined earlier, the risk assessor calculates an incremental cancer risk* for benzo[*a*]pyrene for each of the following age groups.

Receptor Age Groups	Incremental Cancer Risk
Infants (0 to 6 months incl.)	5.45×10^{-6}
Toddlers (7 months to 4 years incl.)	5.43×10^{-5}
Children (5 years to 11 years incl.)	6.32×10^{-5}
Teens (12 years to 19 years incl.)	3.22×10^{-5}
Adults (20 to 80 years incl.)	2.54×10^{-5}

* Incremental cancer risks are calculated by multiplying the estimated daily dose rate for each age group by the slope factor.

What are the estimated cancer risks over a lifetime?

Solution

Straight arithmetical weighting can be used (sometimes referred to as a composite receptor). The fraction of the entire lifetime that each age group represents is determined. This is done by dividing the time frame of each age group by the duration of a life (e.g. 80 years). The non-weighted risk estimate for each age group is then multiplied by the fraction of the lifetime that age group represents to obtain a weighted estimate for each age group. The total incremental lifetime cancer risk estimate is then estimated as the sum of the individual weighted estimates. The following table illustrates this process.

Receptor Age Groups	Fraction of an 80 Year Lifetime That Age Group Represents	Non-Weighted Cancer Risk Estimate	Weighted Cancer Risk Estimate
0 to 6 months	0.006	5.45×10^{-6}	3.4×10^{-8}
7 months to 4 years	0.06	5.43×10^{-5}	3.1×10^{-6}
5 to 11 years	0.09	6.32×10^{-5}	5.5×10^{-6}
12 to 19 years	0.1	3.22×10^{-5}	3.2×10^{-6}
20 to 80 years	0.75	2.54×10^{-5}	1.9×10^{-5}
Total Weighted Incremental Lifetime Cancer Risk Estimate			3.1×10^{-5}

Thus, the total numerical cancer risk for benzo[*a*]pyrene would be estimated as 3.1 in 100,000. The straight arithmetic weighting approach may underestimate the importance of early life exposures and, as a result, may be underconservative for some substances.

Example 6.4 Estimate of Total Lifetime Risks from Less-Than-Lifetime Exposure to Threshold-Response Chemicals

Chronic Exposure Scenario

Problem

A person works at a site for a period of 1 year. The site contains measurable levels of polychlorinated biphenyls (PCBs), which are threshold-response chemicals. The exposure assessment results indicate that an adult worker receives exposure of 0.02 mg/kg bw/d of PCBs for the 1 year spent at the site, but then no further exposure from the site for the rest of the person's lifetime. The PCBs have a TRV of 0.00013 mg/kg bw/d. What are the lifetime risks?

Solution

For a threshold chemical, the dose is **not** amortized over the entire lifetime. The daily exposure that occurred over the 1 year exposure period exceeds the toxicity reference value that is deemed to be protective for chronic exposure scenarios. By definition, the 1 year period was a chronic exposure scenario. The fact that exposures are expected to last only 1 year is communicated qualitatively as part of the risk assessment report.

Acute/Subchronic Exposure Scenario

Problem

While at a site, a person will be exposed to formaldehyde in air at a concentration of 25 $\mu g/m^3$ for a period of 30 minutes. What are the risks?

Solution

The risk assessor determines that Health Canada has identified a 1 hour exposure limit of 123 $\mu g/m^3$ for formaldehyde based on eye irritation (HC, 2006). Thus, the risk assessor concludes that exposure to formaldehyde at this level for a period of less than 1 hour is unlikely to be associated with adverse health effects.

Example 6.5 Estimate of Risks from Exposure to Mixtures of Chemicals

Non-Threshold-Response Chemicals

Problem

A site investigation has indicated that three carcinogenic polycyclic aromatic hydrocarbons (PAHs) are present at a site. A subsequent risk assessment has estimated the incremental lifetime cancer risks (ILCRs) as in the following table.

Chemical	ILCR
Benzo[*a*]pyrene	2.3×10^{-6}
Benzo[*a*]anthracene	9.1×10^{-6}
Benzo[*g*,*h*,*i*]perylene	1.0×10^{-7}

What is the ILCR estimate from exposure to this mixture of PAHs?

Solution

Review of the information provided in the toxicity assessment indicates that these PAHs may act on similar target tissues and cells and cause similar toxicological endpoints. As a result, it is assumed that they may act additively to cause cancer. Therefore, to obtain a risk estimate for the total chemical mixture, the lifetime cancer risk estimates of the individual chemicals are added together. The ILCR estimate for the PAHs as a group is calculated as 1.15×10^{-5} (i.e. $2.3 \times 10^{-6} + 9.1 \times 10^{-6} + 1.0 \times 10^{-7} = 1.15 \times 10^{-6}$).

Threshold-Response Chemicals

Problem

A site investigation has indicated that three threshold-response chemicals are present at a site. A subsequent risk assessment has estimated that the risks associated with the site are as in the following table.

Chemical	Hazard Quotient
Ethylbenzene	0.031
Toluene	0.055
Xylene	0.019

What is the lifetime risk from exposure to this mixture of petroleum hydrocarbons?

Solution

Review of the information provided in the toxicity assessment indicates that these chemicals may act on similar target tissues and cells and cause similar toxicological endpoints. As a result, it is assumed that they act additively to cause toxicity. Therefore, to obtain a risk estimate for the total chemical mixture, the lifetime risk estimates of the individual chemicals are added together. Using the hazard quotient (HQ) value approach, the lifetime risk estimate for the chemicals as a group is calculated as an HQ value equal to 0.105 (i.e. 0.031 + 0.055 + 0.019 = 0.105).

6.3.3 Exposure to mixtures

For simultaneous exposure to multiple chemicals of potential concern, non-cancer HQs should be assumed to be additive and summed for those substances determined by the risk assessor to have similar target organs/effects/mechanisms of action. For purposes of DQRA, exposures associated with this total HQ ≤ 1.0 will be deemed negligible. Risks for chemicals with unique target organs/effects/mechanisms of action should be shown individually.

For carcinogens with the same target organ and form of cancer, the risks should be assumed to be additive and summed. The total cancer risk in such cases will be deemed to be "essentially negligible" where the estimated total ILCR is ≤ 1 in 100,000 (1×10^{-5}).

6.3.4 Risks from locally acting chemicals

For locally acting chemicals that cause effects at their point of contact with the body rather than after absorption into the body (e.g. irritants of the skin or respiratory system), risk estimates are calculated based on the estimated exposure concentration and the acceptable concentration in the exposure medium (i.e. the TRV). The TRV is often presented as a maximum acceptable or TC. It is important to recognize that for locally acting chemicals, the duration of exposure is often important in judging the potential severity of effects that may occur. Therefore, the risk assessor must determine if use of a TRV that is derived using a specific exposure duration is appropriate (i.e. its exposure duration is at least as long as the duration of exposure predicted from the site). Risk estimates for locally acting chemicals do not account for bioavailability because absorption is not required to elicit toxic response.

The risk estimate for a locally acting chemical is obtained by dividing the estimated concentration in the environmental medium by the TRV. The resulting ratio is termed either a HQ similar to those calculated for systemically acting threshold-response chemicals.

(6.2)

$$HQ = \frac{Estimated\ Exposure}{TRV}$$

The risk assessor must ensure that the exposure concentrations and TRVs are expressed in the same units, as well as ensuring that the exposure durations for the site exposure and TRV are compatible.

6.3.5 Risks from multiple exposure pathways/routes

At many federal contaminated sites, human receptors may experience simultaneous exposure to a chemical through multiple exposure pathways. Normally, these multiple pathway exposures will have been addressed during the exposure assessment (see section 4.0). As explained in section 4.8.2, exposure estimates for multiple pathways within the same exposure route (e.g. soil ingestion, water ingestion, and ingestion of backyard produce) should be summed; subsequent risks will be estimated from these summed exposures.

Exposure estimates or risk estimates may likewise be summed across exposure routes (ingestion, dermal, and inhalation routes) if there is evidence that the same mechanisms of toxicity occur or the same target organs are affected. Any time that the same TRV is applied for multiple exposure routes, risks should be determined for the summed exposures (adjusted for relative bioavailability).

The risk assessor should only sum exposures or risk estimates if they may occur simultaneously to the same receptor; unnecessarily combining exposure pathways/routes for different and unique receptors will lead to an overestimation of risks.

6.4 Interpretation of Risk Estimates

6.4.1 Interpretation of deterministic estimates

As noted in section 6.6.3 and elsewhere in this document, risk estimates are subject to considerable uncertainty and, potentially, bias. Consequently, risk estimates can be applied with greatest confidence when they are used to make comparisons between two or more different exposure scenarios, using similar methodologies (e.g. comparing the risk to a receptor exposed to a contaminated site with the risk to one who is affected only by background exposure levels). Because the same TRVs, receptor characteristics, and exposure estimation methods are used for each scenario, uncertainties associated with parameter estimation are often effectively cancelled out. This approach is particularly effective for communication of risks to the public.

Threshold compounds with HQ values less than one 1 indicate that the estimated exposure is less than the TRV; therefore, the estimated risk is generally not considered to pose a health risk to exposed individuals as long as background exposure, exposure through multiple pathways/routes, and exposure to chemical mixtures have all been accounted for in the risk estimates.

Threshold compounds with HQ values greater than one1 represent scenarios that may be cause for concern because the rate of exposure is predicted to exceed the acceptable level of exposure. However, this situation does not necessarily indicate that health effects will occur; a TDI does not distinguish between health and disease. The risk estimate should be compared to background exposure because, in some cases, background exposure alone may exceed a TDI and result in a calculated HQ greater than one. Also, because of uncertainties in some of the parameters used in the estimation of risk, there may be conservative factors in the risk estimate. An evaluation of the uncertainty in the risk estimate (HQ) (see section 6.6.3) may reveal areas where further investigation can reduce uncertainty, and allow for refinement and improved accuracy of risk estimates.

In cases where background exposure alone exceeds the TDI, the risk assessment approach should be discussed on a chemical-specific basis with Health Canada.

If the risk estimates are based on input parameters that are intended to represent the extremes of the various numerical parameters (i.e. worst-case values or values representative of "reasonable maximum exposure"), then the incorporation of numerous conservative assumptions in the exposure assessment, combined with uncertainty factors and other conservative assumptions within the toxicity assessments, means that the absolute degree of risk will generally be overestimated to a substantial, but indefinite, degree. Therefore, deterministic estimates of the absolute risks from a particular chemical or chemicals under a particular exposure scenario may be exaggerated, and should be interpreted with caution. If the point estimates of risk are based on input parameters that are intended to represent average or mean values, then resulting risks will be less conservative.

The lack of quantitative indicators of uncertainty and of variability in deterministic risk assessments precludes any reliable estimate of the degree of conservatism or confidence in the point estimate. However, many regulators are comfortable with the results obtained from deterministic risk assessments using average or mean values, particularly if conservatism is introduced in other parts of the risk assessment (e.g. conservative TRVs and/or modelling assumptions).

In summary, a deterministic estimate of risk should be interpreted with reference to both background and generally acceptable levels of risk, in the context of the level of conservatism incorporated into the exposure assumptions (e.g. reasonable maximum exposure versus average exposure).

6.4.2 Interpretation of probabilistic estimates

The interpretation of risk estimates based on a probabilistic risk assessment is facilitated by a probability distribution of estimated risk. The probability distribution enables the probability of exceedance of a specified risk level to be determined; this can be used to estimate the percentage of a population that may be exposed to risks greater than a target level. In addition, a risk estimate corresponding to a desired quantitative level of conservatism (e.g. 95th percentile) can be selected as a basis for a risk management decision.

For probabilistic risk assessments, Health Canada proposes that the 95th percentile dose estimate (dose estimate for the 95th percentile receptor) should have an HQ ≤ 1.0 (i.e. exposure ≤TDI) or an estimated cancer risk ≤ 1×10^{-5}. It is recognized that it is never possible to have all possible receptors at or below the target risk level at all times. However because probabilistic exposures represent, at best, the distribution of exposures on any given random day (see Richardson, 1997), it is believed that day-to-day and year-to-year variations in individual exposures over a life stage or over a lifetime will result in the vast majority of individual risks being essentially negligible if the 95th percentile risk estimate is essentially negligible.

Further discussion on the interpretation of probabilistic risk assessments is presented in section 7.0.

6.4.3 Uncertainty and sensitivity analyses

The description and evaluation of uncertainty and variability associated with risk estimates and the sensitivity of the risk estimates to this uncertainty and variability are critical parts of the risk assessment process. Uncertainty and sensitivity analyses are performed for all risk estimates, regardless of whether they have been derived from deterministic or probabilistic approaches.

Further discussion of probabilistic approaches is presented in section 7.0.

6.4.3.1 Uncertainty and variability

Although it is often loosely used to describe natural variability, the term "uncertainty" strictly refers to a lack of knowledge or insufficient data. Uncertainty may exist in model parameters and assumptions, resulting from insufficient data, characterization, or measurements of the values. There is also uncertainty in the models themselves because of an imperfect understanding of the processes being modelled and the necessary simplifications of reality made within models. Because all risk assessments are based on models, uncertainty is associated with all risk estimates. This uncertainty does not necessarily invalidate the model output or the risk estimates; however,

acknowledgement and description of uncertainty and the quality of the input assumptions are important factors in interpreting the risk estimates. Uncertainty can be reduced given sufficient time and resources to expand and refine the data available.

Variability refers to the diversity and heterogeneity within a population. For example, air inhalation rates vary among individuals, and also between different times or activity levels for the same individual. Likewise, many soil properties vary considerably across a site. Variability cannot be reduced because it is an intrinsic property of the population. However, more data can allow more confident quantification of the bounds and amplitude of variability.

6.4.3.2 Sensitivity analysis – deterministic

It is impossible to predict the proportion of a population that might remain at risk when an average, reasonable maximum, or worst-case deterministic exposure scenario is used to quantify risks. Uncertainty and variability (and subsequent confidence in the results) are normally only qualitatively characterized in a deterministic risk assessment. A sensitivity analysis for a deterministic risk assessment should consist, at a minimum, of a qualitative summary of the uncertainties and variability associated with each input variable, and a prediction of how these uncertainties are expected to affect the risk estimates. For example, the risk assessor might indicate that the mean of the measured values for organic carbon content in soil was used for fate and transport modelling calculations, and that the mean value is believed to be representative of average conditions at the site, but the predicted exposure concentrations are inversely proportional to the assumed organic carbon content, so that the consequences of being wrong in the assumption can be indicated and discussed.

The influence of uncertainty and variability on a risk assessment can be assessed to a limited degree by varying individual model parameters (e.g. calculate risk estimates using low-end, high-end, and mean values) to evaluate the sensitivity of the model output to the varied parameter. This procedure is sometimes referred to as a discrete sensitivity analysis, where the term “discrete” implies independent variation of single parameters. The process may yield valuable information on which parameters are critical to the risk assessment calculation, so that any efforts to reduce uncertainty (collect more data) can be focussed on variables that truly matter.

The completed sensitivity analysis allows the risk assessor to comment on the degree of their confidence in the risk estimates, considering the degree of uncertainty and variability in each model parameter, determining which parameters influence the results to the greatest extent, and identifying key data gaps. However, as noted above, the overall uncertainty in the risk assessment and the degree of confidence can be evaluated qualitatively only in a deterministic risk assessment.

6.4.3.3 Sensitivity analysis – probabilistic

A probabilistic risk assessment allows for a quantitative assessment of uncertainty and the conservatism of the risk assessment. Conservatism is quantified by selecting a risk estimate corresponding to a specific percentile of the distribution of predicted risk estimates (e.g. the 95th percentile risk). The statistical parameters of the risk estimate distribution (e.g. mean, standard deviation, percentiles) allow for the quantitative evaluation of uncertainty. Software used for probabilistic risk assessments generally allows for the testing of the influence of individual parameters on the variability of the risk estimate. It should be noted, however, that the quantification of uncertainty and conservatism becomes meaningless if inappropriate or incorrect distributions are assigned to model input parameters. Further details on sensitivity analysis for probabilistic risk assessments are presented in section 7.0.

6.4.3.4 Communication of uncertainty and variability

It is important that the risk assessor be able to communicate the uncertainty and variability (and degree of confidence) associated with the risk estimates, so they can be adequately considered in the decision-making process for any further action at the contaminated site. An assessment of uncertainty and variability may result in a decision to collect additional data and information. Alternatively, the uncertainty and variability may be considered to be within acceptable bounds and would be accounted for within the risk management decision process.

The risk assessment should clearly document the potential uncertainty and variability in model input parameters. For deterministic risk assessments, the rationale for the selected point estimate should be presented; for probabilistic risk assessments, the rationale for the selected distribution should be provided. Parameters that have a significant effect on the risk estimate should be identified.

The overall uncertainty in the risk assessment and degree of confidence and conservatism should be discussed. Key assumptions that may affect the degree of conservatism should be highlighted. Where appropriate, significant data gaps should be identified, along with recommendations for addressing these data gaps if they are considered significant.

6.5 Recommended Deliverables

Using the procedures and considerations discussed in this section, estimates of the risk are obtained for all chemicals and human receptors of concern. Risk estimates can be presented as point estimates in the case of deterministic analysis techniques or as distributions in the case of

probabilistic analysis techniques. Depending on how the chemicals act, risk estimates will be expressed as:

- **hazard quotient (HQ)** values (threshold-response chemicals only)
- **incremental lifetime cancer risk (ILCR)** values (non-threshold-response chemicals only)

The risk estimation results should be presented in a concise and readable format, preferably including the use of tables and/or figures. Chemicals should be grouped logically, and presented in the same order as the exposure estimates and TRVs determined in sections 4.0 and 5.0.

Health Canada recommends that a risk characterization follow the principles of transparency, clarity, consistency, and reasonableness, defined as follows by the U.S. EPA (2000).

- Transparency includes clearly describing the approach, assumptions, extrapolations, and models; identifying data gaps, distinguishing science from policy; and describing uncertainty.
- Clarity includes avoidance of unnecessary technical terms, employing brevity, and using simple tables, graphics and equations.
- Consistency involves following regulatory guidance, placing the risk assessment in context, and defining the level of effort.
- Reasonableness involves using the best available scientific information, good judgment, and plausible alternatives.

When the risk estimation is completed, the risk assessor should determine whether or not all significant and dominant risk-driving chemicals have been evaluated. This component of the risk description process is based on a phenomenon termed the "Pareto principle" (Wadsworth, 1990). The Pareto principle simply states that a relatively large proportion of the problems being addressed (in this case, a large proportion of the site-attributable risks) in any given situation is found to be caused by only a few factors (in this case, only a few chemicals).

In many cases, the risk characterization will lead to a subsequent iteration of the earlier stages of the risk assessment, particularly if unacceptable risks have been predicted but data gaps or uncertainties that significantly influence the results are identified.

6.6 References

Health Canada (HC). 2006. *Residential Indoor Air Quality Guideline. Formaldehyde.* Accessed August 8, 2010 from http://www.hc-sc.gc.ca/ewh-semt/pubs/air/formaldehyde-eng.php).

HC. 2010. *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA*), *Version 2.0*. Contaminated Sites Division, Safe Environments Directorate, Ottawa.

Richardson, G.M. 1997. *Compendium of Canadian Human Exposure Factors for Risk Assessment*. O'Connor Associates Environmental Inc., Ottawa.

United States Environmental Protection Agency (U.S. EPA). 2000. *Risk Characterization Handbook.* Science Policy Council, U.S. EPA., Washington, DC. EPA 100-B-00-002.

Wadsworth, H.M. Jr. (ed.). 1990. *Handbook of Statistical Methods for Engineers and Scientists*. McGraw-Hill Publishing Co., New York.

7.0 PROBABILISTIC RISK ASSESSMENT

7.1 Overview

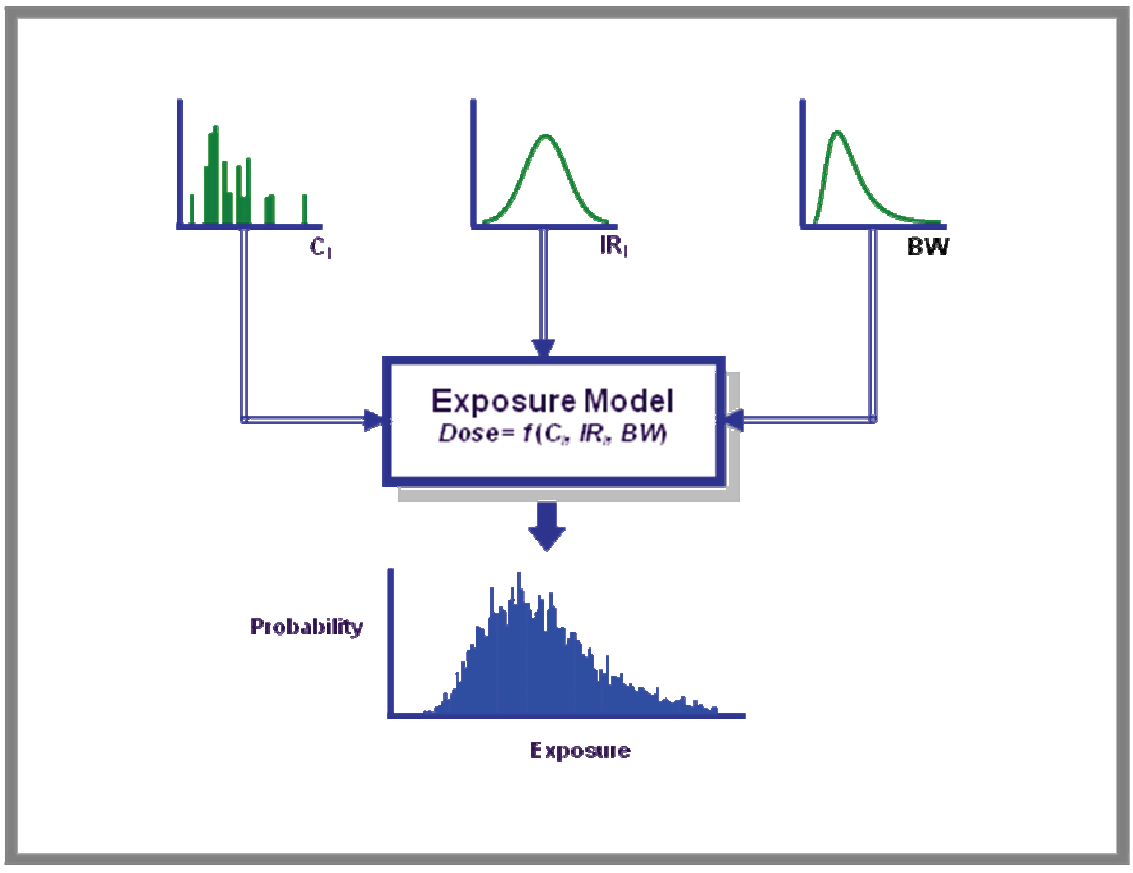
7.1.1 What is probabilistic risk assessment?

Quantitative HHRA generally involves assigning numerical values to a variety of input parameters that are used, in combination with appropriate exposure and/or risk models, to obtain a quantitative estimate of risk. Numerical values are required for parameters describing contaminant concentrations in environmental media, contaminant fate and transport, human exposure, and toxic response. These values may be measured, assumed, prescribed or based on published literature. Variability and uncertainty in these parameters result in variability and uncertainty in the estimate of risk.

Traditional deterministic methods of quantitative risk assessment use single, or "point estimate," values for input parameters and produce a single estimate of risk or hazard. Although input parameters may be selected with some knowledge of their variability or uncertainty, a deterministic analysis does not normally provide any information on the variability of the resulting risk estimate. For example, although input values are often selected to represent either average or reasonable maximum exposure conditions, the location of the point estimate of risk in the context of its potential range and distribution cannot be determined directly. A discrete, or deterministic, sensitivity analysis may provide some indication of the potential range of estimated risk values, but the actual variability of, and hence confidence in, the risk estimate remains unknown.

Probabilistic risk assessment uses probability distributions to characterize variability and uncertainty in key parameters and produces a probability distribution of estimated risk. The resulting distribution provides not only a description of the variability and uncertainty in the calculated risk, but also a basis for selecting a risk estimate for decision-making purposes whose likelihood of exceedance can be quantified. Probabilistic risk assessment is shown conceptually in Figure 7.1.

Figure 7.1 Conceptual Representation of Probabilistic Exposure Analysis



7.1.2 Uncertainty and variability in risk assessment

Uncertainty arises as a result of lack of knowledge or insufficient data. For example, uncertainty in determination of the mean concentration of a contaminant in a particular environmental medium (e.g. a soil unit) may be due to insufficient measurements as well as measurement error. This is referred to as parameter uncertainty. Uncertainty is also present in modelling; model uncertainty is due to inadequate characterization of a process that is being modelled (e.g. contaminant transport or uptake), resulting from lack of knowledge of the process, simplification of the process for modelling purposes, or insufficient data to fully define parameters governing the process. Uncertainty can generally be reduced by the collection of additional or more accurate data.

Variability refers to diversity or heterogeneity within a population, such as the spatial variation of a contaminant concentration within a soil unit or the variation of body weight within the human population. Such variability is an attribute of the population and cannot normally be reduced.

The potential variation in a measured or assumed parameter in risk assessment may be a result of both uncertainty and variability. An example would be the concentration of a contaminant in a soil unit. As a result, the estimated distribution of risk may also reflect both uncertainty and variability. Where possible, uncertainty and variability should be distinguished in order to facilitate interpretation and communication of the results of a risk assessment.

7.1.3 Simulation methods

Probabilistic risk assessments are commonly conducted using a method referred to as Monte Carlo analysis. This method uses computer simulation to combine the probability distributions of the input variables in a model. The model is run a large number of times, each time using a discrete value for each input parameter that is sampled from its respective probability distribution. Each step or iteration produces a single estimate of risk; the resulting population of calculated values defines the probability distribution of risk.

Simulations that consider all input distributions concurrently within each iteration are referred to as one-dimensional simulations. In some cases, it is necessary to distinguish the influence of two sets of input distributions, for example where one set represents variability and the other uncertainty. In these cases, a simulation may be run in the form of nested loops, whereby a family of distributions of risk is obtained — each distribution based on iterations of the "inner" loop for a given sampling step of the "outer" loop. The outer loop is repeated to generate the family of distributions. Such a method is referred to as two-dimensional simulation.

7.1.4 Relationship of probabilistic risk assessment to deterministic risk assessment

Both deterministic risk assessment and probabilistic risk assessment have merit in different situations; neither method should be considered superior to the other. Instead, the data required and information provided by each approach should be evaluated in the context of each specific application and the objectives of the risk assessment. For contaminated site risk assessment, a deterministic analysis would almost always be conducted at the outset, although in some cases it may be a preliminary or screening stage prior to the completion of a more detailed probabilistic risk assessment.

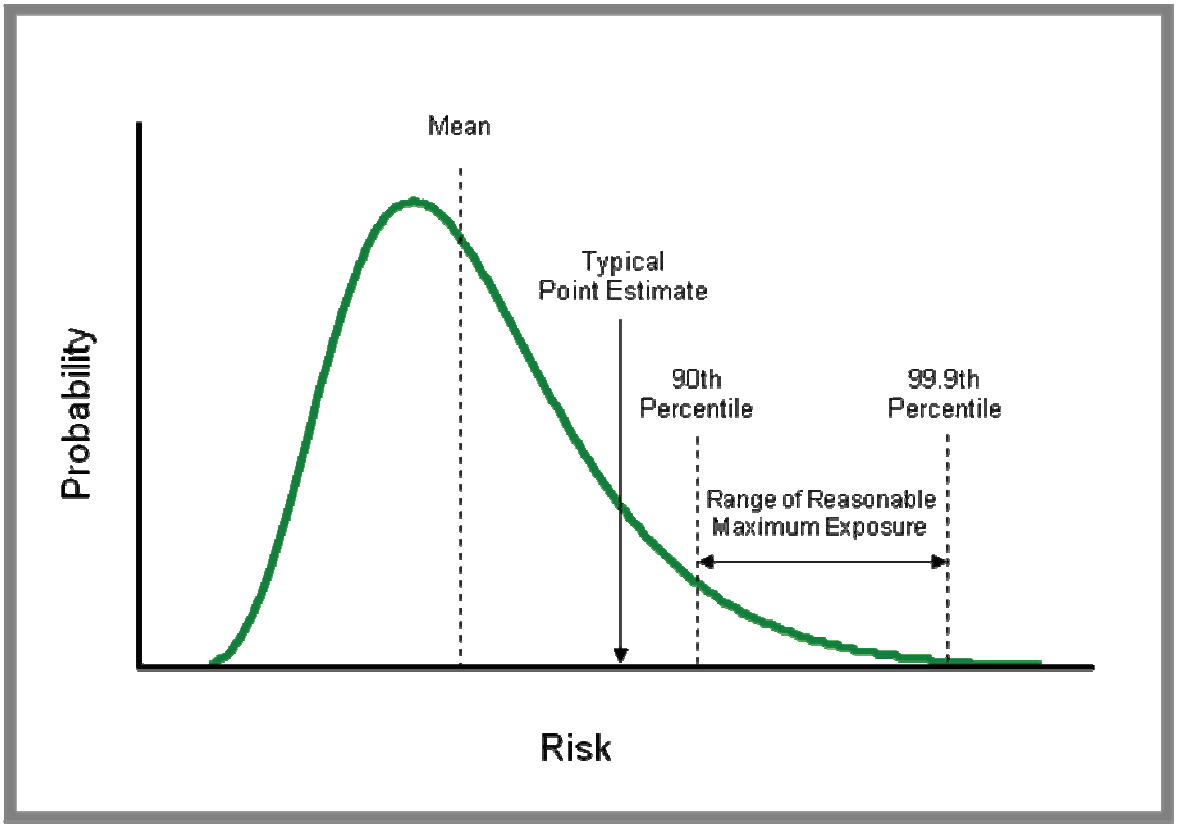
A deterministic risk assessment represents a single point estimate of risk for a specific set of exposure assumptions. The assumptions may be selected to represent average, worst-case, or reasonable maximum exposure, or an alternate protection level, and may be prescribed by regulatory policy or by guidance documents intended to ensure consistency and transparency in the conduct of risk assessments. Because most risk-based environmental quality guidelines and standards in Canada are derived using deterministic or point estimate methods, a site-specific deterministic risk assessment represents a direct extension of, and point of comparison to, the use of generic numerical guidelines. The link between the use of generic risk-based numerical guidelines, or a deterministic site-specific risk assessment, and a probabilistic risk assessment is less evident; this is because the level of protection inherent in the assumptions in the deterministic and probabilistic approaches cannot be directly compared. Furthermore, while each iteration in a probabilistic risk assessment is a deterministic calculation of the exposure or risk assessment model using randomly sampled input values, a given deterministic risk assessment does not necessarily represent a repeatable subset or sample of the probabilistic risk assessment results. In other words, the result of a probabilistic assessment, depending upon the statistic or percentile selected as the risk estimate, cannot be related back to a specific set of point estimate exposure assumptions, either for tracing or review purposes or for comparison with a deterministic analysis.

Deterministic analysis is often carried out as an initial step in a detailed probabilistic risk assessment for the purpose of assessing sensitivity of exposure or risk to key assumptions (through a discrete sensitivity analysis) or for obtaining preliminary estimates of risk in order to screen receptors, pathways, or contaminants.

The result of a deterministic analysis (expressed as a point estimate of exposure or risk) would normally lie within a probability distribution for exposure or risk obtained from a probabilistic analysis (Figure 7.2). However, the location of

the point estimate with respect to the statistical parameters of the probabilistic distribution is a function of the level of protection targeted by the deterministic assumptions. For example, a combination of assumptions representing reasonable maximum exposure would result in a deterministic risk estimate near the upper tail of the distribution of risk; a combination of "average" assumptions would be closer to the mean of the distribution. However, the exact location of the estimate within the continuum of possible values cannot be ascertained solely from a deterministic analysis. On the other hand, the probability distribution of risk permits the percentile of any individual risk estimate to be determined—in practical terms, enabling the risk assessor to select a risk estimate with a desired or specified probability of exceedance. When managing for reasonable maximum exposure, it is common to select a risk estimate corresponding to a percentile between 90 and 99.9 (i.e. with a probability of exceedance of between 10% and 0.1%, respectively) (Figure 7.2).

Figure 7.2 Point Estimate Versus Probability Distribution



A complete assessment of the effect of variability and uncertainty on the distribution of risk, including a sensitivity analysis, also provides insight into the value of information, and assists the risk assessor in deciding if the resources allocated to additional data collection would be justified by a reduction in uncertainty associated with the risk estimate.

7.1.5 Advantages and disadvantages of deterministic and probabilistic risk assessments

Deterministic and probabilistic risk assessments differ in complexity and data requirements; they also differ with respect to the level of information provided to the risk assessor and their respective value to the risk management decision process. Advantages and disadvantages of the deterministic approach are summarized in Table 7.1. Advantages and disadvantages of the probabilistic approach are summarized in Table 7.2.

Table 7.1 Advantages and Disadvantages of a Deterministic Method for Risk Assessment

Advantages	Disadvantages
• Calculations are generally simple and require fewer resources. • Guidance is available for standard assumptions. • Method is useful for screening purposes – may allow risk management decisions with no additional work. • Discrete sensitivity analysis provides semi-quantitative assessment of variability. • Method is easily described and communicated. • Method is readily amenable to tracing and peer review. • Method is compatible with deterministic methods used for existing regulatory guideline development.	• Results are often interpreted as definitive; importance of uncertainty is sometimes lost. • Information from sensitivity analysis is generally limited to dominant exposure pathways; other key exposure variables and uncertain parameters may be overlooked. • Method does not provide measure of probability that risk exceeds a regulatory level of concern or level of confidence in risk estimate. • Method provides fewer incentives for collecting better of more complete information. • Simplified approach may preclude use of available data for characterizing variability and uncertainty.

Source: Adapted from U.S. EPA, 2001.

Table 7.2 Advantages and Disadvantages of a Probabilistic Method for Risk Assessment

Advantages	Disadvantages
• Method makes more complete use of available data when defining input parameters. • Method provides more comprehensive characterization of variability and uncertainty in risk estimates; this may provide added insight to confidence of risk assessment and facilitate risk communication. • Sensitivity analysis can help identify exposure variables and model parameters that influence risk estimates. • Sensitivity analysis, in combination with probability distribution of risk, can guide decisions regarding the value of information. • Approach allows a consistent selection of risk estimate for a chosen or prescribed level of protection.	• Method is more complex and less amenable to standardization. • Input requires more comprehensive data on exposure variables and site conditions. • Method places greater burden on the risk assessor with respect to resources and documentation, as well on regulator and/or peer reviewer. • Complexity of approach may obscure errors or inaccuracies in underlying assumptions and models. • Communication of results is more complex, and may convey false sense of accuracy or sophistication when data are sparse; if communication is ineffective, may create mistrust of method and results. • Method is not directly compatible with deterministic methods used for guideline derivation, and may result in different levels of protection.

Source: Adapted from U.S. EPA, 2001.

7.1.6 When to conduct a probabilistic risk assessment

As noted previously, a deterministic analysis is almost always undertaken as part of a site-specific quantitative HHRA. A preliminary deterministic analysis may be conducted as a first step in either a detailed deterministic risk assessment or a probabilistic risk assessment. Its purpose may be one or more of the following: to screen contaminants, exposure pathways, and/or receptors; to determine the need for further, more detailed risk assessment; to determine the sensitivity of the risk estimate to key assumptions (by discrete sensitivity analysis); and/or to assess the requirement for additional data collection (which may also lead to a more detailed risk assessment). In many cases, the risk assessment may not proceed beyond the preliminary deterministic stage, perhaps because the risks determined on the basis of a set of conservative assumptions are found to be below the level of concern. Often, however, the preliminary analysis serves as a scoping stage for a more detailed risk assessment, whether deterministic or probabilistic.

Before proceeding with a probabilistic risk assessment, the risk assessor should consider whether a probabilistic analysis is necessary and/or appropriate, given the objectives of the assessment and the availability of data. The primary reason for undertaking a probabilistic analysis is to determine the possible range and distribution of the estimated risk in cases where a single point estimate of risk is insufficient. Other reasons for a probabilistic analysis may include quantifying the influence of uncertainty and communicating the resulting confidence in the risk estimate; quantifying the selection of a risk estimate in terms of the portion of the population potentially receiving greater exposure; decision making about the value of information and additional data collection; and cost-benefit analysis and allocation of resources for remediation or risk management strategies.

A probabilistic analysis necessarily involves a greater commitment of resources to the collection of data, the risk assessment modelling, and the reporting and presentation of results. Sufficient data should be available to adequately characterize the probability distributions for all key input variables; it should be recognized that the results of a probabilistic risk assessment based on insufficient data may be misleading. In practice, probabilistic analyses are more commonly conducted in the case of large and/or complex sites, where the additional resources required are expected to be justified by a more complete estimate of potential risks, and possibly by the informed selection of a more cost-effective risk management strategy.

A probabilistic analysis is not appropriate in some situations—for example, where the availability of data is insufficient to define probability distributions for input variables; where policy or available guidance prescribes the selection of point estimate values for input assumptions; or where the purpose is to calculate a site-specific deviation from a generic risk-based guideline, derived deterministically using a prescribed protocol (e.g. a “Tier 2” soil remediation objective). The decision to proceed with a detailed probabilistic analysis would normally be made following an initial deterministic assessment, typically at or following completion of the problem formulation stage of the risk assessment.

7.2 Procedures for Probabilistic Risk Assessment

Steps in the completion of a probabilistic risk assessment are described in the following sections, which are divided according to the four main stages of risk assessment. The following sections describe procedures or modifications in procedures that are specific to probabilistic analysis, and do not reiterate many of the aspects of each stage that are common to both types of assessment. The following sections therefore supplement the corresponding ones of the preceding guidance, to which the reader should refer when conducting a probabilistic risk assessment. The reader is also referred to other guidance documents (e.g. Oregon DEQ, 1998; U.S. EPA, 2001) for additional information.

7.2.1 Problem formulation

The approach to problem formulation for a probabilistic risk assessment is essentially the same as for a deterministic assessment, the main objectives being the review of existing information; the identification and screening of contaminants, receptors, and exposure pathways; and the development of a CSM. Defining the scope of the risk assessment and determining the need for probabilistic analysis as opposed to deterministic analysis are normally part of this stage. Factors considered when selecting between probabilistic and deterministic risk assessment are described in section 3.3.3 as well as in section 7.1.6. The decision to proceed with probabilistic risk assessment is a function of the goals of the risk assessment, the societal issues and degree of public concern, the complexity of the site, and the availability of sufficient data to permit probabilistic characterization of relevant parameters. If sufficient data are not available, additional data collection requirements would be identified. The optimization of further data collection activities, if necessary, may require that a deterministic risk assessment together with a sensitivity analysis be carried out in order to identify the most significant parameters. Thus, the process may be somewhat iterative.

7.2.2 Exposure assessment

Probabilistic exposure assessment, like deterministic exposure assessment, involves the estimation of the intake of each COPC for the receptor and exposure pathway combinations identified in the CSM. Exposure assessment includes the characterization of contaminant concentrations in applicable media, the characterization of receptors and exposure factors, and the estimation of intake. The estimation of intake may involve fate, transport, and exposure modelling, necessitating the selection of parameter values required for the modelling.

Probabilistic exposure assessment uses probability distributions instead of point estimate values to reflect variability and, where appropriate, uncertainty in key assumptions and parameters used in the assessment. Although the nature and form of the distributions selected may differ depending on the type of parameter being considered (e.g. contaminant concentrations, hydrogeological parameters, human exposure factors), the principles involved in selecting distributions are similar, and are described in general in the following sections. Considerations specific to contaminant concentrations, fate and transport parameters, and human exposure factors are addressed in the respective subsequent sections.

7.2.2.1 Characterization of input variables for probabilistic exposure assessment

General

A number of methods are available for characterizing input variables for probabilistic exposure and risk assessment. The appropriate method depends on the nature of the particular variables to be characterized, as well as the availability of data. Also, as noted previously, variability and uncertainty are two factors that contribute to variation in an exposure or risk estimate; it is normally recommended that these be separated in a probabilistic analysis. The methods of characterizing variability and uncertainty are generally different.

A typical human health exposure or risk assessment incorporates a variety of input parameters and assumptions, including source and exposure point concentrations, receptor characteristics or exposure factors, and absorption factors. In addition, depending on the exposure pathway, models may be used to describe contaminant fate and transport, and are themselves dependent on assumptions and parameters that describe the transport processes.

Input variables in a probabilistic analysis may be characterized by point estimate values, ranges of values, and data sets or probability distributions, depending on availability of data and the influence of the variable on the outcome of the analysis. The influence of each variable is generally determined at the outset, based on a discrete sensitivity analysis. Parameters that do not significantly influence the result are commonly input as single point estimates, although ranges of values (e.g. upper and lower bounds) may be used in the sensitivity analysis itself. If the parameter is subject to uncertainty, this is frequently reflected in the selection of the point estimate (e.g. upper 95% confidence limit of the mean), although a more complex analysis may subsequently be conducted to quantify the effect of that uncertainty.

Input variables that are influential on the outcome of the analysis would typically be characterized as ranges of values or probability distributions. In most cases, it is recommended that these ranges or distributions be selected to reflect variability across a population or space–time continuum, and not to reflect uncertainty arising from lack of information, measurement errors, or model uncertainty. The results of a risk assessment in which variability and uncertainty were considered together would be a single probability distribution, in which the effects of uncertainty and variability could not be distinguished. Uncertainty may be represented by statistics pertaining to the estimated mean (e.g. percentiles or confidence limits), or it may also be represented by a probability distribution; however, the effects should ideally be assessed in an iterative manner or using a two-dimensional simulation as described previously.

Methods for characterizing input variables are discussed in the following sections under the overall headings of variability and uncertainty. It may not be appropriate to consider both variability and uncertainty for all types of input variables. For example, point-of-exposure concentrations are often specified as point estimate values representing average concentrations to which a receptor or population is exposed, in which case the point values selected (e.g. upper 95% confidence limit of mean) would account for uncertainty. However, in some cases, the spatial or temporal variability of a concentration may be of interest and significance to a risk estimate. Population exposure factors, on the other hand, may be defined on the basis of a very large data set that minimizes uncertainty; variability across the population becomes the issue of concern.

Variability

Methods for representing variability in a probabilistic risk assessment generally fall into two categories: parametric and non-parametric (U.S. EPA, 1997a; 2000; 2001; and others). Selection of the appropriate method will depend on the available data, the nature of the variable and, to some extent, the preferences of the risk assessor. Selection may also be dictated by available guidance with respect to default parameters, such as for human exposure factors. The ultimate aim in selecting a method is to ensure that the population of possible values for a parameter is adequately

represented by the chosen distribution over the entire range, including the extreme values or tails, of the distribution.

Parametric methods are generally subjective in that they employ a degree of judgment. Non-parametric methods are more objective because they use measured data directly and are therefore empirical, although these methods also may utilize the judgment of the assessor and thereby become somewhat subjective. A subjective method should not be considered inferior, because the applied judgment enables prior knowledge or understanding of the underlying behaviour of a variable to be taken into account.

Parametric methods

Parametric methods involve selecting a distribution that is considered appropriate to the variable and/or available data set, and specifying the distribution in terms of its defining statistical parameters. The process of selecting a distribution will consider the underlying mechanistic processes that govern variability of the parameter, as well as the ability of the available data to define the bounds or constraints of the variable.

The following factors should be considered in selecting a distribution (after U.S. EPA, 1997a, 2001).

- Is there any 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 symmetrical and, if skewed, in which direction?
- What other aspects of the shape of the distribution are known?
- How well do the tails of the chosen distribution represent the observations?

It is widely known that many natural phenomena may be represented by normal and lognormal distributions. The basis for this is the central limit theorem; a variable that is influenced by the sum of a number of random processes is generally normally distributed, whereas a variable that is influenced by the product of a number of random processes is typically lognormally distributed. Other distributions may be appropriate to other types of data. Examples of distributions that are commonly used in probabilistic HHRAs are illustrated in Figure 7.3, and are summarized along with their defining parameters (or constraints) and other information in Table 7.3.

Figure 7.3 Examples of Commonly Used Probability Distributions

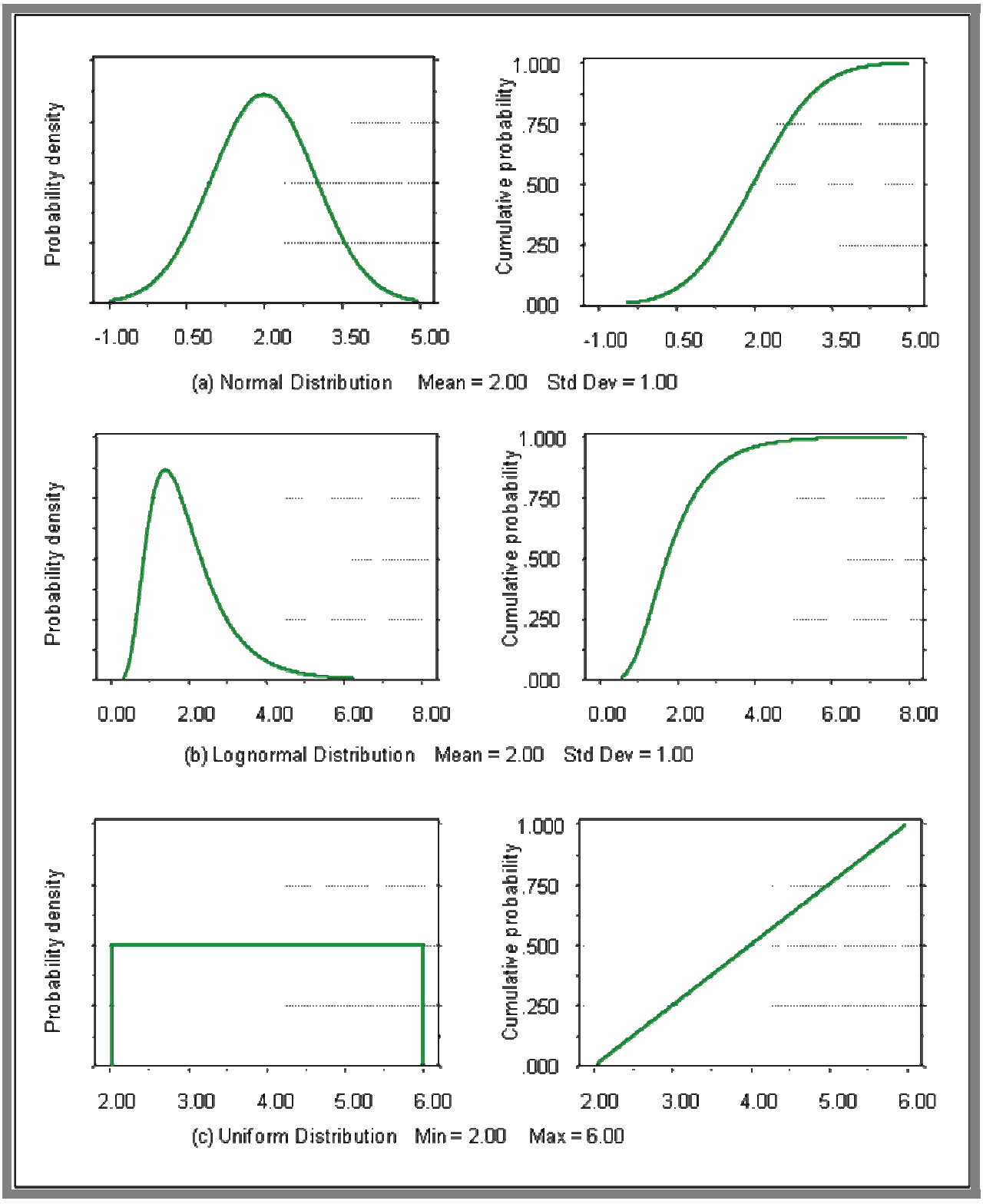


Figure 7.3 Examples of Commonly Used Probability Distributions (continued)

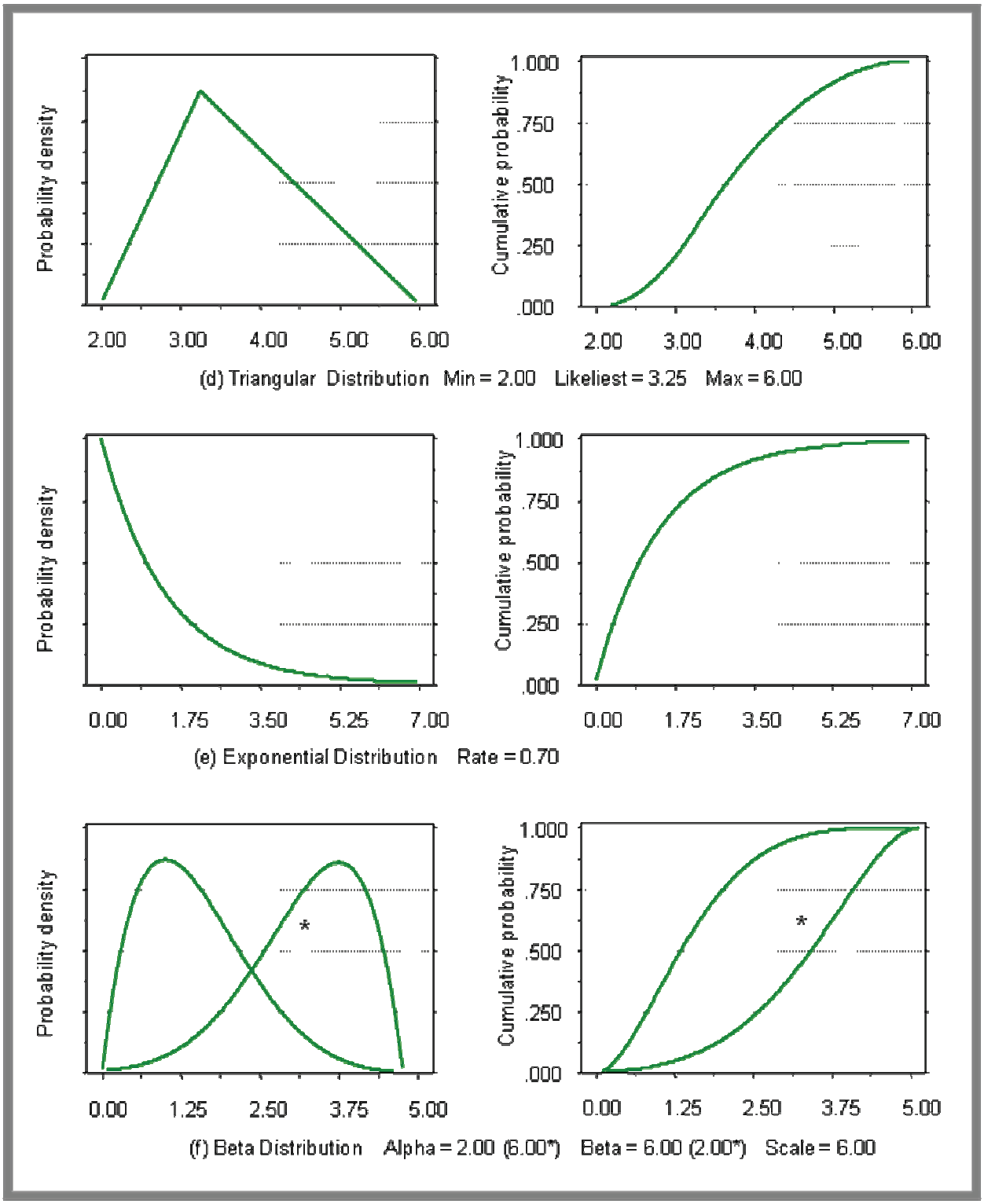


Table 7.3 Commonly Used Probability Distributions

Distribution	Constraints	Mechanistic Basis	Application
Normal	Mean, standard deviation	Variables influenced by the additive effect of a large number of independent processes; describes many natural processes	Measurement errors; uncertainty in sample means; some environmental processes; can assume negative values; may require truncation
Lognormal	Mean, standard deviation	Variables influenced by the multiplicative effect of a large number of independent processes; describes many natural processes	Contaminant concentrations in environmental media; human exposure factors; other natural phenomena and environmental processes
Uniform	Maximum, minimum	Variables that have an equal probability of attaining any value within a range; rough distribution with no mechanistic basis	Variables with known upper and lower limits, but no information about nature of variation; exposure frequency and other behavioural assumptions; assumptions related to mechanical systems (e.g. air-exchange rates)
Triangular	Maximum, mode, minimum	Variables that can attain any value within a range, but with central tendency; rough distribution with no mechanistic basis	Variables with known upper and lower limits and most likely value, but no other information about nature of variation; exposure frequency and other behavioural assumptions
Exponential	Rate	Exponential decay function; also represents time between random events (inverse of frequency of random events)	Time between random events such as storms, floods; processes subject to rates of decay; representation of tails of empirical data distributions
Beta	Scale, alpha. beta (shape factors)	Mathematical function that can vary between upper and lower bounds, but no mechanistic basis; very flexible shape	Variables with known upper and lower limits, with or without known central tendency

A variable often can be adequately characterized by more than one distribution; in fact, in some cases a non-intuitive distribution may actually be ranked higher on the basis of a goodness-of-fit test. The preferred choice is generally the simplest distribution that fits the observations and is consistent with the underlying mechanistic basis.

It is important that the form of the distribution fit the observed data or the expected variability in the population represented by the data. This is particularly so with respect to the upper and lower limits of the distribution. The selected distribution must not only be consistent with the theoretical upper and lower bounds of the variable, but the tails of the distribution should reflect the likely variability within these regions because these are often the regions of interest in risk assessment (e.g. the 90th or 95th percentile of the distribution of risk). These issues can be addressed by selecting the distribution type (e.g. selection of a lognormal distribution to represent a variable that is positively skewed and cannot be negative, such as an exposure concentration), and/or by truncating a distribution at its theoretical upper and lower bounds (e.g. selection of a normal distribution truncated at 0 and 1 to represent an absorption factor). In some cases, truncation may be applied to avoid the generation of unrealistic values, such as excessively low or high adult body weights.

If insufficient data are available to select and define a distribution, or in the absence of an understanding of the underlying mechanism of the variable, the risk assessor should consider whether the variable should be characterized probabilistically or whether the use of a conservative point estimate value would be more appropriate. If a distribution is still desirable in order to reflect the expected variability, use of the simplest distribution that is consistent with the known constraints of the variable is recommended. For example, a uniform distribution may be appropriate to describe a parameter that is known to vary between an upper and lower bound, but where little is known about the nature of that variation. Similarly, a triangular distribution can be used, with or without truncation at the upper and lower limits, when the variable is known to exhibit a degree of central tendency around a modal value.

An alternative approach, when little is known about a variable, is to use a parametric distribution that reflects maximum uncertainty within a given set of constraints, such as a beta distribution. A disadvantage of this approach is that the specification of such a distribution can be complex. Furthermore, the form of the resulting distribution may be non-intuitive and create the impression of a greater degree of knowledge of the variable, whereas a simple distribution immediately conveys to the audience or reviewer the limitations in the data on which it is based. Caution is advised in establishing the upper and lower bounds for a variable that is not well characterized; this is because the output of a probabilistic analysis is generally more sensitive to the range than to the shape of a distribution.

The procedure for selecting or assigning a probability distribution using a parametric approach may be summarized as follows.

- Determine the appropriate form of distribution, based on the underlying mechanism(s) of the variable.
- Establish the constraints of the data set or population (i.e. known statistical parameters, bounds, etc.).
- Select and define the appropriate distribution in terms of its constraints.
- Examine and adjust the distribution as required to ensure that the variable is adequately represented over its entire range, including the extreme values of the distribution.

Note that if statistical parameters, such as mean and standard deviation, are used to define the shape of a distribution that is subsequently modified (e.g. by truncation), the statistical parameters will generally not be representative of the resulting population.

Non-parametric methods

Non-parametric methods of establishing probability distributions include the use of actual data sets, empirical distributions, and graphical or statistical methods of distribution fitting. Although methods based on empirical data tend to be largely objective and therefore defensible, it is noted that prior knowledge and judgment are often introduced to these methods in order to address data limitations; the methods therefore can become somewhat subjective.

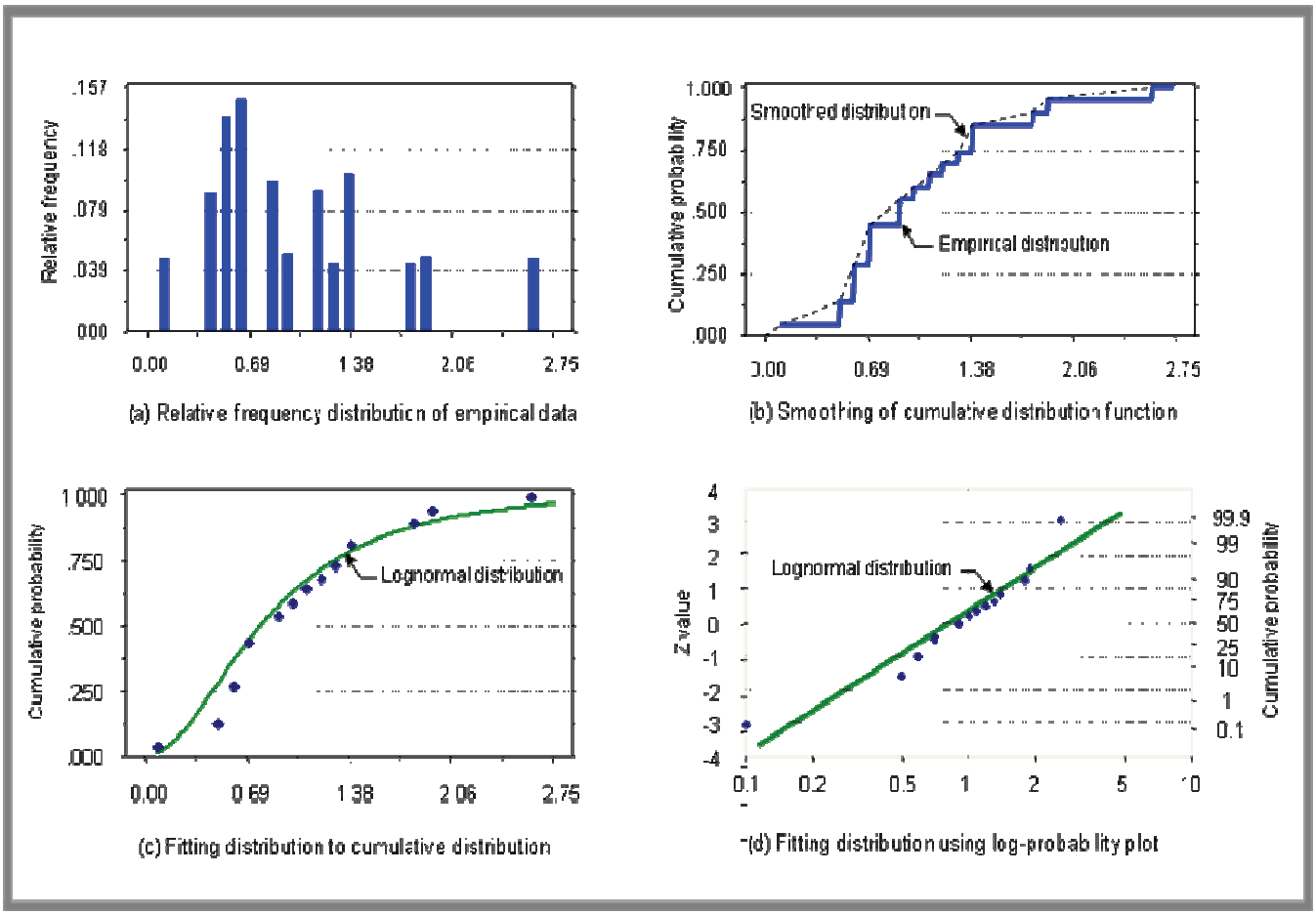
1. Use of actual data sets

Most computational methods for probabilistic risk assessment allow input variables to be specified in the form of actual data sets known as empirical or custom distribution functions. Although such methods are objective and provide complete representation of the data with no loss of information, a significant limitation to the approach is that a data set must be large and based on random sampling in order to adequately represent a population. Data representing the tails of the population distribution may be limited, and values outside the range of actual data cannot be represented. Empirical distribution functions are useful, however, for describing complex variables, such as bimodal distributions that may be associated with distinct subgroups within a population.

In the absence of a large random data set, the use of empirical data can be enhanced by incorporating certain

modifications to the empirical distribution function. For instance, the empirical distribution may be smoothed using a smoothing method such as linearized interpolation or a least-squares method (Figure 7.4). In addition, the tails of the distribution may be extended by adding upper and lower bounds and/or a parametric distribution (e.g. an exponential distribution) to the tails.

Figure 7.4 Fitting Distributions to Empirical Data Sets Using Graphical Techniques



2. Distribution fitting

Distribution fitting may be done graphically (visually) and/or statistically. Visual and graphical methods involve overlaying a hypothesized probability distribution onto an empirical distribution using a suitable graphing format, and estimating the statistical parameters of an appropriate distribution. This is best performed using a graphing space that is transformed such that a given distribution type plots as a straight line. For example, a normal distribution would be a straight line on a rank probability or *z*-value plot; similarly, the log-transformed values from a lognormal distribution would also be a straight line on a rank probability plot. Other graph types that can be used to facilitate visual distribution fitting include the cumulative distribution function (CDF) or a form of the rank probability plot in which the values from the CDF for the hypothesized distribution are replaced with either probabilities or percentiles. In the latter case, the hypothesized distribution would be a straight line and the empirical data can be readily compared. These approaches are illustrated in Figure 7.4.

Statistical distribution-fitting methods essentially use goodness-of-fit tests to evaluate the differences between the hypothesized distribution and the empirical data. Methods include the use of residual plots to display and assess differences between empirical data points and a candidate distribution, least-squares regression methods in conjunction with graphical plots to determine a best-fit line, and significance or hypothesis testing in conjunction with a statistical goodness-of-fit test such as the chi-square test. The latter approach yields a **p-value**, which is a measure of the likelihood that the data set represents an independent sample from a population represented by the candidate distribution. A p-value of 0.05 or 0.1 is often used as a cut-off; a p-value below this level suggests that it is unlikely that the empirical data and candidate distribution represent the same population. Caution should be exercised in interpreting p-values; several distributions may yield a high p-value when compared to different data sets. The p-value should not be used as an absolute indicator for selecting the most appropriate distribution; rather, its purpose should be to reject distributions that are statistically significantly different from the empirical data.

Uncertainty

As noted previously, uncertainty should be addressed separately from variability where possible. Data sets of measured values, such as contaminant concentrations across a spatially extensive soil unit, commonly exhibit both variability and uncertainty that cannot readily be distinguished without a detailed assessment of uncertainty in sampling and analytical techniques. The resulting distribution of estimated risk would also reflect variability and uncertainty, both of which contribute to the overall confidence of the risk assessment, but of which the influence cannot be separately evaluated and communicated.

Where sampling and measurement errors are known or can be independently determined, uncertainty may be separated from variability. Also, in the case of parameters that would not be expected to exhibit significant variability, or for which variability is unimportant (e.g. contaminant concentrations across a small area that are to be averaged for the purposes of exposure assessment), variations in a data set of measured values can be interpreted as primarily reflecting uncertainty.

Uncertainty may be characterized in a number of ways in probabilistic risk assessment, depending on the source of uncertainty. If it is due primarily to sample size effects, it is frequently accounted for by the determination and use of confidence intervals (e.g. upper 95% confidence limit of the mean) or an alternate appropriately conservative statistic (e.g. 90th percentile). If a measurement procedure is known to be subject to an error of $\pm x$%, this can similarly be represented by a conservative upper (or lower) bound. In these cases, the appropriate parameter values may be used in the form of constant point estimates or as ranges of values that can be varied discretely.

Uncertainty can also be represented by probability distributions. Examples include the distribution (generally normal) of the estimated mean of a population based on a given sample, or a representative distribution for a measured parameter determined from the distribution of a sample data set. Specific methods are available to determine probability distributions for key population parameters that cannot be reliably estimated using statistics from small samples. One example is **bootstrapping**. Bootstrapping is a method of repeated re-sampling of a population to generate probability distributions for key parameters, based on the samples. Because the population itself is not available for re-sampling, it is assumed that the population exhibits the same characteristics as the actual sample or data set, and the population can be generated by reproducing the sample a large number of times. In practice, re-sampling of the population is actually achieved by re-sampling the original sample with replacement. The result is a set of probability distributions reflecting uncertainty in any number of key parameters, such as the mean, standard deviation, or a specified percentile. These distributions can then be used directly, if desired, to characterize an uncertain term in a probabilistic risk assessment.

If uncertainty is accounted for using a single conservative statistic for a given parameter, no variation due to uncertainty is reflected in the resulting distribution of estimated risk. If the uncertainty in a parameter is reflected by a range of values or a probability distribution, the effect of this uncertainty could be assessed within the “outer loop” of a two-dimensional probabilistic analysis, or through discrete rerunning of the

one-dimensional probabilistic model for different values of the key uncertain parameter(s) (essentially equivalent to a sensitivity assessment).

The foregoing sections described general principles for the characterization of variability and uncertainty in input parameters for probabilistic exposure assessment. The approaches presented can be applied to the characterization of contaminant concentrations, physical and chemical parameters governing contaminant fate and transport, and receptor exposure characteristics, depending on the nature of the variability within the respective parameters. Additional considerations specific to contaminant and receptor characterization are discussed in the following sections.

7.2.2.2 Characterization of chemical concentrations

The way in which contaminant concentrations are represented in probabilistic risk assessment depends on the nature of the variation in concentration and the purpose of the risk assessment. Risks associated with contaminated sites are commonly assessed for receptor exposure based on average concentrations in source media (e.g. soil and groundwater). Those average concentrations are represented by parameters such as the upper 95% confidence limit of the mean. If spatial or temporal variability is important, then probability distributions may be used to describe the concentrations. For example, if receptors are potentially exposed to contaminated soil on a random basis over a relatively wide area, perhaps as a result of random deposition of a contaminant, a more complete representation of the variability of the soil concentrations is appropriate. Similarly, if dissolved contaminant concentrations in groundwater are found to vary with time as a result of fluctuating infiltration rates and groundwater flow conditions, a probability distribution based on historical measurements may be used.

The form of distribution used to represent variability in contaminant concentrations would be determined on the basis of available data and knowledge of the processes of contaminant deposition and movement of the contaminated media. Contaminant concentrations in the environment often exhibit normal or lognormal distributions; however, such distributions may require modification to exclude negative values and values in excess of theoretical solubility or saturation limits or other constraints. Also, if the intent of a risk assessment is to assess exposure to spatially variable contaminant concentrations, the distribution employed must reflect unaffected areas (i.e. background or zero concentrations), if applicable. In this situation, or in situations involving multiple contaminant sources, a multi-modal probability distribution may be required.

The risk assessor should exercise professional judgment in determining how contaminant concentrations are best represented in a given situation. It is noted that the use of a complete probability distribution to estimate exposure, when average exposure is of interest, would increase the variability of the estimated exposure or risk that, in turn, could result in the selection of a more conservative value as the basis for risk management in order to achieve a given level of protection.

7.2.2.3 Characterization of assumptions governing contaminant fate and transport

Contaminant fate and transport modelling generally requires the use of a large number of assumptions related to factors such as soil conditions, hydrogeological parameters, contaminant physical and chemical properties, meteorological conditions, and time-dependent chemical or biological transformation processes. Many of these factors are inherently variable and uncertain.

Most relevant physical and chemical properties of contaminants are relatively well established and consistently presented in the literature, although some experimentally determined contaminant-specific properties, such as organic carbon and octanol–water partitioning coefficients, may vary over orders of magnitude. Soil and hydrogeological properties exhibit natural variability; some parameters, such as hydraulic conductivity, can also vary by orders of magnitude. Soil conditions and hydrogeology, however, would normally be assessed on a site-specific basis with the aim of establishing representative values, or ranges of values, for use in fate and transport modelling. Factors governing time-dependent chemical or biological processes, such as biodegradation rates, and other processes that may be affected by meteorological conditions or seasonal phenomena may be assessed based on a combination of published information, site-specific data, and historical records (in the case of meteorological data).

Where such factors are likely to be significant to the estimation of risks, it may be appropriate to account for variability and uncertainty in key parameters. Because of the large number of parameters used in fate and transport modelling, a discrete sensitivity analysis would normally be conducted to identify those parameters that have the greatest influence on the results of the risk assessment, relative to their respective uncertainty and variability. This would assist in limiting the complexity of the resulting probabilistic risk assessment, as well as in focusing efforts for additional data collection. In accounting for variability and uncertainty in model parameters, it should also be recognized that the models themselves are subject to uncertainty, as mentioned previously.

Given the large number and diversity of variables affecting contaminant fate and transport, it is beyond the scope of this document to provide guidance for the selection of specific distribution types to represent these parameters. However, the approaches outlined previously in section 7.2.2 would normally be followed, with the overall goal being to select the simplest distribution that adequately describes the data or meets the constraints while being consistent with the underlying mechanistic basis for the variable.

7.2.2.4 Characterization of human exposure factors and receptor characteristics

Human exposure factors and receptor characteristics are generally obtained from published sources for the population as a whole, but they may also be determined on a site-specific basis for receptor groups and subpopulations of interest. Statistical parameters for use in deterministic risk assessment are widely available for many of the key variables, such as body weight and intake rate of various environmental media, as well as for behavioural activity patterns such as time spent outdoors or time spent working. In many cases, corresponding parameters and associated distributions for use in probabilistic risk assessment are also available, although for some factors it may be necessary to consult the original studies or surveys to obtain complete information on variability of the specific characteristics.

Health Canada recommends the use of Canadian data where possible (Richardson, 1997). Richardson (1997) presents a compilation of information obtained from Statistics Canada data and other sources regarding exposure factors and related characteristics applicable to the Canadian population. As discussed previously in section 4.5, recommended mean values are provided for a number of relevant factors for use in deterministic risk assessment. Richardson (1997) also provides recommended parameters and distributions for use in probabilistic risk assessment. For most human exposure factors, population data are most closely represented by normal or lognormal distributions. Specifically, Richardson (1997) recommends the use of lognormal distributions to describe the exposure factors recommended therein. The reader is referred to Richardson (1997) for data related to food consumption (including the characteristics of Native populations of Canada) and selected other activity patterns.

Other sources of published information include the U.S. EPA (1997b); although specific to the U.S. population, considerable information is given on activity patterns, including cumulative distribution functions for time spent at various activities. The U.S. EPA, through its website, maintains links to updated studies on various human exposure characteristics.

Although lognormal distributions are recommended for many human exposure factors, caution should be exercised in the interpretation and use of such distributions for the consumption of certain foods. Statistics for consumption rates of foods that are subject to personal taste, such as fish, are often presented in the context of the subpopulation of consumers of the specific food type (e.g. fish eaters), as opposed to the population as a whole. An assessment of exposure to a chemical through consumption of fish would therefore pertain only to the portion of the population that consumes fish. The corresponding probability distribution for fish consumption by the population as a whole would be a bimodal distribution comprising a lognormal distribution with a “lumped mass” at a consumption rate value of zero (representing non-eaters of fish).

A further caution is provided on the use of probability distributions for food consumption. Most food consumption surveys are based on a daily or weekly “recall” of individual consumption patterns across a sample of the population. The data therefore represent the consumption of the population as a whole at a point in time, rather than the variability in consumption rate for an individual through a prolonged exposure period (e.g. a lifetime). It is therefore theoretically inappropriate to use such data for the assessment of long-term risks to an individual due to consumption of foods (Richardson, 1997).

Several assumed behavioural characteristics relate to the “exposure term” in the generalized exposure equation. These characteristics include activity patterns such as time spent per day at a particular activity or location, frequency of exposure, and exposure duration (e.g. number of years at a given residence or workplace). While population data are available for a number of these parameters (e.g. U.S. EPA, 1997b), simplified probability distributions, such as uniform or triangular distributions, are commonly adopted given the lack of a mechanistic basis for the variability of the parameters. Professional judgment should be used in assigning distributions to these variables, and the risk assessor should provide a rationale for the distribution selected.

7.2.2.5 Correlations among variables

In a probabilistic analysis, parameter values are sampled randomly from their respective distributions and used as input for each iteration of the analysis. Unless otherwise specified, variables are sampled independently of one another. In reality, however, key variables in the exposure model may be correlated (i.e. a mathematical association may exist between two or more variables). For example, in human exposure assessment, skin surface area would be expected to increase to some degree as body weight increases. On the other hand, air inhalation rate may decrease as body weight increases in response to decreased activity levels. A positive correlation is one where two variables increase or decrease together; a negative

correlation exists if one variable increases as the other decreases. Note that a causal relationship between variables is not required for a correlation to exist.

If correlated parameters are allowed to vary independently, the variability in the risk estimate may be either understated or overstated, depending on the effect of combining extreme values. Therefore, most computational methods for probabilistic analysis permit correlations between variables to be specified and hence reflected in the simulation sampling, resulting in a more realistic distribution of estimated risk. The effect of correlating variables is likely to be more evident in the tails of a distribution than in the mean or median values.

Correlations may be specified as mathematical correlation coefficients that measure the strength of a linear relationship between two variables. Variables exhibiting different probability distributions are unlikely to be linearly related; therefore, the use of a **rank** correlation (describing the correlation between rank or relative magnitude of two series of values) may be more appropriate. Since considerable data may be required to establish correlations between variables, a sensitivity analysis to determine the relative significance of correlations between variables on the model output should be conducted.

7.2.3 Toxicity assessment

Of the key assumptions and parameters used in a quantitative risk assessment, the toxicity and bioavailability of a contaminant are among those with the highest degree of uncertainty and variability. Intra-species variability, low dose and inter-species extrapolation, determination of absorbed dose, and other factors contribute to considerable variability and uncertainty in the establishment of a TRV. Although comprehensive statistical data are often produced in a toxicity study and may be sufficient to characterize at least some of the variability and uncertainty in the TRV, it is usually standard practice in HHRA for contaminated sites to use a conservative point estimate value as a TRV. For example, cancer SFs are commonly determined from the upper 95% confidence limit of the dose-response relationship, and threshold TRVs are typically established by applying uncertainty factors to conservatively determined "no-effects" levels.

In most probabilistic risk assessments, therefore, the toxicity of a given contaminant is represented by a single constant TRV, whereas other parameters governing human exposure may be considered probabilistically as appropriate. The resulting probabilistic estimate of risk does not, therefore, explicitly account for uncertainty and variability in the toxicity. However, in principle, the use of a conservative TRV ensures that the uncertainty and variability in toxicity will not lead to an underestimate of exposure or risk.

From some perspectives, the effort required to characterize and account for uncertainty and variability in other areas of the risk assessment process may be open to question, given the potential relative magnitude of uncertainty in toxicity. However, from the risk assessor's standpoint, the probabilistic treatment of other variables in the risk assessment model permits site conditions and population characteristics to be represented more realistically while still maintaining a certain "margin of safety" afforded by the conservative TRV.

7.2.4 Risk characterization

The risk characterization stage of a risk assessment includes the quantitative estimation of risk, the interpretation of the results of the assessment, and an evaluation of uncertainties in the assessment.

7.2.4.1 Estimation of risk

As in a deterministic risk assessment, the estimation of risk in a probabilistic assessment involves combining the estimated exposure from the exposure assessment stage with the appropriate TRV from the toxicity assessment in order to assess HQ and/or cancer risk. Because of the use of integrated risk assessment models, particularly in a probabilistic analysis, the combination of exposure and toxicity is not necessarily a discrete step, but is commonly conducted as an integral part of each iteration of the risk assessment model. Therefore, the output of a probabilistic analysis typically comprises a probability distribution of the calculated result (i.e. the HQ and/or cancer risk).

The probability distribution of estimated risk provides a complete description of the variability in risk as a result of the specified variability and/or uncertainty in the input parameters. Key statistics can be obtained from the distribution, such as mean, standard deviation, range, and any desired percentile. A single value of estimated risk may be required for communication to stakeholders or as a basis for subsequent risk management decisions; the value selected is usually that corresponding to a predetermined statistic that may reflect regulatory policy and/or the risk tolerance of the risk manager and other stakeholders. For instance, the estimated risk may be expressed as the mean risk (also referred to as "expected" risk) or as the 90th percentile of the distribution of risk (a risk level that may be exceeded in 10% of the possible scenarios encompassed by the probabilistic analysis).

7.2.4.2 Interpretation of risk assessment results

In addition to the informed selection of a single risk estimate, the results of a probabilistic analysis provide insight into the possible range and variability of risk. In practical terms, this information can be interpreted in a number of ways. The

ability to determine risk levels corresponding to various percentiles facilitates the selection of a risk value corresponding to a desired level of conservatism (e.g. worst-case, 90th or 95th percentile). Conversely, the likelihood of a target risk level being exceeded can also readily be assessed. For example, if a 90th percentile of risk is considered to be an appropriate basis for a risk management decision, a situation may arise where the 90th percentile risk estimate is considered acceptable (i.e. below the value determined by the regulatory authority to be essentially negligible), but where there is still a 5% chance (say) that the regulatory target could be exceeded. Such information is valuable in placing the risk estimate into context.

The likelihood of exceedance of a given risk value reflects the percentage of possible scenarios (combinations of exposure factors, site conditions, and contaminant source concentrations) that give rise to a higher level of risk. Depending on the assumptions for which variability and uncertainty are considered, the interpretation given to likelihood of exceedance may range from the probability of underestimating risk (where the predicted variation is due primarily to parameter uncertainty) to the percentage of the population that may be exposed to unacceptable risks (where the variation is primarily due to variability in exposure factors across a population). Limiting the use of probability distributions to certain selected categories of parameters (e.g. exposure factors) would facilitate the latter type of interpretation.

Finally, the results of a probabilistic analysis provide a basis for the assessment of conservatism inherent in a point estimate of risk.

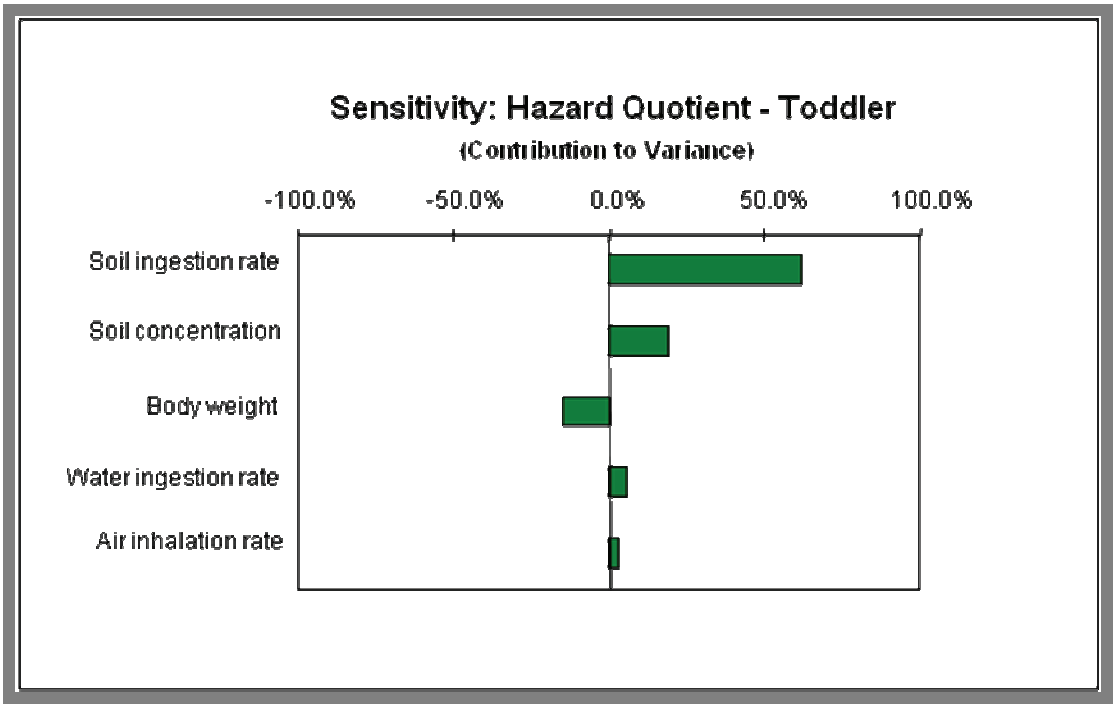
7.2.4.3 Uncertainty and sensitivity analysis

The results of a probabilistic risk assessment, like those of a deterministic analysis, are affected by uncertainty and variability in key assumptions. Parameters that exhibit natural variability may be both variable and uncertain; however, variability cannot generally be reduced by the collection of additional data, and the effects of variability are considered in the interpretation of the risk assessment results as discussed above. One reason for assessing uncertainty is to provide a measure of the level of confidence in the results of the risk assessment; another common reason is to assess the sensitivity of the results to uncertainty, so as to focus further data collection efforts. Different approaches may be employed to assessing uncertainty in each of these contexts.

Both variability and uncertainty affect the level of confidence in the results of a risk assessment. However, as stated previously, variability and uncertainty should, to the extent possible, be addressed separately in a probabilistic analysis. Uncertainties are commonly accounted for by the use of conservative point estimates for uncertain parameters; however, the effects of these uncertainties can only be quantified by varying the respective parameters within their ranges of uncertainty. In a deterministic assessment, this is normally accomplished by a discrete sensitivity analysis; in a probabilistic assessment, uncertain parameters would be assigned probability distributions but would be removed to the outer loop of a two-dimensional Monte Carlo simulation, whereby the effect of variations in the uncertain parameter(s) on the distribution of estimated risk, or key statistics thereof, could be assessed either step-wise or through a complete probabilistic simulation. Further discussion on two-dimensional Monte Carlo analysis can be found elsewhere (e.g. U.S. EPA, 2001); commercially available tools for probabilistic risk assessment, such as Crystal Ball®, accommodate such two-dimensional analyses.

A sensitivity analysis can be valuable for focusing data collection efforts aimed at reducing uncertainties. A discrete sensitivity analysis, normally conducted in conjunction with a deterministic assessment, indicates the relative influence on the estimated risk of independent or discrete variations in individual input parameters; the variations considered can be due to either uncertainty or natural variability. Information from a discrete sensitivity analysis is often used, as noted previously, to identify key parameters both for consideration in a subsequent probabilistic analysis, as well as for the collection of additional data. A limitation of the discrete sensitivity analysis is that the influence of a parameter on risk is heavily dependent on the potential range of values of that parameter; the analysis does not provide information on the relative contributions of key parameters within a realistic scenario, particularly where correlations exist among parameters. A sensitivity analysis may be performed in conjunction with a probabilistic assessment. Sometimes termed a **model** or **simulation** sensitivity analysis, this analysis provides information on the relationship between the variation in estimated risk and the variations in all key parameters, varied concurrently within their respective specified distributions. An example of the output of such an analysis is presented in Figure 7.5. The results are presented in the form of the contribution of each parameter to the variance of the estimated risk, but can also be presented in terms of rank correlation between each parameter and estimated risk. A key output of this type of sensitivity analysis is a clear ranking of the relative contribution of different input parameters. Note, however, that this does not distinguish between the effects of uncertainty and variability, except to the extent that the distributions defined for the input parameters reflect primarily one or other source of variation.

Figure 7.5 Sensitivity Chart from Probabilistic Analysis



7.3 Use of Probabilistic Methods for the Derivation of Risk-Based Remediation Objectives

Discussion in the foregoing sections has addressed the interpretation of probabilistic risk assessments conducted in a **forward** manner, whereby contaminant source concentrations are used to estimate risk to human receptors at a site and to determine the need for risk management. Risk assessments are also commonly carried out by **backward** calculation to determine acceptable concentrations corresponding to a certain target risk level (e.g. site-specific risk-based remediation objectives for soil and groundwater). However, the use and interpretation of probabilistic analyses conducted in conjunction with backward calculations are not necessarily as transparent and straightforward as deterministic methods, and therefore merit some discussion.

The methods are discussed herein with reference to determination of soil remediation objectives, but the concepts are broadly applicable to acceptable concentrations in any source medium. Methods for back-calculating a soil concentration that corresponds to a target level of risk fall into two categories: (i) rearranging the terms of the risk model such that the specification of target risk results in a calculated soil concentration, and (ii) iterative forward calculation until a soil concentration is found that results in the target risk level. The two methods are equivalent in deterministic assessment, but their equivalency in probabilistic analysis depends on the way soil concentration and risk are expressed.

In a forward probabilistic analysis, soil concentration may be specified as a point estimate or as a probability distribution; in both cases, the output is typically a probability distribution of risk. On the other hand, in a backward calculation, the target risk would be specified as a single value and not as a distribution; the resulting soil concentration would still be in the form of a distribution. Using a single soil concentration to estimate a distribution of risk is generally equivalent to using a single risk value to estimate a distribution of soil concentration, provided that the other variable parameters in the model are independent of one another and of the soil concentration and risk terms. Also, the appropriate statistics must be used to select values from the output distributions. For example, the 90th percentile of risk in a forward calculation corresponds to the 10th percentile of soil concentration in a backward calculation (i.e. represents the same level of protection). Other factors associated with representation of the

concentration and risk terms, and with the modelling itself, have also been found to result in the two methods not being equivalent (U.S. EPA, 2001). It is recommended that, if acceptable soil concentrations are determined by backward calculation, a forward calculation be performed to confirm that those concentrations do in fact result in a risk estimate equal to or less than the target risk. Use of iterative forward calculation avoids these concerns but may be computationally more intensive. If soil concentration is specified as a probability distribution in the latter type of calculation, uncertainty and variability should be adjusted appropriately with the soil concentration in each iteration.

It is important to note that an acceptable soil concentration corresponding to a target risk level, obtained from either a forward or backward probabilistic analysis, is not necessarily the same as a remediation objective, depending on how the soil concentrations are represented in the model and how the remediation objectives are implemented. Soil concentration may be specified as a probability distribution (reflecting uncertainty and/or variability) or as a suitable upper confidence limit of the mean (reflecting uncertainty). A remediation objective is generally a maximum acceptable concentration above which soil removal or remediation is conducted. Theoretically, this would be the maximum value of a distribution of concentrations that results in an acceptable level of risk, and would not be the same as the upper confidence limit of the mean. Furthermore, the post-remediation probability distribution of soil concentrations will likely be different from that used to characterize the pre-remediation situation because of the removal of higher concentrations. Some of these considerations are not unique to probabilistically derived soil concentrations, but should be borne in mind when establishing and implementing any form of risk-based remediation objectives. Further discussion of the development of site-specific remedial objectives is presented in section 8.0.

7.4 References

Oregon Department of Environmental Quality (Oregon DEQuality). 1998. *Guidance for Use of Probabilistic Analysis in Human Health Risk Assessments*. Portland, OR.

Richardson, G.M. 1997. *Compendium of Canadian Human Exposure Factors for Risk Assessment*. O'Connor Associates Environmental Inc., Ottawa.

United Stated Environmental Protection Agency (U.S. EPA). 1997a. *Guiding Principles for Monte Carlo Analysis*. U.S. EPA, Washington, DC. EPA/630/R-97/001.

U.S. EPA. 1997b. *Exposure Factors Handbook*. U.S. EPA, Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 2000. *Options for Development of Parametric Probability Distributions for Exposure Factors.* U.S. EPA, Washington, DC. EPA/600/R-00/058.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund: Volume III – Part A, Process for Conducting Probabilistic Risk Assessment*. U.S. EPA, Office of Emergency and Remedial Response, Washington, DC. EPA 540-R-02-002.

8.0 DEVELOPMENT OF SITE-SPECIFIC RISK-BASED REMEDIAL OBJECTIVES

8.1 Introduction

The traditional application of contaminated site risk assessment methods is to estimate the risks associated with a particular level of contamination. However, the same methods are also frequently used to develop objectives (target concentrations) for remediation or risk management (i.e. the maximum acceptable chemical concentrations in exposure media). Instead of estimating risks based on specified chemical concentrations, the chemical concentrations that would lead to a specified degree of (acceptable or negligible) risk are estimated.

Many of the available generic environmental quality guidelines have been developed using risk assessment methods, including the *Canadian Environmental Quality Guidelines* (CCME, 1999) and similar guidelines and standards developed by several provinces. These guidelines are generally developed using protocols (e.g. CCME, 2006) that specify exposure equations, receptor characteristics, fate and transport models, and input parameters for models and equations. In most cases, a few common land use scenarios (e.g. agricultural, residential, commercial, industrial) are defined, and generic screening-level guidelines are developed for each of these land uses. The generic guidelines are often referred to as "Tier 1" guidelines.

Generic guidelines are developed to reflect a limited number of common land uses and exposure scenarios, but they cannot encompass every possible situation. Land uses and/or exposure scenarios at specific sites may be very different from any of the "generic" land uses and/or exposure scenarios specified for guideline development, or default soil or other characteristics assumed for guideline derivation may differ greatly from those of a specific site. In either case, it may be appropriate to develop site-specific remedial objectives that are established to achieve the same level of risk reduction/health protection, but reflect the unique characteristics of the site that affect exposure. Also, generic guidelines are normally developed using relatively conservative assumptions and "screening-level" models; the use of more sophisticated approaches may allow for the development of remedial objectives for a site that reflect greater "realism."

8.2 Levels of Complexity

Several different approaches can be used to develop site-specific remedial objectives, depending on the site and the needs of the proponent.

Elimination of incomplete exposure pathways: Generic (Tier 1) risk-based guidelines typically provide values for, or simultaneously incorporate, several exposure pathways. If one or more of these exposure pathways is not applicable at the site in question, the values or dose contributions derived for these non-operational pathways can be eliminated. For example, if there is no potable groundwater aquifer beneath the site, the guideline values calculated for the protection of potable groundwater or the dose from ingestion of contaminated groundwater could be eliminated, given that other relevant pathways (e.g. vapour intrusion to indoor air, if relevant) are considered and included.

Recalculation of guidelines using site-specific parameter values: By using site-specific values for model input parameters (e.g. soil properties, depth to groundwater, building characteristics) with the same models and methods used to develop the generic guidelines, remedial objectives can be developed which reflect the specific site. This is often referred to as a "Tier 2" approach.

Development of remedial objectives using a site-specific risk assessment: The equations, scenarios, and parameters of a more complex site-specific risk assessment can be used to develop remedial objectives by solving for media concentrations, given an acceptable or negligible chemical dose. This is often referred to as a "Tier 3" approach, and is the main focus of this section.

Most Canadian jurisdictions require the same level of human health protection to be demonstrated, regardless of which approach is used.

8.3 Calculation of Site-Specific Risk-Based Remedial Objectives

The approach used to calculate site-specific risk-based remedial objectives is essentially the same as that used to perform a site-specific risk assessment, as documented in earlier sections. However, the calculations are performed in a "backwards" mode; a target risk is defined, the exposure dose that would lead to that risk is determined, and the chemical concentration in each relevant exposure medium that would lead to that exposure dose is calculated.

Site-specific remedial objectives can be calculated using either deterministic or probabilistic methods. When probabilistic methods are applied, a point estimate is selected from the probability distribution of calculated objectives, such as the 5th percentile, that would ensure that at least 95% of the receptor population receives a dose ≤ the acceptable or negligible dose.

Both background exposure to the COPCs and exposure through additional contaminated media at the site should be considered. The CCME (2006) protocol for deriving soil

quality guidelines accomplishes this in generic guidelines for non-threshold chemicals by subtracting estimated background exposure to the chemical from the TDI to determine a residual tolerable daily intake (RTDI), apportioning the RTDI equally between potential site exposure media (soil, water, air, food, and consumer products) and adding the background soil concentration to the calculated soil quality objective. The CCME (2006) approach is shown in Example 8.1.

Example 8.1 Canadian Council of Ministers of the Environment Method

A site-specific risk-based remedial objective for the protection of human health is being derived for copper in soil at an industrial site. The main exposure pathways are expected to be direct contact with contaminated soil (incidental soil ingestion, dermal contact, and inhalation of soil particulate). Only adults are expected to be exposed to the soil, with an exposure time of 10 hours per day, 5 days per week, 50 weeks per year. Workers wear long-sleeved clothing at all times, so only hands are considered to be exposed for the dermal contact pathway. The soil particulate concentration in air is approximately 200 µg/m^3.

An assessment of background exposure indicates that a typical non-smoking adult in the area would have an estimated background exposure of 0.022 mg/kg bw/d. The background copper concentration in soil is determined to be 85 mg/kg. The toxicity assessment indicates that the *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, *Version 2.0.* (HC, 2010) The tolerable daily intake (TDI) for copper (0.1 mg/kg bw/d) is appropriate. A bioavailability assessment indicates that a relative absorption factor (RAF) of 0.06 for dermal contact is appropriate.

The receptors are assumed to have the following characteristics:

body weight = 70.7 kg
soil ingestion rate = 0.02 g/d
inhalation rate = 16.6 m^3/d
skin surface area – hands = 890 cm^2
soil loading to skin – hands = 1 x 10^{-4} g/cm^2/event
soil exposure events/day = 1

A remedial objective can be calculated based on the CCME approach to soil quality guidelines as follows (adapted from CCME, 2006).

$$PSQG_{HH} = \frac{(TDI - EDI) \times SF \times BW}{[(RAF_G \times SIR) + (RAF_L \times IR_S \times ET_2) + (RAF_S \times SR)] \times ET_1} + BSC$$

where,

$PSQG_{HH}$	=	preliminary human health-based soil quality guideline (mg/kg)
TDI	=	tolerable daily intake (mg/kg bw/d)
EDI	=	estimated daily intake (background exposure assessment) (mg/kg/d)
RSD	=	risk specific dose (mg/kg/d)
SF	=	soil allocation factor (unitless)
BW	=	body weight (kg)
BSC	=	background soil concentration (mg/kg)
RAF_G	=	relative absorption factor for gut (unitless) = 1
RAF_L	=	relative absorption factor for lung (unitless) = 1
RAF_S	=	relative absorption factor for skin (unitless)
SIR	=	soil ingestion rate (kg/d)
IR_S	=	soil particulate inhalation rate (kg/d)
	=	air inhalation rate (m^3/d) x soil particulate concentration in air (kg/m^3)
SR	=	soil dermal contact rate (kg/d) – see below
ET_1	=	exposure term (unitless) accounting for days/week and weeks/year exposed
ET_2	=	exposure term (unitless) accounting for hours/day exposed

Soil Dermal Contact Rate

$SR = (SA_H DL_H + SA_o DL_o)\ EF$

where,
SA_H = exposed surface area of hands (m^2)
SA_O = area of exposed body surfaces other than hands (m^2)
DL_H = dermal loading of soil to hands (kg/m^2/event)
DL_O = dermal loading of soil to other surfaces (kg/m^2/event)
EF = exposure frequency (events/day)

Therefore,

$SR = (890\ cm^2 \times 10^{-4}\ g/cm^2/event) \times 1\ event/d = 0.089\ g/d = 8.9 \times 10^{-5}\ kg/d$

$IR_S = 16.6\ m^3/d \times 200\ \mu g/m^3 = 3320\ \mu g/d = 3.3 \times 10^{-6}\ kg/d$

$$PSQG_{HH} = \frac{(0.1\ mg/kg/d - 0.022\ mg/kg/d) \times 0.2 \times 70.7\ kg}{[(1 \times 0.00002\ kg/d) + (1 \times 3.3 \times 10^{-6}\ kg/d \times 10/24) + (0.06 \times 0.000089\ kg/d)] \times 5/7 \times 50/52} + 85\ mg/kg$$

$$PSQG_{HH} = \frac{1.1\ mg/d}{0.00000367\ kg/d} + 85\ mg/kg = 300804\ mg/kg$$

The site-specific risk-based remedial objective for the protection of human health would therefore be approximately 300,800 mg/kg when rounded down to two significant figures. If additional pathways were found to be active (e.g. migration of wind-blown dust to more sensitive nearby properties or migration to potable groundwater), these would also have to be evaluated.

The following main steps in the development of site-specific remedial objectives are essentially the same as in other risk assessments.

Problem formulation

- Define the objectives.
- Determine the scope/level of complexity.
- Identify COPCs.
- Identify potential human receptors.
- Identify potential exposure pathways.
- Develop the CSM.
- Establish the target risk and hazard levels.

Exposure assessment

- Characterize receptors.
- Conduct any necessary fate and transport modelling.
- Conduct exposure averaging and amortization.
- Assess bioavailability.
- Define exposure equations.

Toxicity assessment

- Classify chemicals based on toxicological action.
- Determine TRVs.
- Assess bioavailability.

Risk characterization

- Calculate site-specific remedial objectives.
- Conduct uncertainty and sensitivity analyses.
- Compare measured concentrations to remedial objectives.

There are a few key differences in the approach used to develop site-specific remedial objectives. As noted above, equations are rearranged to solve for the chemical concentration in the medium of interest that leads to a specified (acceptable or negligible) risk. Also, although some investigation of chemical concentrations is undertaken before the risk assessment, it is not always essential that chemical concentrations are fully characterized prior to developing the site-specific remedial objectives (physical characteristics of the site and characteristics of the receptor population must still be adequately defined, however).

The target risk and hazard estimates are typically the levels that would be deemed acceptable by the regulatory authority for a risk assessment. However, regulatory authorities often require a target HQ of lower than 1 (e.g. 0.2) to account for exposure from multiple contaminated media, background exposure, and/or toxic interactions among different chemicals at the site, unless these are explicitly considered during the derivation of the remedial objectives.

An illustration of the procedures used in the determination of site-specific remedial objectives is presented in Example 8.2.

Example 8.2 Site-Specific Remedial Objectives

Problem

Gasoline from a leaking underground storage tank has contaminated soil and groundwater beneath a service station in a small town. Concentrations of benzene, toluene, ethylbenzene, and xylenes (BTEX) considerably in excess of generic remediation guidelines were measured in soil at depths between 3 m and 7 m below grade, and in a groundwater aquifer with a water table approximately 6 m below grade. Several residences are located in the vicinity of the service station, and the local water supply is obtained primarily from private water wells. As part of an ongoing risk management plan, the property owner elects to have site-specific remedial objectives developed for BTEX. Based on discussions with the regulatory authority, an incremental lifetime cancer risk of 1 in 100,000 would be deemed essentially negligible, and a target hazard quotient of 1 would be acceptable if all contaminated media, background exposure, and potential chemical interactions are considered.

Solution

Receptors of concern based on the current land use would include workers at the service station (adults) and inhabitants of nearby residences (all ages). If the remedial objectives are intended for unrestricted future use, all ages would likely have to be considered on site as well. Direct exposure to contaminated soil is unlikely in this case; exposure pathways to be considered would include inhalation of vapours at the service station, inhalation of vapours in nearby residences (migrating offsite either as vapours or in groundwater), and ingestion of contaminated groundwater in nearby water wells.

Inhalation toxicity reference values have been published by Health Canada in *Federal Contaminated Site Risk Assessment in Canada, Part II: Health Canada Toxicological Reference Values (TRVs) and Chemical-Specific Factors*, Version 2.0 (HC, 2010). Because the only ingestion-based pathway is ingestion of groundwater, the BTEX concentrations are screened against the *Guidelines for Canadian Drinking Water Quality*.

Receptor characteristics are based on the *Federal Contaminated Site Risk Assessment in Canada, Part I: Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA), Version 2.0* (HC, 2010) guidance. Exposure amortization would be conducted based on commercial (on-site) and residential (off-site) exposure scenarios.

Although current exposure to BTEX could be evaluated using point-of-exposure measurements (e.g. indoor air and drinking water samples), fate and transport models would likely be necessary to evaluate future exposure and predict acceptable soil and groundwater concentrations. Models could be used to evaluate:

- biodegradation of BTEX over time,
- transport of BTEX through groundwater to offsite residences (for vapour inhalation) and water wells,
- migration of BTEX vapours in soil and groundwater into the onsite building and offsite residences, and
- volatilization of BTEX from water used for domestic purposes (showering, washing, etc.).

The specific models used would be selected based on the available site data, level of realism desired, and applicability to the scenario. Some models and software packages may allow for direct calculation of soil and groundwater remedial objectives based on the allowable point-of-exposure concentrations. If this is not the case, initial concentrations in source media may need to be specified. If the source and point-of-exposure concentrations are directly proportional (i.e. doubling the source concentration results in the point-of-exposure concentration also doubling), then it may be possible to solve for the remedial objective using the following equation:

$$\text{Remedial objective} = \frac{\text{Target hazard or risk} \times \text{Initial source concentration}}{\text{Calculated hazard or risk}}$$

If the source and point-of-exposure concentrations are not directly proportional for the model selected, then the model may need to be applied iteratively to determine the appropriate remedial objectives.

8.4 Site Use Restrictions

Depending on the assumptions made during the development of site-specific risk-based remedial objectives, there may be a need for site use restrictions or a risk management plan. Development of remedial objectives with these conditions is often undertaken when the proponent is retaining care and control of the site.

Although requirements may vary based on the regulatory authority involved, in general any assumptions that do not relate to fixed stable site conditions may lead to site use restrictions and/or risk management requirements. For example, if the remedial objectives are based on an assumption that human receptors would spend no more than 4 hours per day at the site or that any building on the site would include a vapour management system, measures must be in place to ensure that these assumptions remain valid. On the other hand, specifying a site-specific value for a soil property, such as the organic carbon fraction, based on-site measurements, would not normally lead to any long-term restrictions.

Likewise, excluding potential exposure pathways may also lead to site use restrictions or risk management requirements. For example, excluding exposure from the ingestion of contaminated groundwater may require restrictions on groundwater use at the site.

8.5 Recommended Deliverables

The presentation requirements for the derivation of site-specific remedial objectives is essentially the same as those for other risk assessments. The major risk assessment stages (problem formulation, exposure assessment, toxicity assessment, and risk characterization) must be fully documented. Presentation of the final risk-based remedial objectives is normally in tabular form, organized in a similar manner to the exposure assessment and toxicity assessment results. If multiple sets of remedial objectives are derived for a site (e.g. for different parts of the site, or for current and potential future land uses), each set should be clearly identified.

8.6 References

Canadian Council of Ministers of the Environment (CCME). 1999. *Canadian Environmental Quality Guidelines. Canadian Soil Quality Guidelines for the Protection of Environmental and Human Health*. Updated online : http://ceqg-rcqe.ccme.ca. Winnipeg.

CCME. 2006. *A Protocol for the Derivation of Environmental and Human Health Soil Quality Guidelines*. The National Contaminated Sites Remediation Program, Winnipeg. CCME-EPC-101E.

9.0 DUE DILIGENCE ISSUES RELATED TO FEDERAL SITES

Federal departments and consolidated Crown corporations that assess potential human health risks from a contaminated site, and subsequently use that assessment to design and implement risk management or remedial plans, should consult their agency's legal counsel concerning due diligence responsibilities on a site-by-site basis.

Canadian federal contaminated sites are a legacy of past management practices for which environmental consequences were not appreciated at the time. These sites are under the custodial care and responsibility of federal departments and consolidated Crown corporations. They include harbours and ports, military bases, Distant Early Warning (DEW) line sites, certain sites on First Nation reserve lands, and abandoned mines in the North, among numerous other categories. Federal contaminated sites represent an estimated financial liability to the federal government of $3.5 billion. Some of these contaminated sites pose significant human health risks, as well as environmental risks to flora, fauna, and habitat.

Federal departments and consolidated Crown corporations must exercise due diligence when sampling, assessing, managing, and/or remediating sites under their responsibility. Due diligence requires that the reasonable management and care of a site be undertaken at all times. Due diligence practices and activities must minimize potential adverse effects associated with the management and stewardship of the site. Due diligence practices must be consistent with the nature of the site, its setting with respect to its surroundings, the nature of potential hazards and hazard scenarios associated with the site, and the nature and extent of consequences associated with potential risks ranging from reasonably possible events to extreme and rare events.

In the broad sense, due diligence practices must address worker health and safety, public health and safety, and ecological protection. In developing remedial plans, reference should be made to environmental engineering industry practice and standards, as well as standards that have evolved through civil and criminal law. In its simplest terms, due diligence requires the timely notification of other jurisdictions (provincial, territorial, municipal), adjacent property owners, and/or local affected communities (neighbourhoods, municipalities, Aboriginal communities) if off-site impacts are known or suspected, or if access to the site by local populations (via recreation, hunting, fishing, trespassing, etc.) is known (and that access does or may present risks of unacceptable exposures).

It is incumbent on the site custodian to ensure that due diligence measures and programs are developed and implemented to prevent potential harm to humans or the environment, thus preventing so-called "toxic torts" in which legal action may be initiated against the site manager, custodial department, and/or federal government for failure to take reasonable action to inform those potentially at risk or failure to take reasonable action to prevent, reduce, or terminate that exposure in a timely manner.

10.0 GLOSSARY

Absorbed dose: The amount of a chemical penetrating the absorption barriers (the exchange boundaries) of an organism via either physical or biological processes. For the purpose of this document, this term is synonymous with internal dose.

Absorption: The process involving the taking up of chemicals by the skin, mucous surfaces, or absorbent vessels.

Acute exposure: Short-term exposure usually involving a single dose or exposures that are short in duration.

Acute TRV: The adverse effect occurring within a short time of administration of a single dose of a chemical or multiple doses given within 24 hours. See also **Chronic** and **Subchronic toxicity**.

Administered dose: The amount of chemical given to a test subject (human or animal).

Background exposure: The exposure of a receptor to a chemical of concern from sources not related to the contaminated site.

Baseline risk assessment: A health risk assessment conducted on the basis of prevailing (current) conditions at a site. Does not consider potential future site conditions following remediation.

Bioavailability: The tendency of a chemical to enter the general systemic circulation following administration or exposure; generally expressed as the fraction of the chemical that enters general systemic circulation.

Cancer risk: A numerical cancer risk estimate that is used by some regulators to estimate risks associated with exposures to non-threshold-response chemicals (i.e. genotoxic carcinogens). A numerical cancer risk value is calculated by multiplying the estimated exposure by the slope factor.

Carcinogen: An agent that is reactive or toxic enough to act directly to cause cancer.

Chronic exposure: Long-term exposure.

Chronic toxic effects: The development of adverse effects after an extended exposure (conventionally, at least one-tenth of the expected life span of an organism) to a chemical.

Conceptual site model (CSM): A qualitative model of how site-specific health risks may develop based on hypotheses describing contaminant source, release, environmental transport, and biological uptake.

Confidence limits: The bounds within which a population parameter is known to lie, to a specified degree of certainty. For example, it is 95% certain that the true mean of a population lies within the 95% confidence limits of the mean. See also Appendix A).

Contaminant (Chemical) of Potential Concern (COPC): A chemical that is not excluded as a result of screening procedures and is retained for further risk assessment.

Correlation: The degree to which two variables are related, often expressed as a value between 0 (no relation) and 1 (completely related); a negative value for correlation implies an inverse relationship. Correlation does not necessarily imply a cause–effect relationship between the variables.

Deterministic analysis: An analysis where point-estimate values are used to represent all variables in calculations.

Deterministic approach: An analytical approach to modelling which employs point estimates of input and output parameters and does not address uncertainty (variability) in the parameters.

Dose rate: Dose per unit time (e.g. mg/day), sometimes also called dosage. Dose rates are often expressed on a per-unit-bodyweight basis, yielding units such as mg/kg/d expressed as averages over some time period (e.g. a lifetime).

Dose-response: The relationship between the dose of a chemical administered or received and the incidence of an adverse health effect in exposed populations.

Environmental media (or medium): One of the major categories of material found in the physical environment that surrounds or contacts organisms (e.g. water, soil, or air) and through which chemicals can move and reach organisms.

Exposure pathway: The combination of contaminant release and transport via various media that results in contact with a receptor. Examples of exposure pathways include the ingestion of water, food, and soil, the inhalation of air and dust, and dermal absorption.

Exposure route: The physiological means by which a chemical enters the body. Conventionally taken to mean ingestion, inhalation, or dermal uptake.

Exposure scenario: A set of facts, assumptions, and inferences about how exposure takes place that aid the exposure assessor in evaluating, estimating, or quantifying exposures.

Hazard: The adverse impact on health that can result from exposure to a substance.

Hazard quotient (HQ): The form of risk estimate computed for threshold-response chemicals, also known as the exposure ratio. Derived by dividing the estimated environmental exposure rate (mg/kg bw/d) by a TRV (mg/kg bw/d).

Incremental lifetime cancer risk (ILCR): The increase in lifetime cancer risk above the normal risks associated with background exposures.

Maximum acceptable concentration (MAC): Exposure concentration not to be exceeded under any circumstances.

Microenvironments: Well-defined surroundings within a site that can be treated as homogeneous (or well characterized) with regard to the concentrations of a chemical or other agent.

Monte Carlo simulation (or analysis): A method of performing a probabilistic analysis whereby point-estimate values are selected at random from the probability distributions for each variable to obtain a point estimate of the calculation result; this is repeated many times to obtain a probability distribution of calculation results. See also Appendix A.

Non-threshold-response chemical: A chemical that is believed, in theory, to have the potential to elicit a toxic effect at any level of exposure greater than zero. Genotoxic carcinogens are generally included in this category. Generally the toxicity reference value is expressed as a RSD or a slope factor.

No observable adverse effects level (NOAEL): The highest dose of a chemical administered in a toxicity test, at which no adverse health effects are observed in the test organisms.

Potency factor: See **Slope factor**.

Potential dose: The amount of a chemical contained in material ingested, air breathed, or bulk material applied to skin.

Probabilistic analysis: An analysis where probability distributions are used to represent at least some variables in calculations. See also Appendix A.

Probabilistic approach: An analytical modelling approach that employs probability distribution functions to describe input and output parameters, thereby addressing uncertainty (variability).

Probability distribution : A statistical distribution of numerical values used to represent variability in a parameter.

Receptor: The people or other organisms that may be exposed to substances that are elevated at a contaminated site.

Risk: The likelihood, or probability, that the toxic effects associated with a chemical may be produced in populations of individuals under actual conditions of exposure. Risk is usually expressed as the probability of occurrence of an adverse effect (i.e. the expected ratio between the number of individuals who would experience adverse effects in a given time and the total number of individuals exposed to the risk factor). Risk is expressed as a fraction, without units, and takes values from 0 (absolute certainty that there is no risk, which can never be shown) to 1.0.

Risk analysis: The process and techniques used to identify and evaluate the nature and magnitude of a risk, as well as methods to best use the resulting information. Risk analysis includes risk assessment, risk communication, and risk management.

Risk-based remedial objective: A site-specific environmental quality criterion developed from basic exposure principles (equations) and using a pre-defined target risk estimate.

Risk estimation: The integration of the exposure assessment and the toxicity assessment in order to evaluate the likelihood of adverse human health effects associated with exposure to an environmental chemical.

Risk-specific concentration (RSC): The maximum average exposure concentration for non-threshold compounds. A RSC is a function of the unit risk and a defined risk level (e.g. 1 in 1 million). For example, a RSC for a risk of 1 in 1 million (10^{-6}) is RSC = 1 H 10^{-6} ÷ unit risk.

Risk-specific dose (RSD): The TRV determined for chemicals assumed to act as genotoxic non-threshold carcinogens. A RSD is a function of carcinogenic potency (slope factor) and a defined risk level (e.g. 1 in 100,000). For example, a RSD for human exposure at a risk of 1 in 100,000 (10^{-5}) is RSD = 1 H 10^{-5} ÷ slope factor.

Slope factor: A measurement of carcinogenic potency. The slope of the low-dose region of the dose-response model for the estimation of risk following exposure to a carcinogen.

Subchronic exposure: An exposure of a duration that is intermediate between acute and chronic.

Subchronic toxicity: The adverse effects resulting from the repeated daily exposure to a chemical for a short time (e.g. between 14 and 90 days). See also **Acute** and **Chronic toxicity**.

Threshold-response chemical: A chemical that elicits a toxic effect only at or above some threshold of exposure and manifests toxicity via a threshold-response mechanism. These chemicals have toxicity expressed as RfD, ADI, or TDI.

Tolerable Concentration (TC): An estimate of the maximum concentration to which the human population (including sensitive subgroups) could be exposed on a continual basis without an appreciable risk of adverse health effects.

Tolerable Daily Intake (TDI): An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure of the human population (including sensitive subgroups) that is likely to be without an appreciable risk of adverse health effects over a lifetime.

Toxicity: The production of any type of damage, permanent or impermanent, to the structure or functioning of any part of the body. The conditions of exposure under which toxic effects are produced: the size of the dose and the duration of the dosing needed vary greatly among chemicals. See also **Acute, Chronic,** and **Subchronic toxicity**.

Toxicological reference value (TRV): A value such as a TDI, TC, RSD, RSC, slope factor or UR used to represent the toxicity of a chemical.

Uncertainty: The lack of certainty due to lack of knowledge about an item or parameter. General types of uncertainty in risk assessments are related to parameter uncertainty, model uncertainty, and decision-rule uncertainty. This relates to the lack of knowledge about the true value of a particular parameter, which may result from insufficient data, sampling error, model limitations, etc.

Unit risk (UR): The amount of risk predicted per unit concentration (e.g. risk per mg/m^3 in air) to which a human receptor is exposed on a continual basis. The unit risk multiplied by the amortized exposure concentration is the estimated risk.

Uptake: The process by which a chemical crosses an absorption barrier and is absorbed into the body. See also **Absorption**.

Variability: A measure of the inherent diversity or heterogeneity within a population.

APPENDIX A
PROBABILISTIC RISK ASSESSMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

A1.0 OVERVIEW

In cases where risks to human health are clearly not negligible or not acceptable, a probabilistic risk assessment may be useful to better characterize risk. A probabilistic risk assessment takes account of uncertainties and data variability to produce estimated probabilities of exceeding toxicity benchmarks or probabilities of effects of differing magnitude. Generally, uncertainty relates to the lack of knowledge (incertitude) that interferes with accurately defining input variables and their underlying distributions. This uncertainty leads to a lack of knowledge about true exposure and risk with regard to the estimates generated in the risk assessment. Variability, on the other hand, relates to known quantifiable stochastic (often natural and inherent) differences across a population of data (e.g. body weights in a population, concentrations in a water body). Variability cannot be reduced. Uncertainty can be reduced through further research and data collection. Evaluating, calculating, and conveying the degree and magnitude of variability and uncertainty in each of the components of the risk assessment process provides decision makers and the public with a strong scientific foundation for understanding risk and evaluating the credibility of the final risk estimates.

Probabilistic risk assessment is more than Monte Carlo analysis, more than a listing of things uncertain, and more than the choice of a favourite statistic. The analysis is a process by which a degree of belief or probability is inferred about the possible values of a risk endpoint. In most real-world problems, uncertainty and variability are inherent in most steps of a risk analysis. For example, the choice of data set, treatment of outlying data points, choice of model, choice of spatial and temporal scales, etc., are decisions that impart uncertainty to the risk estimates. Some uncertainty is quantifiable, some is not. Each source of uncertainty has a role in the interpretation of risk and in the expectation that decisions made on the basis of the risk assessment will turn out as expected. A properly conducted probabilistic risk assessment leaves the reader with both a quantitative and qualitative set of information from which the severity, validity, robustness, and usefulness of the risk estimates can be judged.

The objective of this appendix is to describe the process used to conduct a probabilistic risk assessment (PRA) for human health endpoints at a contaminated site. Section A1.0 provides a brief overview of PRA, and includes topics such as when to do a PRA, why do a PRA, sources of uncertainty in human health risk assessment (HHRA), general mechanics of a PRA, and PRA techniques and outputs. Section A2.0 describes and provides guidance on each of the steps in a PRA, with a focus on the exposure portion of HHRA. Subsequent sections describe how PRA methods can be used in effects assessment (section A3.0) and in decision making (section A4.0). The appendix concludes with a description of the principles of good practice for PRA (section A5.0). A glossary is provided in section A7.0.

A1.1 When to Do a Probabilistic Risk Assessment

A deterministic analysis is almost always undertaken as part of a site-specific quantitative HHRA. Its purpose may be one or more of the following: to screen contaminants, exposure pathways, and/or receptors; to determine the need for a probabilistic risk assessment; to determine the sensitivity of the risk estimate to key assumptions (by sensitivity analysis); and/or to assess the requirement for additional data collection (U.S. EPA, 2001). In many cases, the risk assessment may not proceed beyond the deterministic analysis, either because the risks were shown to be negligible based upon a deliberately conservative analysis, or were shown to be obviously unacceptable. Often, however, the deterministic analysis serves as a scoping stage for a more detailed PRA.

Before proceeding with a PRA, the risk assessor should consider whether a probabilistic analysis is necessary and/or appropriate, given the objectives of the assessment and the availability of data. A probabilistic analysis necessarily involves a greater commitment of resources to collect required input data, to conduct the analysis, and to report and present the results. In practice, probabilistic analyses are more commonly conducted with large complex sites, where the costs and consequences of an incorrect decision are unacceptable (e.g. overlooking possibility of catastrophic events, or spending millions of dollars on unnecessary cleanup). In these cases, the additional resources required of a PRA are justified to ensure a complete understanding of risk and the uncertainties surrounding risk, and ultimately to ensure that a cost-effective risk management strategy is developed.

A1.2 What Is a Probabilistic Risk Assessment?

Risk is generally defined as the probability of a given hazard (health effect or exceedance of a benchmark deemed protective against health effects) occurring at a particular level of exposure. However, most contaminated site risk assessments conducted in Canada (as elsewhere) do not define that probability because deterministic methods are typically used to estimate risk. Quantitative HHRA generally involves assigning numerical values to input variables in an appropriate exposure or risk model or equation to obtain a quantitative estimate of risk. Numerical values are required for parameters describing contaminant concentrations in environmental media, human intake of and interaction with those media, contaminant fate and transport, human exposure, and toxic response. These values may be measured, assumed, prescribed, or based on published literature. Variability and uncertainty in the input parameters

or risk model result in variability and uncertainty in the resulting estimates of risk.

Traditional deterministic methods of quantitative risk assessment use single, or "point estimate," values for input parameters and produce a single estimate of risk or hazard. Although input parameters may be selected with some knowledge of their variability or uncertainty, a deterministic analysis does not normally provide any information on the variability of the resulting risk estimate. For example, although input values are often selected to represent either average or reasonable maximum exposure conditions, the location of the point estimate of risk in the context of its potential range and distribution cannot be determined directly. A discrete, or deterministic, sensitivity analysis may provide some indication of the potential range of estimated risk values, but the variability of, and hence confidence in, the risk estimate remains unknown.

PRA uses probability distributions to characterize variability and uncertainty in input parameters, and produces a probability distribution of estimated exposure or risk. The exposure distribution can be directly compared to a toxicity benchmark to estimate the probability of exceedance. Alternatively, the exposure distribution may be combined with a dose-response curve to generate a risk curve that indicates probabilities of effects of differing magnitude.

A1.3 Why Do a Probabilistic Risk Assessment?

Uncertainty is a widely recognized aspect of HHRA, but it is often ignored in regulatory applications. In decision making for contaminated sites, there are compelling reasons to conduct a PRA to avoid the mistaken impression that model results are precise and well understood (Finkel, 1994; Reckhow, 1994).

- Risk managers need to know the expected uncertainties in the model predictions, so that they can adjust their responses accordingly (e.g. ask for more experimentation or monitoring, hedge decisions away from large losses). In particular, knowledge of uncertainties in risk assessment is an essential part of the judicious application of the precautionary principle.[1]
- An uncertainty analysis can pinpoint the priorities for obtaining new information, so that uncertainty can be reduced, and the decision maker can have increased confidence in the decision ultimately made.

The traditional approach to dealing with uncertainties is to make the deterministic risk assessment conservative through the use of extreme assumptions and point estimates, and large safety factors. There are, however, costs to this approach (Moore and Elliott, 1996). In regulatory programs in which worst-case assumptions are the norm, expensive risk mitigation measures may be enacted for chemicals that pose little threat to human health or the environment. Conversely, in programs that rely on best-guess assumptions or so-called reasonable conservative assumptions, chemicals having low but real likelihoods of causing effects may be ignored. This would be a mistake if the effects were potentially catastrophic (e.g. stratospheric ozone depletion).

Uncertainty analysis makes clear what is known and what is not known about a particular variable, overall exposure, or risk—a huge advantage over the use of simple conservative assumptions and safety factors. Thus, uncertainty analysis provides an objective and transparent means of comparing assumptions, models, and data put forth by stakeholders in an assessment of a contaminated site. After a PRA, it may still be agreed that it would be prudent to be conservative in the selection and implementation of risk management measures. This is appropriate given that the place for applying issues, such as "what is an acceptable risk," is during the risk management stage (the stage at which societal interests are normally considered). Use of conservative assumptions and safety factors in an analysis has the effect of blurring the distinction between science and decision making. The task of assessors is to come up with estimates of what is likely to happen, what might happen, and what is not likely to happen, and to identify possible risk management options, but not to blur the distinction between science and policy. Extending an analogy by Reckhow (1994), a forecast of "it will very likely rain" when rain is highly unlikely is not helpful; rather, one would like to know the true odds and act according to public attitudes toward the risk in question. Thus, rather than bring an umbrella to work every day, one may choose to bring it only when the probability of rain is greater than 30%. The PRA approach does not negate a conservative approach, but rather moves it to the more appropriate risk management stage.

Chao et al. (1994) have provided an excellent example of how the consideration of uncertainties about the consequences of ground-level ozone can lead to a more cost-effective decision-making process. In their example, they considered uncertainties in emissions inventories, ozone formation processes, the transport of ozone and its precursors, and impacts on human health and ecological systems. Then they created a decision analytic tool to assess the effects of these uncertainties in the development of an optimal abatement strategy. Their analysis showed that a

[1] Precautionary principle (defined by the Government of Canada): This principle is put into practice in the context of regulatory risk assessment by implementing assumptions and exposure scenarios that are anticipated not to underestimate exposure. In other words, in the face of uncertainty, exposure and subsequent risk estimates overestimate actual exposure and risk.

flexible strategy, involving the use of less capital-intensive measures initially and taking advantage of new information in the future, reduces the expected total costs for meeting air quality goals when compared with the inflexible strategies initially considered. Thus, uncertainty analysis helps discriminate among management options, identifies critical information needs, and, as shown in this example, "can spur on the iterative search for new decision options that may outperform any of the initial ones offered" (Finkel, 1994).

A1.4 Sources of Uncertainty in Human Health Risk Assessment

Generating a list of the various sources of uncertainties that affect an HHRA is the first step en route to conducting a successful uncertainty analysis. Such a list will help structure the analysis and ensure that major sources of uncertainty are either quantified or explicitly excluded from the study (Finkel, 1990). Uncertainty can be classified in many ways (e.g. Finkel, 1990; Hoffman and Hammonds, 1994; Rowe, 1994; Hora, 1996; Cullen and Frey, 1999; Paté-Cornell, 2002). In this section, common sources of uncertainty in HHRAs of contaminated sites are described and classified according to type of uncertainty.

There are many sources or components of uncertainty in a typical risk assessment. In an HHRA of a contaminated site, one may be uncertain about the identity of the subpopulation at highest risk of exposure, possible routes of exposure, the appropriate multimedia exposure model, ingestion rates, concentrations of chemicals in different media, sensitivity of different age groups to the chemical of interest, the importance of modifying factors (e.g. diet, genetics, health), etc. Despite the long list of possible sources of uncertainty, they all belong to one or several of four general types of uncertainty: variability, incertitude arising from lack of knowledge about parameter values, model structure, and decision rules. For a more in-depth discussion of these types of uncertainty, see Finkel (1990).

Variability refers to observed differences in a population or parameter attributable to true heterogeneity (Warren-Hicks and Moore, 1998; U.S. EPA, 2001). It is the result of natural random or stochastic processes and stems from, for example, environmental, lifestyle, and genetic differences. Variability can be quantified, but it cannot be eliminated through collection of more data. Variability is an inherent characteristic of biological systems and processes. Examples of variability include variation among individuals in size (e.g. height, weight) and physiology (e.g. metabolic rate, food intake rate), and among environments (e.g. soil type, climate, chemical concentration).

Parameter uncertainty refers to our incertitude about the true values of the parameters or variables in a model (Warren-Hicks and Moore, 1998; U.S. EPA, 2001). Parameters are often estimated from laboratory, field, or other studies, or based on professional judgment or "best guesses." This type of uncertainty is introduced because the estimated value typically relies on insufficient, unreliable, or only partially relevant information for the parameter of interest. Several processes contribute to parameter uncertainty, including measurement errors, random errors, and systematic errors (Finkel, 1990). Measurement error often arises from the imprecision of analytical devices and methods use (e.g. to quantify chemical levels in different media or measure levels of detoxifying enzymes in humans). Errors in measurement, however, are not necessarily restricted to analytical hardware. Reconstructing past releases at a contaminated site may be subject to measurement error because historical data can be faulty or ambiguous. Random error or sampling error is a common source of incertitude in HHRA, arising when one tries to draw an inference about a quantity from a limited number of observations. For sample means, one can examine the importance of sampling error by calculating the standard deviation of sample means (Sokal and Rohlf, 1981). Sample means based on 3,000 observations will have a standard deviation only one-tenth that of means based on 30 observations. Systematic error occurs when the errors in the data are not truly random, such as might occur when the sample population is not representative of the entire population (e.g. when sampling is biased toward more contaminated areas). Systematic error, unlike random error, does not decrease with more observations and is not accounted for when calculating sample statistics (e.g. arithmetic mean, standard deviation). When systematic error is pervasive, sample statistics such as 95% confidence intervals can be quite misleading. For example, nearly half of the 27 measures of the speed of light measured between 1875 and 1958 had 95% or 99% confidence intervals that did not bracket the most accurate value available today (c = 299,792.458 km/s) (Henrion and Fischhoff, 1986).

In risk assessment, mathematical models or equations are used to determine which variables to measure, specify how they relate, and to estimate the values of variables that cannot be measured directly. **Model uncertainty** is a serious challenge in risk assessment (Finkel, 1990; Reckhow, 1994). Different dose-response models, for example, commonly lead to 2-fold or more differences in estimated low toxic effects doses (e.g. ED_{05} or LD_{10}), even when the list of models is restricted to those that fit the data equally well and are theoretically plausible (Moore and Caux, 1997). In cancer risk assessment, model uncertainty is further exacerbated by the need to extrapolate to very low levels of effect. Cothern et al. (1986) observed that a concentration of 50 µg/L trichloroethylene in drinking water provides a risk estimate of 1×10^{-2} with a Weibull dose-response model and 1×10^{-10} with a probit model. These estimates provide a range of uncertainty "equivalent to not knowing whether one has enough money to buy a cup of coffee or pay off the national debt" (Cothern et al., 1986). This example illustrates the difficulty of choosing an appropriate model equation even

with a simple system—one medium, one species, and constant laboratory conditions. The problem of model uncertainty is likely to be much more serious with complex models such as regional-scale fate and transport models. Most applications of uncertainty analysis in HHRA do not propagate uncertainties associated with model structure, rather the model structure is assumed reasonable and only parameter uncertainties (within-model uncertainties) are propagated. Beck (1987), Reckhow (1994), Oreskes et al. (1994), and others discuss the issue of model uncertainty, and describe the process for selecting, evaluating, calibrating, and validating models that, if followed, can substantially reduce this source of uncertainty in a risk assessment.

Decision rule uncertainty comes into play during risk management (i.e. after a risk estimate has been generated). This type of uncertainty arises when social objectives, economic costs, value judgments, etc., are part of the decision-making process for deciding on what actions to take to remediate or mitigate a problem. Individual decision makers (e.g. politicians) are likely to be uncertain about how to best represent the complex preferences of their constituents. Such uncertainty can be quantified by collection of empirical data (e.g. opinion polls) and formally treated via decision analysis, but it rarely is. Even with the availability of formal analytical tools, controversial judgments remain about how to value life, distribute costs, estimate benefits and risks among individuals and groups, and decide whether to reduce risks now or some time in the future (Finkel, 1990).

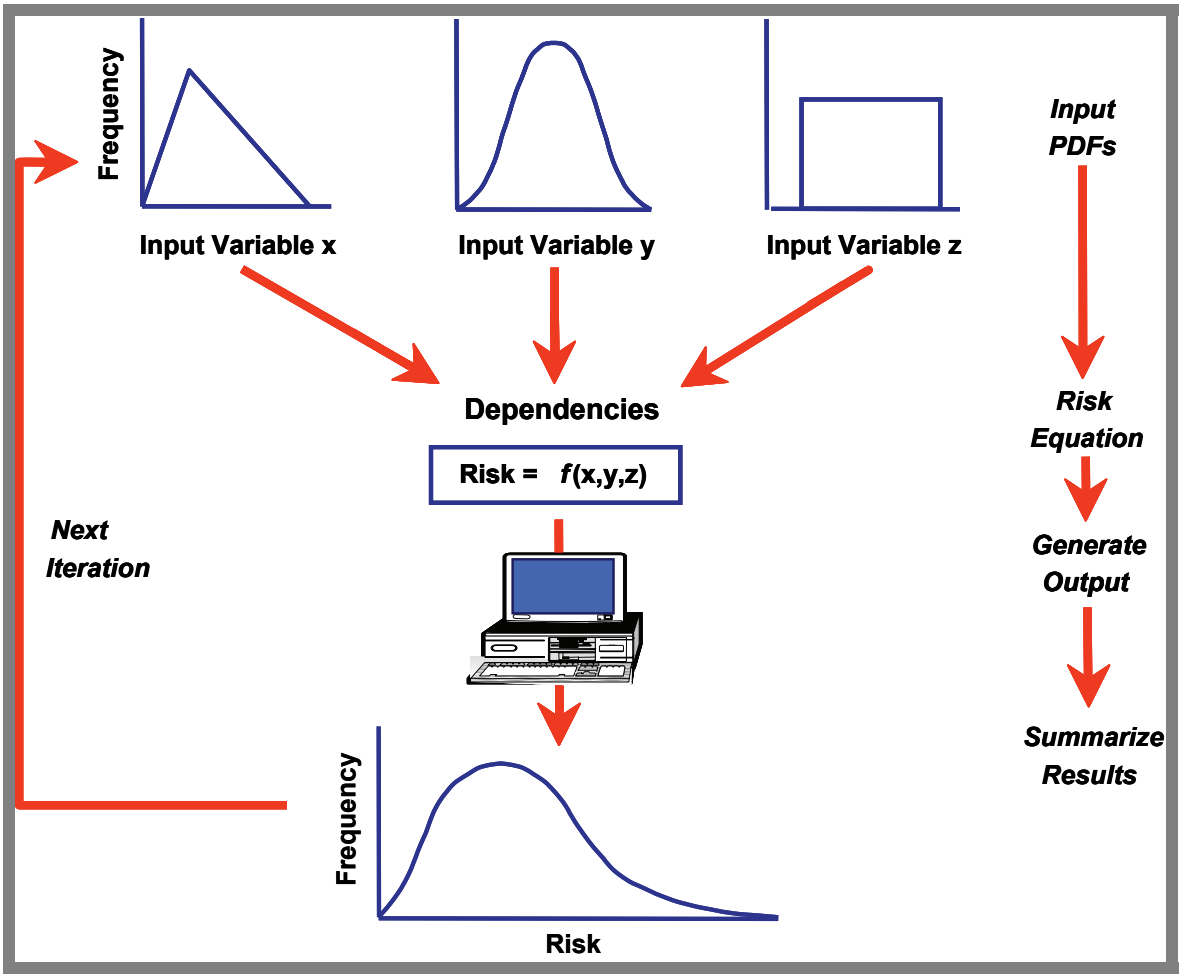
A1.5 General Mechanics of a Probabilistic Risk Assessment

The general mechanics of a PRA (see Figure A1) typically involves the following seven steps (see Finkel, 1990; Hammonds et al., 1994; Moore and Elliott, 1996):

1. **Specify the risk model equation.** The model equation specifies how the inputs will be combined to estimate exposure, effects and/or risk. It can range from the very simple (e.g. probabilistic quotient = exposure/no effects dose) to the complex (e.g. toxicokinetic model for trichloroethylene: Banks and Potter, 2004). The risk equation in a multiple chemical assessment must specify how the chemicals will combine to exert their effects on human receptors. This can be difficult when the chemicals interact in a non-additive fashion, compete for absorption or sites of cytotoxic action, or when the chemicals cause different types of effects (e.g. reduced growth, neurotoxicity) to exposed individuals. Reckhow et al. (1990) and a number of papers in Volume 15 of *Advances in Water Research* (Flavelle, 1992) discuss methods for calibrating, evaluating, and validating a model equation. These topics are also briefly discussed in A2.1.

2. **List all variables that will be specified as distributions.** In general, it is preferable to keep this list as short as possible by specifying input distributions only for those variables that are likely to have an important influence on the output (Seiler and Alvarez, 1996; U.S. EPA, 2001)

Figure A1 General Steps of a Probabilistic Risk Assessment



Note: PDF, probability density function; *f, function*

3. **Generate a distribution for each input variable in the model equation.** These are often referred to as probability density functions (PDFs) or probability mass functions (PMFs). The choice of distribution depends on (i) the form (distribution) of the observed data, which may be determined by graphical or goodness-of-fit statistical techniques, and (ii) a basic understanding of the input variable, so that theory about distributions can be used to best describe the underlying reality. Sharp (in Morgan and Henrion, 1990; Haimes et al., 1994; Hattis and Burmaster, 1994; Ott, 1995; Seiler and Alvarez, 1996; Cullen and Frey (1999); U.S. EPA, 1997, 2001) and others discuss distributions commonly used in risk assessment. Methods for parameterizing input distributions and guidance on their use are provided in section A2.2.

4. **Determine and account for dependencies among input variables.** This is an often-overlooked aspect of PRA. Ignoring correlations among important input variables (e.g. spatial correlations of multiple stressors) can lead to under- or overestimates of risk (Ferson and Burgman, 1995). Methods for accounting for dependencies among input variables are described in section A2.4.

5. **Generate the output distribution by combining the input distributions as specified in the model equation.** This step often involves Monte Carlo analysis, but there are other methods (see sections A1.6 and A2.5).

6. **Fine-tune the analysis.** Sensitivity analysis can be used to determine important input variables by identifying those input variables that, for example, have the highest correlations with the output variable. The available methods for conducting a sensitivity analysis are described in section A2.6. If re-examination of these variables reveals that they have little scientific support, additional empirical data and expert knowledge should be obtained. Once the input variables and, if necessary, the model equation have been fine-tuned, the analysis is

repeated. Fine-tuning of a PRA typically involves numerous iterations.

7. **Summarize the results, highlighting important implications for risk managers.** A variety of graphical and statistical techniques can be used, the choice of which depends on the outputs and the statistical sophistication of the audience. Warren-Hicks and Moore (1998) discuss various means of communicating uncertainty to lay and scientific audiences. Managers and interested parties should also be informed of unresolved scientific controversies and sources of uncertainty that could not be included in the quantitative analysis (Finkel, 1990; Covello and Merkhofer, 1993).

A1.6 Available Techniques for Probabilistic Risk Assessment

Essentially five approaches to probabilistic analysis have been used in HHRA:

- summary statistics
- first-order moment propagation
- first-order Monte Carlo simulation
- probability bounds analysis
- second-order Monte Carlo simulation

The first three approaches are more commonly used in HHRA. The latter two approaches are needed when analysts and managers want to separate variability (known as natural heterogeneity) from uncertainty due to lack of knowledge. Bayesian methods are also sometimes mentioned in such lists, but whether an analysis is Bayesian is really a separate consideration. Any of the five approaches to calculation can be used within a Bayesian framework, or outside of one. For a general overview of Bayesian methods, see Berger (1985).

Summary statistics may be used to characterize uncertainty in data sets where the focus is on a single variable (e.g. concentration of a contaminant in drinking water). A variety of summary statistics are available to estimate centrality (e.g. median, arithmetic mean, geometric mean, mode) and spread (e.g. standard deviation, absolute deviation, quartiles, range) in a data set (Sokal and Rohlf, 1981).

First-order moment propagation is a distribution-free approach that uses the elementary laws of probability to estimate the means and variances of sums, products, differences, and quotients based on the means and variances of the input variables (Slob, 1994). This approach is useful when it is hard to specify the statistical distributions of the input variables, but their means and variances are known. This is obviously a fairly crude approach, but it can be useful with simple models.

First-order Monte Carlo simulation is an increasingly widely used approach to PRA. The method requires the specification of the statistical distributions of each of the input variables and their interdependencies as measured by correlations. Computer software packages such as Crystal Ball® or @Risk® are used to "sample" from the distributions, and compute the risk expression many times so as to build up a histogram that serves as the estimate of the full distribution of risks (explicitly including the tail risks of extreme events). Monte Carlo is an approximate but very general strategy for such problems.

Probability bounds analysis is an exact numerical approach (not based on simulation) that takes as input the same probability distributions used in Monte Carlo simulation or, when they are difficult to specify precisely, the bounds on these distributions, and rigorously computes bounds on the output cumulative distribution functions (CDFs) (Ferson, 2002). Probability bounds analysis is also useful when independence assumptions are untenable (such as between body mass and inhalation rate), or when sparse empirical data make it difficult to quantify the correlations among variables (Ferson et al., 2004). This approach is closely similar in spirit with so-called robust Bayesian methods (Berger, 1985).

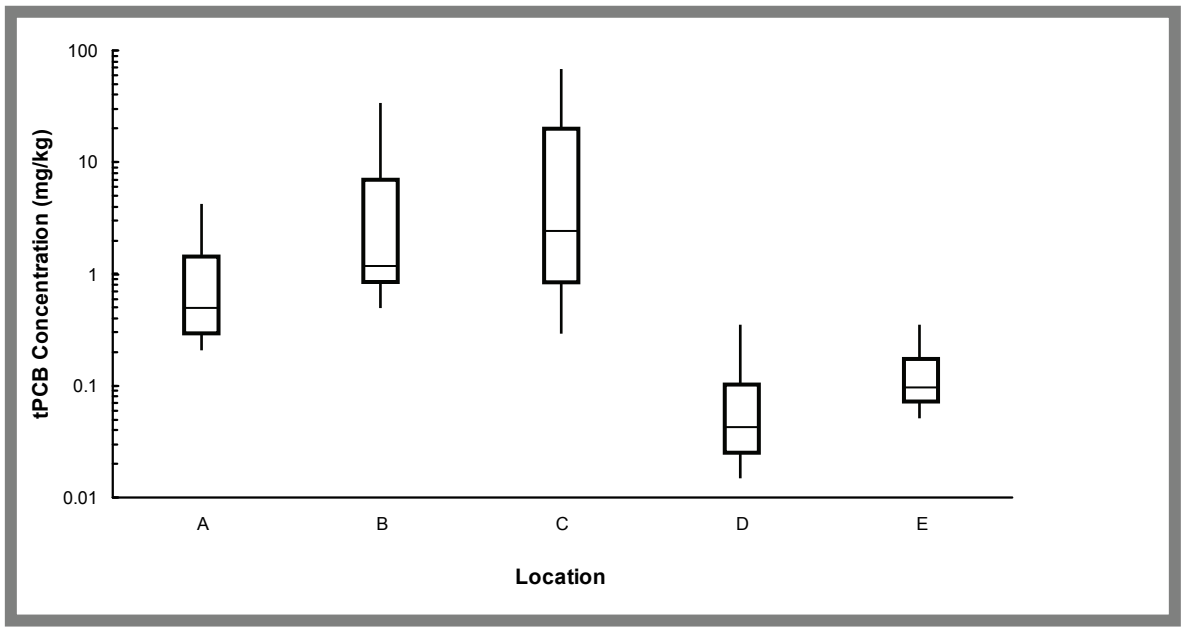
Two-dimensional Monte Carlo simulation is another approach, similar to probability bounds analysis, that is designed to handle both incertitude and variability in a comprehensive way that does not confound the two. It is a Monte Carlo simulation nested within a separate Monte Carlo simulation (U.S. EPA, 2001). Even though this approach has squared computational cost, it can still be performed on current desktop computers. The idea is that the inner Monte Carlo simulation represents variability whereas the outer simulation represents the analyst's incertitude about the values of the parameters of the distributions that describe that variability (U.S. EPA, 2001). The approach has also been used to explore the effects of uncertainty about distribution shape and correlations.

A1.7 Displaying the Outputs of a Probabilistic Risk Assessment

Three basic quantitative techniques may be used to summarize the output of a probabilistic analysis: (i) as a PDF, (ii) as a CDF, or (iii) by displaying selected quantiles such as in a Tukey or box-and-whisker plot (Figure A2) (Morgan and Henrion, 1990). For most exposure and risk assessments, an output PDF expresses the probability that the random variable (estimated exposure or risk, for example) falls within some very small interval (Warren-Hicks and Moore, 1998). The PDF excels at showing the relative probabilities of different values, and clearly presents the mode as peaks in the curve (Morgan and Henrion, 1990). The CDF expresses the probability that a random variable X assumes a value

less than or equal to some value *x* (e.g. the probability that exposure dose is less than a benchmark) (Warren-Hicks and Moore, 1998). For continuous random variables, the CDF is obtained from the PDF by integration. If the probability that a quantity lies within a specific interval or is above a specified value is of interest, then the CDF is more useful than the PDF (Morgan and Henrion, 1990). The standard Tukey box or box-and-whisker plot shows a vertical line from the 10th to the 90th percentiles, a box from the 25th to 75th percentiles, a horizontal line in the box at the median, and points at the minimum and maximum observed values (Morgan and Henrion, 1990). This simplified presentation method emphasizes percentiles and the median, and is the easiest of the three methods to explain (Morgan and Henrion, 1990).

Figure A2 Example Box-and-Whisker Plot for Total PCBs in Whole Fish at Five Locations



Note: The horizontal line is the median, the box spans the 25th to 75th percentiles, and the vertical lines span the 10th to 90th percentiles. The lower dot is the minimum and the upper dot is the maximum concentration observed.

A2.0 PROCEDURES FOR PROBABILISTIC RISK ASSESSMENT

This section describes the available procedures for each of the steps in a PRA for a contaminated site, beginning with selection of the model through to the fine-tuning of the analysis. The general steps involved in a probabilistic analysis are shown in Figure A1.

A2.1 Evaluating, Calibrating, and Validating Models

Despite their many uncertainties, regulatory decisions are often guided by the predictions derived from models. The precautionary principle further requires that regulators faced with uncertainty err on the side of caution by selecting or designing models that do not underestimate risks. The risk of making a wrong decision will depend on the reliability of the model predictions. Therefore, there is a strong need to establish the validity of a given model used in a PRA of a contaminated site or, at the very least, to be confident that exposures and risks will not be underestimated.

The modelling exercise should be viewed in formal quality assurance (QA) terms. QA consists of a plan to ensure that the model meets defined standards of quality with a stated level of confidence. The QA approach will provide guidelines for the project and establish a baseline against which the success and validity of the model can be judged. The model quality objectives in a QA plan are established a priori, and generally involve statements about the desired accuracy of the model and the levels of uncertainty considered acceptable for decision making on contaminated sites.

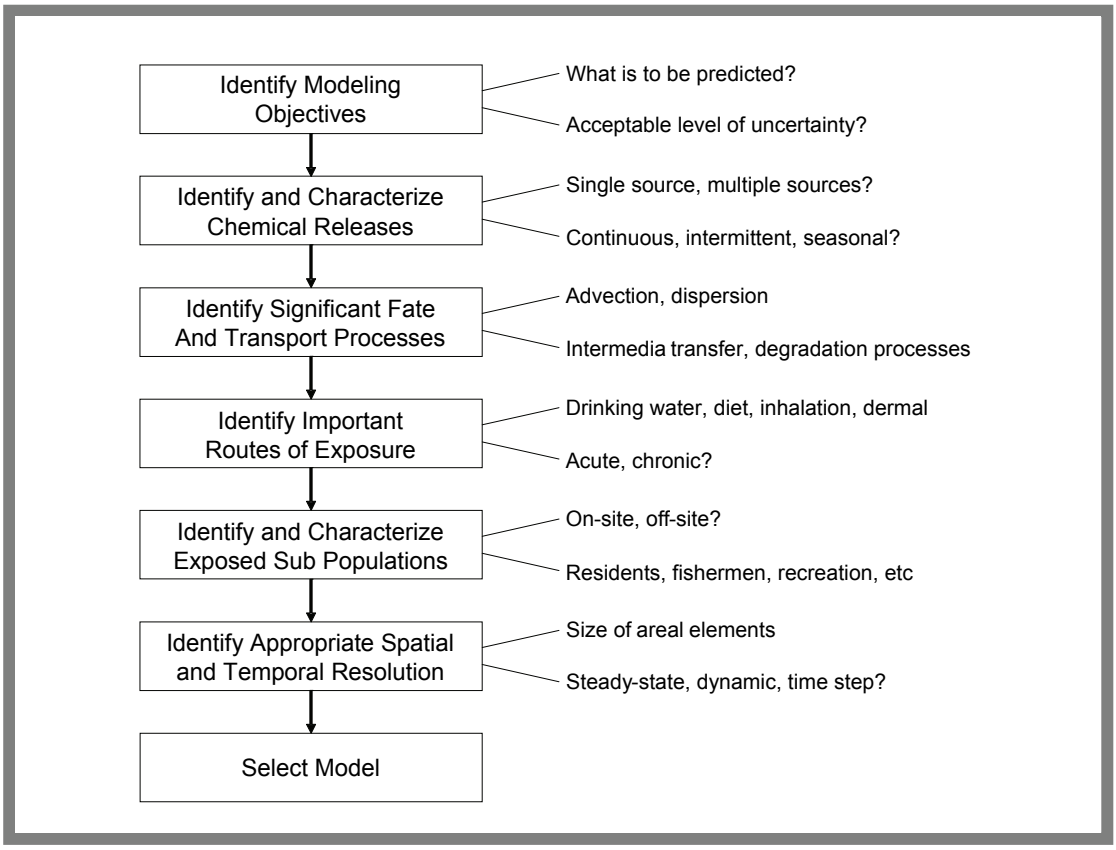
The next step is to choose the appropriate model. Although this step is of critical importance, there is no tried-and-true approach for choosing models because:

- there is no such thing as a "perfect" model—instead, a model that will perform adequately for the task at hand must be chosen;
- the modelling exercise will likely have objectives and constraints that conflict with one another (e.g. inclusion of many variables conflicts with resource limitations); and
- there may be uncertainty about what aspects of the analysis are critical, and those that can be safely left out of the model.

It is useful to break the process of choosing a model into two steps: model identification and model selection. In the first step, the assessor identifies those aspects of the chemical(s) and receiving environment that must be included in the model (i.e. a conceptual model is constructed, see Figure A3). In the second step, the modeller chooses an existing analytical or computer model that expresses the conceptual model, or develops a new one. Both of these steps must consider issues such as availability of data, the appropriate aggregation level for model inputs, spatial and temporal scaling, initial condition sensitivity, applicability, and whether the model has been appropriately calibrated and validated (Beck, 1987).

Before proceeding with a model simulation, credibility of the model should be assessed (Suter and Barnthouse, 1993). The assessment of credibility is different from model calibration and validation, because it does not yet consider the fully parameterized form of the model. The goal is to develop a qualitative measure of the predictive accuracy of the model or, at the very least, regulatory acceptance of the model's basis and predictions. The following are indications that the model of interest is credible: experimental testing (e.g. comparison of predictions to measured concentrations in blood from other contaminated sites) that indicates the model performs adequately; has been published in a peer-reviewed journal; has a long history of use; has been subjected to carefully designed validation studies; is supported by government agencies and/or independent experts.

Figure A3 Procedure for Model Selection



Assuming that the model is found credible, the next steps are model calibration and validation. The purpose of model calibration is to take the generic model and, by specifying the "correct" parameter values, turn it into a predictive tool for the system of interest. Validation tests the adequacy of the calibration exercise on an independent set of data. Numerous statistical and graphical techniques may be used to assess how well model performance matches observations. These include lumped measures of average model goodness-of-fit (Reckhow et al., 1990), correlation measures (Lin, 1989), parametric and non-parametric statistical tests (Venkatram, 1982; DeGroot, 1986; Gilbert, 1987; Parrish and Smith, 1990; Flavelle, 1992), spatial analysis of goodness-of-fit (Gilbert, 1987), and Bayesian measures of estimation error (Kitanidis, 1986; Butcher et al., 1991). Discussion of these and other quantitative methods for testing model validity can be found in Reckhow et al. (1990) and in Volume 15 of *Advances in Water Research* (Flavelle, 1992), which is dedicated to the discussion of validation of computer models.

Many measures and tests are available to evaluate model performance. The optimal measure is depends on the model quality objectives established prior to model selection. For example, if the objective is to evaluate average concentration over a broad area, then a lumped measure of average model goodness-of-fit is the appropriate statistic. For a set of paired measurements, x, and model predictions, y, the general form of the equation is:

(A1)

$$R = \frac{1}{n}\sum_{i=1}^{n}\left|\left(x_i - y_i\right)\right|^j$$

where, R is the mean of the residuals, and j is generally 0.5, 1, or 2 depending on whether it is most important to fit outliers (j=2) or the median (j=0.5) (Reckhow et al., 1990). Alternatively, if the objective is to predict spatial point concentrations, other techniques that correct for spatial correlations, such as kriging, should be used in testing model performance (Gilbert, 1987). Finally, if one cares more about certain model predictions, but not about others, decision analysis can explicitly incorporate such goals in evaluating model performance. For example, estimating ground-level ozone concentrations over urban areas may be of high importance whereas estimating concentrations over rural areas may be of low importance. Further, the (type I) error associated with stating that ozone concentrations will remain below a specified limit, when they will not, is likely greater than the (type II) error associated with predicting a violation when none will occur. A Bayesian decision analysis provides a direct means of incorporating both risk management

objectives and uncertainty in model estimates in the evaluation of model performance (e.g. Chao et al., 1994).

The objective of the validation exercise is not to derive a yes or no answer. In regulatory decision making, one must often accept or contend with approximations, as long as one is aware of the domain of the model's applicability (i.e. the areas, sites or situations in which the model works successfully) (Caswell, 1976). In the evaluation of two long-range dispersion models, ADPIC and EXPRESS, Rodriguez et al. (1995) found that the models consistently overestimated concentrations close to the source (<1,000 km), but provided unbiased and reasonably accurate predictions (observed and predicted concentrations within a factor of two) at greater distances. A binary decision on the validity of these models would have missed the point, because all models are invalid under some sets of conditions (Oreskes et al., 1994). The objective is to determine when the model performs adequately, and when it does not.

A2.2 Steps for Selecting and Parameterizing Input Distributions

A2.2.1 Selecting an input distribution

The important factors to consider when selecting input probability distributions are presented in Box A1 (U.S. EPA, 1997, 2001). Selecting a distribution is not a purely computational exercise. Sometimes more than one probability distribution function may adequately characterize variability or uncertainty. In general, the preferred choice of a probability distribution function is not the one that ranks the highest in a goodness-of-fit test with the data set.

Box A1. Selecting an Input Distribution

1. Will the variable have an important influence on the output? If not, do not worry about it.
2. Is the distribution known for the input variable?
3. If not, are there sound theoretical reasons for assigning a specific distribution to the input variable (see Ott, 1995, for a discussion of the theory underlying several key distributions)?
4. If not, are the data adequate for fitting a distribution?
5. If not, do appropriate surrogates exist? If yes, repeat 2 to 4.
6. If not, do data exist addressing components of the variable? If yes, repeat 2 to 4.
7. If not, solicit expert opinion.

For some variables, there may be enough empirical information to fit parametric distributions or even specify empirical histograms. More commonly, there is little or only partial empirical evidence to support the distributions selected as inputs. In this circumstance, best judgment or a convention among practitioners often suggests which distributions should be used (e.g. lognormal distribution for concentration data in environmental media). As a result, the analysis usually requires assumptions that cannot be justified by appeal to evidence. The consequences of this may be substantial, because the results of probabilistic risk analyses are known to be sensitive to the choice of distributions used as inputs (Bukowski et al., 1995), an effect that tends to be even stronger for the tail probabilities. The difficulties of developing and justifying input distributions are well known in the field of risk analysis, and have been the subject of considerable attention (e.g. Finley et al., 1994; Haimes et al., 1994). Although there is a large literature on the subject of estimating probability distributions from empirical data (e.g. Morgan and Henrion, 1990; Cullen and Frey, 1999), standard approaches are of limited practical effectiveness when few data exist. The following is a simple hierarchy of decision criteria that may be used to characterize input distributions (Moore, 1996, originally adapted from Haimes et al., 1994).

1. **Will the variable have an important influence on the output?** If not, do not worry about it. For example, the inhalation route is often a trivial route of exposure for humans exposed to persistent and bioaccumulative compounds (e.g. methylmercury). Thus, inhalation rate and concentration in air can be treated as point estimates or ignored in an exposure analysis for humans living near a contaminated site.

2. **Is the distribution known for the input variable?** There are relatively few instances in HHRAs where distributions and parameters are known for an input variable. Examples of known distributions are most likely for variables defined by well-understood physical processes that lead to variability (e.g. Poisson distribution for radionuclide counts within a short time interval) (Seiler and Alvarez, 1996).

3. **If not, are there theoretical reasons for assigning a specific distribution to the input variable (see Ott, 1995, for a discussion of the theory underlying several key distributions)?** Normal and lognormal distributions may be inferred from the structure of the variations in a random variable. If the variability of a quantity arises as a sum of contributions of many variations, each with a mean and variance, then the distribution of the sum will be normal (Ott, 1995). Error terms in allometric regression models often have an underlying normal distribution. If the variability of a quantity arises as a product of contributions of many variations, each with a mean and variance, then the

distribution of the sum will be lognormal (Ott, 1995). Environmental concentrations in different media often have an underlying lognormal distribution.

4. **If not, are the data adequate (e.g. sample size large enough, data directly relevant to the contaminated site in question) for fitting a distribution?** This will often be the case for well-measured variables at a site (e.g. chemical concentration in water) or variables that are common to many sites (e.g. body weight by gender and age class) (e.g. Hope, 1999).

5. **If not, do appropriate surrogates exist?** If yes, repeat steps 2 to 4. An example would be the use of surrogate data from rat studies for toxicokinetic variables in an exposure model for humans.

6. **If not, are data available to address components of the variable?** If yes, repeat steps 2 to 4. For example, dose-response data may be available for high-dose treatments, but the resulting dose-response relationship must then be extrapolated to lower doses.

7. **If not, solicit expert opinion.**

Choice of distribution becomes inherently more subjective as one moves down the list. When data are severely limited and specifying PDFs becomes an exercise in "best guessing," one may want to consider a second-order Monte Carlo analysis or probability bounds analysis (see section A2.5).

Probability distribution functions often arise from the fundamental properties of the quantities one is attempting to represent. An understanding of the mechanistic basis (Table A1) of the processes that generate variability is essential when selecting a distribution (Hattis and Burmaster, 1994). This will give rise to at least preliminary ideas about what distributional forms are likely to describe the underlying reality. Analyses of uncertainty and variability estimates will sometimes benefit from taking into account plausible alternative mechanism-based theories of the sources of variability and their interrelationships. Thus, expert judgment is likely the initial step when selecting a distribution. Another important step in selecting a distribution is 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. weight, height, concentration). A mathematical function describes the probability for each value across an interval for a continuous variable. Discrete variables take either a finite or (at most) a countable number of values that have only integers (e.g. mortality, number of newborns).

Table A1 Common Input Distribution Types and Applications

Distribution	Example Applications
Beta	Modelling environmental concentrations and proportion and percent variables; rough model in absence of data
Binomial	Number of deformities in a sample of specified size
Chi-square	Sum of weights of objects, each following a normal distribution
Exponential	Time between events; lifetime of device with constant probability of failure
Gamma	Time to complete task; modelling environmental concentrations
Geometric	Number of trials until success is achieved
Lognormal	Product of a large number of other quantities; modelling environmental concentrations; distribution of physical quantities in nature
Normal	Size of quantities that are the sum of other quantities; regression model error
Poisson	Number of events in a given unit of time (e.g. accidental releases)
Triangular	Rough modelling when only mode, minimum, and maximum are known
Uniform	Distribution when only a range is known
Weibull	Modelling toxicity test results with continuous data; lifetime of a device

Known physical or biological processes may dictate the shape of the distribution. For example, normal distributions result from processes that sum random variables (e.g. regression model error) whereas lognormal distributions result from multiplication of random variables (e.g. concentration of a chemical in soil). A Poisson distribution is used to characterize the number of independent and randomly distributed events in a unit of time or space (e.g. number of children per family over a selected time and space). Whichever distribution is selected, it is important to justify the underlying choice of the distribution, given the variable of interest.

Although a mechanistic basis often exists when selecting a distribution, it is always a good idea to plot the existing data (if available) to determine if they support the underlying reality. Plotting of data and visual inspection can be used in conjunction with exploratory data analysis. For example, if a large number of data (n>20) exist, a histogram of the observations can be developed to provide a reasonable idea of the underlying shape (distribution) of the data. For small sample sizes, a quantile-quantile plot (i.e. Q-Q plot) can be used to verify whether the data arise from a normal distribution or one that is skewed. Other graphical methods include frequency distributions, stem-and-leaf plots, and scatter plots (Tukey, 1977; Conover 1981; Morgan and Henrion, 1990). Gilbert (1987) and Ott (1995) provide descriptions of the use of probability plotting to derive parameter estimates for distributions.

If sufficient data exist, they may be used directly to define the distribution (empirical probability distribution function) as opposed to fitting the data to a probability distribution function. Use of an empirical data distribution provides a complete representation of the data with no loss of information as empirical distributions do not depend on assumptions associated with estimating parameters for other probability models. Methods exist to linearize the data for interpolating purposes. However, empirical probability distribution functions may not adequately represent the tails of the distributions due to limitations of data acquisition. Methods exist to extend the tails, but this will likely introduce uncertainty. Advantages and disadvantages of this approach are provided in U.S. EPA (2001).

When limited information is available for a random variable, maximum entropy inference (MEI) can be used to maximize the uncertainty in the input distributions (Lee and Wright, 1994; Vose, 1996). The technique is conservative, and uses a formal set of rules to specify input distributions according to the amount of available information. For example, if estimates of the lower bound, upper bound, and mean exist, the MEI solution would be to choose a beta distribution as opposed to the more commonly used triangular distribution. The formalism of MEI has several advantages compared with subjective judgments by individuals (e.g. avoids human bias and mitigates against unfounded confidence in our predictive capabilities). The credibility of the distribution will, obviously, depend on the accuracy of the information.

A2.2.2 Evaluating the fit of the distribution

Once a distribution has been chosen to represent the data, and its parameters (e.g. mean, variance) have been determined, it is appropriate to test how good the fit is (i.e. goodness-of-fit tests). There are several standard approaches to testing whether a set of data is consistent with a proposed distribution (Box A2).

Box A2. Goodness-of-Fit Tests

Chi-squared Test – Can be used to test a continuous or discrete distribution and for data that are ordinal. Chi-square is a measure of the normalized difference between the square of the observed and expected frequencies.

Shapiro-Wilk Test – Can determine whether or not a small data set (n<50) is normally or lognormally distributed.

Kolmogorov-Smirnov Test – Nonparametric test that compares the maximum absolute difference between the step-wise empirical cumulative distribution function (CDF) and the theoretical CDF. This test is of little use when the tails of the distributions are of concern.

Anderson-Darling Test – Places more emphasis on fitting the tails of the distribution. Uses a weighted average of the squared differences between the observed and estimated cumulative densities.

Goodness-of-fit tests are statistical tests of the hypothesis that the data represent an independent sample from an assumed distribution. These tests involve a comparison between the data and the theoretical distribution under consideration. However, the results of goodness-of-fit tests can be misleading, particularly when sample size in the data set is large. In such cases, goodness-of-fit tests routinely fail (i.e. a statistically significant deviation from the prescribed distribution is detected) despite visual evidence that the fit is reasonable or excellent.

A2.2.3 Accuracy of the tails of the distribution

From a regulatory perspective, the tails of a distribution are most generally of concern when characterizing risk (U.S. EPA, 2001). To ensure that the shape of a given probability distribution function does not significantly affect the distribution of possible outcomes, a sensitivity analysis should be performed using various plausible shapes of the

input probability distribution function. For example, where goodness-of-fit tests do not differentiate between normal and lognormal distributions, a lognormal distribution could be substituted by a normal distribution for a particular input variable. The sensitivity analysis would quantify the effect of this substitution on a chosen percentile of the output distribution (e.g. 95th percentile).

In a Monte Carlo analysis, it is important to generate a sufficient number of iterations to obtain a good representation of the tails of the output distribution. This is usually accomplished when the output distribution becomes "stable." Typically 10,000 iterations are sufficient to capture most of the variability in the tails of the input distributions, although some stochastic variability will always remain. A Latin hypercube sampling approach is also recommended, and ensures that sampling will occur across the whole range of each input distribution.

A2.3 Characterizing Chemical Concentrations

A2.3.1 Simple methods

Concentrations of chemicals vary spatially and temporally in the media to which humans are exposed (e.g. indoor dust, soil, air, drinking water, diet). During long-term exposures, humans may move over large portions of a contaminated area, or in and out of a contaminated area. As a result, individuals tend to integrate spatial and temporal variation in the chemical concentrations to which they are exposed. Therefore, estimates of the central tendency (e.g. arithmetic means) are generally used in human health exposure models as an expression of the spatial and temporal averaging of chemical concentrations in different media (U.S. EPA, 1992, 2001). There is, however, uncertainty associated with estimating the true average concentration of a chemical at a site, usually because of limited sample size. Uncertainty regarding the arithmetic mean concentration of a chemical arising from limited sample size can be estimated by calculating confidence limits on the mean. The estimated arithmetic mean and associated confidence limits may then be used to parameterize a normal distribution (when the confidence limits are symmetric about the mean), lognormal distribution (generally, the case with asymmetric confidence limits), or other distribution as appropriate (Box A3).

Several parametric and non-parametric methods have been developed to estimate lower and upper confidence limits (Singh et al., 1997, 1999, 2002a, 2002b, 2004). The parametric methods include:

- Student's *t*
- approximate gamma confidence limits using chi-square approximation
- adjusted gamma confidence limits
- Land's H statistic for lower and upper confidence limits
- Chebyshev inequality based confidence limits

The non-parametric methods include:

- central limit theorem
- modified *t* statistic (adjusted for skewness)
- adjusted central limit theorem (adjusted for skewness)
- Chebyshev inequality based confidence limits (using sample mean and sample standard deviation)
- jackknife procedure
- standard bootstrap
- percentile bootstrap
- bias-corrected accelerated bootstrap
- bootstrap *t*
- Hall's bootstrap

The appropriate choice of a method depends on the underlying distribution of the data, spread in the data, and sample size. Guidance for method selection is provided in Singh et al. (2004). Note that the guidance in Singh et al. (2004) is for calculating upper confidence levels. The guidance and methods, however, also apply to calculating lower confidence levels. Briefly,

- The data set should be tested to determine if the data have an underlying normal, lognormal or gamma distribution. Quantile-quantile plots and the Shapiro-Wilk or Lillifors statistical tests may be used to test for normality (using non-transformed data) or lognormality (using log-transformed data). Quantile-quantile plots and the Anderson-Darling or Kolmogorov-Smirnov statistical tests may be used to test for an underlying gamma distribution.
- Parametric methods should be used for data sets with underlying normal, lognormal, or gamma distributions. Otherwise, non-parametric methods should be used to calculate the confidence limits.
- Student's *t* is the appropriate method for estimating the confidence limits when the data set is normally distributed, lognormally distributed but with low skewness and standard deviation, or symmetrically distributed.
- For data sets that follow a gamma distribution, the confidence limits should be computed using adjusted gamma confidence limits (when the shape parameter, *k*, for the gamma distribution is >0.1, but <0.5) on the mean or approximate gamma confidence limits (when $k>0.5$) on the mean. For values of $k < 0.1$, confidence limits may be

obtained using the bootstrap *t* method or Hall's bootstrap method when the sample size is small (n<15), and for larger samples, the confidence limits of the mean should be computed using the adjusted or approximate gamma method.

- For data sets that follow a lognormal distribution but not a gamma distribution or normal distribution, the confidence limits should be calculated using the Land's H or Chebyshev methods. The Land's H method produces unreasonably high upper confidence limit values when skewness is high and sample size is low. For data sets with high skewness and low-to-moderate sample size, the Chebyshev method should generally be used. The Land's H method should be used to estimate the confidence limits with moderate skewness and moderate to high sample size. In cases of extreme skewness, non-parametric methods such as Hall's bootstrap should be used.
- For skewed data sets that are neither gamma nor lognormal, non-parametric Chebyshev or Hall's bootstrap confidence limits (for small data sets) of the mean may be used.

.

Box A3. Procedure for Deriving Concentration Distribution

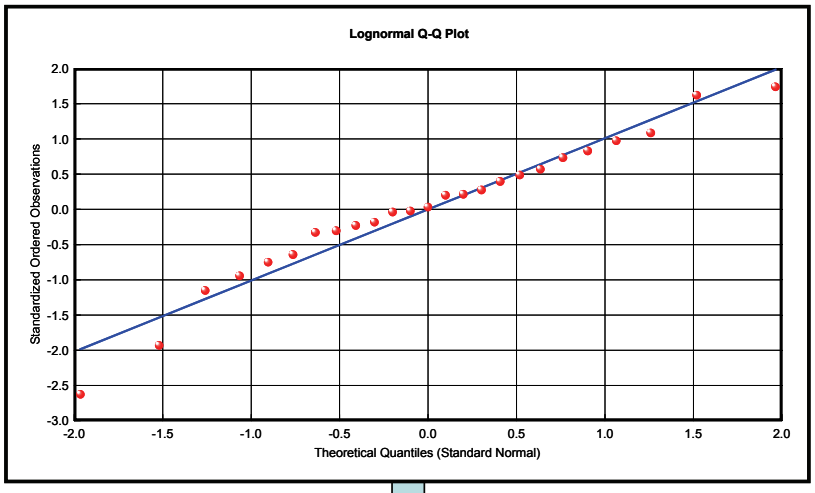
1. Data

12.26
5.12
4.79
7.16
6.81
4.87
6.87
10.35
3.89
5.28
7.75
11.38
1.65
8.24
5.87
2.77
6.09
17.51
3.18
5.81
9.70
8.70
3.62
1.04
18.95

2. Test for Underlying Distribution

Shapiro-Wilk Statistic = 0.962
Critical Value (0.05) = 0.918
Data are Lognormal

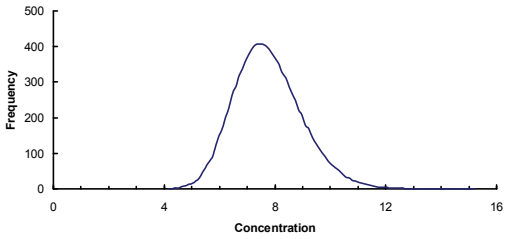
3. Check Fit of Data to Lognormal Distribution Using Graphical Method



4. Use Land H-statistic to Calculate UCL and LCL

Land H 95% UCL = 9.91
LCL = 5.93

5. Use UCL and LCL to Parameterize Distribution



Note: LCL, lower confidence limit; UCL, upper confidence limit

A2.3.2 Dealing with non-detects

Chemical analysis of water, soil, sediment, or tissue samples may result in the concentration in some samples being reported below laboratory detection limits (referred to as non-detects). The traditional approach for dealing with non-detects in many contaminated site assessments has been to assign values equal to half the detection limit (DL) when calculating summary statistics or deriving distributions. This approach can severely skew summary statistics and distributions, particularly when non-detects are frequent in the data set. This section briefly describes the methods available to deal with non-detects in a data set.

The three methods that are most often cited for handling non-detects are described below.

- **The substitution method** assigns a constant value to non-detects. Three commonly used conventions are (i) assume non-detects are equal to zero, (ii) assume non-detects are equal to the DL, or (iii) assume non-detects are equal to one-half the DL (the midpoint).
- **Maximum likelihood methods** use estimates from an assumed distribution to estimate values for non-detects. Given a distribution, estimates of summary statistics are computed that best match the observed concentrations above the DL and the percentage of data below the limit.
- **Probability plot methods** "fill in" values for non-detects using the following steps (see Gleit, 1985, for details).
    - Select an initial estimate for the mean and variance using a normal or lognormal probability plot.
    - Using the assumed distribution and the current value of the mean and variance estimate, impute expected values for the first *m* values, where *m* is the number of censored observations.
    - Calculate the mean and standard deviation or fit a distribution as usual from the constructed data set.
    - Repeat second and third steps until the mean and variance are stable.

There are variations on the probability plot method for multiply censored data and for different distributions.

A review of the published literature on censored data indicates that all of the methods have some advantages and limitations. When there is a large fraction of non-detects, or when sample size is small, no method works particularly well. Substitution by one-half the DL has been found to provide acceptable results for hypothesis testing when the censoring is not excessive (Clarke and Brandon, 1996). However, replacement of non-detects by a constant will produce biased estimates of the mean and variance (El-Shaarawi and Esterby, 1992). Therefore, substitution methods should be used only for data sets with a low proportion of non-detects.

If the distribution of the data is known, or if there are enough data to reliably estimate the distribution, maximum likelihood estimation provides superior estimates of percentiles. However, the maximum likelihood method still produces bias in the estimates of means and variances for lognormally distributed data (Helsel, 1990), and typical data sets will have insufficient data to reliably estimate the underlying distribution. Nevertheless, the maximum likelihood method is a useful approach with large data sets.

Gleit (1985) used simulations to compare the performance of the three methods listed above for small sample sizes (n=5) and singly censored normal or transformed lognormal data. Based on observed performance, he recommended the probability plot method. The probability plot method (referred to as a robust method because the reliance on a particular data distribution affects only the non-detected data) is also recommended by Helsel (1990). For data sets of modest to large sample size, the probability plot method is recommended as the means of dealing with non-detects in calculating summary statistics or fitting distributions.

A2.3.3 Advanced methods

In some situations, more advanced methods may be required to estimate distributions for spatially and temporally averaged chemical concentration. For example, one may need to

- use weighting to account for non-random spatial sampling, or
- use weighting to account for non-uniform habitat use by human receptors.

Geostatistical techniques (e.g. inverse distance weighting, universal kriging) and area-weighted bootstrapping are examples of techniques that can be used to correct for biases in sampling intensity (e.g. where sampling was focused on the most contaminated areas) in developing distributions (Ginevan and Splitstone, 1997; Goovaerts, 1997; Brakewood and Grasso, 2000; Barabas et al., 2001; Okabe et al., 2001). Random walk models may be used to account for non-uniform habitat use by, for example, assigning utility and area weightings to different habitats and then calculating running spatial averages as each individual moves about the environment (Hope, 2000). Geostatistical techniques and random walk models are beyond the scope of this appendix. The reader is referred to the references cited herein for additional information on these topics.

A2.4 Dependencies Among Input Variables

In most HHRAs, dependencies among input variables are likely to occur. For example, risk models for estimating exposure to contaminants via dermal contact typically require information on human body surface area and body weight (Smith et al., 1992). Surface area reflects the skin area available for dermal absorption of a contaminant. The dermal intake thus estimated is divided by body weight to obtain a measure of whole body dose that can be compared to the appropriate toxicity benchmark. Body weight and surface area are correlated variables, and thus an assessment that models inter-individual variability in exposure would have to account for this dependency. Other likely examples of dependencies in HHRA include: (i) positive correlations among inhalation rate, drinking water intake rate, and food ingestion rate, (ii) positive correlations between the various intake rates and body weight, (iii) positive correlations in space between contaminant concentrations in soil and home-grown produce, (iv) negative correlations among consumption of different dietary items (e.g. as consumption of beef increases, consumption of fish decreases), and (v) many others.

Many risk assessors assume independence among all random variables, even when there is no justification for doing so. It is improper, however, to assume independence among random variables unless there is convincing evidence (e.g. scatter plot indicates no relationship) or a compelling argument (e.g. contaminant concentration in home-grown produce is unlikely to be related to contaminant concentration in fruit imported from another country) that this is a reasonable assumption. If a dependency is ignored, the answer obtained will be wrong (Ferson et al., 2004). Under certain conditions, the central tendency of the output distribution could be approximately correct (Smith et al., 1992). However, the spread of the output distribution and the tail probabilities can be highly inaccurate, particularly with strong dependencies among influential input variables (Bukowski et al., 1995; Ferson and Burgman, 1995; Ferson et al., 2004). If, for example, one had dependencies between inhalation and ingestion rates, and among contaminant concentrations in soil, home-grown vegetables, and indoor air concentrations in an exposure model for toddlers, ignoring the dependencies would lead to a gross underestimate of exposure for the most highly exposed individuals. It is the latter group that is generally of most concern in a contaminated site risk assessment.

A variety of strategies for dealing with dependencies in a PRA follow (for a detailed discussion of the topic, see Ferson et al., 2004).

- In cases where the dependencies are monotonic and sufficient data are available to estimate the correlation coefficient, software packages such as Crystal Ball® and @Risk® generate joint distributions for the correlated marginal distributions using Spearman rank correlation coefficients for each pair of dependent variables. Other software packages may use Pearson (only appropriate if input variables are normally distributed), Kendall, or other types of correlation coefficients (Ferson et al., 2004).
- For pairs of input variables that have correlation coefficients that are close to one or other nearly perfect dependencies, replace one of the input variables with a formula in the model of choice. This approach is equivalent to assuming a perfect (linear or non-linear) relationship between the two input variables and was the approach used by McKone and Bogen (1992) to relate human body surface area to body weight.
- In cases where dependencies are suspected but data are lacking, "what if" analyses may be conducted to determine the possible influence that the suspected dependencies could have on estimated exposure or risk. For example, with a suspected positive relationship (e.g. concentration in soil and vegetables), one could vary the correlation coefficient from zero (i.e. assume independence) to plus one (assume perfect dependence). This range may be narrowed if information is available to support this decision. If the "what if" analyses indicate that exposure or risk estimates (particularly, the tail values) are not sensitive to the value of the correlation coefficient, then this source of uncertainty is not of concern. Otherwise, one could collect new data to better specify the dependency or hedge the eventual decision to account for this source of uncertainty (e.g. choose the correlation coefficient that produces the most conservative estimate of risk). When "what if" analyses are used to explore the influence of a suspected dependency, the results of the exercise should be presented in the risk assessment.
- Varying correlation coefficients to explore the influence of a suspected dependency assumes that the dependency is monotonic. In cases where non-monotonic relationships could occur (e.g. no effect at zero dose, stimulation at low dose, toxicity at high dose) and data are lacking, more complex strategies are required to explore the influence of the suspected dependency (e.g. using different correlation coefficients for different portions of the input variable values, using different non-monotonic equations with varying error terms to link independent variables). Such strategies require much professional judgment and are unlikely to cover the full range of possibilities for suspected dependencies.

In cases where there is insufficient knowledge to precisely specify a suspected dependency, bounding methods such as Dempster-Shafer theory and probability bounds analysis (later described in section A2.5) may be used. With these methods, what is known about the suspected dependency and what is not must be specified (e.g. dependency is positive, but correlation coefficient is unknown). Given this

state of knowledge about a suspected dependency, the bounding methods compute the lower and upper bound probability distributions for the output variable (Figure A5). Between these two bounds lies the "true" output distribution. The "true" output distribution cannot be outside the bounds, assuming that the state of knowledge has been accurately represented in the analysis. When bounds are close together, uncertainty about the suspected dependency is having little influence on the distribution of the output variable. When the bounds are far apart, uncertainty about the suspected dependency is having much influence on the distribution of the output variable.

A2.5 Uncertainty Propagation Methods

Morgan and Henrion (1990) provide an excellent discussion of many of the probabilistic methods used in risk analysis. The following text reviews probabilistic risk methods, including several that have emerged since the publication of Morgan and Henrion=s text.

A2.5.1 Interval analysis

Interval analysis (Dwyer, 1951; Moore, 1966; Alefeld and Herzberger, 1983; Neumaier, 1990) is the simplest comprehensive method for uncertainty propagation through mathematical equations. Each input variable is expressed as an interval bounded by the lowest and highest possible values. The mechanics of interval analysis are very simple. If $A = [a_1, a_2]$ and $B = [b_1, b_2]$ are two intervals, then their sum $A+B$ is the interval $[a_1+b_1, a_2+b_2]$. Interval addition is therefore just the element-wise addition of the minima and maxima. For positive intervals, the product $A \times B$ is likewise $[a_1 \times b_1, a_2 \times b_2]$. Their difference, however, $A-B$ is the interval $[a_1-b_2, a_2-b_1]$. In the case of the difference, the minima and maxima are combined anti-element-wise. This ensures that the result encloses all possible values of the difference. Interval division similarly combines the minima and maxima anti-element-wise to ensure the result encloses all possible values of the quotient.

Interval analysis is applicable whatever the nature or source of the uncertainty. It yields reliable results whether the uncertainty arises from measurement error or stochastic variability. Interval analysis is insensitive to correlations or other dependencies among input variables. No matter what dependencies might exist, interval analysis will yield results that enclose the true value or values. As long as the analyst can specify bounds that are certain to contain the uncertain quantities, the subsequent application of interval analysis is free of further assumptions.

Interval analysis often yields results that are quite wide. Indeed, worst-case analysis, which if properly conducted amounts to interval analysis, has often been accused of hyperconservatism. This result is usually the consequence of ignoring information about the probabilities of values between the minima and maxima and dependencies among the input variables.

A2.5.2 Variance propagation

Analytical methods may be used to estimate exposure or risk with relatively simple model equations. The analytical method most commonly used is variance propagation (Morgan and Henrion, 1990; Hammonds et al., 1994; Slob, 1994). If one has a simple additive model and the input variables are independent, the mean value of the output distribution is the sum of the input means. Similarly, the variance of the output distribution is the sum of the variances of the input variables. That is:

(A2)

$$\mu_R = \sum_{i=1}^{p} \mu_i$$

$$\sigma_R^2 = \sum_{i=1}^{p} \sigma_i^2$$

where p is the number of variables in the model. The shape of the resulting output distribution will tend to be normal even if the distributions assigned to the inputs are not normal.

A similar approach can be taken with multiplicative models after first converting the model to its additive form by logarithmically transforming the input variables.

(A3)

$$Y = a \times b \times c$$

$$1n\,(Y) = 1n\,(a) + 1n\,(b) + 1n\,(c)$$

For multiplicative models, the geometric mean is the exponential term of the sum of the mean values of the logarithms of each input variable. The geometric standard deviation is found by taking the square root of the sum of variances of the transformed variables and exponentiating. That is:

(A4)

$$X_{g,R} = e^{\mu,R}$$

$$S_{g,R} = e^{\sqrt{\sigma_R^2}}$$

where $X_{g,R}$ is the geometric mean of the resulting output distribution, μR is the sum of the means of the logarithms of the input variables, $S_{g,R}$ is the geometric standard deviation of the resulting output distribution, and σ_R^2 is the variance of the logarithms. The distribution of the output distribution for a multiplicative model will tend to be lognormal even when the input distributions have different shapes. Hammonds et al. (1994) describe variance propagation in more detail.

A2.5.3 First-order Monte Carlo analysis

The most commonly used PRA technique is Monte Carlo simulation. The basis for a Monte Carlo analysis is straightforward; point estimates for variables in a model equation are replaced with probability distributions, samples are randomly taken from each distribution, and the results tallied, usually in the form of a PDF or CDF. Several variations of the Monte Carlo technique for sampling from input distributions are available (Morgan and Henrion, 1990). One variation is importance sampling, where values of particular importance (usually the tails of the input distributions) are sampled more often and then given reduced weight to improve resolution in the tails of the output distribution. In stratified sampling, the input distributions are divided into intervals and input values obtained by random sampling from within each interval. The most popular version of stratified sampling is Latin hypercube sampling that divides input distributions into equiprobable intervals. Latin hypercube sampling is more precise than conventional Monte Carlo sampling because the entire ranges of the input distributions are sampled in a more even and consistent manner (Iman and Helton, 1988).

First-order Monte Carlo techniques excel when large quantities of data and theory exist to specify the model equation and the input distributions. Difficulties arise when insufficient information exists to specify input distributions or the relationships among them (Ferson, 1996; Moore, 1996). Dependencies among input distributions can exaggerate or reduce the predicted probabilities of exposure (or effects) compared to the uncorrelated case (Smith et al., 1992; Ferson and Long, 1994; Bukowski et al., 1995). In cases where the dependency relationships are monotonic and the data exist to specify the correlation coefficients, @Risk®, Crystal Ball®, and other software packages have the capability to induce the dependencies in the analyses. If important dependencies are suspected, but insufficient data exist to specify the relationships, then the analysis becomes problematic. In such cases, other techniques such as probability bounds analysis should be considered because the results of such analyses do not depend on knowledge about the covariance among input variables (Ferson and Long, 1994; Ferson et al., 2004).

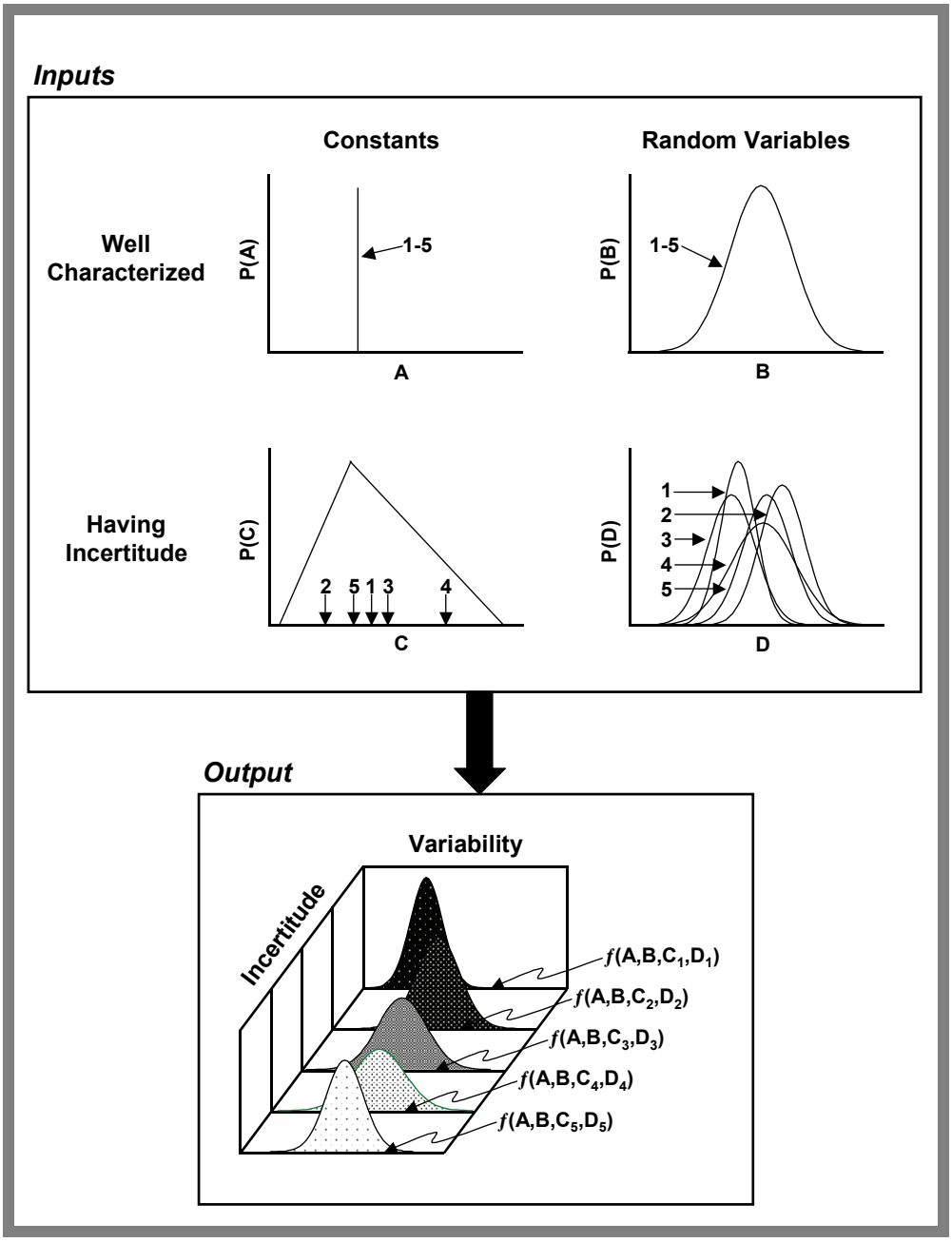
A2.5.4 Second-order Monte Carlo analysis

Second-order Monte Carlo analysis consists of two loops; the inner loop represents variability (known and quantifiable variation such as the case for individual body weights, concentration measurements, and so forth), and the outer loop represents parameter incertitude. To conduct an analysis, the following steps are required (also see Figure A4).

- Specify the model equation and identify which model inputs are (i) well-characterized constants (e.g. solubility of a chemical in water where there is little variation among a number of well-conducted studies), (ii) constants that have incertitude (e.g. water solubility of a chemical where only limited or poor quality data are available), (iii) well-characterized random variables (e.g. chemical concentration in a playground from which numerous soil samples have been collected and analyzed), and (iv) random variables for which there is incertitude about the shape and/or parameters of the distribution (e.g., chemical concentration in a playground for which limited or poor quality data are available).

- In software systems such as Crystal Ball®, well-characterized constants and well-characterized random variables are assigned to the inner loop. Constants with incertitude and poorly understood constants or uncertain random variables are assigned to the outer loop. For both situations of parameter distributions, a distribution is selected (e.g. normal or lognormal distribution for body weight of active men) and the parameter values for the distribution specified (e.g. mean = 75 kg, standard deviation = 15). For constants with incertitude, the distribution and parameter values will likely be based on considerable professional judgment. Data-fitting techniques may be used to parameterize well-characterized random variables. Random variables with incertitude must be included in both loops. To do this, a distribution is selected for the random variable (e.g. lognormal distribution for chemical concentration in soil) for the inner loop. Instead of specifying exact parameters for the random variable, however, distributions are assigned. In the case of a lognormal distribution in the inner loop, one would assign a distribution for the mean and/or a distribution for the standard deviation. These latter distributions would reflect our incertitude about what the true mean and/or standard deviation are for the random variable of interest.

Figure A4 Use of Second-Order Monte Carlo Approach to Distinguish Between Variability and Incertitude for Mathematical Expressions Involving Constants and Random Variables



Note: Five hypothetical values or distributions from the outer loop simulation are shown for the inputs and output. For the well-characterized input constants and random variables, the values and distributions, respectively, do not change from one outer loop simulation to the next.

- Specify the number of inner and outer loop simulations for the second-order Monte Carlo analysis. In the first outer loop simulation, values for the parameters with incertitude (either constants or random variables) are randomly selected from the outer loop distributions. These values are then used to specify the inner loop constants and random variable distributions. The analysis then proceeds for the number of simulations specified by the analyst for the inner loop. This is analogous to a first-order Monte Carlo analysis. The analysis then proceeds to the second outer loop simulation and the process is repeated. When the number of outer loop simulations reaches the value specified by the analyst, the analysis is complete. The result is a distribution of distributions, a "metadistribution," that expresses uncertainty both from

incertitude and from variability (Figure A4). The slopes of the distributions represent variability while the spread between distributions represents uncertainty due to lack of knowledge.

Some issues are associated with second-order Monte Carlo analyses. Computational time can be a problem because the necessary number of replicates is squared (i.e. number of inner loop simulations times number of outer loop simulations). In practice, specifying variability and incertitude with random variables is a difficult exercise because the analyst is essentially trying to quantify what is not known or only partially understood.

Issues involving dependencies become more complex in a second-order Monte Carlo analysis (Hora, 1996). As with first-order Monte Carlo analysis, dependencies can arise among different input variables (e.g. intake rates for air, water, and food) in a second-order Monte Carlo analysis. In a second-order Monte Carlo analysis, however, dependencies may also need to be specified among distribution parameters of a particular random variable. For example, means and standard deviations are typically correlated in nature. Thus, for a normally distributed random variable, analysts must not only quantify what they do not know about the mean and standard deviation but also what they do not know about the relationship among these parameters.

The major benefit of second-order Monte Carlo analysis is that it allows analysts to propagate their incertitude about distribution parameters in a probabilistic analysis. The analyst need not specify a precise estimate for an uncertain parameter value simply because one is needed to conduct the simulation. The relative importance of the inability to precisely specify values for constants or distributions for random variables can be determined by examining the spread of distributions in the output. If the spread is too wide to promote effective decision making, additional research is required to replace incertitude with variability (i.e. replace lack of knowledge with knowledge).

A2.5.5 Probability bounds analysis

Many researchers have argued that lack of empirical information implies that probability distributions of input variables cannot be precisely specified (e.g. Ferson and Ginzburg, 1995). Probability bounds analysis represents an uncertain input distribution with an entire class of probability distributions that conform to the available empirical information about the variable. Sometimes this class is small, and might be a single distribution when information is abundant. Other times, the class can be large, reflecting a poor state of knowledge about the variable. By using probability bounds analysis, an analyst can propagate the entire class through the risk model.

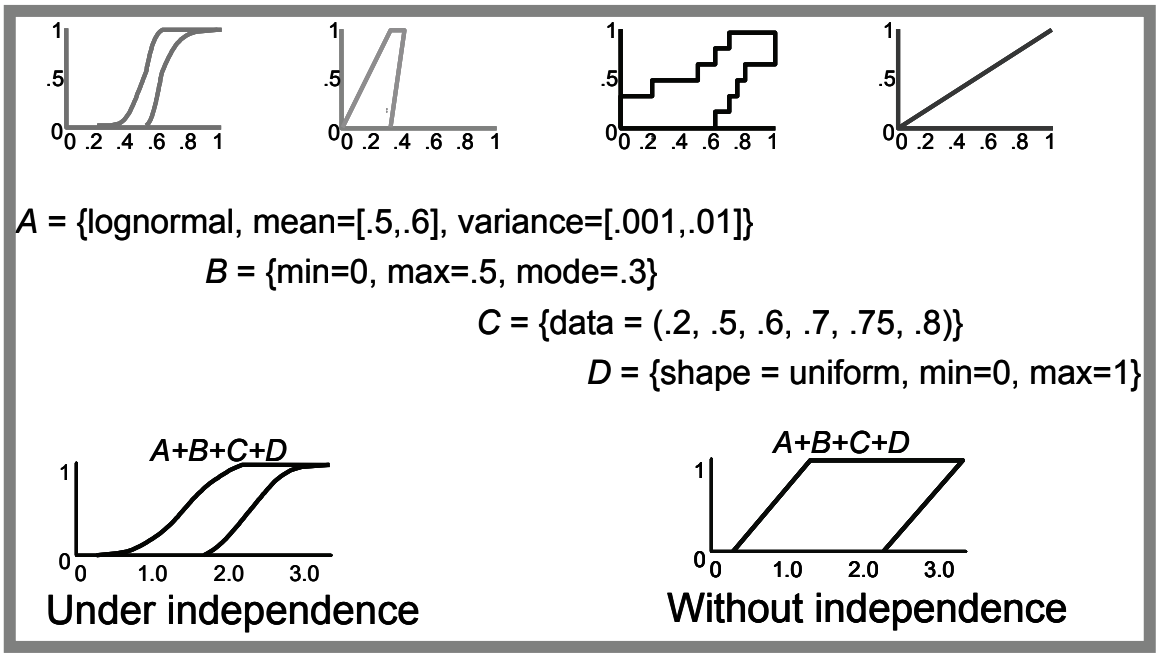
With efficient numerical algorithms (e.g. Williamson and Downs, 1990; Berleant, 1993), bounds can be computed on the output CDF when the input distributions are represented by probability bounds. In risk assessment, the use of probability bounds analysis reveals how much larger (or smaller) the probability of a result of a given magnitude might be. Probability bounds analysis offers a means for determining the reliability of risk estimates that is more comprehensive than "what-if" or interval approaches and yet computationally faster than Monte Carlo methods.

When each variable in the risk model appears only once, the results from a probability bounds analysis are guaranteed to be optimally narrow. In other words, the bounds could not be any tighter unless there was more information (or more assumptions) about the inputs.

The problem of not knowing the dependencies among input variables can be addressed using the probability bounds approach with dependency bounds analysis (Glaz and Johnson, 1984; Frank et al., 1987; Williamson and Downs, 1990; Ferson and Long, 1994; Ferson and Burgman, 1995; Ferson, 1996; Berleant and Goodman-Strauss, 1998). For each arithmetic operation, an analyst may assume independence between random variables, or the analyst may choose to make no assumption at all about this dependence.

Probability bounds analysis is much more efficient than sensitivity analysis when many variables are involved and, for many estimation problems, it is quite easy to implement. However, it provides only upper and lower bounds without any indication about the relative likelihoods within the range. It is more comprehensive than "what-if" or sensitivity studies, cheaper computationally, and easier to interpret than analogous second-order Monte Carlo methods. Figure A5 illustrates how bounded input distributions are combined to *produce output bounds.*

Figure A5 Hypothetical Probability Bounds Analysis Showing Results, Assuming That Input Variables Are Independent or That There Is No Knowledge Concerning the Dependencies That May Exist Among Input Variables



A2.5.6 Method selection

The choice of an appropriate method for propagating uncertainty depends on the complexity of the risk analysis, the available information, and the expertise of the assessor. The following are general points of guidance for choosing an appropriate probabilistic risk analysis method.

- At an early stage of the assessment or when little information is available, interval analysis can be used to generate risk estimate bounds (analogous to best- and worst-case scenarios). If the upper bound of exposure is well below the dose or risk benchmark, no further analysis is required.
- For simple additive models and multiplicative models involving logarithmic data (where log-transformation results in addition of exponents), and where all the variables are independent, an analytical approach such as variance propagation is a simple and effective tool for propagating uncertainty.
- For more complex models, first-order Monte Carlo simulation methods may be the most appropriate uncertainty propagation tool, but only if sufficient information exists to adequately characterize the input distributions and the relationships among them (i.e. random variability dominates and incertitude is relatively low in comparison).
- When information is limited and incertitude prevails, other methods with less restrictive requirements should be considered (e.g. probability bounds analysis).
- To estimate bounds on probability estimates (i.e. as an expression of our confidence in the estimated values), second-order Monte Carlo analysis and probability bounds analysis are useful techniques.

The principles of best practice developed by Burmaster and Anderson (1994) for Monte Carlo analysis should be followed when conducting any type of probabilistic risk analysis. These principles are discussed in section A5.0.

A2.6 Sensitivity and Elasticity Analysis Methods

The purpose of a sensitivity analysis is to identify how variation in the output of a model (e.g. total daily intake of a chemical) is influenced by uncertainty in the input variables. If the output variance precludes effective decision making, sensitivity analysis may be used to identify the input variables that contribute the most to the observed output variance. Subsequently, research efforts may be initiated to reduce uncertainty in those key input variables (if that is possible). Sensitivity analysis can also be used to simplify model structure by identifying those input variables that contribute

little to the output (e.g. a minor route of exposure) and thus can be removed from the analysis.

Uncertainty and sensitivity analysis both focus on the output of a model and are therefore closely related. The purposes of the two types of analyses, however, are different. An uncertainty analysis assesses the uncertainty in model outputs that arises from uncertainty in the inputs. A sensitivity analysis assesses the absolute and relative contributions of the inputs to the total uncertainty in the output.

Sensitivity analysis methods may be classified into three groups: screening methods, methods for local sensitivity analysis, and methods for global sensitivity analysis. Screening methods are generally used to separate influential input variables from non-influential ones, rather than to quantify the impact that an input variable has on the output of the model. Screening methods are useful for models with large numbers of input variables. They are able to identify important input variables with little computational effort, but at a cost of losing quantitative information on the importance of the input variables. In contrast, local and global sensitivity measures provide quantitative estimates of the importance of each input variable. The difference between them is that the former focuses on estimating the impact of small changes in input variable values on model output, whereas the latter addresses the contribution to model output variance over the entire range of each input variable distribution. This section reviews screening, local, and global sensitivity analysis methods.

A2.6.1 Screening methods

Most screening methods revolve around the idea of "what if" analyses. That is, how would the output value change if the value of a selected input variable was changed? With large models, this exercise needs to be systematic to be useful. Factorial designs, for example, are used to measure the influence of input variables on the output by taking into account both additive effects and interactions. The design involves selecting combinations of input variable values that provide the most information on the relationships between input variables and output distributions. With a factorial design and a large model, however, the number of model runs (n^k, where k is the number of input variables, and n is the number of values for each variable) quickly becomes unmanageable. If one can assume that higher-order interactions (e.g. dependencies) can be ignored, fractional factorial designs may be used (e.g. Andres, 1997). Fractional factorial designs involve much fewer model runs than do factorial designs.

Sequential bifurcation is another screening method for sensitivity analysis (Bettonvil and Kleijnen, 1997). The steps involved in this method follow.

- Begin by assigning lowest possible values to the input variables that are expected to be positively correlated with the output computation and the highest possible values to the input variables that are expected to be negatively correlated with the output computation. The signs for each of the input variables can be determined outside of the analysis (e.g. using correlation analyses). Run the model to obtain the "low" response.
- Repeat the above analysis, but using the highest possible values for positively correlated input variables and the lowest possible values for the negatively correlated variables. Run the model to obtain the "high" response.
- If the "high" response is well above the "low" response, one or more of the input variables has an important influence on the output. If this is the case, continue with the following steps.
- To determine which of the input variables is important, another model run is performed for which half of the input variables are set equal to their values in the "low" response model run and half are set equal to their values in the "high" model response run. If the resulting output response is again well above the "low" response, then one or more of the input variables that was switched to "high" response has an important influence on the output. If this is the case, the set of input variables switched to high can be split again. If not, the analysis focuses on the input variables that have yet to be switched to high.
- The process of setting half of the selected input variables to their "high" response level continues until all important input variables have been identified.

Because sequential bifurcation applies only two values for each input variable, the outcome is highly influenced by the values selected for the input variables. The results of the method are also influenced by the grouping of input variables, given that some variables will be important only if other variables are included in their group (the situation when two variables are multiplied together in the model equation). Although easy to perform, the sequential bifurcation method for sensitivity analysis is an error-prone design.

One way to overcome the shortcomings of the sequential bifurcation method is to set all input variable values to achieve the low output response and only increase one input variable to its high level at a time (Cotter, 1979). Theoretically, one-at-a-time designs require $k + 1$ runs for a model with k input variables. This is really just a variation on local sensitivity analysis methods (see later in the text). To evaluate the sensitivity of the model to each input variable and reduce the dependence of the starting point, other designs are required. The Cotter (1979) design applies one run with all input variables set to their low response level, one run with all input variables set to their high response level,

and 2 x *k* runs in which (i) each input variable is set to high response level while the other input variables are held to their low response level, and (ii) vice versa. The Cotter design is not sensitive to the order in which input variables are changed from their low to their high response level as is the case with the sequential bifurcation design. The Cotter design also does not require knowledge of whether input variables are positively or negatively correlated with the output. The Cotter design, however, requires more model runs than does the sequential bifurcation method, and is also sensitive to the values chosen for the input variables. More sophisticated one-at-a-time designs in which intermediate values are included for each input variable are available (e.g. Morris, 1991).

A2.6.2 Local sensitivity methods

Another measure of sensitivity is called elasticity. Elasticity seeks to measure the proportional rate of change in the model output resulting from a given input variable. The elasticity ε_{ij} of a model output O_i measured with respect to an input variable value P_j is:

(A5)

$$\varepsilon_{ij} = \frac{\Delta O_i / O_i}{\Delta P_j / P_j}$$

where Δ expresses the change in the output and/or change in the input variable (Evans and Dempson, 1986; Morgan and Henrion, 1990). Elasticity or local sensitivity analysis is generally carried out to determine the input variables for which it is important to obtain highly reliable estimates. The standard approach for a local sensitivity analysis is to perturb the values of the input variables by a small amount and observe the results on the values of the output (Warren-Hicks and Moore, 1998). This may be done using mathematical tools (e.g. computing partial derivatives of the output function with respect to the input variables), but is more commonly done using computational techniques (Helton and Davis, 2002). For example, Moore and Bartell (2000) conducted a local sensitivity analysis on an aquatic food web model with 220 input variables by assigning each input variable a normal distribution with a coefficient of variation of 1%. Parameter values were randomly selected from each input distribution using a stratified random sampling procedure (Latin hypercube sampling). A sensitivity index was then calculated as the square of the simple correlation between each input variable and model output.

A2.6.3 Global sensitivity methods

A simple and effective method to determine whether an output value depends on an input variable is to generate scatter plots. In a scatter plot, each input variable sample from a Monte Carlo simulation is plotted against the corresponding output values. Important input variables will have scatter plots showing strong relationships, while unimportant input variables will have scatter plots showing no discernible relationships. Scatter plots can be used to help identify which of the quantitative techniques below would be most effective in a quantitative global sensitivity analysis. Scatter plots, however, are not quantitative, and may be less useful with large complex models. Quantitative methods that may be used in a global sensitivity analysis are reviewed in the following text.

A2.6.3.1 Correlation coefficients

A common method to determine the linear relationship between an input variable and output values is to calculate a Pearson correlation coefficient. The Pearson correlation coefficient is a measure of the covariance between two variables. Unlike scatter plots, however, Pearson correlation coefficients assume that the relationship between two variables is linear and monotonic (Sokal and Rohlf, 1981). If scatter plots have shown that this is not the case, other statistics (see later in the text) should be used for the global sensitivity analysis.

If the relationship between an input variable and an output variable is non-linear but monotonic, rank transformation may be a solution for quantifying their relationship. This is achieved by ordering the values from the Monte Carlo simulation for each input variable and output in ascending order. The rank of a value is the location of the value after ordering of the data. Thus, a data set with values {3.1, 4.4, 2.9, 2.4, 5.3} would be replaced with the ranks {3, 4, 2, 1, 5}. Rank transformation linearizes non-linear monotonic relationships. Spearman's correlation coefficient may then be calculated for each combination of the transformed input and output datasets (Sokal and Rohlf, 1981). Crystal Ball® calculates sensitivity by computing rank correlation coefficients between every input variable and the output variable.

The correlation coefficient methods described previously have limitations; (i) the sensitivity calculations will be inaccurate for correlated input variables, and (ii) the calculations will be inaccurate for relationships that are non-monotonic. For example, if an important input variable was highly correlated with an unimportant input variable, the latter would likely have a high sensitivity with regard to output values. Turning off correlations in a Monte Carlo analysis is one way to avoid this problem. The tornado chart option in Crystal Ball® may be used to determine if any input variables have a non-monotonic relationship with the output variable of interest.

A2.6.3.2 Non-monotonic and non-random patterns

If scatter plots or tornado charts indicate that relationships between values of an input variable and respective output values may not be monotonic, several methods may be applied to quantify sensitivity. These methods are reviewed by Saltelli et al. (2000). Following a Monte Carlo analysis, the tests involve splitting the parameter space for an input variable into equiprobable parts (e.g. 10 classes with 1,000 sample values each, class 1 having the lowest 1,000 values, class 2 having the next lowest 1,000 values, and so on). If an input variable and resulting output values are unrelated, the means of the output values associated with the classes will be approximately equal. This can be tested directly using an *F*-test (analogous to an ANOVA). If the requirement for normally distributed output data cannot be met, a chi-square test for equal medians may be done. A Kruskal-Wallis test may be performed on rank-transformed data if the assumption of homogeneity of variance among classes for output values cannot be met.

Another option to detect non-monotonic relationships is to subdivide both input and output values into classes (e.g. 10x10 = 100 classes). This is analogous to dividing the scatter plot into 100 grid units. The number of data points in each class are then determined and a test for non-random pattern carried out.

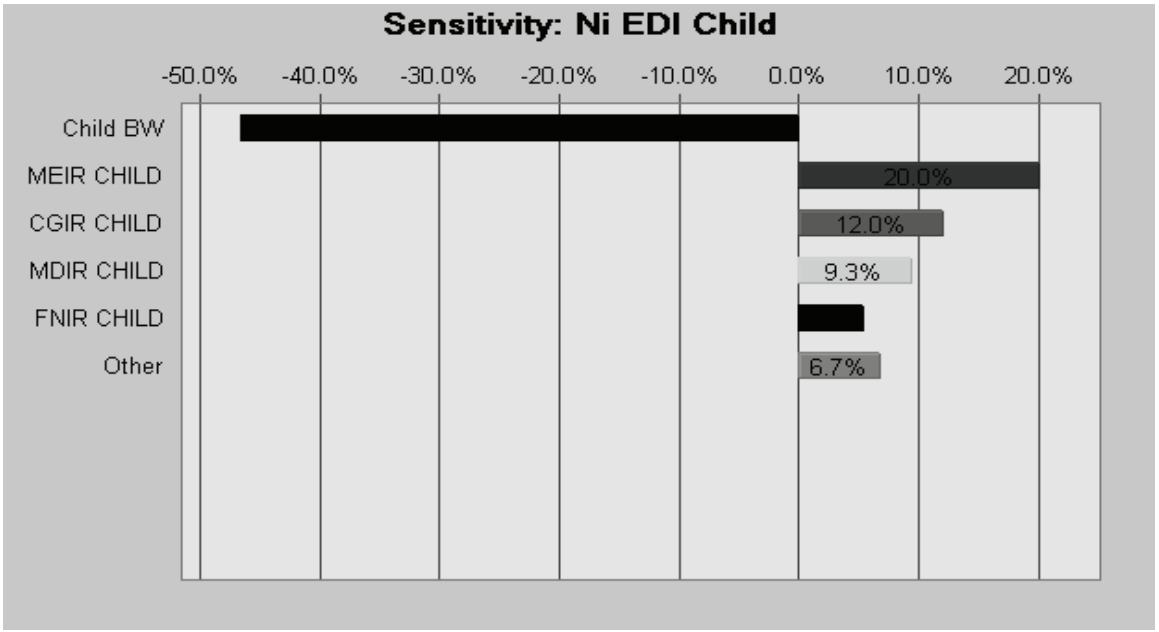
A2.6.3.3 A practical example of global sensitivity analysis

Software packages (such as Crystal Ball® and @Risk®) have the ability to identify those input variables (e.g. body weight, inhalation rate, food intake rate) that have the greatest influence on a particular forecast of interest (e.g. the estimated daily intake). The following example illustrates how the sensitivity analysis function (provided in the software package Crystal Ball®) can be used to identify those input parameters that are highly correlated (either positively or negatively) with the forecast of interest (i.e. the estimated daily intake [EDI] of nickel from food).

In this example, Crystal Ball® calculates sensitivity by determining rank correlation coefficients between input variables (e.g. body weight, food intake rate) and the forecast of interest (EDI of nickel from food). Crystal Ball® determines the contribution to variance by squaring all of the rank correlation coefficients and normalizing them to 100%. The "contribution to variance" provided by Crystal Ball® is an approximation and does not necessarily represent the true variance apportionment, particularly when there are dependencies among input variables.

Figure A6 is a sensitivity chart indicating that approximately 46.5% of the variance observed in the EDI of nickel from food in children is the result of the body weight input variable.

Figure A6 Crystal Ball® 7 Directional Sensitivity Chart Showing Contribution to Variance, Assuming Uncorrelated Inputs



Note: BW, body weight; MEIR, meat and eggs intake rate; CGIR, cereals and grains intake rate; MDIR, milk and dairy intake rate; FNIR, fats, nuts, and oil intake rate.

Figure A6 is a directional sensitivity chart indicating not only the magnitude but the direction of the effect each assumption has on the distribution representing the EDI of nickel. The directional component of the sensitivity chart simply indicates whether an assumption is positively or negatively correlated with a particular forecast. Figure A6 was generated assuming uncorrelated assumptions between body weight and food intake rate. In other words, the probability distribution function describing the EDI rate of nickel did not use food intake rate data normalized for body weight. As a result, potential dependencies between body weight and food intake rate were not captured in this Monte Carlo analysis. Some databases report food intake rates for humans on a per kg body weight basis (i.e. mg/kg bw/d). These databases divide each reported consumption rate by the reported body weight, thereby reducing concern with regards to dependencies that may exist between body weight and food intake rate.

The impact of assuming a correlation between food group-specific intake rates and body weight on the EDI of nickel are demonstrated by comparing figures A7 and A8.

Figure A7 Probability Density Function (PDF) Showing Variability in Estimated Daily Intake (EDI) of Nickel, Assuming Correlated Assumptions

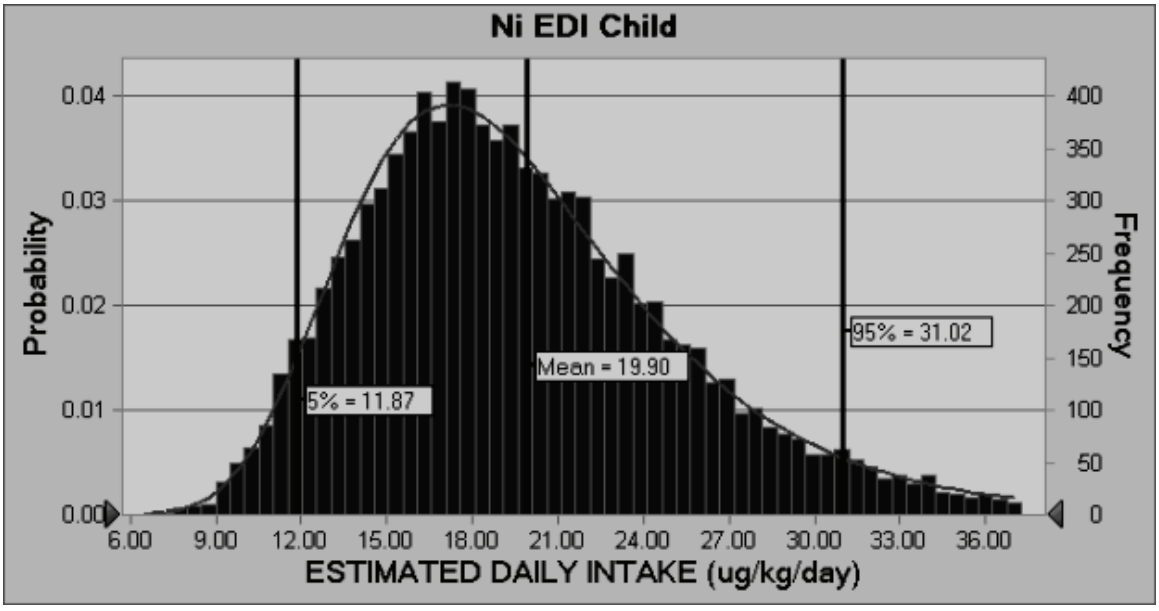
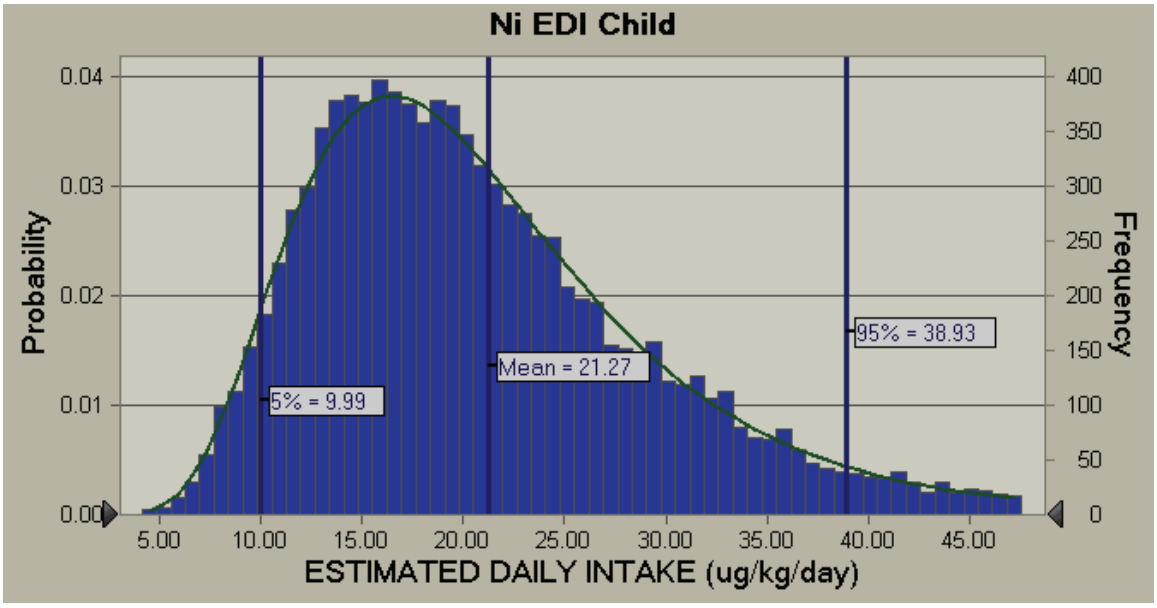


Figure A8 Probability Density Function (PDF) Showing Variability in Estimated Daily Intake (EDI) of Nickel, Assuming Uncorrelated Assumptions



Assuming a positive correlation coefficient between food intake rate and body weight assumptions resulted in a "tightening" (i.e. 5th and 95th percentiles are closer in value in Figure A7 than is the case in Figure A8) of the overall PDF.

A2.6.3.4 Other methods

Numerous other methods may be used in a global sensitivity analysis, including partial correlation coefficients, regression-based methods, and variance-based methods (e.g. correlation ratio, importance measures, and the Fourier amplitude sensitivity test). Discussion of these methods can be found in Sobol (1993), Saltelli et al. (1993, 2000), McKay (1995), Homma and Saltelli (1996), and Helton and Davis (2002).

A3.0 EXTENDING UNCERTAINTY ANALYSIS TO DERIVATION OF EFFECTS METRICS

Effects data can be characterized and summarized in a variety of ways, ranging from benchmarks designed to be protective of human health to dose-response curves for the surrogate species of interest (rodents, for example, if toxicity has been determined largely from rodent bioassay studies). Most jurisdictions rely on conservative reference doses (RfD) and mathematical models (unit risk approach) as effects metrics when characterizing risks for threshold acting (non-carcinogenic) and non-threshold acting (carcinogenic) chemicals, respectively. The unit risk approach is used for chemicals that show a non-threshold type dose-response relationship, and where there is evidence of damage to genetic material to evaluate carcinogenic risk through the use of cancer potency estimates (q1*) or other (TD_{05}) dose metrics. These metrics specify the dose or exposure level of the chemical that would result in some specified level of carcinogenic risk. The approach assumes "absolutely no chance of adverse effects" would be observed only when the exposure level or dose is zero. The dose-response curves for such chemicals are considered not to show an exposure threshold because the lesions produced are self-replicating. That is, the damage to genetic material (e.g. the mutation) can be passed on from one cell generation to the next during normal cell division. This means that once DNA damage has occurred, the presence of the chemical is no longer required for the expression of the adverse effects. Because there is a finite possibility of one molecule of a genotoxic chemical causing a mutation that results in a self-replicating lesion, there would not be an exposure threshold below which no risk of adverse effects would occur. The assumption of the absence of an exposure threshold in such circumstances results in a conservative hazard assessment because the damage to genetic material may be repaired, or the damaged cell(s) may die and not reproduce.

Mathematical models are used for carcinogenic chemicals to estimate an exposure level commensurate with risks acceptable to the individuals or groups involved (e.g. the dose associated with a risk of 1 in 100,000 or 1 in 1 million). A variety of mathematical models and approaches are available for estimating unit risk values for chemicals (Van Ryzin, 1980; Krewski et al., 1982; Miller et al., 1983; Crump and Howe, 1984; Carr, 1985; Crump and Crockett, 1985; U.S. EPA, 2005). None of these mathematical models have or can be biologically verified due to the low levels of risk deemed acceptable for humans (e.g. 1 in 100,000 or 1 in 1 million) (Clayson, 1987). Generally, the models used cannot distinguish among responses of about one order of magnitude, and consequently the lower 95% confidence limit of the exposure estimate at a specified risk is used in an attempt to err on the side of safety. Alternate procedures could include the use of the upper bound of the dose-response relationship to estimate a unit risk value, or the estimation of a risk specific dose from the maximum likelihood exposure estimate, rather than the lower 95% confidence limit.

There can be large differences in the exposure limit estimated with different mathematical models on the same data sets. Further, the difference in exposure limit derived by dose-response extrapolation using a mathematical model compared with the no observed adverse effect level (NOAEL) extrapolation factor approach can be three or more orders of magnitude (over 1,000 times). Therefore, the choice of procedures for estimating exposure limits can have wide-reaching ramifications on the hazard assessment of a chemical.

Although there are no plans to change the approaches currently used by Health Canada (see main report) to characterize effects and risk in the near future, this section undertakes a discussion of other approaches that may be used to characterize effects in a PRA. Risk assessors are routinely asked to predict the proportion of an exposed population that may suffer ill effects from chemical exposures. Although this is routinely answered for carcinogens, it is not generally answered for non-carcinogenic substances. Assessors should contact Health Canada before considering any of the approaches described below.

In higher-tier ecological risk assessments (e.g. U.S. EPA, 2002, 2004), effects characterization has often relied on concentration- or dose-response curves, but defaulted to benchmarks or other estimates of effect (e.g. NOAEL, or lowest observed adverse effect level [LOAEL]) when insufficient data were available to derive dose-response curves with any certainty. This section provides an overview of these and other procedures that have been used for characterizing effects information.

A3.1 Dose-Response Relationships

Most PRAs previously conducted estimated the probability that exposure exceeded a reference dose that is protective of human health (for non-cancer endpoints). An alternative approach is to estimate the probabilities of effects of varying magnitude. To do this, a concentration- or dose-response model is required. Generally, five or more treatments are required to develop concentration- or dose-response relationships, either from a single study or from multiple studies that used a similar methodology. The Generalized Linear Model (GLM) framework described by Kerr and Meador (1996) and Bailer and Oris (1997) is a useful one for deriving these relationships. It involves using link functions to transform effects metrics (e.g. probit or logit link functions for

quantal responses such as mortality) and assigning appropriate error distributions (e.g. binomial distribution for quantal responses). Linear regression can then be conducted on the transformed data to derive the dose-response relationship. Thus, the framework can be used for all available types of response variables (e.g. Moore et al., 2000). By adding a quadratic term to the linear model, the framework can be adapted to incorporate stimulation at low doses. The framework can also be easily expanded to include other dose-response models used in HHRAs for non-cancer endpoints (e.g. Weibull model) (Cothern et al., 1986).

Dose-response relationships may be combined with the corresponding exposure distribution in risk characterization to derive risk curves that characterize the relationship between probability and magnitude of effect.

A3.2 Hypothesis Testing to Determine LOAEL and NOAEL

Analysis of variance (ANOVA) is the most common method of estimating low-level toxic effects from subchronic and chronic tests. Reasons for this include the wide availability of software capable of performing ANOVA and related non-parametric tests, and the familiarity of regulators with the technique. Until recently, most toxicity-testing protocols specified experimental designs more suited to hypothesis-testing methods such as ANOVA than to regression-based approaches. However, hypothesis testing as an approach for estimating low-level toxic effects has some limitations.

- NOAELs and LOAELs are test doses that do not correspond with specified effects levels from one test to the next.
- Poor experimental design may mistakenly indicate that a contaminant is less toxic than it really is.
- Most information available from the toxicity test is not used (Pack, 1993; Stephan and Rogers, 1985; Suter, 1996).

As a result, hypothesis testing may not be the preferred method for analysis of toxicity data within PRAs conducted for contaminated sites.

In many cases, toxicity studies with five or more treatment levels are not available for mammalian species that are reasonable human surrogates. In those cases, the use of hypothesis testing may be necessary to estimate the NOAEL and LOAEL. In many toxicological studies, these endpoints were previously determined and reported. Such studies should be evaluated to determine that proper statistical procedures were followed. When the data can be obtained from the reports or directly from the authors, the data should be re-analyzed. In cases where a re-analysis is conducted, information regarding the minimal difference required to give a significant result must be reported (e.g. number of replicates, test variance, α, β, test-dose intervals). The percent effect associated with the LOAEL, relative to the control, should also be reported.

A3.3 Reference Dose Ranges and Distributions

Reference doses, tolerable daily intakes, maximum acceptable concentrations, and other human health toxicity benchmarks are designed to be conservative (i.e. err on the side of being overly protective of human health). In fact, they are generally derived as the best estimate of the human threshold dose or exposure level (i.e. the dose or exposure that is free of detrimental effects). For example, Rai et al. (2002) used the following equation to estimate tolerable daily intake (TDI) for tetrachloroethylene:

(A6)

$$TDI = \frac{NOAEL}{UF_1 \times UF_2 \times UF_3} = 0.0.14 \text{ mg/kgbw/d}$$

where UF_1, UF_2, and UF_3 are uncertainty factors to account for interspecies differences (10-fold extrapolation from Sprague Dawley rats to humans), intraspecies variation (10-fold to account for variation in sensitivity among individual humans), and extrapolation from a subchronic study duration to chronic exposures (also 10-fold), respectively. For a given human, however, it is possible that *UF*1 and *UF*2 could just as easily be as low as 0.01 and 2, respectively (Baird et al., 1996). Similarly, *UF*3 could be much smaller than 10, perhaps as low as 0.1 (Baird et al., 1996). Thus, if the goal of a PRA is to predict the proportion of humans for which exposure exceeds their "true" individual tolerable daily intake, then an unbiased assessment would involve the following steps.

- Derive an exposure distribution for the population of interest using the methods described in section A2.0.
- Derive a distribution for TDI by estimating distributions for the toxicity study effects metric (easily accomplished with, for example, an ED_{05} and confidence limits, and less easily accomplished with a NOAEL), and each of the uncertainty factors. Figure A9 illustrates this approach using the distributions specified by Rai et al. (2002) for tetrachloroethylene. The results indicate that there is a 97.8% probability that the "true" TDI for a given human is higher than the conservative point estimate TDI of 0.014 mg/kg bw/d. If only ranges (rather than PDFs) can be specified for the effects metric and uncertainty factors, then use the minima and maxima to parameterize uniform distributions. The result will be an unbiased TDI range. From a regulatory perspective, the "protective" (conservative) nature of the TDI for tetrachloroethylene may seem beneficial. When combined with the

conservative assumptions in the exposure assessment to characterize risk, the proportion of the population that may be "protected" approaches 100%. Unfortunately, it is impossible to define to what extent the risk assessment is overpredicting risks, and subsequently requiring site remediation or exposure mitigation when none is actually required.

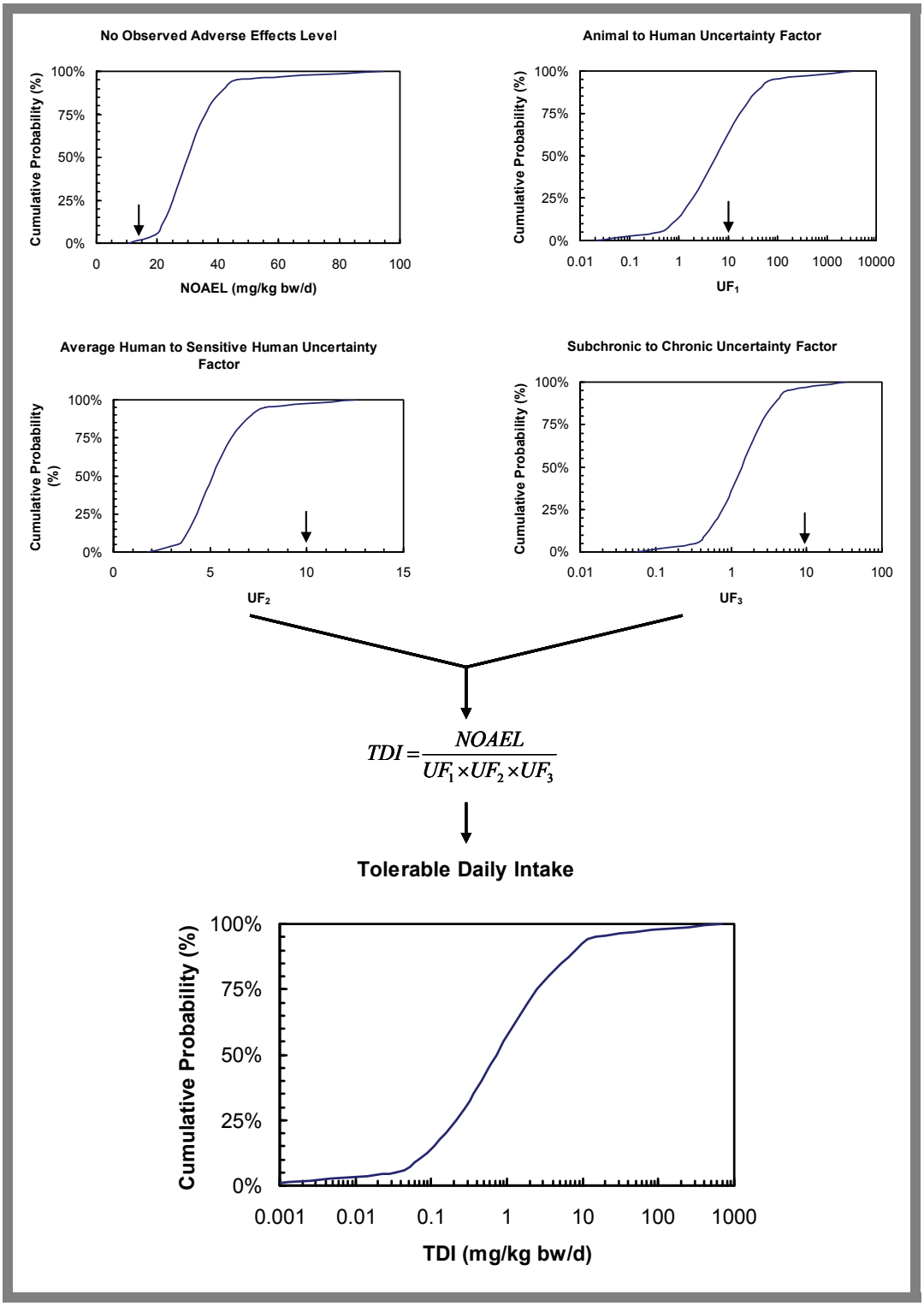
- Divide the exposure distribution by the TDI distribution.

A3.3.1 An approach for non-threshold chemicals

The effects metric for genotoxic non-threshold carcinogenic chemicals is a statistically derived value based on the upper 95% confidence limit of the unit risk cancer potency estimate, slope (SF) factor, or q_1^* value. Each of these is calculated by low-dose extrapolation of the dose-response curve using the linearized multistage model. To develop a distribution function for each q_1^*, the upper 95% confidence limit and the maximum likelihood estimate of the q_1^* slope-potency estimates can be calculated with the linearized multistage model (Crump and Howe, 1984). If a distribution shape is assumed for carcinogenic potency at low doses (e.g. lognormal), then uncertainty bounds can be placed around the q_1^* value used in risk estimation.

The issue of deriving effects metrics in a PRA is provided here for discussion purposes only. The above approaches to effects characterization have not been approved by Health Canada and should be discussed on a case-by-case basis. Proposals to incorporate a probabilistic treatment of the toxicity or effects metric will be considered on a case-by-case basis. However, proponents should be prepared for delays associated with the resolution of policy issues associated with such proposals.

Figure A9 Monte Carlo Analysis to Estimate Tolerable Daily Intake (TDI) for Tetrachloroethylene: Distributions and Point Estimates (see arrows



Note: *NOAEL*, no observed adverse effect level; UF_1, UF_2, and UF_3, uncertainty factors to account for interspecies differences,
Source: Rai et al. (2000).

A4.0 USES OF PROBABILISTIC RISK ASSESSMENT IN DECISION MAKING

The results of a PRA can contribute to decision making in a number of the following ways.

- Comparison of the results of the deterministic risk assessment (e.g. a hazard quotient) to the results of the PRA (e.g. quotient distribution) can be used to reveal the degree of conservatism in the former.
- A PRA reveals the understanding of risk. This understanding can have a dramatic impact on decision making. If, for example, an exposure distribution is narrow (i.e. uncertainty is low) and above a toxicity benchmark, then the decision to require a cleanup is readily justified. Alternatively, if the exposure distribution is narrow and below a toxicity benchmark, then the decision to not clean up the site is equally readily justified. If the exposure distribution is wide (i.e. uncertainty is high) and straddles the toxicity benchmark, then the decision on whether or not to clean up is less clear; in this case, further data collection and analysis may be required before an effective decision can be made, or a policy decision must be rendered on the "acceptable" probability of exceeding that benchmark.
- A PRA and accompanying sensitivity analysis can be used to identify research priorities for the site of interest and for Health Canada or other research programs. Those model variables with relatively high incertitude, but also with significant influence on the model output, are easily identified for further investigation to replace incertitude with known measured variability (knowledge).
- The PRA used to estimate risk can also be modified to estimate cleanup levels. The approach requires defining the criterion for acceptable risk (e.g. >95% of individuals in a population have exposure below the TDI, or would experience a cancer risk $<10^{-5}$). This decision is one of policy not science, and is not an easy one. Consultation with Health Canada and potentially other jurisdictions and interested parties would be required. Once acceptable risk has been defined, different values may be entered for the concentration term in an iterative series of probabilistic risk analyses until the output results in the agreed upon definition of "acceptable" risk.
- The alternative approach of rearranging the risk quotient equation to isolate the concentration term on the left cannot be used, because Monte Carlo analysis cannot be used to compute backcalculations (Ferson, 1996). If, for example, one had the equation *C* (risk) = *A* (exposure) ÷ *B* (effects), one might solve for *B* by re-arranging the equation to get *B* = *A* ÷ *C*. In a deterministic analysis, this backcalculation rearrangement works fine. In a Monte Carlo analysis, however, the answer is incorrect, as can be demonstrated by putting the distribution for *B* back into the original equation and computing *C*. The output distribution for *C* will be significantly different from the distribution for *C* when used as an input to deriving *B*. The reason for the discrepancy is that *B* was computed by assuming independence between *A* and *C*. Risk (*C*), however, is not independent of exposure (*A*). As Ferson (1996) states, "one is a function of the other, so independence is manifestly an inappropriate assumption." Solving such backcalculation problems either requires the iterative approach suggested here or a special operation called deconvolution (Burmaster and Thompson, 1995; Ferson, 1995; Ferson et al., 2004).

A PRA provides crucial information required for decision making. Such an analysis can be used to justify hedging away from large financial investment in remediation in the face of significant uncertainties, or to identify which aspects of the analysis are most uncertain or apt to affect the decision. This information can then be used to determine if it is wise to invest in collection of more information to reduce uncertainty before a remedial decision is made or a precise remedial plan is devised.

A PRA can make clear what is known and what is not, and what variables may be worth knowing more about—a huge advantage over the use of deterministic methods with conservative assumptions and safety factors. Thus, uncertainty analysis provides an objective and transparent means of comparing assumptions, models, and data put forth by interested parties in an environmental dispute. After such an analysis, it may still be agreed that it would be prudent to be conservative. This is appropriate given that the place for applying issues, such as defining what an acceptable risk should be, is during the risk management stage (the stage at which societal interests and policies are normally considered). Use of conservative assumptions and safety factors in an analysis has the effect of blurring the distinction between science and decision making (Burmaster, 1996). The task of risk assessors is to provide estimates of what is likely to happen, what might happen, what is not likely to happen, and to identify possible risk management options. The PRA approach does not negate the precautionary principle or a conservative approach, but rather moves it to the more appropriate risk management stage.

A5.0 PRINCIPLES OF GOOD PRACTICE

The following principles of good practice (adapted from Burmaster and Anderson, 1994) should be followed for all PRAs conducted for contaminated sites.

Describe the Underlying Model

- Show all formulae.
- Distinguish components of the models based upon scientific consensus, those with less acceptance, those based on professional judgment, and those based on policy.

Describe the Basis for Input Variable Distributions

- Describe the data used to construct the distributions. Discuss how well they represent the target variable. Evaluate data quality.
- Discuss methods and report the goodness-of-fit for distributions fit to data.
- Provide detailed information and graphs for each input distribution.
- Discuss the potential consequences of any extrapolations used.
- Discuss the effects or limitations of alternate distributions.
- Describe possible dependencies between input variables. If possible, invoke dependencies in probabilistic analyses. Otherwise, conduct "what if" analyses to determine potential consequences of dependencies on predictions of exposure or risk.

Describe the Analytical Procedures

- Explain how uncertainty propagation techniques were chosen.
- Describe how stochastic variability and uncertainty were handled.
- Conduct sensitivity analyses to identify which input variables need to be well quantified, or to develop a reduced set of input variables for further study.

Present Output

- Provide detailed information and graphs for the output distributions.
- Discuss stochastic variability, parameter uncertainty, and model uncertainty.

Many other standard QA procedures will enhance the accuracy, credibility, and transparency of a PRA. These include peer review of models and input data, and testing of models against performance criteria. Double entry should be used to detect errors in data or formula entry. Computer source codes and statistical models must be made available so that reviewers may repeat the analysis. The results of the uncertainty analysis need to be discussed to give reviewers an understanding of the strengths and weaknesses of the analysis. Scientific judgments or default assumptions used to bridge information gaps should be discussed. Analysts need to identify which uncertainties can be reduced, discuss the importance of uncertainties not included in the analysis, and make recommendations for additional data collection. Finally, it is important to discuss how the uncertainty analysis could influence and ultimately improve regulatory decision making.

A6.0 REFERENCES

Alefeld, G., and J. Herzberger. 1983. *Introduction to Interval Computations*. Academic Press, New York.

Andres, T.H. 1997. Sampling methods and sensitivity analysis for large parameter sets. *J. Stat. Comput.* 57: 77–110.

Bailer, A.J., and J.T. Oris. 1997. Estimating inhibition concentrations for different response scales using generalized linear models. *Environ. Toxicol. Chem.* 16: 1554–1559.

Baird, S.J.S., J.T. Cohen, J.D. Graham, A.I. Shlyakhter, and J.S. Evans. 1996. Noncancer risk assessment: A probabilistic alternative to current practice. *Human Ecol. Risk Assess*. 2: 79–102.

Banks, H.T., and L.K. Potter. 2004. Probabilistic methods for addressing uncertainty and variability in biological models: Application to a toxicokinetic model. *Math. Biosci.* 192: 193–225.

Barabas, N., P. Goovaerts, and P. Adriaens. 2001. Geostatistical assessment and validation of uncertainty for three-dimensional dioxin data from sediments in an estuarine river. *Environ. Sci. Tech*. 35: 3294–3301.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Res*. 23: 1393–1442.

Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Berleant, D. 1993. Automatically verified reasoning with both intervals and probability density functions. *Interval Comput.* 1993: 48–70.

Berleant, D., and C. Goodman-Strauss. 1998. Bounding the results of arithmetic operations on random variables of unknown dependency using intervals. *Reliable Comput*. 4: 147–165.

Bettonvil, B., and J.P.C. Kleijnen. 1997. Searching for important factors in simulation models with many factors: Sequential bifurcation. *Eur. J. Oper. Res*. 96: 180–194.

Brakewood, L.H., and D. Grasso. 2000. Floating spatial domain averaging in surface soil remediation. *Environ. Sci. Tech*. 34: 3837–3842.

Bukowski, J., L. Korn, and D. Wartenberg. 1995. Correlated inputs in quantitative risk assessment: The effects of distributional shape. *Risk Anal.* 15: 215–219.

Burmaster, D.E. 1996. Benefits and costs of using probabilistic techniques in human health risk assessments – with emphasis on site-specific risk assessments. *Human Ecol. Risk Assess*. 2: 35–43.

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 assessments. *Risk Anal*. 14: 477–481.

Burmaster, D.E., and K.M. Thompson. 1995. Backcalculating cleanup targets in probabilistic risk assessments when the acceptability of cancer risk is defined under different risk management policies. *Human Ecol. Risk Assess*. 1: 101–120.

Butcher, J.B., M.A. Medina, and C.M. Marin. 1991. Empirical Bayes regionalization methods for spatial stochastic processes. *Water Resources Res*. 27: 7–15.

Carr, C.J. 1985. Risk in decision making. *Reg. Toxicol. Pharmacol*. 5: 121–122.

Caswell, H. 1976. The validation problem. In: *Systems Analysis and Simulation in Ecology*. Volume IV. B.C. Patten (ed.). Academic Press, New York.

Chao, H.-P., S.C. Peck, and Y.S. Wan. 1994. Managing uncertainty: The tropospheric ozone challenge. *Risk Anal*. 14: 465–475.

Clarke, J.U., and D.L. Brandon. 1996. *Applications Guide for Statistical Analyses in Dredged Sediment Evaluations*. U.S. Army Corps of Engineers. Technical Report D-96-2.

Clayson, D.B. 1987. The need for biological risk assessment in reaching decisions about carcinogens. *Mutation Res.* 185: 243–269.

Conover, W.J. 1981. *Nonparametric Statistics*. John Wiley, New York.

Cothern, C.R., W.A. Coniglio, and W.L. Marcus. 1986. Estimating risk to human health. *Environ. Sci. Technol*. 20: 111–116.

Cotter, S.C. 1979. A screening design for factorial experiments with interactions. *Biometrika* 66: 317–320.

Covello, V.T., and M.W. Merkhofer. 1993. *Risk Assessment Methods: Approaches for Assessing Human and Environmental Risks.* Plenum Press, New York.

Crump, K.S., and P.W. Crockett. 1985. Improved confidence limits for low-dose carcinogenic risk assessment from animal data. *J. Haz. Mat*. 10: 419–431.

Crump, K.S., and R.B. Howe. 1984. The multistage model with a time-dependent dose pattern: Applications to carcinogenic risk assessment. *Risk Anal.* 4: 163–176.

Cullen, A.C., and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment*. Plenum Press, New York.

DeGroot, M.H. 1986. *Probability and Statistics*. Addison Wesley, Reading, MA.

Dwyer, P. 1951. *Linear Computations*. John Wiley, New York.

El-Shaarawi, A.H., and S.R. Esterby. 1992. Replacement of censored observations by a constant: An evaluation. *Water Res*. 26: 835–844.

Evans, G.T., and J.B. Dempson. 1986. Calculating the sensitivity of a salmonid population model. *C. J. Fish. Aqua. Sci*. 43: 863–868.

Ferson, S. 1995. Using approximate deconvolution to estimate cleanup targets in probabilistic risk analyses. In: *Hydrocarbon Contaminated Soils*. P.T. Kostecki, E.J. Calabrese, and M. Bonazountas (eds.). Amherst Scientific Publishers, Amherst, MA. pp. 245–254.

Ferson, S. 1996. What Monte Carlo methods cannot do. *Human Ecol. Risk Assess*. 2: 990–1007.

Ferson, S. 2002. *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis Publishers, Boca Raton.

Ferson, S., and M. Burgman. 1995. Correlations, dependency bounds and extinction risks. *Biol. Conserv*. 73: 101–105.

Ferson, S., and L.R. Ginzburg. 1995. Hybrid arithmetic. In: *Proceedings of the 1995 Joint ISUMA Third International Symposium on Uncertainty Modeling and Analysis and NAFIPS Annual Conference of the North American Fuzzy Information Processing Society*. IEEE Computer Society Press, Los Alamitos, CA. pp. 619–623.

Ferson, S., and T.F. Long. 1994. Conservative uncertainty propagation in environmental risk assessments. In: *Environmental Toxicology and Risk Assessment – Third Volume*. J.S. Hughes, G.R. Biddinger, and E. Mones (eds.). American Society for Testing and Materials, Philadelphia. pp. 97–110. ASTM STP 1218.

Ferson, S., R.B. Nelsen, J. Hajagos, D.J. Berleant, J. Zhang, W.T. Tucker, L.R. Ginzburg, and W.L. Oberkampf. 2004. *Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis*. Prepared by Sandia National Laboratories under contract to U.S. Department of Energy, Oak Ridge, TN. SAND2004-3072.

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

Finkel, A.M. 1994. Stepping out of your own shadow: A didactic example of how facing uncertainty can improve decision-making. *Risk Anal*. 14: 751–761.

Finley, B., D. Proctor, P. Scott, N. Harrington, D. Paustenbach, and P. Price. 1994. Recommended distributions for exposure factors frequently used in health risk assessment. *Risk Anal*. 14: 533–553.

Flavelle, P. 1992. A quantitative measure of model validation and its potential use for regulatory purposes. *Adv. Water Resour*. 15: 5–13.

Frank, M.J., R.B. Nelson, and B. Schweizer. 1987. Best-possible bounds for the distribution of a sum – a problem of Kolmogorov. *Probability Theory and Related Fields* 74: 199–211.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Ginevan, M.E., and D.E. Splitstone. 1997. Improving remediation decisions at hazardous waste sites with risk-based geostatistical analysis. *Environ. Sci. Technol*. 31: 92A–96A.

Glaz, J., and B.M.K. Johnson. 1984. Probability inequalities for multivariate distributions with dependence structures. *J. Am. Stat. Assoc*. 79: 436–440.

Gleit, A. 1985. Estimation for small normal data sets with detection limits. *Environ. Sci. Technol*. 19: 1201–1206.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.

Haimes, Y.Y., T. Barry, and J.H. Lambert. 1994. When and how can you specify a probability distribution when you don't know much. *Risk Anal.* 14: 661–706.

Hammonds, J.S., F.O. Hoffmann, and S.M. Bartell. 1994. *An Introductory Guide to Uncertainty Analysis in Environmental and Human Health Risk Assessment*. Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN.

Hattis, D., and D.E. Burmaster. 1994. Assessment of variability and uncertainty distributions for practical risk analysis. *Risk Anal*. 14: 713–730.

Helsel, D.R. 1990. Less than obvious: Statistical treatment of data below the detection limit. *Environ. Sci. Technol*. 24: 1766–1774.

Helton, J.C., and F.J. Davis. 2002. Illustration of sampling-based methods for uncertainty and sensitivity analysis. *Risk Anal.* 22: 591–622.

Henrion, M., and B. Fischhoff. 1986. Assessing uncertainty in physical constants. *Am. J. Phys.* 54: 791–797.

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: 707–712.

Homma, T., and A. Saltelli. 1996. Importance measures in global sensitivity analysis of model output. *Rel. Eng. Sys. Safety* 52: 1–17.

Hope, B.K. 1999. Distributions selected for use in probabilistic human health risk assessments in Oregon. *Human Ecol. Risk Assess*. 5: 785–808.

Hope, B.K. 2000. Generating probabilistic spatially-explicit individual and population exposure estimates for ecological risk assessments. Risk Anal. 20 (5): 573-589.

Hora, S.C. 1996. Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management. *Rel. Eng. Sys. Safety* 54: 217–223.

Iman, R.L., and J.C. Helton. 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Anal*. 8: 71–90.

Kerr, D.R. and J.P. Meador. 1996. Modeling dose response using generalized linear models. *Environ. Toxicol. Chem.* 15: 395–401.

Kitanidis, P.K. 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Res*. 22: 499–507.

Krewski, D., J. Kovar, and M. Bickis. 1982. Optimal experimental designs for low dose extrapolation. In: *Topics in Applied Statistics*. T.W. Dwivedi (ed.). Marcell Dekker, New York.

Lee, R.C., and W.E. Wright. 1994. Development of human exposure-factor distributions using maximum-entropy inference. *J. Expos. Anal. Environ. Epidemiol*. 4: 329–341.

Lin, L.I.-K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45: 255–268.

McKay, M.D. 1995. *Evaluating Prediction Uncertainty*. US Nuclear Regulatory Agency and Los Alamos National Laboratory. Technical Report NUREG/CR-6311.

McKone, T.E., and K.T. Bogen. 1992. Uncertainties in health-risk assessment: An integrated case study based on tetrachloroethylene in California groundwater. *Reg. Toxicol. Pharmacol*. 15: 86–103.

Miller, C.T., D. Krewski, and I.C. Munro. 1983. Conventional approaches to safety evaluation. In: *Proceedings of International Conference on Safety Evaluation and Regulation of Chemicals*. F. Homburger, S. Karger and A.G. Basel (eds.). Karger, Basel. pp. 66-76.

Moore, D.R.J. 1996. Using Monte Carlo analysis to quantify uncertainty in ecological risk assessment: Are we gilding the lily or bronzing the dandelion? *Human Ecol. Risk Assess*. 2: 628–633.

Moore, D.R.J., and S.M. Bartell. 2000. Estimating ecological risks of multiple stressors: Advanced methods and difficult issues. In: *Multiple Stressors in Ecological Risk and Impact Assessment: Approaches to Risk Estimation*. J. Foran and S. Ferenc (eds). SETAC Press, Pensacola, FL. pp. 117–168.

Moore, D.R.J., and P.-Y. Caux. 1997. Estimating low toxic effects. *Environ. Toxicol. Chem.* 16: 794–801.

Moore, D.R.J., and B.J. Elliott. 1996. Should uncertainty be quantified in human and ecological risk assessments used for decision-making? *Human Ecol. Risk Assess*. 2: 11–24.

Moore, D.R.J., W. Warren-Hicks, B.R. Parkhurst, R.S. Teed, R.B. Baird, R. Berger, D.L. Denton, and J.J. Pletl. 2000. Intra- and inter-treatment variability in reference toxicant tests: Implications for whole effluent toxicity testing programs. *Environ. Toxicol. Chem.*19: 105–112.

Moore, R.E. 1966. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, NJ.

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

Morris, M.D. 1991. Factorial sampling plans for preliminary computation experiments. *Technomet*. 33: 161–174.

Neumaier, A. 1990. *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge, U.K.

Okabe, A., B. Boots, K. Sugihara, and S.N. Chiu. 2001. *Spatial Tesselations: Concepts and Applications of Voronoi Diagrams*. John Wiley, New York.

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Sci*. 263: 641–646.

Ott, W.R. 1995. *Environmental Statistics and Data Analysis*. Lewis Publishers, Boca Raton.

Pack, S. 1993. *A Review of Statistical Data Analysis and Experimental Design in OECD Aquatic Toxicology Test Guidelines*. Organisation for Economic Co-operation and Development. Paris.

Parrish, R.S., and C.N. Smith. 1990. A method for testing whether model predictions fall within a prescribed factor of true values, with an application to pesticide leaching. *Ecol. Modelling* 51: 59–72.

Paté-Cornell, E. 2002. Risk and uncertainty analysis in government safety decisions. *Risk Anal*. 22: 633–646.

Rai, S.N., S. Bartlett, D. Krewski, and J. Paterson. 2002. The use of probabilistic risk assessment in establishing drinking water quality objectives. *Human Ecol. Risk Assess*. 8: 493–509.

Reckhow, K.H., J.T. Clements, and R.C. Dodd. 1990. Statistical evaluation of mechanistic water-quality models. *J. Environ. Eng*. 116: 250–268.

Reckhow K.H. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecol. Modeling* 72: 1–20.

Rodriguez, D., H. Walker, N. Klepikova, A. Kostrikov, and Y. Zhuk. 1995. Evaluation of two pollutant dispersion models over continental scales. *Atmos. Environ*. 29: 799–812.

Rowe, W.D. 1994. Understanding uncertainty. *Risk Anal.* 14: 743–750.

Saltelli, A., T.H. Andres, and T. Homma. 1993. Sensitivity analysis of model output: An investigation of new techniques. *Comput. Stat. Data Anal*. 15: 211–238.

Saltelli, A., K. Chan, and E.M. Scott. 2000. *Sensitivity Analysis*. John Wiley, Chichester, UK.

Seiler, F.A., and J.L. Alvarez. 1996. On the selection of distributions for stochastic variables. *Risk Anal*. 16: 5–18.

Singh, A.K., A. Singh, and M. Engelhardt. 1997. *The Lognormal Distribution in Environmental Applications*. Office of Research, U.S. EPA, Washington, DC. EPA/600/R-97/006.

Singh, A.K., A. Singh, and M. Engelhardt. 1999. *Some Practical Aspects of Sample Size and Power Computations for Estimating the Mean of Positively Skewed Distributions in Environmental Applications*. Office of Research, U.S. EPA, Washington, DC. EPA/600/S-99/006.

Singh, A., A.K. Singh, M. Engelhardt, and J.M. Nocerino. 2002a. *On the Computation of the Upper Confidence Limit of the Mean of Contaminant Data Distributions*. Office of Research, U.S. EPA, Washington, DC.

Singh, A., A.K. Singh, and R.J. Iaci. 2002b. *Estimation of the Exposure Point Concentration Term Using a Gamma Distribution*. Office of Research, U.S. EPA, Washington, DC. EPA/600/R-02/084.

Singh, A., A.K. Singh, and R.W. Maichle. 2004. *ProUCL Version 3.0 User Guide*. Office of Research, U.S. EPA, Washington, DC. EPA/600/R04/079.

Slob, W. 1994. Uncertainty analysis in multiplicative models. *Risk Anal*. 14: 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: 467–474.

Sobol, I.M. 1993. Sensitivity analysis for non-linear mathematical models. *Math. Modeling Comput. Exper.* 1: 407–414.

Sokal, R.R., and F.J. Rohlf. 1981. *Biometry*. W.H. Freeman and Company, NewYork.

Stephan, C.E., and J.W. Rogers. 1985. Advantages of using regression analysis to calculate results of chronic toxicity tests. In: *Aquatic Toxicology and Hazard Assessment*. R.C. Bahner and D.J. Hansen (eds.). STP 737. American Society for Testing and Materials, Philadelphia. pp. 377–387.

Suter, G.W. 1996. Toxicological benchmarks for screening contaminants of potential concern for effects on freshwater biota. *Environ. Toxicol. Chem*. 15: 1232–1241.

Suter, G.W., and L.W. Barnthouse. 1993. Assessment concepts. In: *Ecological Risk Assessment.* G.W. Suter (ed.). Lewis Publishers, Chelsea, MI.

Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley, New York.

United States Environmental Protection Agency (U.S. EPA). 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 9285.7-081.

U.S. EPA. 1997. *Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund (RAGS): Volume 3 – (Part A, Process for Conducting Probabilistic Risk Assessment).* Office of Solid Waste and Emergency Response, Washington, DC. EPA 540-R-02-002.

U.S. EPA. 2002. *Calcasieu Estuary Remedial Investigation/Feasibility Study (RI/FS): Baseline Ecological Risk Assessment (BERA)*. Appendix E1: Evaluation of the Predictive Ability of Effects-Based Sediment Quality Guidelines in the Calcasieu Estuary. U.S. EPA, Region 6, Dallas.

U.S. EPA. 2004. *Ecological Risk Assessment for General Electric (GE)/Housatonic River Site, Rest of River.* U.S. EPA, New England Region, Boston. DCN: GE-100504-ACJS.

U.S. EPA. 2005. *Guidelines for Carcinogenic Risk Assessment*. Risk Assessment Forum, U.S. EPA, Washington, DC. EPA/630/P-03/001F.

Van Ryzin, J. 1980. Quantitative risk assessment. *J. Occup. Med.* 22: 321–326.

Venkatram, A. 1982. A framework for evaluating air quality models. *Boundary Layer Meteorol*. 24: 371–385.

Vose, D. 1996. *Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling*. John Wiley, Chichester, U.K.

Warren-Hicks, W., and D.R.J. Moore (eds.). 1998. *Uncertainty Analysis in Ecological Risk Assessment.* SETAC Press, Pensacola, FL.

Williamson, R.C., and T. Downs. 1990. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *Int. J. Approx. Reasoning* 4: 89–158.

A7.0 GLOSSARY

Absolute deviation: The sum of the absolute differences between observed and predicted values divided by sample size.

Arithmetic mean: A measure of central tendency that is calculated as the sum of all the values of a set of measurements divided by the number of values in the set.

Bayesian: The Bayesian or subjective probability view holds that the probability of an event suggests the level of belief or state of knowledge warranted by the data in hand. In the classical or frequentist view, the probability of an event is the frequency with which an event occurs given a long sequence of identical and independent trials. The decision as to the appropriateness of either approach (Bayesian or classical) is based on the available data and the extent of subjectivity deemed appropriate.

Bias: The difference between the expected value of a statistic and the population value it is intended to estimate.

Bootstrap method: A statistical technique using a sample (e.g. 5,000) obtained from an original data set by randomly drawing and replacing the same number of values (e.g. 5,000) from the original sample or a distribution estimated for that sample.

Bound: An upper bound of a set of real numbers is a real number that is greater than or equal to every number in the set. A lower bound is a number less than or equal to every number in the set. In this report, the bounds on functions are also considered. These are not bounds on the range of the function, but rather bounds on the function for every function input. For instance, an upper bound on a function $F(x)$ is another function $B(x)$ such that $B(x) \geq F(x)$ for all values of x. $B(x)$ is the lower bound on the function if the inequality is reversed. If an upper bound cannot be any smaller or a lower bound cannot be any larger, it is called a best possible bound.

Coefficient of variation (also coefficient of variance or coefficient of variability): An estimate of relative variability that is equal to the standard deviation divided by the mean. It is independent of the unit of the measurement and is typically expressed as a percentage.

Confidence interval: The numerical interval constructed around a point estimate of a population parameter and associated with a particular probability level. A 95% confidence interval has the property that 95% of such intervals contain the true value.

Correlation coefficient: A measure of the closeness of the relationship between two variables. A correlation coefficient of +1 indicates perfect positive correlation. A correlation coefficient of -1 indicates perfect negative correlation. A correlation coefficient of 0 indicates no correlation. Widely used measures include the linear correlation coefficient (also called the product-limit or Pearson correlation coefficient) and non-parametric measures such as the Spearman rank-order or Kendall's tau correlation coefficients.

Correlation analysis: An investigation of the measure of statistical association among random variables based on samples. When the data are non-linear, non-parametric correlation is generally considered to be more robust than linear correlation.

Cumulative distribution function (CDF): The CDF is also referred to as the distribution function, cumulative frequency function, or the cumulative probability function. The cumulative distribution function, $F(x)$, expresses the probability that a random variable X assumes a value less than or equal to some value x, $F(x)$ = probability ($X \leq x$). For continuous random variables, the cumulative distribution function is obtained from the probability density function by integration (or by summation in the case of discrete random variables).

Dependency: A relationship between random variables. If one random variable is unrelated to another random variable, they are said to be independent. A dependency may be linear, non-linear, monotonic, or non-monotonic, and thus can take on a variety of shapes. The traditional measures of a dependency (correlation coefficients) are ill suited to non-monotonic relationships.

Elasticity: A term related to sensitivity that is defined as the relative change in model prediction over the relative change in the parameter value.

Expert judgment: A source of information based upon the experience of one or more scientists or experts. For Bayesians, expert judgment is frequently used to form the prior distribution, thus formally incorporating an expert's degree of belief into statistical procedures.

Expert: A person who has (i) training and experience in the subject area resulting in superior knowledge in the field, (ii) access to relevant information, (iii) an ability to process and effectively use the information, and (iv) is recognized by their peers or those conducting the study as qualified to provide judgments about assumptions, models, and model parameters at the level of detail required.

Geometric mean: The antilogarithm of the mean of the logarithms of values in a data set.

Goodness-of-fit: A set of mathematical tests performed to determine the fit between a standard probability distribution and a data set.

Hypothesis testing: In classical statistics, a formal procedure for testing the long term expected truth of a stated hypothesis. The statistical method involves comparison of two or more sets of sample data. On the basis of an expected distribution of the data, the test leads to a decision about whether to accept the null hypothesis (usually that there is no difference between the samples) or to reject that hypothesis and accept an alternative one (usually that there is some difference between the samples).

Incertitude: The kind of uncertainty arising from imperfect knowledge. Incertitude is also known as epistemic uncertainty, ignorance, subjective uncertainty, Type II or Type B uncertainty, reducible uncertainty, non-specificity, and state-of-knowledge uncertainty.

Kriging: An exact interpolation routine that depends on the probabilistic nature of surface changes with distance.

Latin hypercube sampling (LHS): In Monte Carlo analysis, one of two sampling schemes is generally employed: simple random sampling or Latin hypercube sampling. Latin hypercube sampling may be viewed as a stratified sampling scheme designed to ensure that the upper or lower ends of the distributions used in the analysis are well represented. It is considered to be more efficient than simple random sampling, that is, it requires fewer simulations to produce the same level of precision. Latin hypercube sampling is generally recommended over simple random sampling when the model is complex or when time and resource constraints are an issue.

Measurement error: The difference between a measured quantity and its actual or true value is called measurement error. The term is also used to refer to the imprecision or uncertainty about a measurement, although the term measurement uncertainty is now preferable for this meaning.

Median: The middle value for an ordered set of *n* values dividing a frequency distribution into two halves. It is represented by the central value when n is odd and by the mean of the two most central values when n is even.

Mode: That value that, if it exists, occurs most often in a data set.

Monotonic: A relationship between two variables that is either positive or negative through the entire range of the independent variable.

Non-monotonic: A relationship between two variables that changes between positive and negative over the range of the independent variable.

Non-parametric method: A body of statistical methods that do not require the estimation of population variance or mean. As these methods typically do not make any distributional assumptions (e.g. assume a normal distribution) of the sampled population, they are sometimes called distribution-free methods.

Parameter: Two distinct, but often confusing, definitions for parameter are used. (i) A constant or an independent variable in a mathematical equation or model. For example, in the equation $Z = X + 2Y$, the independent variables (X, Y) and the constant are all parameters. (ii) The constants characterizing the probability density function or cumulative distribution function of a random variable. For example, if the random variable W is known to be normally distributed with mean M and standard deviation F, the constants F and M are called parameters.

Parametric method: A statistical technique that depends on the assumption that the data are drawn from a specific distribution, such as the normal.

Point estimate: A value chosen to represent a constant or distribution.

Population: In statistics and sampling design, the total universe addressed in a sampling effort.

Probability: The Bayesian or subjective view is that the probability of an event is the degree of belief that a person has, given some state of knowledge that the event will happen or hold true. In the classical or frequentist view, the probability of an event is the frequency of an event occurring given a long sequence of identical and independent trials.

Probability density function (PDF): The probability density function expresses the probability that a continuous random variable falls within some very small interval. For discrete random variables, the term probability mass function (PMF) is preferred.

Probability mass function (PMF): The probability mass function expresses the probability that a discrete random variable takes a specific value.

Probability plot: Plot of observed versus fitted quantiles for a chosen distribution.

Quantile: The value in a distribution that corresponds to a specified proportion of the population distribution or distribution function. Quartiles (25th, 50th, and 75th percentiles) and the median (50th percentile) are special cases of quantiles.

Quantile-quantile (Q-Q) plot: A graphical data analysis technique for comparing the distributions of two data sets that plots the quantiles of the sample data against the quantiles of another data set or of a theoretical distribution.

Random variable: A random variable is a quantity that can take on any number of values but whose exact value cannot be known before a direct observation is made. For example, the outcome of the toss of a pair of dice is a random variable, as is the height or weight of a person selected at random from a city phone book.

Range: The difference between the largest and smallest values in a data set.

Representativeness: The degree to which a sample is characteristic of the population for which the samples are being used to make inferences.

Risk: The likelihood and severity of an adverse effect or event occurring to man or the environment following exposure, under defined conditions, to a risk source(s).

Sampling error: That part of the total variance of an estimator that is attributable to the sampling process.

Sensitivity (sensitivity analysis): Mathematical technique for determining the relative influence of an individual model input parameter on the output (related to elasticity). In a broader sense, sensitivity can refer to how conclusions change if models, data, or assessment assumptions are changed.

Simulation: In the context of Monte Carlo analysis, simulation is the process of approximating the output of a model through repetitive random application of a model algorithm.

Standard deviation: A measure of variation about the mean of a distribution that is expressed as the square root of the variance.

Stochastic: A process involving a random variable.

Systematic error: Occurs when the errors in the data are not truly random, such as might occur when the sample population is not representative of the entire population.

Uncertainty: Uncertainty refers to lack of knowledge about specific factors, parameters, or models. For example, one may be uncertain about the mean concentration of a specific pollutant at a contaminated site or one may be uncertain about a specific measure of uptake (e.g. 95th percentile fish consumption rate among all adult males in the United States). Uncertainty includes parameter uncertainty (measurement error, sampling error, systematic error), model uncertainty (uncertainty due to necessary simplification of real-world processes, mis-specification of the model structure, model misuse, use of inappropriate surrogate variables), and scenario uncertainty (descriptive errors, aggregation errors, errors in professional judgment, incomplete analysis). Generally, uncertainty can be reduced with further information and knowledge.

Variability: Variability refers to observed differences attributable to true heterogeneity or diversity in a population or exposure parameter. Sources of variability are the result of natural random processes and stem from environmental, lifestyle, and genetic differences among humans. Examples include human physiological variation (e.g. natural variation in body weight, height, breathing rate, drinking water intake rate), weather variability, variation in soil types and differences in contaminant concentrations in the environment. Variability is usually not reducible by further measurement or study (but can be better characterized).

Variance: A measure of the dispersion of a set of values or of a distribution. Sample variance is calculated by (i) calculating squares of differences between observed and predicted values, (ii) summing the squares of differences, (iii) dividing the sum by sample size minus one, and (iv) calculating the square root of the result from step (iii).

APPENDIX B
GUIDANCE FOR DEVELOPMENT OF TOXICOLOGICAL REFERENCE VALUES (TRVS) FOR FEDERAL CONTAMINATED SITE RISK ASSESSMENTS, IN THE ABSENCE OF PUBLISHED REGULATORY TRVS

TABLE OF CONTENTS

B1.0 INTRODUCTION

Despite the efforts of national and international regulatory agencies, it is still common to encounter chemicals at contaminated sites for which those agencies have not derived estimates of toxic potency. This is understandable because, when the list of chemicals is combined, these regulatory agencies provide the rationale for toxicological reference value (TRVs) used for fewer than 1,000 individual chemicals. Meanwhile, over 100,000 chemicals are estimated to be used in global commerce (with more than 1,000 new chemicals entering the global environment each year). Also, the chemical form or species of a substance at a contaminated site may differ from the form or species for which a TRV has been prescribed. Consequently, it is not uncommon to identify a chemical for which no TRV is available for evaluation of human health risks at contaminated sites.

This report provides general guidance for development of TRVs, also known as regulatory exposure levels (RELs), when specific values are not available from recognized regulatory agencies. Guidance is provided on how a toxicologist may approach the development of proposed TRVs under circumstances where no TRVs are available from recognized agencies. In addition, the guidance provided in this report may be useful in circumstances where new toxicological data become available but have not yet been considered by major health agencies. This report outlines the approaches generally employed by Health Canada in the development of TRVs at the current time. It provides references to other more detailed guidance documents and highlights key considerations to be made in data interpretation. However, this document does not, and was not intended to provide a precise "recipe" for TRV development. The precise approach, subtleties, and nuances of TRV derivation are very much defined on a chemical-by-chemical basis.

Several sources of toxicological reference values from recognized regulatory agencies should be investigated before initiating the de novo derivation of a TRV. These may include (but are not limited to) the following:

- Health Canada
- Canadian Council of Ministers of the Environment (CCME)
- United States Environmental Protection Agency (U.S. EPA)
- World Health Organization (WHO)
- Agency for Toxic Substances and Disease Registry (ATSDR)
- Netherlands National Institute of Public Health and the Environment (RIVM)
- European Chemicals Agency (ECHA)

In addition to these sources, other international and, occasionally, provincial regulatory agencies may offer guidance for various contaminants of potential concern (COPCs).

The methodology in this document is intended to give general guidance on development of TRVs for which no values are provided by major regulatory agencies. **It is recommended that development of such TRVs should be completed only by experienced toxicologists with an understanding of the process by which TRVs are developed for protection of human health in Canada and elsewhere.** TRV development is a particularly sensitive area of human health risk assessment and improper analysis may result in a drastic over- or underestimation of actual risks. Care must be taken to ensure that a properly trained and experienced toxicologist has completed the evaluation in a scientifically defensible and thorough manner.

References are made to major guidance documents and to key papers that relate to issues in risk assessment. The reference list is not intended to be exhaustive, but provides the reader with initial sources of information that can help to define key issues and guidance for consideration in TRV development.

It must be stressed that the derivation of de novo or unique TRVs following this guidance cannot be construed in any way as providing Health Canada endorsement, acceptance, or approval of the new or revised TRV. The process of regulatory adoption and acceptance of TRVs involves a process of extensive internal review, external peer review, and often federal/provincial/territorial discussion. Any new or revised TRV proposed for the risk assessment and/or risk management of federal contaminated sites in Canada must be submitted to Health Canada for discussion, review, and consideration as may be appropriate.

Finally, it should be noted that, in cases where TRVs currently exist for a given substance, the review by Health Canada of proposed revised TRVs will be considered only under significant and extenuating circumstances, including:

- appreciable changes to the quantity and quality of data published since the release of the existing TRV;
- overwhelming evidence for a change in classification as carcinogenic or non-carcinogenic; and/or
- a major shift in the chemical description or analysis of a substance (for example, change from total polychlorinated biphenyls (PCBs) to non-co-planar and co-planar congeners).

B2.0 GUIDANCE ON METHODOLOGY FOR DEVELOPING TOXICOLOGICAL REFERENCE VALUES

B2.1 Introduction

In all cases, it is stressed that the toxicologist must use a "weight of evidence" approach in deriving proposed TRVs. Where possible, the weight of evidence approach considers all of the applicable data in the process before reaching conclusions on the toxic potency of chemicals (HC, 1994; U.S. EPA, 2002, 2005a). For example, in the designation of a chemical as either a threshold-response[2] or non-threshold-response[3] chemical, the entire data set must be considered (i.e. genotoxicity assays, animal studies, human studies, structure activity relationships, etc.). In the case of chemicals with existing TRVs, all available data, not just the most recent data, must be considered. With the above noted, it is also important that where uncertainties in the interpretation of toxicological data exist, the toxicologist follow the precautionary principle[4] and err on the side of conservatism to ensure that the selected TRV is more likely to overstate rather than understate the chemical's toxic potency, thereby being adequately protective of the health of the Canadian public. In particular, the importance of protecting potentially sensitive subgroups of the population, including the foetus and young children, the elderly, people with pre-existing health conditions, pregnant women, etc. must be considered.

An overview of the methodology that can be considered in developing TRVs is provided in this section of the report. The Health Canada approach to TRV derivation (HC, 1994, 1995, 1996) is generally consistent with the methods recommended by various regulatory agencies, including the U.S. EPA (U.S. EPA, 2002, 2005a) and the WHO (WHO, 1999). The approach recommended by Health Canada incorporates flexibility to ensure the best possible scientific decisions are made in developing TRVs. With the types of data that must be considered and the various approaches that can be used, it is not possible to provide precise "one-approach-fits-all," step-by-step guidance on TRV development. Instead, guidance is provided herein on the general approaches that can be considered. Literature references are identified to more specific guidance. Key guidance documents follow.

- Health and Welfare Canada (1991): *Carcinogen Assessment. A Research Report to the Department of National Health and Welfare*
- Health Canada (1994): *Human Health Risk Assessment for Priority Substances*
- Health Canada (1995): *Approach to the Derivation of Drinking Water Guidelines. Part I* in: *Guidelines for Canadian Drinking Water Quality – Supporting Documents*
- Health Canada (1996): *Health-Based Tolerable Daily Intakes/Concentrations and Tumorigenic Doses/Concentrations for Priority Substances*
- U.S. EPA (1995): *The Use of the Benchmark Dose Approach in Health Risk Assessment*
- U.S. EPA (2000): *Benchmark Dose Technical Guidance Document*
- U.S. EPA (2002): *A Review of the Reference Dose and Reference Concentration Processes*
- U.S. EPA (2005a): *Draft Final Guidelines for Carcinogen Risk Assessment (Final Report)*
- U.S. EPA (2005b): *Supplemental Guidance for Assessing Cancer Susceptibility from Early-Life Exposure to Carcinogens*
- WHO (1999): *Principles for the Assessment of Risks to Human Health from Exposure to Chemicals*

In addition to these, the following guidance documents provide information on specific areas of toxicological assessment.

- U.S. EPA (1991). *Guidelines for Developmental Toxicity Risk Assessment*
- U.S. EPA (1996): *Guidelines for Reproductive Toxicity Risk Assessment*
- U.S. EPA (1998): *Guidelines for Neurotoxicity Risk Assessment*

---

[2] Threshold response toxicants are "those for which the critical effect is not considered to be cancer or a heritable mutation" (HC, 1994, p. 10). It is noted that there are certain chemicals that may cause cancer via a threshold response mechanism (i.e. non-genotoxic carcinogens) that may be considered threshold toxicants.

[3] Non-threshold toxicant is used for substances "for which the critical effect is assumed to have no threshold (i.e. currently restricted to mutagenesis and genotoxic carcinogenesis)" (HC, 1994, p. 8).

[4] The precautionary principle follows a framework that is flexible and responsive to particular circumstances. The application of the precautionary principle "recognizes that the absence of full scientific certainty shall not be used as a reason for postponing decisions where there is a risk of serious or irreversible harm" (HC, 2003).

- WHO (2001): *Principles for Evaluating Health Risks to Reproduction Associated with Exposure to Chemicals*

Overall, guidance for development of TRVs is very similar among Health Canada, the WHO, and the U.S. EPA. It is noted that these different regulatory agencies may have different TRVs for the same chemicals, despite often having access to the same toxicological data. The main reasons for differing TRVs among these regulatory agencies tend to be chemical-specific considerations as opposed to differences in overall approach. Most differences in TRVs can be classified according to the following considerations:

- identification of different pivotal studies for TRV derivation
- differing classification of threshold versus non-threshold response on a chemical-specific basis
- differing use of uncertainty factors on a chemical-specific basis

The latter two considerations reflect differing policies regarding the interpretation of scientific data (e.g. how much and what kind of data are required to conclude a substance is carcinogenic) rather than being scientific differences per se. A comparison of WHO and U.S. EPA TRVs for 27 pesticides revealed similarities between the two agencies for the prescribed TRVs for most of the chemicals (a difference of 10% or less), but they differed by approximately 25% for three of them (the U.S. EPA TRV was lower than the WHO TRV in all cases). This comparison illustrates how the complexity of the scientific database and the application of individual (agency) scientific judgment and policies can lead to disparity, even when the science evaluated and the general agency principles and approaches are the same (Lu and Dourson, 1992).

B2.2 Determination of Exposure Parameters of Concern

Once it has been ascertained that an appropriate TRV is not available from a recognized regulatory agency, the toxicologist should first define the exposure parameters that are of primary concern in the assessment. Key exposure parameters include the nature of the contaminated site receptor (adult, child, worker, etc.), the pathways of exposure (ingestion, inhalation, dermal absorption), and the duration of exposure (acute, subchronic, chronic). If data permit, the TRV should be tailored to the specific exposure situation that most suits the contaminated site scenario of concern. For example, if exposure will be from childhood and throughout life, then data from juvenile animal studies and lifetime studies should be considered. If exposure will be only to adults, data from studies using adult animals may be more appropriate. If worker exposure is the key area of concern, human epidemiology studies using similarly exposed workers may be most appropriate.

The anticipated exposure routes must be identified because some chemicals act differently depending on exposure route. Where this is the case, it is important to consider studies with exposure routes relevant/equivalent to anticipated routes of human exposure for TRV development. For example, a chemical may cause toxicity in the respiratory tract following inhalation exposure, but may be relatively non-toxic following ingestion—and vice versa. If the exposure route of the subject risk assessment is ingestion, ideally the TRV should be based on the results of ingestion toxicology studies and not inhalation studies. The toxicologist should recognize that situation-specific TRVs may be appropriate in some cases.

The exposure duration of concern should also be identified so that relevant studies can be considered in TRV development. For example, it would not be desirable to compare exposures of long-term (chronic) duration to TRVs developed to be protective in short-term (acute, subchronic) exposure scenarios. It is recognized that in most situations involving contaminated site human health risk assessment, that long-term exposure durations do require protection. Thus, in the vast majority of situations, it is anticipated that exposure duration for the TRV will be chronic. Nevertheless, this step is recommended to affirm that the proper duration of exposure is being evaluated.

As previously noted, it is typically desirable to compare an exposure estimate to a TRV developed for that duration of exposure; however, in many cases, TRVs may be developed only for a chronic exposure duration. In most cases, if the exposure estimate is based on a duration that is shorter than that used to develop the TRV (i.e. the risk of a subchronic exposure is characterized against a chronic TRV), it is unlikely that human health risks will be underestimated (in fact, risks may be substantially overestimated). This is because humans can generally tolerate a higher level of exposure over a shorter time period; consequently, acute TRVs are generally greater in magnitude (value) than subchronic TRVs that are, in turn, greater in value than chronic TRVs. On the other hand, risks may be underestimated if the exposure estimate is based on a duration that is appreciably longer than that upon which the TRV was developed. This is typically countered by the application of an additional uncertainty factor when shorter duration studies are used as the basis for TRV development.

In summary, the first steps involved in developing a situation-specific TRV where no TRV currently exists are (i) identify receptors of concern, (ii) identify exposure routes of concern, and (iii) identify anticipated exposure duration. Identifying these key exposure parameters will guide the choice of data to be used in TRV development.

B2.3 Literature Search

The next step of TRV development involves the toxicologist conducting a thorough search of the relevant toxicological and scientific literature. Relevant literature may include the following:

- in vitro toxicology studies (cytotoxicity, mutagenicity, genotoxicity, etc.)
- in vivo toxicology studies in laboratory animals
- epidemiological studies
- case reports
- controlled clinical studies
- absorption, distribution, metabolism, and excretion studies
- toxicokinetics studies
- mechanistic studies

On-line databases are the most common tools for completing toxicological literature searches. Some of the important on-line databases providing toxicological literature include the following:

- PubMed database maintained by the U.S. National Institutes of Health (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed)
- MedlinePlus database maintained by U.S. National Library of Medicine and U.S. National Institutes of Health (http://medlineplus.gov/)
- TOXNET database maintained by the U.S. National Institutes of Health (http://toxnet.nlm.nih.gov/)
- TOXLINE database maintained by the U.S. National Institutes of Health (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE)
- EMBASE database (http://www.embase.com/about)

In an assessment of search efficiency using multiple bibliographic databases in the fields of occupational and environmental toxicology, it was found that searching two or more of the major databases will provide the most exhaustive search. A combination of EMBASE and TOXLINE, or to a lesser extent EMBASE and Medline, provided the best balance of time consumed versus search efficiency (Gerhanno et al., 1998). However, the failure to identify important studies can seriously undermine the credibility and validity of any proposed TRVs. Therefore, the literature search must be reasonably exhaustive.

Key words and terms that could be included in the search include (but not limited to):

- chemical name(s), Chemical Abstract Service (CAS) registry number, toxicity, adverse effects, mutagenicity, genotoxicity, carcinogenicity, acute toxicity, subchronic toxicity, chronic toxicity, reproductive toxicity, developmental toxicity, immunological toxicity, neurotoxicity, teratogenicity, epidemiology, clinical trial, case report, and all synonyms and related terms of these

It is noted that some chemicals may have a number of different names so it will be important to complete the literature search using the various synonyms for the chemical and/or the CAS registry number. The CAS registry number is a unique identifier that will indicate whether or not a chemical has multiple synonyms. A good source of information on chemical names and synonyms can be found at http://chemfinder.cambridgesoft.com.

All relevant literature, not just cited studies, must accompany the submission to Health Canada. In some cases, abstracts are relatively easy to obtain and can often provide sufficient information to discount the importance of certain papers. However, for all papers that are considered to be potentially relevant, copies of the papers must be obtained. If a proposed TRV is to be submitted to Health Canada for discussion and consideration, it is generally not acceptable to rely on third-party accounts or secondary references of what the paper states, demonstrates, or concludes. If information must be cited from a secondary source, this must be clearly noted.

In some cases, the manufacturer of the chemical may have data from in-house testing, and research reports may be available upon request. Also, academic researchers may share preliminary or unpublished data. Although such reports and data might be used as supporting or confirmatory evidence, studies that have not been subjected to independent peer review (such as through publication in a relevant scientific journal) generally should not be used as the key or critical basis for TRV derivation.

B2.4 Review of the Toxicological Studies

Once the literature search has been completed and the literature acquired, the toxicologist will need to determine the adequacy of the studies for development of TRVs. Because toxicological literature is published by a wide variety of sources (e.g. government agencies, universities, industry), data quality can vary substantially. Literature published since 1980 may be of greater quality (especially papers based on Good Laboratory Practice or GLP); however, all relevant literature must be obtained because there are numerous

examples of older literature being very relevant to the derivation of TRVs. Papers that have been published in peer-reviewed scientific journals are more reliable than those that have not been peer reviewed. As previously mentioned, studies that have not been subjected to independent peer review (such as through publication in a relevant scientific journal) should generally not be used as the key or critical basis for TRV derivation.

In all cases, the acquired literature should be critically reviewed and scrutinized to determine that it is adequate to support the development of a TRV. In this review, the toxicologist must ascertain to the extent possible that sound scientific principles were followed in completing the study. The U.S. EPA (2002) provides a list of questions that could be considered in the evaluation of toxicological data from animal and human studies. Some of the general issues that the toxicologist must be particularly concerned with follow:

- study design (including selection of proper dose groups, adequate control groups, number and type of animals, route of exposure, duration of the study, use of GLP, etc.)
- shape/validity of the dose-response relationship
- biological plausibility of the chemical to cause the effect(s)
- validity of the no observed adverse effect level (NOAEL) and/or the reported effects
- potential species differences in toxic response
- pharmacokinetic and metabolism considerations
- validity of the statistical analyses
- validity of the conclusions

In all cases, the toxicologist must be confident that the toxicological data are well supported by the science and are relevant to the issue of concern.

B2.4.1 Minimum data requirements

For a toxicologist to derive a TRV for a chemical, there must be valid toxicological data providing dose-response information applicable to the exposure route and duration of concern. The toxicological database must provide information on sublethal effects (i.e. a database that provides only mortality or other extremely severe toxicity is not appropriate for TRV development). If such toxicological information is not available, it will generally not be possible to propose a new or revised TRV with the available information to the satisfaction of Health Canada.

Under certain circumstances, alternative approaches may be considered. Occasionally, chronic TRVs are derived from subchronic (but never acute) toxicity data, through the application of additional uncertainty and modifying factors that account for the discrepancy in exposure duration. Other approaches have also been proposed (e.g. toxicological surrogate approach, chemical of unknown potency approach) but these are generally not desirable. See section B3.0 for a discussion of these alternate approaches.

B2.5 Determination of Whether the Chemical Elicits a Threshold or Non-Threshold Response

Once the relevant literature has been obtained and reviewed, the next step of the process involves determining whether or not a chemical may display a threshold or non-threshold dose-response relationship. As described by Health Canada (1994), threshold-response toxicants are “those for which the critical effect is not considered to be cancer or a heritable mutation” (although certain chemicals that may cause cancer via a non-genotoxic mechanism may be considered threshold toxicants). More specifically, threshold-response chemicals are chemicals that are believed to cause toxicity only when a certain level of exposure (or threshold) is exceeded. For these chemicals, a single molecule of the chemical is essentially assumed to have no possibility of causing adverse health effects. Instead, toxicity is assumed to occur only after a certain level of exposure is exceeded that overwhelms the cells of the organisms. Common examples of threshold-response chemicals (whose primary effects are not considered to be associated with genotoxic cancer initiation or heritable mutation) include various irritant chemicals (e.g. sulfur oxides), various systemic toxicants (e.g. barium), neurotoxins (e.g. mercury), reproductive toxicants/most teratogens (e.g. xylenes), and certain carcinogens that are not considered to be genotoxic (e.g. epigenetic carcinogens that act via mechanisms not involving direct interaction with DNA, such as PCBs).

In contrast to threshold-response toxicants, there is a separate group of chemicals assumed to act via a non-threshold dose-response relationship. According to Health Canada (1994), non-threshold toxicant is used for substances “for which the critical effect is assumed to have no threshold (i.e. currently restricted to mutagenesis and genotoxic carcinogenesis).” For these chemicals, it is essentially assumed that any exposure other than zero is associated with some risk of effect. In other words, even one molecule could theoretically cause an adverse effect (although the risk would be very low). For these chemicals, it is conservatively assumed there is no level of exposure that is completely without some level of risk (aside from zero dose); thus, risk from these chemicals is evaluated using the concept of de minimis risk levels. Genotoxic carcinogens are considered to act by a non-threshold mechanism. Chemicals that can interact with germ cell DNA leading to reproductive toxicity (genotoxic teratogens) may also potentially act through a non-threshold mechanism.

As suggested in the discussion above, the terms "threshold response" and "non-threshold response" are preferable to the terms "non-carcinogen" and "carcinogen." Certain chemicals may cause cancer via a non-genotoxic response mechanism and, thus, not all threshold-response chemicals are considered to be non-carcinogens.

A weight of evidence analysis is required for classification of the potential human carcinogenicity of a chemical. Health and Welfare Canada (1991), Health Canada (1994), and the U.S. EPA (1999) provide some of the factors to consider in evaluating the potential carcinogenicity of a chemical. Briefly, to determine if a chemical should be treated as a threshold-response or non-threshold-response chemical, the toxicologist should answer the following questions.

1. Has the chemical or its metabolites been associated with relevant malignant tumours in animals or people? (i.e. Did it increase the number or type of tumours? Did it decrease the development time for tumours to occur?)
2. Has the chemical or its metabolites been found to interact with DNA? (i.e. Does the weight of evidence from a battery of genotoxicity tests indicate direct mutagenic action?)
3. Is the chemical structure of the substance very similar to any other chemical that is considered to be a genotoxic carcinogen?
4. Do supporting or corroborative occupational or other studies on human subjects exist?

If the toxicological data indicate a "yes" response to questions (1) and either (2) or (3), it is probably most appropriate (and conservative) to assume that the chemical is a non-threshold-response chemical, particularly in the absence of clear evidence to the contrary. It is noted that this is a general rule, and it is possible that sometimes a chemical is considered to be a threshold-response chemical under some circumstances, even when a positive response is made to the preceding questions. For example, sometimes tumours may occur in laboratory animals via mechanisms that are not relevant to humans because of species differences in metabolism or tissue distribution, species differences in anatomy, or species differences in DNA repair efficiency. In some cases, the doses used in the animal studies may be too high to be extrapolated to doses anticipated for human exposure. Similarly, in some cases, genotoxicity may occur only at concentrations that cause cytotoxicity, and may not be relevant to low concentration exposures. Finally, in some cases, cancer may occur only at dose levels that are much greater than other toxic endpoints; thus, protection against non-cancer effects will also protect against cancer risks. A weight of evidence assessment of all the available data is required to support the classification of a carcinogenic mechanism as being threshold or non-threshold based.

Although the preceding discussion primarily concerned genotoxic carcinogens (i.e. the most common type of non-threshold toxicant), a similar approach can be used to determine if a chemical should be considered to be a genotoxic teratogen.

The classification of a chemical as threshold or non-threshold acting dictates the overall approach to be taken in the establishment of the TRV. In the past, two completely different approaches were taken to derive TRVs for threshold- versus non-threshold-acting chemicals. In recent years, however, regulatory agencies have come to the conclusion that a single approach is suitable in defining the dose-response relationship in the measurable part of the dose-response curve, regardless of the mechanism of toxic action. This new approach involves dose-response modelling to determine a dose level at which it is anticipated that a certain percentage of subjects (usually 1%–10%) will respond. This is called the benchmark dose (BMD). When air concentration data are used instead of doses per unit body weight, the corresponding benchmark concentration (BMC) is derived.

The BMD or BMC becomes the starting point for further development of the TRV. If the chemical is threshold acting, uncertainty factors are applied to the BMD or BMC to arrive at the TRV. If the chemical is non-threshold acting, a low-dose linear extrapolation is performed to determine a TRV associated with a given level of risk. These approaches are described in more detail in later sections of this document. Because the development of the BMD or BMC is common to both types of assessment, this is described first in section B2.6, and specific discussions of TRV development for threshold-acting and non-threshold-acting chemicals follow in sections B2.7 and B2.8, respectively.

B2.6 Key Issues to Consider in Weight of Evidence Assessment and Extrapolation to Humans

B2.6.1 Use of human versus animal data

Data relevant to the derivation of TRVs may come from experimental animal studies, experimental human studies, or epidemiology studies. The issues to be considered in choosing the most appropriate data set for use in risk assessment have been discussed by Kimbrough (1995).

Occasionally, data from a human study are available and suitable for the purpose of dose-response assessment and TRV derivation. However, such cases are rare for substances lacking existing regulatory TRVs. Therefore, animal studies will most often be used for de novo TRV

development. Nevertheless, when data are collected from human studies in a sufficient and robust manner, this type of data has several advantages over animal studies. Suitable human studies would likely meet the following criteria: (i) involve large numbers of persons in the study group as compared to animals, (ii) involve exposures to dose rates and concentrations that are closer to the levels associated with the contaminated site, and (iii) may involve all segments of the population including the young, the elderly and the infirmed.

On the other hand, human epidemiology studies are typically less controlled than animal toxicity studies and often suffer from limitations that include: (i) poor quantification and unclear understanding of actual concentrations that persons were exposed to, (ii) concomitant exposures to multiple substances, (iii) often involve less-than-lifetime exposure durations, (iv) do not properly consider the impact of confounding lifestyle or socio-economic factors, (v) are often from workplace exposures that could be much higher than environmental exposures, and (vi) frequently suffer from reporting bias. Controlled human studies are sometimes available for short-term exposures and can provide valuable information, but the data generally cannot be extrapolated to long-term risk assessment.

Animal studies often are the most suitable for TRV development but this needs to be evaluated on a case-by-case basis. Data from animal studies are often limited by the use of relatively small numbers of animals, exposure to doses far in excess of the expected human dose range from environmental exposure, and the possibility of species differences in response. On the other hand, the doses and all other environmental factors can be closely controlled, exposure to other chemicals is prevented, the possibility of inter-individual differences in response is minimized, and effect parameters can be monitored precisely and consistently. Data from animal studies also often allow assessment of the mechanism of action and serve as a basis for evaluation of possible mechanisms in humans. Overall, both human and animal data, if available, must be considered in the weight of evidence assessment.

B2.6.2 Selection of the key target organ for toxicity

A weight of evidence assessment should be conducted to determine the key target organ for toxicity. The target organ is usually the organ system that is adversely affected at the lowest observed adverse effect level (LOAEL). However, in determining the key target organ, consideration should be given to dose route, dosing vehicle, duration of exposure, and species differences in anatomy, metabolism, distribution, or sensitivity where data allow.

With respect to cancer as a target organ endpoint, estimates of potency are generally restricted to tumours that have been found to have a statistically significant increase over controls and indicate a dose-response relationship. All of these tumours are then considered in the TRV development process. In some cases, benign and malignant tumours in the same organ are combined; in some cases, these are modelled separately. The natural history of the tumour development must be considered. It is also noted that in some cases laboratory animals may not be considered to be adequately representative of tumour response in people (often because of anatomic or metabolic differences). For example, some chemicals that cause tumours in rats (e.g. peroxisomal proliferators) are not considered to have the same effect in people. Thus, when laboratory animal studies are used, the toxicologist must attempt to ensure that a representative animal model has been used.

B2.6.3 Selection of the critical animal study (or studies)

One or more key toxicological studies using the animal model that is most relevant to humans may be selected as the basis for TRV development. In some cases, the toxicologist may exclude certain studies that are not considered to be relevant or appropriate to the issue of concern. Selection of the most relevant animal model usually involves consideration of data from several animal studies that may involve various animal species. Accordingly, based on a defensible biological rationale, the toxicologist should identify the animal studies most relevant to humans. Considerations in selection of the appropriate study may include: (i) route and vehicle of administration, (ii) species/strain/sex of animals, (iii) tumour types or target organ response, (iv) number of animals evaluated, (v) statistical relevance (both power and significance), (vi) quality of the study, (vii) existence of a dose-response relationship, and (viii) any information on metabolism, toxicokinetics, and mechanism of action for animals relative to humans.

Ideally, the process of selecting a key study should maximize the biological correlations between animal species and humans. However, if no animal model is clearly considered to be the most relevant, it is then recommended that the most sensitive study (i.e. the study indicating toxicity at the lowest administered dose or concentration) be selected as the critical study, after carefully weighing the mechanistic evidence and evaluating the dose-response and quality of the study.

B2.6.4 Consideration of dose metric, dose route, and bioavailability/toxicokinetics

For systemic toxicants, the internal dose to the target tissue may not be directly related to the total body dose. However, the vast majority of dose-response data relate to the administered dose, not to target tissue dose, absorbed dose, or other measures whereby bioavailability and/or tissue distribution and toxicokinetics have been considered. Therefore, the dose metric for oral studies will usually be

based on the quantity of chemical ingested. For inhalation studies, the exposure metric will typically be the air-borne concentration to which the animals were exposed.

In some cases, the dose route used in the most suitable animal study may not be the same as the dose route through which humans may be exposed. In other cases, the dose route may be the same, but the rate of metabolism, type of metabolic conversion, excretion, and/or tissue distribution may be different for humans and experimental animals. Differences may also exist among animal species and among life stages in the same species. In cases where supporting data exist, it is important to consider these factors and to base TRV development on the dose measure most closely associated with the effect. Various issues surrounding route-to-route extrapolation are discussed in some of the guidance documents previously listed, and a good summary prepared by scientists from the U.S. EPA and Health Canada has been published (Pepelko and Withey, 1991). The Health Canada perspective on consideration of pharmacokinetics (toxicokinetics) and mechanism of action in dose extrapolation has also been described (Matula et al., 1994).

B2.6.5 Dose scaling to estimate human equivalent doses/concentrations

Health Canada considers it unnecessary to scale the dose when extrapolating from animals to humans except in the case of chemicals whose metabolism is tied directly to basal metabolic rate, in which case a dose-scaling factor of body weight normalized to the 3/4 power would be used (Health and Welfare Canada, 1991). The U.S. EPA (2002, 2005a) may also use dose scaling to estimate human equivalent doses or human equivalent concentrations. Essentially, for oral exposures, the U.S. EPA recommends that doses should be scaled from animals to humans on the basis of mg/kg $bw^{3/4}$/d unless toxicokinetic and toxicodynamic data are available to suggest otherwise. In other words, the mg of chemical administered to the animal is normalized by the 3/4 power of the body weight in some cases. In the case of inhalation exposures, the U.S. EPA (2002) has generally recommended that dose scaling is not necessary except in cases whereby the animal's respiratory physiology indicates appreciably less susceptibility than a human's respiratory tract.

B2.6.6 Adjustments for discontinuous exposure

For studies where the chemical is not administered on a continuous basis, it is usually recommended to adjust the effective dose for the reduced exposure that is occurring. For example, if a chemical in a chronic study was administered to animals via food at a rate of 70 mg/kg bw/d but only for 5 days per week, this exposure should be adjusted to a rate of 50 mg/kg bw/d to account for the daily exposure on a weekly basis (i.e. 70 mg/kg bw/d x 5 d/7 d = 50 mg/kg bw/d).

Similarly, if a group of animals was exposed to an air concentration equal to 60 mg/m^3 for 6 hours per day, 7 days per week, the adjusted air concentration would be 15 mg/m^3 when expressed on a daily basis (60 mg/m^3 x 6 hours/24 hours x 7 days/7 days = 15 mg/m^3).

With respect to studies that are less than a lifetime, care must be taken to ensure that the less-than-lifetime exposures are adequately predictive (and therefore protective) of the effects that may occur over greater durations. It is not uncommon to use less-than-lifetime occupational exposure to establish a chronic (lifetime) cancer potency value (slope factor or unit risk) or to establish a chronic TRV for threshold effects. However, care must be taken to avoid underestimation of actual effects.

Exposure amortization is an issue that requires sound scientific judgment and expertise in toxicology. Care must be taken to ensure that the approach used to amortize exposures to estimate dose levels is justifiable.

B2.6.7 Sensitive life stages

In developing TRVs, it is important that the toxicologist determine whether or not all sensitive life stages have been reasonably evaluated in the toxicological data set. Failing to account for all life stages may result in an underestimation of toxic potency. According to the U.S. EPA (2005a, 2005b) and other researchers (referenced therein), the cancer risk attributable to early-life exposure can be appreciably higher than exposures of similar duration that occur in adult life stages. Similarly, various non-cancer toxic effects have been noted during both prenatal and perinatal life stages that may not occur at later stages in life. Consequently, it is important that the toxicologist evaluate and quantify any uncertainties regarding sensitive life stages and the developed TRV.

B2.7 Calculation of Toxicological Reference Values for Threshold-Response Chemicals

For threshold-response chemicals whereby the route of exposure is oral or dermal, TRVs in Canada are most often expressed as the tolerable daily intake (or TDI). According to Health Canada (1994, p. 3; 1996, p. 4):

> "The Tolerable Daily Intake is the intake to which it is believed that a person can be exposed daily over a lifetime without deleterious effect."

In other words, it is the amount of exposure that is considered to be unlikely to cause adverse health effects in the general population, including sensitive individuals, but excluding those with allergy or other hypersensitivity. The TDI is, effectively, the best estimate of the human threshold dose, considering uncertainties and variability in intra-species

(inter-individual) toxic response, inter-species toxic response, and limitations of the toxicological database. TDIs are usually provided as daily dose rates in units of mass of chemical per kilogram of body weight of a person per day (e.g. μg/kg bw/d). Other terms that are analogous to the tolerable daily intake include acceptable daily intake (ADI) used by the World Health Organization, reference dose (RfD) used by the U.S. EPA, and minimum risk level (MRL) used by the ATSDR.

In the case of airborne chemicals, the TRV is usually expressed as a tolerable concentration (TC) by Health Canada or a reference concentration (RfC) by the U.S. EPA. The TC can be interpreted as the air concentration to which it is believed that a person can be exposed to on a continuous basis over a lifetime without deleterious effect.

Health Canada-derived TDIs and TCs are meant to protect the vast majority of members of the general public. To the maximum extent that is possible, it is important that TRVs developed for use in human health risk assessment offer similar scope and protection to members of the general public (including susceptible life stages such as the embryo, foetus, infant, the young child, and the elderly). In addition, the TRV should be derived to be protective of all types of adverse effects for the given duration of exposure. This is usually achieved by basing the TDI or TC on that toxicological study reporting biologically significant and/or statistically significant toxic response at the lowest doses or exposure levels. Where uncertainties exist, it is important that these are identified and accounted for in the TRV development process.

There are two basic approaches to TRV development for threshold-response chemicals that are acceptable to Health Canada. The first is the no observed effect level (NOEL)[5]/uncertainty factor approach; the second is the BMD/uncertainty factor approach. The BMD approach is discussed in section B3.1. For the NOEL/uncertainty approach, the NOEL is divided by a series of uncertainty factors to arrive at a dose level that is considered to be essentially without risk of causing adverse effects (see section B2.7.2 for discussion of application of uncertainty factors), considering variability in toxic response among individuals and among species, and accounting for other potential uncertainties.

In determining a NOEL, the key studies must be identified and all considerations made as previously discussed in section B2.6 with respect to inter-species and dose-route extrapolation. Considerations in selecting an appropriate NOEL are discussed in the following section.

[5] The no observed effect level (NOEL) is the highest dose in a toxicity study that results in no observed effects (HC, 1994).

B2.7.1 NOEL selection

The NOEL represents the highest level of exposure that has not been associated with observable effects in either humans or laboratory animals. More specifically, the NOEL is the highest dose at which there is no statistically or biologically significant indication of effects due to a chemical. It is usually expressed in units of mass of chemical per body weight per day (e.g. μg/kg bw/d). The NOEL is typically identified from the most appropriate study and for the most appropriate toxic endpoint. It is noted that in some cases a no observed adverse effect level (NOAEL)[6] may be identified in contrast to a NOEL. NOAEL is a term used to identify a level of exposure where no adverse effects were observed but it is possible that non-toxic or irrelevant effects occurred, whereas NOELs relate to any and all observed effects, irrespective of biological relevance or significance. Thus, in some cases, changes may be observed in animals that are not considered to be adverse. For example, a chemical that caused a change in enzyme levels or blood chemistry that is not considered to adversely threaten health may sometimes be considered as a non-adverse effect. Consequently, NOAELs are often considered to be acceptable for use in TRV development.

The data from all relevant toxicological studies are critically evaluated to identify whether they support the identification of a NOEL or NOAEL and the rationale for this selection must be documented.

In some cases, the data may not be sufficient to identify either a NOEL or NOAEL. In these cases, the lowest observed effect level (LOEL)[7] or the lowest observed adverse effect level (LOAEL)[8] may be considered. Use of a LOEL or LOAEL to estimate a TRV is generally not desirable because it implies no threshold has been identified; however, sometimes use of a LOEL or LOAEL is unavoidable on account of data set limitations. The LOEL or LOAEL should be used only when the adverse effects are considered to be reasonably non-severe, the LOEL or LOAEL of the selected key study is equal to or less than the NOAELs from valid

[6] The no observed adverse effect level (NOAEL) is the highest dose in a toxicity study that does not result in any observed adverse effect. An adverse effect is a change in morphology, physiology, growth, development, or lifespan of an organism that results in impairment in its capacity to compensate for additional stress or in an increase in its susceptibility to the harmful effects of other environmental influences (HC, 1994).

[7] The lowest observed effect level (LOEL) is the lowest dose in a toxicity study that results in an observed effect (usually one dosage level above the NOEL (HC, 1994).

[8] The lowest observed adverse effect level (LOAEL) is the lowest dose in a toxicity study that results in an observed adverse effect (usually one dosage level above the NOAEL (HC, 1994).

toxicological studies with other effects, and there is some confidence that the NOEL or NOAEL is likely within an order of magnitude of the LOEL or LOAEL. As discussed below, use of a LOEL or LOAEL mandates incorporation of an additional uncertainty factor to calculate the TRV (see section B2.7.2).

For inhalation exposure studies, it may be appropriate to express the potential toxic effects as an air concentration rather than as a dose rate. This avoids the potential errors and uncertainties inherent in converting the air concentrations to doses (including uncertainties in assumed inhalation rates, respiratory absorption rates, and body weights of the test animals). When a NOEL is expressed as an air concentration, it is termed a no observed effect concentration (NOEC). Use of a NOEC is most often appropriate when the route of exposure is inhalation and/or the toxicity assessment is concentration dependent rather than dose dependent (e.g. a local versus a systemic toxicant). This often requires the judgment of an experienced toxicologist to determine which units should be used to express toxic potential. The same general rules for development of a TRV apply whether or not the NOEL is expressed as an air concentration or a dose rate. In cases where the toxic effects of the chemical are related to internal dose and not directly related to the concentration of the chemical in the air, the NOEC should be converted to a dose to the test animal in units of mg/kg bw/d to assure that species differences in the breathing rate to body weight ratio are considered. In addition, if the NOEC/TRV approach does not account for protection of infants and young children, the TRVs for inhaled chemicals that act through a systemic mechanism should be expressed on a dose per unit body weight basis to ensure that the safety of infant and child receptors (with lower body weights and greater intakes per unit body weight than adults) are adequately protected.

B2.7.2 Application of uncertainty factors

Once the NOEL (or equivalent) has been identified, uncertainty factors are applied to derive the proposed TRV. The selection of uncertainty factors involves careful consideration of the toxicological data available. Although there are no true default uncertainty factors (uncertainty factors are determined on a case-by-case basis), it has become common practice to employ factors of 3 to 10 to address each of a series of unknowns or uncertainties in the toxicological data. In practice, Health Canada generally applies uncertainty factors as follows:

- typically a 10-fold factor to account for inter-species differences (i.e. When the NOEL or equivalent is based on laboratory animal data, Health Canada typically applies a 10-fold uncertainty factor.)
- typically a 10-fold factor to account for intra-species (inter-individual) differences (i.e. When the NOEL or equivalent is based on laboratory animal data or limited human data, Health Canada typically considers a 10-fold uncertainty factor to account for sensitive individuals within the general public.)
- typically a 3- to 10-fold factor to account for deficiencies in the toxicological data set deficiencies (e.g. lack of reproductive or developmental studies, lack of chronic studies, lack of identification of a NOAEL)
- an additional 3- to 10-fold factor may also be employed to account for the nature and severity of the potential toxic effects (e.g. concern over irreversible or life-threatening impacts)

Discussions of the basis for the Health Canada uncertainty factors (previously called safety factors) have been published (McColl, 1990; Kroes et al., 1993; Meek et al., 1994). It is important that the rationale provided for the selected uncertainty factor is clearly provided. Uncertainty factors exceeding 5,000-fold are generally not applied even if the total of the individual uncertainty factors exceeds this value.

B2.7.3 Calculation of tolerable daily intake

Once the NOEL and uncertainty factors have been determined, a TDI can then be calculated according to the following formula:

(B1)

$$TDI = \frac{\text{NOEL or NOAEL or LOEL or LOAEL}}{\text{Total Uncertainty Factor}}$$

For chemicals where it is appropriate to report the TRV as an air concentration, the TRV is termed “tolerable concentration” (TC) by Health Canada whereas the U.S. EPA prefers the term “reference concentration” (RfC). The TC is expressed in units of mass of chemical per unit air concentration (e.g. $\mu g/m^3$) and assumes a continuous lifetime exposure (i.e. 24 hours per day, 365 days per year for a 75-year lifetime). Similar in concept to the TDI, the TC is calculated as the NOEC or no observed adverse effect concentration (NOAEC) or lowest observed effect concentration (LOEC) or lowest observed adverse effect concentration (LOAEC) divided by the total uncertainty factor according to the following formula:

(B2)

$$TC = \frac{\text{NOEC or NOAEC or LOEC or LOAEC}}{\text{Total Uncertainty Factor}}$$

B2.8 Calculation of Toxicological Reference Values for Non-Threshold-Response Chemicals

For chemicals considered to be non-threshold-response chemicals (e.g. genotoxic carcinogens or germ cell mutagens) following a weight of evidence analysis, it is typically assumed that there is some probability of harm to human health at any level of exposure; thus, it is not possible to calculate a dose or concentration below which adverse effects are not expected to occur. For genotoxic carcinogens, TRVs are usually expressed as cancer potency factors that can include: (i) cancer slope factors, (ii) unit risks, or (iii) risk specific dose/concentration levels. According to Health Canada (1994), cancer potency factors are typically developed for chemicals that are classified as Group I (carcinogenic to humans) or Group II (probably carcinogenic to humans), and are considered to act via a genotoxic mechanism. Health Canada carcinogen risk assessment guidelines have been published (Health and Welfare Canada, 1991; HC, 1994).

As described in Health Canada (1994), for Group III chemicals (possibly carcinogenic to humans), cancer potencies are not generally derived. Instead, an additional uncertainty factor, to account for uncertainty in the potential for human carcinogenicity, is applied to establish an interim TDI or TC. Those interim TDIs/TCs are then re-evaluated at the earliest possible opportunity, as the database on carcinogenic bioassays and occupational and epidemiological studies expands.

Considerations for weight of evidence data evaluation, inter-species extrapolation, and dose scaling apply in cancer dose-response characterization as for threshold dose-response characterization and have been previously discussed in section B2.6.

In the past, when animal data were used, the linearized multi-stage model was widely used for estimation of the potency of chemicals whereas a multi-stage model with a linear term was often used to estimate the potency of chemicals when epidemiological information was used. The BMD approach is now preferred (HC, 1994, 1996; WHO, 1999; U.S. EPA, 2005a) with linear extrapolation to a specified risk level. The linearized multi-stage model or other dose-response extrapolation model may still be used to determine the BMD. The reader is referred to section B3.1 and the various guidance documents listed at the beginning of this report for specific guidance in cancer risk assessment and the use of the BMD approach.

B2.8.1 Development of Health Canada toxicological reference values for carcinogens

The names for BMDs or BMCs for carcinogens used by Health Canada are tumorigenic dose (TD_{05}) or tumorigenic concentration (TC_{05}), respectively (these should not be confused with TC, which is used to denote tolerable concentration for threshold chemicals in air). The TD_{05} and TC_{05} are doses or concentrations associated with a 5% increase in the incidence of tumours over control groups, based on either animal or human data; these are the counterparts to the ED_{05} and EC_{05} discussed in section B3.1 for threshold-acting chemicals. As noted in Health Canada (1994) and WHO (1999), any model that fits the empirical data well is likely to provide a reasonable estimate of the TD_{05} or TC_{05}. Health Canada (1994) has noted that the choice of the model may not be critical because the estimated value (TD_{05} or TC_{05}) is typically in the range of observed doses and, thus, removes many of the uncertainties associated with low-dose extrapolation. Health Canada does not consider it necessary to determine a lower confidence limit on the BMD, as does current U.S. EPA practice.

In the case of the TC_{05}, the toxicologist should ensure that the safety of children is adequately considered. Children weigh less and have respiratory rates that are relatively greater than adults on a per kg body weight basis and, as such, are exposed to higher internal doses of inhaled carcinogens at a given airborne concentration. TC_{05}s used to develop TRVs for the protection of adults may not protect children, and TC_{05}s used to develop TRVs to protect children may be overly conservative for use in adult-only risk assessments. For systemically acting carcinogens, converting a TD_{05} to an exposure dose and basing the risk assessment on this internal dose, rather than other exposure concentrations, provides some consideration for the differences in body weight and breathing rates of receptors in various age categories.

Once the TD_{05} or TC_{05} has been estimated, risk-specific doses or risk-specific concentrations can be estimated. Health Canada (1996) recommends linear extrapolation for derivation of risk-specific doses or risk-specific concentrations for protection at an incremental lifetime cancer risk (ILCR) of 1 x 10^{-5}. To illustrate this process, the following equations (B3) can be applied.

Because most risk assessments in Canada rely on potency factors (expressed as either slope factors or unit risks), potency factors may then need to be estimated. Potency factors are related to risk-specific doses and risk-specific concentrations according to the equations (B4) below.

In this manner, slope factors usually in units of $(\mu g/kg\ bw/d)^{-1}$ and unit risks usually in units of $(\mu g/m^3)^{-1}$) can be estimated for use in human health risk assessments.

(B3)

$$\textit{Risk-Specific Dose}\ (\text{for } 1\text{x } 10^{-5} \text{ cancer risk}) = \frac{TD_{05}}{5{,}000}$$

$$\textit{Risk-Specific Concentration}\ (\text{for } 1 \text{ x } 10^{-5} \text{ cancer risk}) = \frac{TC_{05}}{5{,}000}$$

(B4)

$$\textit{Slope Factor} = \frac{\textit{Risk Level}}{\textit{Risk-Specific Dose}} = \frac{1 \text{ x } 10^{-5}}{\textit{Risk-Specific Dose}}$$

$$\textit{Unit Risk} = \frac{\textit{Risk Level}}{\textit{Risk-Specific Concentration}} = \frac{1 \text{ x } 10^{-5}}{\textit{Risk-Specific Concentration}}$$

B3.0 OTHER APPROACHES TO DERIVATION OF TOXICOLOGICAL REFERENCE VALUES

B3.1 Benchmark Dose Determination

Where possible, Health Canada and the U.S. EPA are using the BMD approach for new assessments and for updated assessments (see HC, 1996; U.S. EPA, 2000, 2002, 2005a). In the BMD approach, the dose associated with a given response rate (usually 1%–10%) in a given toxicology data set is estimated. The response rate, called the benchmark response value (BMR), is usually set at 10% or less, depending on the sensitivity of the individual study under evaluation. Health Canada often employs a 5% response rate (HC, 1996); however, alternate response rates (i.e. up to 10%) may sometimes be used. The U.S. EPA is currently developing guidance for the criteria for BMR selection (U.S. EPA, 2000). The toxicologist should consider the issues related to BMR selection as outlined by the U.S. EPA (2000) to ensure that consideration is given to the possibility that a BMR other than 5% should be used. A BMR of 5% is discussed for the remainder of this document because, according to Health Canada precedents at the time of writing, the 5% BMR is considered appropriate for use in Canada (see HC, 1996).

Before a BMD can be determined, the same rigorous evaluation of the weight of evidence of the entire database bearing on the toxicological assessment of the chemical of concern must be undertaken as would be required for TRV development using other accepted methods. The weight of evidence assessment will determine the suitability of the available studies for use in risk assessment, the identification of key studies and relevant target organ endpoints, and the determination of the proper methods to use in extrapolation of test data to humans in the context of the exposure scenarios under consideration in the contaminated site risk assessment. Guidance for the various aspects of the weight of evidence assessment is given in detail in the various guidance documents listed at the beginning of this report. Key issues are briefly reviewed in the following sections.

Once the dose-response relationship has been characterized and appropriate dose scaling and extrapolation to humans has been conducted, the dose-response data can be used in the derivation of the BMD. Using a mathematical model, the response rate for a particular toxic effect (e.g. liver lesions, cancer, reproductive endpoints) is modelled to determine the ED or EC that elicits a 5% response rate (i.e. the ED_{05} or EC_{05}).

The BMD approach has an advantage over simply selecting the NOEL (in the case of threshold chemicals) or using mathematical models to extrapolate to doses far below the observable response range (in the case of non-threshold chemicals) because the 5% response rate is within the observed part of the dose-response relationship. In other words, interpolation within the observed data rather than extrapolation outside of the range of observations is used, thus providing greater statistical and practical confidence in the quantification of the metric (the BMD). Also, the BMD is associated with a known observed response rather than with a hypothetical non-observed response. The NOEL is outside of the observed dose-response relationship by an unknown amount, and the low-dose extrapolation models are not scientifically based in terms of the way they predict the shape of the dose-response curve at low doses. It is anticipated that there is less error and uncertainty inherent in the BMD because the BMD is based on extrapolation within the measured data range. The BMD modelling approach takes into consideration the number of experimental subjects and in doing so accounts for differing degrees of uncertainty associated with greater or lesser numbers of subjects.

Notwithstanding the above, the BMD approach may not be suitable for all data sets. In some (possibly many) cases, there may not be data suitable for dose-response analysis or curve fitting. There may be no models that fit the data, inadequate statistical information, or too few data points. In these cases, the toxicologist may need to consider the NOEL approach in the case of threshold chemicals. In the case of non-threshold chemicals, data inadequacies that would preclude BMD modelling would also preclude low-dose extrapolation modelling; in these cases, statistical modelling is not an option for TRV development. Evaluation of the data fit and a comparison of the NOEL with the predicted 5% response rate (the ED_{05} or EC_{05}) will help support a decision on which method to use. In some cases for threshold-response chemicals, there may be both BMDs and NOELs to be considered in developing a TRV. Different data sets may yield different BMDs, and scientific judgment is required in choosing the most appropriate approach. Guidance is given to help with this decision by the U.S. EPA (2002).

The appropriate dose-response model (curve) is chosen on the basis of goodness of fit to the observed data, the type of endpoint, and experimental and statistical considerations (U.S. EPA, 1995, 2000). Several models may be applied and the results compared. A combination or choice of the most appropriate results may be made. Improved data fit may be obtained by leaving out high doses that may be associated with cytotoxic or other acute effects unrelated to the chronic sublethal effects observed and of primary interest at lower doses. Improved data fit may also be achieved by using internal dose (requiring pharmacokinetic modelling) as a dose surrogate. There are no hard and fast rules for model selection; instead, scientific judgment is used to determine the most appropriate approach. Further guidance on model

selection and interpretation of results is provided by the U.S. EPA (2000).

In some cases, the BMD approach can allow consideration of data from more than one study rather than focusing upon only the most sensitive animal study. Data that are statistically and biologically compatible may be combined. This is considered advantageous because it may allow for greater overall confidence and statistical strength in the characterization of the dose-response relationship. The BMD approach also allows consideration to be given to uncertainties associated with variation among individual responders in the data set. The U.S. EPA (2000) has indicated that the BMD approach offers advantages over the traditional NOEL approach, in that (i) conclusions are not based on one single data point from one single study, (ii) it accounts for variability in dose-response estimates, (iii) it accounts for the slope of the dose-response curve, and (iv) it does not require identification of a NOEL. This latter point applies only when the lowest response rate is not significantly above the 5% response rate.

The drawback to the BMD approach is that it may be more complicated to apply than the traditional approach in the case of threshold chemicals, and data requirements are more rigorous. For example, the BMD approach can be used only if the response rate is given in terms of the number of subjects tested and if the response shows a trend in relation to dose. If confidence limits are required, then mean and variation (standard deviation, standard error or variance) must be available for each group.

In the case of the BMD approach, a 10-fold uncertainty factor is generally not applied in cases where there is no NOEL (i.e. effects occurred at all tested doses) because the uncertainty associated with low-dose extrapolation is implicit in the modelled results. Where the observed NOEL lies close to the actual NOEL, the TRV derived using the BMD/uncertainty factor or NOEL/uncertainty factor approach will be similar. In a series of comparisons of TRVs developed using the BMD approach compared with the NOAEL approach for developmental toxicity data sets, the TRVs were similar using the two methods (U.S. EPA, 1995).

Overall, the BMD approach/uncertainty factor approach is considered to be appropriate for development of TRVs in Canada where data permit. Information that can be found in U.S. EPA (2000) and the U.S. EPA benchmark dose website (http://cfpub.epa.gov/ncea/cfm/bnchmrk/new.cfm?ActType=default) are recommended sources of specific guidance in development of TRVs using the BMD approach. Once a BMD has been derived, different approaches are taken to the development of the final TRV, depending on whether the chemical acts through a threshold or non-threshold mechanism. These approaches are described in sections B2.7 and B2.8, respectively.

B3.2 Development of Toxicological Reference Values Using Data from Toxicological Surrogates

In certain cases where toxicological data are lacking, it may be appropriate to use data from toxicological surrogates. Although the use of toxicological surrogates is not a desired approach, sometimes it can be useful in addressing health issues. For a chemical to be evaluated using a toxicological surrogate approach, it first must be determined that there is an absence of appropriate chemical-specific data to estimate a TRV for the chemical of concern. If this is the case, it may sometimes be appropriate to estimate toxic potency from a chemical of similar structure. For example, PAHs with a common structure (e.g. a “bay” region) or halogenated aliphatics (e.g. chlorinated aliphatics versus brominated aliphatics) may represent possible toxicological surrogates. However, before selecting a toxicological surrogate, care must be taken to ensure that the surrogate is not likely to be more active than the chemical for which the toxicological data exist. If there is an indication that the surrogate is more active, appropriate modifying factors should be included based on sound scientific judgment and expertise.

B3.3 Development of Toxicological Reference Values from Worker Threshold Limit Values

It is generally unacceptable to use threshold limit values (TLVs) developed for workplace exposures as TRVs for the general public except under emergency or other extenuating circumstances. TLVs are based upon available information from industrial experience, human studies, and/or animal studies; when possible, a combination of all three may be used. The basis on which the values are established may vary from substance to substance, and a variety of factors may be considered. Although often based on toxicological principles, the procedures used to develop TLVs for protection of adult workers are quite distinct from the approach used for the general public. For example, TLVs generally assume that healthy adult workers are the persons requiring protection, whereas TRVs need to consider the protection of the health of all members of the general public, including young, pregnant, elderly, and sick receptors. In addition, TLVs do not offer the same level of protection against effects that may range from irritation to cancer. Finally, TLVs generally do not consider that continuous lifetime exposures occur.

Consequently, application of TLVs for protection of the general public is generally not considered acceptable. However, depending on the situation, the chemical, and the adverse health impact of concern, it may sometimes be necessary to extrapolate workplace information to derive TRVs for the general public. It is stressed that this is a case-

by-case issue, and no single factor to use in all circumstances is currently recommended by Health Canada or other national or international environmental health agency.

B3.4 Development of Toxicological Reference Values for Chemicals of Unknown Potency

For certain chemicals, literature searches may indicate that there is no appropriate toxicological data available to estimate the potency of a chemical. In addition, relevant toxicological surrogates may not exist. In these cases (and only these cases), it may be reasonable to use a TDI of 0.02 μg/kg bw/d for chemicals of unknown toxic potency. As summarized by Wilson et al. (2000), this TRV is likely to protect against chronic adverse health effects for chemicals of unknown potency whether or not they may be carcinogens or non-carcinogens. It is stressed that this TDI should be used only after it has been ascertained that no useful toxicological data are available. The most useful information for prediction of potential toxic effects is toxicological data specific to the chemical of concern. Nevertheless, if appropriate data are not available, this TDI may be a useful measure in extenuating circumstances.

B3.5 Development of Toxicological Reference Values Using Structure Activity Relationships

For some chemicals lacking specific toxicological data, there may be an appreciable amount of toxicological data available from structurally related analogues. Several models are available for estimating potential toxic potency based on structure activity relationships. These models usually do not represent a good source of information for quantifying TRVs; however, such models can sometimes be an important tool in qualifying the nature and potential severity of effects or providing additional support for a default TDI of 0.02 μg/kg bw/d (see section B3.4). The U.S. EPA (2005a) discusses some of the considerations in using structure activity relationships.

B4.0 REPORTING

It is vital that any proposed TRV developed for possible application at Canadian federal contaminated sites be presented to Health Canada at the earliest opportunity. The analysis and report must be accurate and thorough. Although reporting structures can vary, the TRV development report should include the following:

- identification of the substance
- verification that no regulatory agency has developed or proposed a TRV for the substance that is currently considered to be relevant or appropriate
- description and review of all of the key studies available for consideration (usually broken down into sections that include: (i) data from in vitro studies, (ii) data from short-term and long-term animal in vivo studies, and (iii) data from controlled human or epidemiological studies)
- all assumptions, rationale, and approach used to estimate the TRV
- discussion of the weight of evidence for threshold or non-threshold classification
- uncertainties and limitations to the data
- qualitative assessment of the effects that may be observed if the TRV was exceeded

B5.0 CONCLUSIONS

Development of TRVs for chemicals that have not been addressed by major health agencies is a complex task that should be completed only by experienced toxicologists. It is important that all of the relevant data are considered, and a weight of evidence approach is followed. For threshold-response chemicals, the approaches currently recommended are the BMD/uncertainty factor approach or the NOEL/uncertainty factor approach if data do not support the development of a BMD. For non-threshold-response chemicals, the BMD (5% response level) with application of a 5,000-fold factor to produce a TRV associated with an ILCR of 1×10^{-5} may be most appropriate. All assumptions and uncertainties must be well documented and err on the side of conservatism. Where possible, consultation with experienced Health Canada toxicologists is recommended.

Finally, it is noted that application of this guidance to the development of TRVs cannot be construed in any way as representing Health Canada acceptance or approval of any proposed TRV value. Complete documentation must be forwarded to Health Canada for internal critical review and evaluation, and for external review, if and as deemed necessary by Health Canada. Any decision from Health Canada, for or against the application of the proposed TRV, and the time required to complete the necessary review rests solely with this department.

B6.0 REFERENCES

Gerhanno, J.-F., C. Paris, B. Tiorion, and J.-F. Caillard. 1998. Assessment of bibliographic databases performance in information retrieval for occupational and environmental toxicology. *Occup. Environ. Med.* 55: 562–566.

Health and Welfare Canada. 1991. *Carcinogen Assessment*. A research report to the Department of National Health and Welfare. Cat H49-69/1991. ISBN 0-662-58750.

Health Canada (HC). 1994. *Human Health Risk Assessment for Priority Substances.* Chemical Evaluation Division, Health Canada, Ottawa. Cat En40-215/41E. ISBN 0-662-22126-5.

HC. 1995. Approach to the derivation of drinking water guidelines. Part I in: *Guidelines for Canadian Drinking Water Quality – Supporting Documents*. Health Canada, Ottawa.

HC. 1996. *Health-Based Tolerable Daily Intakes/Concentrations and Tumorigenic Doses/Concentrations for Priority Substances*. Chemical Evaluation Division, Health Canada, Ottawa.

HC. 2003. *A Framework for the Application of Precaution in Science-Based Decision Making About Risk.* Health Canada, Ottawa.

Kimbrough, R.D. 1995. Determining acceptable risks: Experimental and epidemiological issues. *Clin. Chem.* 40:1448–1453.

Kroes, R., I. Munro, and E. Poulsen. 1993. Workshop on the scientific evaluation of the safety factor for the acceptable daily intake (ADI): Editorial summary. *Food Addit. Contam*. 10: 269–273.

Lu, F.C., and M.L. Dourson. 1992. Safety/risk assessment of chemicals: Principles, procedures and examples. *J. Occup. Med. Toxicol.* 1: 321–335.

Matula, T.I., K. Voith, and C.A. Franklin. 1994. Toxicokinetics – A regulatory viewpoint in Canada. *Drug Info. J.* 28: 159–162.

McColl, R.S. 1990. *Biological Safety Factors in Toxicological Risk Assessment*. Environmental Health Directorate, Health and Welfare Canada, Ottawa. Cat. H49-49/1990E, 90-EHD-154.

Meek, M.E., R. Newhook, R.G. Litelpo, and V.C. Armstrong. 1994. Approach to assessment of risk to human health for priority substances under the *Canadian Environmental Protection Act*. *Environ. Carcinogen Ecotoxicol. Rev.* C12(2): 105–134.

Pepelko, W.E., and J.R. Withey. 1991. Methods for route-to-route extrapolation of dose. In: *Advances in Modern Environmental Toxicology*. Volume XIX. Health Hazards and Risks from Exposure to Complex Mixtures and Air Toxic Chemicals (pp. 179–194). Melman, M.A. (ed.). Princeton Scientific Publishing Co Inc., Princeton, NJ.

United States Environmental Protection Agency (U.S. EPA). 1991. *Guidelines for Developmental Toxicity Risk Assessment*. U.S. EPA, Washington, DC. Federal Register 56(234):63798-63826.

U.S. EPA. 1995. *The Use of the Benchmark Dose Approach in Risk Assessment*. U.S. EPA, Washington, DC. EPA/630/R-94/007.

U.S. EPA. 1996. *Guidelines for Reproductive Toxicity Risk Assessment*. U.S. EPA, Washington, DC. Federal Register 61(212):56274-56322.

U.S. EPA. 1998. *Guidelines for Neurotoxicity Risk Assessment*. U.S. EPA, Washington, DC. Federal Register 63(93):26926-2954.

U.S. EPA. 2000. *Benchmark Dose Technical Guidance Document (External Review Draft).* Risk Assessment Forum, U.S. EPA, Washington, DC. EPA/630/R-00/001.

U.S. EPA. 2002. *A Review of the Reference Dose and Reference Concentration Processes.* Risk Assessment Forum, U.S. EPA, Washington, DC. EPA/630/P-02/002F.

U.S. EPA. 2005a. *Guidelines for Carcinogen Risk Assessment (Final).* Risk Assessment Forum, U.S. EPA, Washington, DC. EPA/630/P-03/001F.

U.S. EPA. 2005b. *Supplemental Guidance for Assessing Cancer Susceptibility from Early-Life Exposure to Carcinogens.* Risk Assessment Forum, U.S. EPA, Washington, DC, EPA/630/R-03/003F.

World Health Organization (WHO). 1999. *Environmental Health Criteria 210: Principles for the Assessment of Risks to Human Health from Exposure to Chemicals*. WHO, Geneva.

WHO. 2001. *Environmental Health Criteria 225: Principles for Evaluating Health Risks to Reproduction Associated with Exposure to Chemicals*. WHO, Geneva.

Wilson, R.M., E.A. Sigal, C.M. Baciagalupo, R.F. Willes, and I.C. Munro. 2000. Derivation of risk management criteria for chemicals of unknown toxic potency at contaminated sites. *HERA* 6(1): 131–139.