Golder Associates Ltd.

195 Pemberton Avenue
North Vancouver, British Columbia Canada V7P 2R4
Telephone 604-986-4331
Fax 604-662-8548

FINAL

GUIDANCE FOR DETAILED ECOLOGICAL RISK ASSESSMENTS (DERA) IN BRITISH COLUMBIA

Submitted to:

Science Advisory Board for
Contaminated Sites in British Columbia

DISTRIBUTION:

1 Copy - Science Advisory Board (electronic PDF)
1 Copy - Golder Associates Ltd.

March 31, 2006 05-1421-040

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

APPENDIX I – DIRECT MEASUREMENT TOOLS

APPENDIX II – MODELLING TOOLS

APPENDIX III – INTERPRETIVE TOOLS

LIST OF ACRONYMS

BCMELP	British Columbia Ministry of Environment, Lands and Parks (now BCMOE)
BCMOE	British Columbia Ministry of Environment
CCME	Canadian Council of Ministers of the Environment
COPC	Contaminant of potential concern
CSR	Contaminated Sites Regulation
CWS	Canadian Wildlife Service
DERA	Detailed ecological risk assessment
DFO	Fisheries and Oceans Canada
DSI	Detailed site investigation
ERA	Ecological risk assessment
HQ	Hazard quotient
LOE	Line of evidence
LOEC/LOAEL	Lowest observed effect concentration/lowest observed adverse effect level
NOEC/NOAEL	No observed effect concentration/no observed adverse effect level
PF	Problem formulation
PSI	Preliminary site investigation
QA/QC	Quality assurance/quality control
ROPC	Receptor of potential concern
SAB	Science Advisory Board
SAP	Sampling and analysis plan
SLRA	Screening level [ecological] risk assessment
SSD	Species sensitivity distribution
TRV	Toxicity reference value
UF	Uncertainty factor
USEPA	United States Environmental Protection Agency
WOE	Weight of evidence

1.0 INTRODUCTION

1.1 Purpose of DERA

Detailed ecological risk assessment (DERA) for a contaminated site provides a framework for the assessment of risks to non-human organisms associated with environmental stressors, within the context of that specific site. DERA guidance applies to sites that cannot be addressed through simple risk screening procedures and that require the application of more sophisticated assessment methods to evaluate environmental risk. DERA guidance is also intended to provide the greatest degree of uncertainty reduction per unit cost, such that DERAs are practical and feasible, while also satisfying regulatory requirements. ***DERAs should strive to be as realistic as possible, replacing "conservative" assumptions with best estimates of exposure, effects and associated uncertainties (Dearfield et al. 2005).***

This DERA framework facilitates a transparent and consistent approach in terms of both assessing and prioritizing risks. Although each site is unique, and the tools and approaches applied (including the use of professional judgment) will vary among sites, it is desirable to maximize consistency among DERAs. The role of this DERA guidance is to streamline the following aspects:

- **Completeness:** DERA guidance increases our confidence that key risk components are not ignored (e.g., relevant pathways, stressors, receptors are all considered);

- **Relevance:** DERA guidance increases our confidence that the tools chosen are biologically relevant to the environmental values of interest;

- **Compliance:** DERA guidance increases the probability that the risk assessment deliverable will satisfy pertinent regulatory requirements and receive positive review by interested parties; and

- **Consistency:** DERA guidance, while not prescriptive, will encourage practitioners to address a common set of risk issues, and to provide rationales for the approaches selected. Identification of key decision points and provision of guidance on commonly encountered risk assessment challenges will also facilitate third party reviews of DERAs.

Risk assessment is a decision-making tool; it should proceed only to the point that an informed risk management decision can be made. The underlying role of DERA is to allow responsible parties (site owners and regulators) to make informed site management decisions. At many sites, risk assessment and remediation are applied concurrently. Site managers may opt to remediate the site (or parts of it) at any point in the risk assessment

process. Site managers contemplating DERAs also face differences in management settings (ranging from voluntary site assessments to remedial orders), data availability, time constraints, and other limiting factors. The challenge of DERA is to facilitate consistency in practice while also allowing the process to be customized to site-specific factors. In this respect, DERA is best viewed as process rather than a cookbook.

1.2 Linkage to Other Provincial Guidance

DERAs for contaminated sites in British Columbia are part of a larger regulatory framework that includes site characterization, remediation, and/or risk management. Often, the objective is the award of a regulatory instrument (e.g., a certificate of compliance) once the DERA has been completed and approved. Although risk assessment is specified as a tool for contaminated site management under the Contaminated Sites Regulation (CSR; a regulation under the Environmental Management Act), it can also be applied outside the purview of the CSR. The following provincial risk assessment guidance documents exist for certain types of ecological risk assessments (ERAs) conducted for contaminated sites.

SLRA Guidance: The Science Advisory Board (SAB 2005) recommended guidance for a screening level risk assessment (SLRA) methodology in BC[1]. The guidance was developed as a "streamlined risk assessment procedure to identify sites where substances exist above the numeric standards, but do not represent an unacceptable risk due to the absence of operable pathways of exposure to receptors." The SLRA guidance identified a two tiered screening process (SLRA Level 1 and SLRA Level 2) that contains simple and prescriptive rules and models for screening potential exposure pathways. As indicated in Figure 1, SLRA Level 1 and SLRA Level 2 are intended to apply to a large proportion of sites but are not appropriate for complex sites requiring detailed analysis. Following screening, many sites will be deemed as having "no pathway and/or no receptor" whereas others will require implementation of a DERA.

Tier-1 Guidance: BCMELP (1997) developed a checklist-based approach for a Tier-1 assessment of ecological risks at contaminated sites in BC. Tier-1 guidance contains many elements of a detailed ERA; however, the guidance is not sufficient for sites with relatively high complexity. Tier-1 guidance is applicable to a subset of ecological risk assessments where the prescriptive checklist-based approach is well suited to commonly observed contaminated site conditions. In contrast, the DERA guidance provided in this document is less prescriptive, requires additional expertise and professional judgment, but is applicable to a broader range of sites. DERA guidance emphasizes general concepts that are applicable to all sites and that should always be considered by ERA practitioners. This DERA guidance also provides additional guidance for specific

[1] For the purpose of this guidance manual, we assume that a provincial screening-level ERA guidance manual based on SAB (2005) will be available for use.

elements of ERAs (e.g., role of toxicity testing) that are not well developed in Tier-1. Figure 1 shows the relationship of Tier-1 guidance to both the proposed DERA guidance and the SLRA guidance. We understand that the Ministry intends that both Tier-1 and DERA guidance will be available, and therefore, this DERA guidance is meant to be compatible with the existing ERA guidance where appropriate.

In terms of a hierarchy of risk assessment guidance for British Columbia, we envision that the DERA problem formulation guidance described in this document should be followed for all sites for which a detailed risk assessment is contemplated. As part of this problem formulation process, the risk assessor will determine whether Tier-1 guidance is sufficient and appropriate, or alternatively that the breadth, depth, or complexity of the ERA requires the use of DERA guidance.

Potential triggers for conducting a DERA (according to SAB 2005) are summarized in Figure 2; in general, DERAs are initiated as a result of one of the following scenarios:

- A SLRA was conducted for the site; however, one or more exposure pathways indicated potential risk that could not be eliminated from consideration. In these cases, DERA focuses on only those remaining potential risks.

- A SLRA or Tier-I risk assessment was conducted for the site, but was incomplete, contained unacceptable uncertainty, or was otherwise deficient. In these cases, DERA focuses on issues that require reconsideration and/or more detailed analysis.

- No previous risk assessments have been conducted. However, the site contains one or more triggers that require a DERA (e.g., site contains contaminated sediment) and/or the site investigations suggest that a SLRA is unlikely to bring closure to the risk issues at the site. Professional judgement is often required to determine whether a screening level risk assessment would help frame the risk issues at the site, or alternatively whether a DERA should be initiated directly.

Irrespective of the scenario that triggers the assessment, all DERAs must identify all remaining relevant risk issues during the problem formulation phase, and systematically eliminate exposure pathways with negligible risk in order to focus on those remaining issues that require detailed analysis. The size, complexity and cost of the DERA are determined by the number of issues that could not be eliminated from consideration during a SLRA, as well as the desired level of certainty in the risk estimates relative to site management goals.

1.3 Scope of Document

This guidance document was prepared using the following guiding principles.

Guidance Manual versus Code of Practice: This document is not intended to be a formal code of practice for conducting detailed ecological risk assessments in British Columbia. It is a technical guidance manual, and as such, does not contain guidance on policy-related content that would be likely be required in a future Ministry guidance manual (e.g., legal and formal regulatory reporting requirements for a DERA). We have acknowledged existing Ministry policies where applicable and appropriate; such policies are discussed only to the extent to which they interact with the scientific positions within this document.

Level of Prescription and Detail: The emphasis of this document is on higher level guidance to promote consistency among practitioners in terms of the "key issues". The guidance is intended to be both flexible (to accommodate the range of possible site conditions and evaluation methods) and sufficiently focused (to ensure that practitioners consistently follow general procedures and thought processes). DERAs are sufficiently complex that they cannot be reduced to a standard "cookbook" that can be universally applied at all sites. However, there are many areas in which the risk assessment process can be harmonized between sites by focusing on similar risk questions. We understand that SAB may be asked to develop additional prescriptive guidance for commonly applied risk tools at a later date; therefore, this document was designed in a "modular" format to facilitate future additions. Additionally, certain aspects of specific risk assessment techniques were explored in greater detail (e.g., the role of toxicity testing) in the current document given the immediate need to supplement the existing provincial risk assessment guidance in high-priority areas (notably aquatic assessment, sediment assessment, and wildlands).

Intended Audience: This document was prepared for experienced risk assessment practitioners. Readers are assumed to be well-versed in risk assessment terminology and concepts; additionally, we assume that they have familiarity and experience with the practical application of ERA in British Columbia (including existing risk assessment policies as established in BCMELP 1997, 2000) and elsewhere.

1.4 Document Hierarchy

This DERA guidance manual is organized in a hierarchical manner; the broadest hierarchy is based on the traditional framework for ERA (i.e., problem formulation, exposure assessment, effects assessment, and risk characterization). In this respect, the document follows the structure of the Tier-1 guidance document. The problem formulation module (Section 2.0) must be completed first to identify the ecosystem types

and land uses that are relevant for a particular site. Following completion of the problem formulation module, the risk assessor should then refer to relevant sections of the exposure and effects assessments (Sections 3.0 and 4.0) and the appendices as applicable to the ecosystem types selected. Risk characterization for all sites should reflect the guidance presented in Section 5.0.

The secondary hierarchy is organized around five broad ecosystem types (i.e., deep aquatic, shoreline, upland wildlands, rivers and streams, upland human-use) (Figure 3). Exposure pathways, measurement endpoints, and risk assessment tools tend to differ considerably among these generic ecosystems. For some sites, multiple ecosystems may be present; however, it is rare that all ecosystems will be applicable at a single site.

Throughout the document, key issues and content have been summarized as follows:

Key Issues for the DERA Practitioner:

- Highlights major information needs or "state-of-the-science" issues that may be encountered during the risk assessment process. Guidance manual does not provide prescriptive guidance on these topics. Risk assessor should determine if issue is appropriate on a site-specific basis and proceed accordingly. Some "state-of-the-science" issues in the DERA should be considered only to the extent that they are needed to develop risk estimates that support appropriate site management planning.

Content for the DERA:

- Highlights specific items for inclusion or consideration in the DERA document.

1.5 The DERA Toolbox

Risk assessments consist of different "tools" selected to meet the needs of the project; the number and complexity of the tools reflects the level of detail required in the risk assessment. A SLRA might be based on a limited number of relatively straight-forward tools, whereas DERA often require multiple tools of higher complexity in order to evaluate all exposure pathways at the desired level of uncertainty. There are four categories of DERA "tools", spanning the range of DERA inputs from raw data to high-level interpretative methods, as follows.

1. **Raw Materials**: This category consists of direct measurements that contribute raw data to support the exposure and/or effects assessment.

2. **Modelling Tools**: This category consists of quantitative methods used to: (a) provide estimates of exposure and/or effects where field data are unavailable (i.e., surrogates

for direct measurements); or (b) simulate the fate, bioavailability, or toxicity of stressors in the environment, using field data as inputs.

3. **Interpretative Tools**: This category includes techniques used to evaluate the ecological significance of specific raw materials and/or model tools. Interpretative tools provide the linkage between data and information; they are usually (but not always) quantitative, and provide information targeted to the evaluation of the measurement endpoints.

4. **Synthesis Tools**: This category includes techniques used to integrate findings from multiple interpretative tools. These tools are applied mainly during the risk characterization phase. Synthesis tools include techniques for weight-of-evidence evaluation (WOE) as well as approaches for assessing uncertainty. Synthesis tools can be either qualitative or quantitative.

Many tools are highly modular (i.e., they remain relatively consistent in terms of implementation and interpretation irrespective of the site at which they are applied). Others are more specialized. Discussion of risk assessment tools is incorporated into the throughout the document. The appendices also provide a summary of different tools from the first three categories, organized using the DERA toolbox described above; a discussion of different synthesis tools is provided in the risk characterization section of this document. This modular format is intended to: (a) facilitate a document format that could be readily updated with additional tools; (b) facilitate future expansion of the technical guidance for some tools (i.e., increased level of detail); and, (c) minimize redundancy within the DERA manual and streamline the organization of tool descriptions.

1.6 Incorporating Land Use

Existing provincial guidance for contaminated sites is strongly influenced by land use. Numerical CSR soil and water standards for use at contaminated sites are organized by land use exclusively, while the Tier-1 guidance manual is organized by the risk assessment paradigm, with land use as the secondary level of organization. Ministry policy regarding permissible levels of impact are also land-use dependent. Land use also largely dictates the outcome of screening-level ERAs following guidance from SAB (2005). Examples of where provincial risk assessment policies are dictated by land use considerations are identified throughout this document.

However, DERAs should be influenced primarily by the ecology of the site rather than land use. This is not to say that land use should not be considered—in fact, the degree of anthropogenic disturbance at a site (as expressed through land development) is a primary factor in terms of habitat availability and quality, which in turn influences selection of

receptors of potential concern (ROPCs) and exposure pathways. ***This is a subtle, yet highly significant, difference from the existing Tier-1 guidance which excludes ROPCs based only on land use without this explicit consideration of how the actual land use of the site causes alteration to the site ecology.*** In the majority of instances, this subtle difference is unlikely to result in major changes in ROPC and exposure pathway selection, since land use classification is relatively straightforward where the site is fully developed and sits within a landscape of other, fully developed properties (e.g., an industrial site within an industrial park; other properties in urbanized areas).

However, land use implications for the design of a DERA are less clear when sites are either partially undeveloped or situated in wildlands (e.g., an undeveloped area zoned for residential, but unlikely to be developed in the near future; a decommissioned mine operation; industrial lands in an otherwise wildlands setting). Conducting a DERA based only on land use is also problematic when the context of the surrounding landscape is considered (e.g., a commercial property in a rural area surrounded by natural areas). Fixed divisions based on a limited number of land uses do not necessarily capture the gradient of anthropogenic influences on ecosystems. Additionally, federal guidance makes no distinction in terms of the level of protection afforded aquatic organisms, irrespective of the surrounding land use. Provincial sediment quality assessment guidance notes that "differences in land use activities do not influence the importance of sediments of benthic organisms" (Macfarlane et al. 2003).

The difficulty in assigning all conceivable sites to discrete land use categories has been noted by members of the SAB task group and others. Examples of where considering land use alone would be inapprorpriate for a DERA include:

- Remote wilderness areas can be highly disturbed by linear industrial developments such as survey lines and pipeline right-of-ways;

- Large contaminated sites, by virtue of their size, can contain very important habitat features; and

- Campgrounds and RV parks are commercial operations that have a residential-level of impact in an otherwise wildlands surrounding.

2.0 PROBLEM FORMULATION

2.1 Introduction

The problem formulation (PF) is the most important phase of any risk assessment. Consideration of problem formulation elements described in this section must be: (a) completed at all sites, including those that have undergone screening-level ERAs; and, (b) completed before exposure and effects assessments have been implemented. The level of effort required for the problem formulation is dependent on the complexity of the site (i.e., a site that is captured by SLRA-1 requires less effort than a site that requires DERA). The guidance provided below is intended for application at all sites for which SLRA guidance is not applicable. Although the problem formulation guidance is organized in a sequential manner, problem formulations are not linear in construction; they often require simultaneous consideration of multiple steps and may entail iterative refinements as site knowledge is obtained.

2.1.1 Problem Formulation Definition

The problem formulation phase is a planning and scoping process that defines the feasibility, scope, and objectives for the risk assessment and provides an opportunity for consensus building. This process includes examination of scientific data and data needs, regulatory issues, and site-specific factors. The problem formulation identifies the ecosystems potentially at risk, the stressors, and the measurement and assessment endpoints. This information is summarized in a conceptual model, which hypothesizes how the stressor(s) might affect the ecological components (i.e., the individuals, populations, communities, or ecosystems of concern). Problem formulations have been defined elsewhere as follows:

- "Problem formulation is a process for generating and evaluating preliminary hypotheses about why ecological effects have occurred, or may occur, from human activities. It provides the foundation for the entire ecological risk assessment" (USEPA 1998);

- The problem formulation "documents the key issues, establishes the breadth and depth of the problem, and initiates the process of prioritization… it documents the background for the decision to conduct an ERA" (CCME 1996); and

- "Problem formulation is a process of defining the nature of the problem to be solved and specifying the risk assessment needed to solve the problem" (Suter et al. 2000).

A review of ERA case studies concluded that the majority of difficulties documented in the case studies might have been avoided had more attention been paid to the problem formulation stage of the ERA (USEPA 1993a).

2.1.2 Why are Problem Formulations Important?

A well-constructed problem formulation reduces the likelihood of the following "fatal flaws" in the resulting DERA:

- **Incompleteness:** Risk assessor incorrectly excludes pathways, receptors, contaminants, or analyses that are required to produce a defensible ERA.

- **Incorrect Study Framework or Evidence of Study Bias:** Risk assessor chooses ERA methods based on what is readily available, personal and/or professional experience, or anticipated outcomes and then tries to build a risk assessment framework around them. In these cases, problem formulations appear (improperly) as window dressing around the technical contents of the ERAs, rather than as a means of guiding the scope and objectives of the ERA

- **Inconsistency or Lack of Objectivity**: Risk assessor develops appropriate conceptual model and study endpoints during the problem formulation, but fails to follow through in an objective manner in subsequent ERA phases (e.g., cherry-picking of effects metrics). Major decisions about how to interpret the data are made *post hoc* or without concurrence from interested parties. In these situations, the measurement and assessment endpoints are poorly aligned, and therefore key issues identified during problem formulation remain unaddressed.

- **Lack of Transparency**: Risk assessor does not provide sufficient rationale for methods, interpretations, or conclusions, as required for external reviews or project audits.

- **Technical Error**: Risk assessor chooses or applies a technical tool incorrectly, or interprets results in a manner inconsistent with the science.

2.2 Step PF-1: Planning Phase

Most existing ERA guidance emphasizes the scientific aspects of risk assessment; however, there are a number of non-scientific risk management factors that can influence the nature of an ERA, including environmental policy considerations, management constraints (e.g., project timelines), and the interests of other parties. Risk assessment should not be conducted in a vacuum from risk management. Rather, the role of risk management issues should be explicitly addressed in a transparent manner during the

problem formulation phase, rather than deferred to the exposure and effects assessments. The respective roles of science and policy should be clear within the document. The following subsections summarize risk management issues for consideration in the DERA problem formulation.

2.2.1 Definition of Management Goals

USEPA (1993a) concluded risk assessments were frequently deficient in their articulation of management goals. Management goals, within the risk assessment framework, are defined as "desired characteristics of ecological values that the public wants to protect" (USEPA 1998). This definition often results in vague narrative statements (e.g., "protect ecosystem integrity") that provide little meaningful direction to a contaminated site ERA.

The purpose of management goals is to act as a practical statement regarding the objectives of the ERA with respect to site management. The term "management goals" is used for consistency with other guidance manuals; however, irrespective of the precise term used[2], management goals should not be defined by the risk assessor in isolation. Rather, they should be defined in collaboration with the risk manager relative to business objectives and the applicable regulatory requirements. The difference between the roles of the risk assessor and risk manager is as follows:

- The risk manager serves as the primary decision maker for a site; he/she uses the result of the risk assessment along with information on technical feasibility and social, economic and political concerns to reach a decision regarding the need for and scale of any management actions (such as remediation) (CCME 1996).

- The risk assessor is responsible for the design and implementation of an ERA that meets the overall management goals for the site.

Risk managers and risk assessors are described as separate parties by other risk assessment guidance manuals; however, in practice, this separation rarely exists. For example, a client-consultant relationship is more common for contaminated site ERAs in British Columbia. Clients and site owners frequently request input from the risk assessor on issues such as technical feasibility and typical regulatory concerns; risk assessments are often bundled within a larger site management or remedial action plan document. A single risk manager is also unlikely: risk manager responsibilities are spread across multiple parties, including site owners, lead consultants, regulatory agencies, or members of the risk assessment study team. In fact, separation of the role of risk manager and risk assessor need not always involve different parties[3], provided that ecological risk

[2] CCME (1996) uses the terms "objectives of the ERA" instead of "management goals"
[3] Assigning the roles of the risk manager and risk assessor to different members of the same study team may be possible in some situations.

estimates are clearly based on the science and not other issues. A transparent framework for how risk estimates were incorporated with other issues with respect to developing site management recommendations should also be provided where applicable.

Questions that risk assessors may ask people with risk manager responsibilities (especially clients) to help formulate practical management goals, include:

- Is the ERA intended to simply determine if an unacceptable risk is currently present? For former industrial sites with historical contamination, the present condition may be the worst-case condition. In other cases:

- Will it be necessary to develop a site-specific risk based standard for a particular contaminant that will be used for remediation?

- What is the range of future potential land uses for which the risk assessment is intended to be applied? How will future site development affect risk estimates?

- What is the desired level of certainty in the risk assessment conclusions? Risk assessments that are linked to compressed development schedules typically require greater certainty earlier in the ERA process because they are less amenable to tiered evaluations. The potential for residual liability may also influence the desired level of certainty.

Dialogue with risk managers (or client) regarding these topics at the beginning of the project is recommended, since it provides an opportunity for:

- The risk manager (or client) to communicate their expectations regarding the risk assessment process, which facilitates an understanding of budgetary and timing constraints as well as the nature of the relationship (if any) between the client and other interested parties; and

- The risk assessor to communicate the regulatory expectations and ecological considerations involved in a detailed ERA. The risk manager (or client) may not be aware of jurisdictional issues or of the need for the ERA to fully document the decision making process (rather than simply focusing on the perceived issues of importance).

Multiple management goals may be viable at the beginning of the risk assessment (i.e., *in situ* management, remediation to numerical standards, remediation to risk-based standards, or a combination of multiple approaches). . Two examples of management goals that guide the development of a risk assessment are provided for illustrative purposes:

- **Terrestrial:** "Determine whether the magnitude of soil contamination at the site requires remediation, or whether the magnitude of soil contamination is amenable to *in situ* management since risks to relevant receptors are found to be acceptable (with a high degree of certainty) for a future industrial land use."

- **Aquatic:** "Determine whether concentrations of COPCs present in the surficial layer of marine subtidal sediments represent an unacceptable risk to aquatic life within the existing provincial and federal regulatory frameworks."

Key Issues for the DERA Practitioner:

- Do I understand why the client is doing this risk assessment, and have these needs been incorporated in the risk assessment design?

- Is the client aware of the legal and regulatory constraints that apply to the site evaluation, and have these requirements/limitations been incorporated in the risk assessment design?

- Does the study design (e.g., level of tiering of study components) correspond to the project schedule, if timelines are a significant limiting factor for risk management?

2.2.2 Obtaining Input from Interested Parties

Most contaminated sites ERAs are conducted by consultants on behalf of property owners in a client-consultant relationship. This relationship defines the primary liaison in the design and implementation of the risk assessment. In addition to client input, risk assessments benefit from interactions with other interested parties, ranging from formal regulatory agency direction and/or advice, informal discussion, or public consultation. Potential interested parties include:

- Provincial regulatory agencies (e.g., BC Ministry of Environment [BCMOE] or their representative[4]);

- Federal regulatory agencies (e.g., Department of Fisheries and Oceans [DFO]; Environment Canada; Canadian Wildlife Service [CWS]);

- First Nations;[5]

[4] The proposed "Approved Professional" system, if implemented, may result in instances where "Ministry" consultation is obtained from an approved professional and/or Ministry staff.
[5] Consultation with First Nations is subject to an evolving legal landscape as well as ongoing government process.

- Administrative bodies (e.g., Vancouver Port Authority, Transport Canada, Fraser River Estuary Management Program, Burrard Inlet Environmental Review Committee);

- Other levels of government (e.g., municipal; regional); and

- Non-governmental organizations (e.g., environmental groups; local community organizations).

All DERAs involve liaison with one or more interested parties, although the magnitude and formality of these interactions tends to be commensurate with the size, scope, and complexity of the project. Not all interested parties are applicable to all sites. For many sites (but not all), it is often sufficient to only solicit input from applicable regulatory agencies (since they act as instruments of public policy). The context for determining the appropriate involvement of interested parties varies depending on the following.

Which jurisdictions are applicable to the site?

In British Columbia, environmental matters pertaining to contaminated sites generally fall under the jurisdiction of the provincial Ministry of Environment. Specific regulations relating to the assessment and remediation of contaminated sites include the CSR (BC Reg. 375/96, last amended in 2004), and the Hazardous Waste Regulation (formerly called the Special Waste Regulation; BC Reg. 63/88). The level of input from other interested parties is influenced by CSR provisions (e.g., public consultation [S55.1], and off-site notification [S57.1]).

Consideration of the federal perspective is also recommended, even for risk assessments conducted under provincial guidance. In practice, most sites require some consideration of federal policy, regulation or legislation through one or more of the following triggers (not an exhaustive list):

- Sites that contain or are adjacent to waterbodies that sustain (directly or indirectly) a fishery may trigger the *Fisheries Act* and thus require input from Environment Canada and/or DFO. Federal *Fisheries Act* Section 36(3) concerns the deposit or permitting the deposit of a deleterious substance into waters frequented by fish or in any place where deleterious substances may enter such waters. The *Fisheries Act* is also relevant to site management with respect to habitat alteration, disruption and destruction, especially the need to obtain Section 35(2) Authorization to cause habitat alteration, disruption and destruction.

- Sites with migratory birds may trigger the *Migratory Birds Act*, and thus require input from Environment Canada; and

- Sites with federally-listed rare or endangered species may trigger the *Species at Risk Act*, and thus require input from Environment Canada and the Canadian Wildlife Service.

A rationale for (or against) the inclusion of federal perspectives, emphasizing the site-specific information available to support the decision is useful. For those sites that appear to have a federal trigger, it may be sufficient to simply document how the federal perspective was accommodated through reference to existing federal policies and regulations. Formal dialogue with regulatory agencies is not mandatory for all sites but is recommended for those sites where management goals require federal regulatory approvals, or for those sites where a significant federal regulatory interest is likely to exist (e.g., the site contains sensitive and/or abundant migratory bird or salmonid fish habitat).

Key Issues for the DERA Practitioner:

- Identify the lead regulatory agency for the risk assessment (e.g., Ministry of Environment);
- Determine whether the site is sufficiently complex to warrant formal dialogue prior to preparation of the problem formulation;
- Does the risk assessment connect with other environmental regulatory issues at the site? If so, is formal liaison required to address these issues?
- What level of documentation will be necessary to solicit input from other interested parties? Examples include a stand-alone problem formulation for technical review, informal site visit, or a "briefing note" summary.

What are the desired project timelines?

Where possible, the project timeline should allow formal input from other interested parties (regulatory or otherwise). Formal input on a site often requires sufficient documentation and provision of a review/comment period. Site visits and kickoff meetings may provide a means to obtain informal input regarding the scope of the risk assessment. Ongoing informal dialogue is also advantageous.

Key Issues for the DERA Practitioner:

- Do project schedule constraints limit the level of interaction with other interested parties, including regulatory authorities?
- If formal review of the problem formulation is not feasible prior to conducting the remainder of the risk assessment, how can involvement of other interested parties be optimized?
- Is the client aware of the uncertainties associated with postponing regulatory interactions until later in the risk assessment process?

2.2.3 Assembling a Study Team

The complexity of the study team and degree of specialization required are project specific, although a multidisciplinary study team[6] is typically required. Not all of the scientists involved need be experienced in risk assessment, provided that an experienced risk assessor is involved in coordination and report preparation. The appropriate level of professional designations (e.g., R.P.Bio., P.Eng), academic credentials (B.Sc.; M.Sc.; Ph.D.), and documented expertise in a given discipline (or subdiscipline) for the study team should be considered.

Key Issues for the DERA Practitioner:

- Identify which specialties will likely be required to successfully complete the DERA, and where possible, involve those people in the preparation of the problem formulation.

[6] Examples include toxicology, ecology, fisheries/wildlife biology, botany, forestry, limnology, geology/hydrogeology, chemistry, environmental modeling, statisticians and geographic information specialists.

2.3 Step PF-2: Review Historical Documentation

The problem formulation provides an opportunity to consolidate and consider all relevant site information, including:

- Stage I and II preliminary site investigations (PSI);
- Detailed site investigations (DSI);
- Environmental impact assessments;
- Physical, chemical, and/or biological monitoring reports; and
- Previous ecological or human health risk assessments (screening-level or other).

PSIs and DSIs are often available for contaminated sites prior to initiation of the DERA. Other documents should be reviewed where available since biological data are often not incorporated in PSIs or DSIs. Biological data may be available in seemly-unrelated documents; for example, a baseline environmental assessment for a development project[7] dealing with regional or watershed-level information may contain relevant ecological and biological information applicable to a contaminated site within the watershed. Other biological data sources include the Burrard Inlet Environment Review Committee project archives; Ministry of Environment reports; and other multi-agency watershed level programs. Institutional libraries (e.g., regulatory agencies; universities) are also potential sources of information.

2.3.1 Review Previous Ecological Risk Assessments

All ecological risk assessments previously conducted for the site must be reviewed during the problem formulation. Several scenarios exist in this regard:

- A SLRA was completed following provincial risk assessment guidance which led to the initiation of the DERA. The risk assessor should review the SLRA in terms of its methodologies and conclusions and agree with its decisions regarding exclusions of receptors, pathways or contaminants from the DERA.

- An ERA was conducted for the site based on provincial Tier-1 or other ERA guidance. The risk assessor should determine which receptors, pathways or contaminants may be screened with confidence from further consideration.

- An SLRA was not completed (i.e., the screening ERA stages were skipped for efficiency). In these instances, the risk assessor is limited to the historical documentation described in Section 2.3, above.

[7] BC Environmental Assessment Office (http://www.eao.gov.bc.ca/).

Content for the DERA:

- A narrative or tabular summary of each previous risk assessment should be provided in terms of receptors, pathways, contaminants (and/or physical stressors), risk assessment tools used, major conclusions, areas of uncertainty and recommendations for future work.

- A summary statement for each previous risk assessment should be provided, indicating agreement with the conclusions (or, if disagreement, a rationale for that determination).

Key Issues for the DERA Practitioner:

- Does the available documentation provide sufficient information about the ecology of the site to support the selection of the exposure pathways and receptors of concern?

- Is a site visit and/or habitat characterization by a professional biologist necessary to confirm or supplement the available ecological information?

- Is the biological characterization of the site limited to the legal site boundaries, or does it include descriptions of habitats in adjacent land parcels?

2.3.2 Determine Applicable Ecosystem Type(s)

Site ecology is the primary factor to consider when developing, implementing and interpreting a detailed ERA. USEPA (1992) comments that "knowledge of the ecosystem[8] potentially at risk can help identify ecological components that may be affected and stress-ecosystem interactions relevant to developing exposure scenarios." The following generic ecosystem types were developed based on commonly observed and broad differences in the biotic communities and exposure pathways (Figure 3):

- **Deep Aquatic:** Deep aquatic ecosystems include subtidal marine areas and lake bottoms. These ecosystem types tend to have relatively stable sediments subject to deposition. Deep Aquatic ecosystems can be found in both freshwater and marine environments.

- **Shoreline:** Shoreline ecosystems include intertidal areas, shallow estuarine environments, wetlands, marshes, and rocky shorelines. These ecosystems typically reflect a dynamic and transitional environment (e.g., freshwater to marine; tidal changes). Groundwater flux from upland areas to the aquatic receiving environment is

[8] Ecosystem is defined as the biotic community and abiotic environment within a specified location in space and time (USEPA, 1998).

often an important exposure pathway for this ecosystem type. Shoreline ecosystems can be found in both freshwater and marine environments.

- **Rivers and Streams:** Freshwater environments with flowing water, often associated with more dynamic substrates.

- **Upland Terrestrial (Wildlands)**: Relatively natural terrestrial ecosystems with minimal direct anthropogenic influence. This ecosystem type can vary greatly in British Columbia (e.g., coastal rainforests; high alpine meadows; semi-arid; montane).

- **Upland Terrestrial (Human Use)**: Terrestrial ecosystems that are significantly influenced by human activities. The degree of anthropogenic influence is reflected by land use considerations. For this ecosystem type, the magnitude and type of human use influences both the ecological setting and the protection goals of the ERA. Land use types are organized based on the prevailing land use classifications specified in both the CSR and the Tier-1 guidance for ERA. The land use types of industrial, commercial, residential, urban park, and agricultural may be viewed as subtypes of the upland terrestrial ecosystem type.

These generic ecosystem types are provided as a starting point—combinations of multiple ecosystem types and transitional subtypes within a single site also exist. In some cases, these transitional ecosystem subtypes may be of significant interest in the DERA (e.g., a riparian setback surrounding a stream may require consideration of study components from each of the "shoreline", "rivers and streams" and "wildlands" ecosystem types). A site-specific conceptual model should incorporate relevant components of one or more of the generic ecosystems above as needed.

Content for the DERA:

- The risk assessor should determine which among the five generic ecosystem types (or transitional ecosystem types) are applicable to the site in terms of quantity and configuration of existing habitat. The proportion of the total site area in each category and proximity of site habitats to habitats on adjacent land parcels is important.

- A brief description of relevant meteorological data (e.g., seasonal trends; temperature ranges; rainfall) and the biogeoclimatic classification should be included since it provides context to the selection of ROPCs. The procedures described in the Tier-1 guidance are generally suitable for this purpose.

Key Issues for the DERA Practitioner:

- There may be insufficient information available to properly evaluate ecosystem types. Other sources of information (site visits; professional judgment based on relevant experience) may be necessary.
- If available, habitat mapping data should be used to supplement the characterization of ecosystem types. For example, habitat inventory and classification maps have been produced for the FREMP[9] that show classes of intertidal and riparian habitat types and rate their biological productivity and suitability for development. Provincial wildlife habitat classification guidance is also available.

2.3.3 Summarize Site History

Site history, with emphasis on historical site uses linked to use or distribution of contaminants, should be summarized in the problem formulation. Site history is generally considered in detail in a DSI; in these cases a brief review of the site history in the problem formulation will suffice. The review should consider:

- Historical subdivision or amalgamation of land parcels (i.e., is the study area made up of many smaller properties, or was the site subdivided from other historical lots?);
- Approximate locations of former buildings and site operations in relation to soil, sediment, water, and biota; and
- Historical activities on adjacent or nearby properties that may result in potential off-site contamination sources.

Information on historical site uses is primarily intended to allow the risk assessor to conduct a "reality check" on the adequacy of the available site information to support an ERA exposure assessment.

Content for the DERA:

- A narrative or tabular summary of site history, along with implications for the design of the DERA;
- Identification of site activities that may have altered the distribution or concentration of contaminants of potential concern (COPCs);

[9] FREMP (Fraser River Estuary Management Program). 2005. *Updating the FREMP Habitat Classifications*. Prepared by the Water and Land Use Committee, Fraser River Estuary Management Program (BIEAP-FREMP), Burnaby, BC. February 2005.

- Identification of COPCs that were not considered in previous site investigations; and
- Identification of regional contamination issues if applicable.

Key Issues for the DERA Practitioner:

- What is the potential for on-site and off-site migration of contaminants at concentrations of potential environmental concern?
- Is there site-specific information relevant to bioavailability and/or mobility of contaminants that is not reflected in bulk chemistry measurements? (For example, PAHs associated with black soot particles and metals associated with grit particles tend to be less bioavailable).
- Does existing information provide sufficient detail to develop a comprehensive list of COPCs?
- Is the pattern of site contamination linked to historical site uses?

2.3.4 Evaluate Applicable Land Use(s)

Current and potential future land use of the site is an additional factor to consider when developing, implementing and interpreting a detailed ERA. Land use governs the process used in screening-level ERAs that follow guidance from SAB (2005), and is dominant in the organization of the existing provincial Tier-1 guidance manual (BCMELP 1997). Land use classifications are particularly important for the uplands [human use] ecosystem type, since land use dictates specific ERA attributes, including level of protection for various receptor types.

Land use is a less significant factor in the design of DERA for the aquatic ecosystem types (deep aquatic, shoreline, or river and stream) (see Section 1.6). Incorporation of land use considerations in the uplands (terrestrial) ecosystem types may also be complicated when a particular site does not "fit" well into the local mosaic of land use types[10]. Consequently, site ecology should be the primary consideration in the design of a DERA. Land use is an important but secondary consideration which should not outweigh the ecological context of a site. For example, the Tier-1 ERA guidance excludes large

[10] Land use considerations are relatively straightforward in cases where the site is fully developed and is situated within a landscape of other, fully developed properties (e.g., an industrial site within an industrial park; other properties in urbanized areas). Land use implications for the design of a DERA are less clear when sites are either partially undeveloped or decomissioned (e.g.., an undeveloped area zoned for residential use, but unlikely to be developed in the near future; a disused industrial property along a river). Land use implications are also problematic when the context of the surrounding landscape is considered (e.g., a commercial property in a rural area surrounded by natural areas).

terrestrial mammals from consideration at industrial sites. Such exclusion is appropriate within the context of an urbanized setting, but inappropriate for small industrial sites surrounded by wildlands that support large mammals.

Specific examples of where provincial risk assessment policies are dictated by land use considerations are identified throughout this document.

Content for the DERA:

- A summary of the current (and likely future) land uses.
- Discussion of land uses beyond the legal boundaries of the site but relevant to mobile receptors that cross site boundaries (i.e., regional ecological setting).

Key Issues for the DERA Practitioner:

- Land use classifications based only on land-use zoning may be inadequate for evaluating the ecological attributes of a particular site (or subareas within a large site). Property boundaries are not the same as ecological boundaries.
- The context of the surrounding land uses should also be considered in terms of its implications. An industrial site bordering on sensitive and valued aquatic habitat (e.g., wetland) does not have the same ecological attributes as an industrial site bordered by other industrial sites.

2.3.5 Summarize Site Chemistry

A summary of the available site chemistry should be included in the problem formulation; it provides a basis for understanding the type and magnitude of contamination, and logically leads to the identification of COPCs (Section 2.4). The following summaries of site chemistry are generally required:

- A narrative or tabular summary of concentrations measured in the different environmental media sampled to date, including description of minimum and maximum concentrations, summary statistics (e.g., 95% upper confidence limit of the mean, 90^{th} percentile, mean, median), percentage of non-detects, and sample size. This site chemistry summary is typically included in the historical document review in order to demonstrate familiarity with previous site investigations, and to document the underlying trends in the available chemistry data.
- An Excel-based or database system containing the results of individual analyses. This data summary is used to identify COPCs and will typically include coordinates to facilitate map or GIS-based presentation. Depending on the site and complexity of the

site, compilation of the data in this format is recommended given their importance elsewhere in the problem formulation.

- A brief narrative describing the spatial and temporal variations in chemistry distributions should be provided, particularly as they relate to representativeness and sampling design for additional investigations.

QA/QC should be reviewed to determine if available site chemistry data are appropriate for the risk assessment. Issues include sample collection and storage methods, selection of analytical methods, performance of analytical QA/QC measures such as laboratory duplicates, matrix spikes and use of certified reference material, and the use of appropriate analytical detection limits. Data without detailed QA/QC documentation may be rejected or utilized (with appropriate discussion of its uncertainty) at the discretion of the risk assessor; however, a data set that consists primarily of unverified data indicates that confirmatory sampling as part of the DERA is likely warranted.

Content for the DERA:

- A narrative, tabular or graphical summary of the available chemistry data for each medium. This overview should be linked to site history and describe potential or suspected contaminant sources.

- A spreadsheet or database containing the individual analytical results for use in screening of COPCs and graphical presentation.

- A brief summary of the spatial distribution of chemistry parameters.

Key Issues for the DERA Practitioner:

- Analytical detection limits for site characterization samples should be reviewed for environmental relevance.

- Ancillary data needed to interpret bulk chemistry data (e.g., pH or hardness for metal concentrations) or facilitate other decision making within the DERA (e.g., grain size and total organic carbon data in sediment to facilitate toxicity test species selection) may not be available. These data gaps will need to be addressed as part of the DERA.

2.3.6 Site Overview Map

A site overview map should be prepared; such a figure is often taken from the DSI and modified as necessary. This map should include the following information:

- Legal site boundaries and identification of adjacent properties. Placement of the specific study area within a regional context (in a smaller map window) is recommended.

- Locations of historical site buildings, areas of potential concern (APECs), and zones of known contamination.

- Locations of individual historical sample locations.

- Locations of other relevant site features, such as transportation corridors, waterbodies, changes in topography and significant habitat features.

Content for the DERA:

- A geographical representation of the data presented in the problem formulation is strongly recommended.

Key Issues for the DERA Practitioner:

- Geographical representations facilitate examination of the adequacy of the existing spatial coverage of chemistry data relative to known or suspected contaminant sources as well as significant ecological features. Assessment of spatial coverage is supported by these geographic representations.

- GIS-based approaches facilitate the integration of data management and mapping, and are advantageous in terms of spatial analyses of chemistry [and other] data as well as risk communication.

2.4 Step PF-3: Identify Contaminants of Potential Concern (COPCs)

COPCs are selected primarily based on a comparison of the available site data to the applicable numerical guidelines, standards or criteria values[11]. The presence of one or more samples with a concentration that exceeds these numerical values results in the selection of that analyte as a COPC. However, analytical chemistry data may not be adequate for COPC selection. Professional judgment may be required to ensure that

[11] This section will use the term "guidelines" in lieu of "guidelines, standards and criteria".

potentially relevant COPCs are not excluded due to lack of data. In general, a COPC should be retained for further evaluation unless sufficient information is available to warrant its exclusion. Examples of how professional judgment should be applied in COPC screening are provided below.

Inadequate Chemistry Characterization: COPC selection requires that the site has been adequately characterized. Completion of a DSI is assumed to represent an adequate characterization in terms of spatial coverage; however, a DSI may still contain data gaps in terms of the adequacy of data relative to the specific exposure pathways. For example, if terrestrial exposure pathways are being evaluated for a site that will not be disturbed under its future land use, then COPC selection should be based primarily on surface soil conditions.[12] DSIs often contain chemistry data that may not be representative of ecologically relevant exposures, in part, due to one or more of the following factors:

- Soil data are primarily from depths greater than one meter or composite soil samples from a range of depths. Non-composite samples from the upper 15 cm are recommended where possible.

- Construction and/or remediation activities may result in a future surface soil horizon that is different than the surface soil characterized in the DSI.

- Sampling density in the DSI is not appropriate relative to the foraging ranges and preferred habitats of site receptors.

Analytical Detection Limits: Chemistry data may have analytical detection limits that exceed the applicable numerical guidelines. Compounds that have analytical detection limits greater than guideline values should be retained as COPCs until confirmatory analyses with appropriate analytical detection limits can be conducted. Tier-1 provincial risk assessment guidance suggests that analytical detection limits should be less than the numerical guidelines by a factor of 10, subject to technical considerations (BCMELP 1997).

Numerical Guideline Value Unavailable: COPCs should not be prematurely excluded based on a lack of CSR standards[13]. If CSR standards are not available, provincial ambient guidelines, numerical guidelines from other jurisdictions (e.g., Canadian Council of Ministers of the Environment [CCME], United States Environmental Protection Agency [USEPA], Washington Department of Ecology), or toxicity reference values from the literature can be adopted. The degree to which the derivation procedures reflect the protection goals of the provincial CSR standards should be considered.

[12] Tier-1 guidance defines the plant root zone as the upper 15 cm (BCMELP 1997); however, in practice, data for the top meter are considered surficial

[13] Note that CSR Schedule 10 requires consideration of potential ecological effects for listed compounds present at elevated concentrations.

Anthropogenic compounds present at quantifiable concentrations but without environmental quality guidelines should be retained as COPCs unless a sufficient technical argument can be made for their exclusion. Potential technical arguments include:

- Some COPCs can be eliminated from consideration for certain pathways based on environment fate properties. For example, volatile organic compounds may be screened out of a food-web bioaccumulation pathway, because these chemicals rarely accumulate in organism tissues at levels of environmental concern. Organic compounds with high Henry's Law Constant values (H) means they readily partition to air, while compounds with low K_{OW} values means they tend to be highly water soluble (and therefore readily excreted).

- Some COPCs can be eliminated from quantitative evaluation provided that a related contaminant with higher toxicity and environmental concentration is available for comparison to environmental quality guidelines. For example, the toxic equivalency (TEQ) model is a technically defensible process for evaluating the combined effects of dioxin-like chemicals (e.g., dioxins, furans, coplanar PCBs). Conservative mixture models may also be applied to address aromatic and aliphatic constituents of petroleum-related organic compounds.

- An ecological relevance check can be conducted to assess whether the list of COPCs can be reduced. In some cases, the relevance check amounts to the application of common sense. For example, chloride may be eliminated from the list of COPCs for marine environments because it is a naturally occurring substance in high concentrations in seawater. In other cases, the relevance check is less intuitive and requires supporting evidence from peer-reviewed literature.

- In general, contaminants should be retained as COPCs if site history or other data indicate concentrations at elevated concentrations relative to background conditions are likely. For example, elevated concentrations of resin acids and fatty acids in the vicinity of pulp mill operations would warrant their inclusion as COPCs even though environmental quality guidelines for these substances are lacking. Metals should be retained if the pattern of their distribution suggests that anthropogenic influences have resulted in increased concentration or mobilization.

"Conventional" parameters (e.g., sediment ammonia and sulphide concentration; water pH or hardness; soil or sediment organic carbon content; soil pH) that may mediate biological responses should be assessed even though these parameters may not have applicable guidelines.

Role of Background Concentrations: Provincial guidance (CSR Protocols 4 and 9)[14] provides methods for the determination of background soil and groundwater conditions. An analyte should not be selected as a COPC if concentrations at the site are less than background (as determined by CSR protocol) and the background determination conducted under CSR Protocols 4 or 9 has been approved by the Ministry of Environment. COPC selection in the DERA should describe that analyte concentrations exceeded the applicable numerical guideline, but not the background concentration. The background determination should be included as an appendix to the DERA or, at a minimum, cited.

Content for the DERA:

- A narrative or tabular summary of each COPC considered during the screening phase, along with a rationale for its inclusion or exclusion.
- Arguments for the exclusion of COPCs based on environmental fate, ecological relevance, or background considerations must be fully documented in the DERA.

Key Issues for the DERA Practitioner:

- It is a technical error to exclude COPCs simply because CSR numerical standards are not available.
- It is preferable to conservatively include a COPC even if professional judgment suggests that potential risks associated with the COPC are low.
- It is also preferable to retain a COPC for which there are scant environmental effects data and discuss the data limitations in the uncertainty assessment, as opposed to eliminating the contaminant based on lack of detailed information.
- CSR Schedule 10 lists generic soil and water standards specific to human health, but notes it "is the responsibility of the responsible person for the site to ensure that the use of the soil or water standards… do not constitute a significant risk or hazard to ecological health." Compounds listed on Schedule 10 should be included as COPCs if present at the site.
- DERAs are frequently tailored to reflect COPC-specific issues. Additional information regarding DERAs for metals, hydrocarbons and other contaminant groups is available in the literature.

[14] http://wlapwww.gov.bc.ca/epd/epdpa/contam_sites/policy_procedure_protocol/index.html.

2.5 Step PF-4: Identify Exposure Pathways of Concern

The following exposure pathways of concern should be considered:

- Soil invertebrates and terrestrial plants are in direct contact with elevated COPC concentrations in soil;

- Mammals, birds, amphibians and reptiles ingest elevated COPC concentrations via consumption of prey items. [Note: relevant prey items vary according to receptor];

- Mammals, birds, and amphibians and reptiles ingest elevated COPC concentrations via water ingestion;

- Mammals, birds, and amphibians and reptiles ingest elevated COPC concentrations via incidental soil/sediment ingestion;

- Aquatic species (macrophytes, plankton, invertebrates, and fish) are in direct contact with elevated COPC concentrations in surface water and/or sediment [Note: the proportion of surface water and sediment contact varies according to receptor]; and

- Some aquatic species (e.g., planktivores, piscivores) ingest elevated COPC concentrations via consumption of prey items.

With respect to inhalation and dermal exposure pathways, BCMELP (1997) notes that:

- Inhalation toxicity data are generally lacking for the majority of contaminants;
- Exposure via ingestion is assumed to be substantially larger than inhalation; and
- Dermal exposure is limited by the presence of fur and feathers that reduce the actual dermal contact of the receptor to soil contaminants.

Although these factors suggest that inhalation and dermal exposure routes are unlikely to be applicable at the majority of sites, unique circumstances may warrant the inclusion of either pathway in the detailed ERA. Examples of unique circumstances include:

- The receptor is completely soaked in water or another carrier liquid that reduces the mitigating effect of fur or feathers (e.g., waterfowl in an oil spill). [Note: this scenario is presented as a "special case" in BCMELP 1997].

- The receptor inhabits subsurface burrows within soil contaminated by high concentrations of volatile compounds. BCMELP (1997) provided this scenario and argued that exposure associated with inhalation was minimal relative to the exposure associated with ingestion. However, explicit consideration of the inhalation pathway

may be warranted if the receptor involved is of special concern in the risk assessment (e.g., it is a rare or endangered species).

- Dermal exposure (direct contact with soil and sediment) is a relevant exposure pathway for amphibians and reptiles; however, detailed guidance on how to assess dermal exposure is not available for all compounds or biota.

Content for the DERA:

- A narrative or tabular summary of each exposure pathway considered in the DERA, along with a rationale for its inclusion.

- Arguments for the exclusion of other exposure pathways must be fully documented.

Key Issues for the DERA Practitioner:

- The risk assessment is incomplete if exposure pathways were inappropriately excluded from consideration. It is preferable to conservatively include all possible exposure pathways at the problem formulation stage, even if professional judgment suggests that the exposure is likely minimal.

- Specific COPCs can increase the priority of different exposure pathways. For example, risks to carnivores via food consumption are a higher priority if the COPCs include biomagnifying compounds. Risks to aquatic life via groundwater flow are a higher priority if the COPCs are highly mobile.

2.6 Step PF-5: Identify Receptors of Potential Concern

The selection of receptors of potential concern (ROPCs) for DERA is based on site ecology and, where applicable, land use. The majority of ROPCs reflect populations of species; however, ecosystem- and community-level ROPCs can also be selected where appropriate (Suter 1996a)[15]. One or more ROPCs should be selected for each receptor group present (or likely to be present) at the site. These receptor groups (Table 1) correspond to trophic levels or feeding guilds, depending on the desired level of assessment in the DERA. The underlying objective of the ROPC selection is that it must match the conceptual model for the site (Section 2.8).

2.6.1 Level of Ecological Detail

Table 1 provides generic examples of potential receptor groups. In general, a greater degree of ecological resolution in ROPC selection is appropriate when:

- **Habitat of high ecological importance is present:** For example, a bog or wetland habitat may require further subdivision of the "terrestrial plant" and "aquatic macrophyte" receptor groups listed on Table 1 into multiple subgroups (e.g., floating macrophytes, emergent aquatic vegetation, carnivorous plants, rushes and grasses, shrubs). Conversely, subdivision of the terrestrial plant receptor group may be unnecessary if the site consists primarily of grasses and shrubs.

- **Rare, endangered or threatened species are present (or likely to be present):** If rare, endangered or threatened species are present (or likely to be present, based on the best-available information regarding species geographic distribution and habitat preferences), then an increased level of ecological resolution is appropriate. For example, if a rare small mammal was present, the detailed ERA should explicitly assess risks to that species' feeding guild as well as other small mammal feeding guilds (instead of simply evaluating risks to the larger small mammal receptor group) Existing provincial risk assessment guidance requires assessment of all species that are rare, endangered or threatened (BCMELP 1997).

Key Issues for the DERA Practitioner:

- Consider all rare or endangered species known to be or likely to be present.

[15] An example of an ecosystem-level receptor would be "the wetland ecosystem", for instances where the measure of effect reflects an ecosystem-level process such as nutrient cycling or productivity. An example of a community-level receptor would be "the benthic community", for instances where the measure of effect is community-level attributes such as diversity or abundance.

2.6.2 Relationships to COPCs and Exposure Pathways

Known species sensitivities to COPCs should be considered in ROPC selection. (e.g., birds are known to be sensitive to certain pesticides due to effects on egg shell thinning; some fish are known to be sensitive to selenium based on reproductive toxicity endpoints). Arguments that a single ROPC was selected as a surrogate for other ROPCs based on relative sensitivity are inappropriate unless a detailed rationale is provided (e.g., it is inappropriate to argue that earthworms should be the only soil invertebrate ROPC unless appropriate and relevant toxicity data are available, or the biology of the earthworm makes it inherently more sensitive to site-specific COPCs).

Information about the exposure pathways under consideration should also influence selection of ROPCs. For example, if groundwater flow to aquatic life is an important fate pathway, this may indicate that hard-bottom intertidal receptors (e.g., mussels; kelp) would be more appropriate than migratory fish. The duration of the potential exposure is also a relevant factor: migrant mammalian and avian species are explicitly excluded from several land uses in the Tier-1 guidance, although migratory birds can be included if present during the breeding season (BCMELP 1997). Consideration of the federal regulatory perspective on this issue is recommended if migratory waterfowl are present that trigger the *Migratory Birds Act*.

2.6.3 Land Use Considerations

Tier-1 risk assessment guidance (BCMELP 1997) selects ROPCs based on land use—this approach is primarily applicable to the Uplands (Human Use) ecosystem type. For DERAs, land use should be considered in terms of its influence on habitat quality and availability; ROPC selection is therefore based on site-specific ecology (which may result in exclusion of several feeding guilds due to a lack of suitable habitat as a result of land development). This is a relatively subtle difference in the interpretation of the existing guidance, but is necessary so that significant ROPCs are not excluded from consideration based on simply on land zoning classifications (Section 1.6).

Content for the DERA:

- A detailed rationale for the selection of ROPCs applicable to the site (and equally important, a detailed rationale for why different feeding guilds or trophic levels that might reasonably be present were excluded).

Key Issues for the DERA Practitioner:

- The risk assessment will likely be deemed incomplete (and thus rejected) if a reviewer determines that ROPCs were inappropriately excluded from consideration.

2.7 Step PF-6: Define Assessment Endpoints, Measurement Endpoints and Risk Hypotheses

2.7.1 Definitions

Assessment and measurement endpoints facilitate translation of management goals into a specific scope of work for the detailed ERA. The specific definitions of assessment and measurement endpoints vary among guidance documents. Commonly used definitions include:

- **Assessment Endpoint:** "The characteristic of the risk assessment that is the focus of the risk assessment" (CCME 1996); also "an explicit expression of the actual environmental value that is protected, operationally defined by an ecological entity and its attributes" (USEPA 1998).

- **Measurement Endpoint:** "An effect on an ecological component that can be measured and described in some quantitative fashion" (CCME 1996); also "a measurable change in an attribute of an assessment endpoint or its surrogate in response to a stressor to which it is exposed"[16] (USEPA 1998).

For each management goal, multiple assessment endpoints may be necessary. For each assessment endpoint, multiple measurement endpoints may be necessary. Risk hypotheses for each measurement endpoint should be developed. Risk hypotheses "clarify and articulate the relationships that are posited through the consideration of available data, information from the scientific literature and the best professional judgment of risk assessors developing the conceptual model. This explicit process opens the risk assessment to peer review and evaluation to ensure the scientific validity of the work" (USEPA 1998). Aquatic and terrestrial examples are provided below for illustrative purposes:

Aquatic:

- Management goal: Develop risk-based groundwater standards for use at a contaminated site.

- Assessment endpoint: Abundance and density of the aquatic macrophyte community along the shoreline of the site.

- Measurement endpoint: Measure the survival and growth of giant kelp (*Macrocystis pyrifera*) gametophytes exposed to groundwater concentrations representative of conditions at the point-of-discharge.

[16] USEPA (1998) uses the term "measures of effect" rather than "measurement endpoint".

- Risk hypothesis: The survival and growth of giant kelp gametophytes exposed to groundwater concentrations are not reduced by more than 20% relative to the performance of reference samples.

- Alternate risk hypothesis: The survival of giant kelp gametophytes is not reduced below a value which previous scientific investigations determined to be the minimum survival necessary to support a viable population.

Terrestrial:

- Management goal: Determine if soil COPC concentrations represent an unacceptable risk to small mammals occupying the grassland portion of the site.

- Assessment endpoint: Assess the viability of the deer mouse population at the site.

- Measurement endpoint: Compare the daily ingested COPC dose for deer mice at the site to a toxicity reference value that represents an acceptable level of effects (e.g., a LOAEL-based TRV).

- Risk hypothesis: The estimated daily ingested COPC dose does not exceed the LOAEL-based TRV.

- Alternate measurement endpoint: Compare the number and average weight of deer mice caught at the site relative to the number and average weight of deer mice caught at a similar nearby grassland without elevated soil COPC concentrations (using the same level of sampling effort).

- Alternate risk hypothesis: The number and average weight of deer mice are consistent between the two sites. Note that this comparison may or may not be made on the basis of statistical significance.

Risk hypotheses are not necessarily equivalent to the statistical testing of a null hypothesis; however, the risk assessor may opt to use statistical considerations depending on the particular assessment and measurement endpoints. In these instances, statistical power should be explicitly considered (e.g., sample size, sample locations and study design, normal variability, appropriate alpha levels).

2.7.2 Importance in the DERA Framework

Assessment and measurement endpoints "provide direction and boundaries for the risk assessment" and "minimize miscommunication and reduce uncertainty" (USEPA 1998). ***There must be a measurement endpoint that addresses each combination of COPC, exposure pathway and ROPC.***[17] Failure to properly define assessment and measurement endpoints was identified as a common limitation by USEPA (1993a). Other common problems in selecting assessment and measurement endpoints include:

- Assessment endpoint provides an ambiguous statement best suited to a management goal that cannot be translated into specific measurement endpoints. [Example: assessment endpoint is phrased as "protect the ecological integrity of the aquatic macrophyte community." The term "ecological integrity" is subject to interpretation.]

- Measurement endpoint provides an ambiguous statement that cannot be translated to a quantifiable property that can be accurately measured. [Example: measurement endpoint is phrased as "measure the productivity of the aquatic macrophytes at the site." The term "productivity" is not specified in sufficient detail, and the parameter of measurement interest is not specified.]

- Measurement endpoint is subject to confounding factors or indirect effects that limit its utility for measuring the specific COPC and exposure pathway under investigation. [Example: measurement endpoint involves comparison of *in situ* percent coverage of aquatic macrophytes at the site relative to reference locations, but fails to consider major differences in substrate types between the locations as a confounding effect.]

Content for the DERA:

- A tabular summary of management goals, assessment endpoints, measurement endpoints and risk hypotheses.

Key Issues for the DERA Practitioner:

- Measurement endpoints must be specified in detail so that they demonstrate that a quantifiable property exists, is relevant to the COPC/ROPC/exposure pathway being evaluated, and is being measured.

- Identification of appropriate measurement endpoints crystallizes the selection of "tools" for inclusion in the DERA.

[17] A different measurement endpoint is not required for each combination. A food chain model for evaluating risks to small mammals would simulataneously address risks associated with soil ingestion, food ingestion and water consumption exposure pathways

2.8 Step PF-7: Development of a Conceptual Model

Although this step is described near the end of the problem formulation process (which is consistent with other guidance manuals), creation of the conceptual model is an iterative and ongoing activity throughout all stages of the problem formulation.

2.8.1 Requirements of a Conceptual Model

A well-constructed conceptual model provides a summary of the site ecology. The development of the conceptual model is useful for communicating the risk assessment to others (especially laypersons unfamiliar with risk assessment terminology and assumptions). Visual depiction of the underlying relationships also facilitates a "reality check" on the scope of the risk assessment and the degree to which simplifying assumption have been made in framing the risk issues. Conceptual models should include (Suter 1996a):

- **Contamination sources:** Risk assessments may involve multiple point or non-point sources of contamination (e.g., free-product zone; contaminated groundwater, soil, sediment, water, or air; effluent point sources) that should be included in the conceptual model. All on-site sources must be included; significant off-site sources should also be included. The purpose of including contamination sources in the conceptual model documents that all relevant sources (which lead to exposure pathway and COPC selection considerations) were addressed.

- **Dominant exposure and fate pathways:** All exposure pathways considered in the DERA should be depicted in the conceptual model. Significant environmental fate pathways (e.g., sediment deposition, microbial degradation, groundwater flux, sorption to organic carbon in soil) should also be indicated. Including exposure and fate pathways in the conceptual model documents that all relevant exposure pathways were addressed.

- **Relevant trophic levels or feeding guilds:** All relevant trophic levels and feeding guilds must be depicted in the conceptual model, along with significant interactions between the different trophic levels and feeding guilds (i.e., the conceptual model should include a food web diagram). The inclusion of a food web diagram documents the ROPC selection process, and also illustrates potential indirect effects that may complicate the assessment. [Example: conceptual model correctly indicates that elevated COPC concentrations in soil may impact both soil invertebrates as well as a small mammal ground insectivore. A potential indirect effect that should be considered if the detailed ERA includes field measurement of small mammal abundance is that food sources may also be depleted by direct toxicity of COPCs on soil invertebrates].

2.8.2 Presentation Format

All conceptual models should be linked to a narrative that provides detailed rationale for the decisions made (e.g., source identification, selection of COPCs, ROPCs and exposure pathways. Two different types of conceptual models are commonly applied, each with certain advantages and disadvantages:

Box Diagrams: A "flowchart" style of conceptual model. An advantage of this approach is that it facilitates a more rigorous examination of the pathways and connections among and between contaminant sources, exposure pathways, major fate processes, and biological units. Although a common symbology can be used to simplify these relationships (e.g., a dotted line to indicate exposure pathways; a solid line to indicate fate processes), a highly complex box diagram conceptual model may be visually cumbersome. An example of a box-style conceptual model is provided in Figure 4.

Pictorial: A cartoon-based conceptual model that incorporates visual representations of the pathways and receptors. This style of conceptual model is well suited to communicating contaminant source, exposure pathways, major fate processes, and feeding guilds/trophic levels to a non-technical audience. A disadvantage is that some fate processes and indirect effects cannot be represented easily in a pictorial fashion. An example of a pictorial-style conceptual model is provided in Figure 5.

Content for the DERA:

- A pictorial or box diagram conceptual model (or both) must be included.

2.9 Risk Assessment Strategy

The problem formulation also provides an opportunity to lay out the overall strategy of the risk assessment. The strategy involves selection of specific risk assessment tools and organization of those tools into appropriate tiers. This strategy evolves throughout the problem formulation stage based on study design considerations (e.g., sample size, sample locations, desired statistical power, and potential risk characterization methods). Often, the strategy is documented in a sampling and analysis plan (SAP), which can be submitted for review and input from interested parties.

2.9.1 Choosing from the DERA "Toolbox"

Technical or financial constraints are invariably an issue. Although these constraints are a part of the reality of establishing measurement endpoints, bias or other errors described in Section 2.1.2 should be avoided. This section outlines operational guidance to assist in the translation of measurement endpoints and conceptual models developed during the

problem formulation into a practical risk assessment strategy. Guidance on the application of DERA tools is provided in the exposure and effects assessment sections.

Four different categories of DERA "tools" are presented, which range from the collection of raw data to high-level interpretative methods. These tool categories are: a) direct measurement; b) modeling, and c) interpretative; and, d) synthesis (Section 1.5). The following factors should be considered when selecting specific tools from the DERA toolbox, and in many respects, reflect the need to consider the potential uncertainty in selected approaches as part of the problem formulation (rather than relegating the uncertainty analyses until after the risk characterization is completed):

- **Specificity:** Specificity refers to the degree to which a tool is tailored to the COPC/exposure pathway/ROPC combination being investigated. Tools should be specific to the relevant exposure pathway to the extent possible.

- **Ecological Realism:** Ecological realism refers to the degree to which a tool incorporates the processes and interactions observed in the field, as opposed to requiring highly simplifying assumptions. ***A DERA should maximize ecological realism wherever possible, subject to practical and scientific constraints.*** A decision to implement a simplified tool may still be correct, provided that the uncertainty inherent in the simplifying assumptions is properly documented.

- **Reliability:** Reliability is the ability of the tool to generate meaningful data for the purposes of the risk assessment. Reliability is improved when the tool has written protocols available, the influence of confounding factors are well-documented, and established decision criteria exist for interpretation of results. Avant garde and non-standard tools may be useful, but typically require an increased effort to generate scientifically defensible data. Note that a high level of effort is <u>not</u> a rationale for excluding a tool from consideration; rather, it is an argument for designing an appropriate tiering strategy (see below).

For example, toxicity testing is based on established regulatory protocols, the influence of common confounding factors is relatively well-understood for most tests, and decision criteria are available (i.e., Tier-1 guidance establishes a 20% reduction relative to the negative control as the permissible level of effects)[18]. Conversely, a fish or wildlife population survey has no established regulatory protocol, requires consideration of statistical power and experimental design in order to generate reliable data, and greatly benefits from the inclusion of appropriate reference locations. However, data from a properly constructed field survey may have equal or greater value provided that potential confounding factors and uncertainties are properly addressed.

[18] Similar decision criteria are also available for soil toxicity testing in the Tier I guidance manual; the ECx varies by land use.

2.9.2 Tiering/Iteration

Ideal tool(s) for a DERA are highly specific, ecologically relevant, reliable, and cost-effective; however, the reality is that the costs and level of effort tend to increase in proportion to specificity and ecological realism. DERA tools are therefore frequently implemented in a tiered or iterative manner with tools of increasing ecological realism (and cost) used only if required to achieve the desired level of uncertainty relative to site management goals. Risk assessments have been described as using a "tiered" or "iterative" approach; regardless of the term, the operational concept of starting the risk assessment with a subset of potential tools and then progressing to more complex tools (or refining existing tools) only as needed remains the same.

To the extent possible, the problem formulation should consider the relationships of various tools to one another, along with the decision points to move through the various tiers or iterations. Flowcharts are valuable for scoping (and communicating) the potential tiers or iterations of the DERA with the client and other interested parties; they also provide a rationale for why (or why not) increasingly complex DERA tools may required relative to consideration of uncertainty and site management goals. Several examples of potential tiering and/or iterative arrangements are provided below for consideration. Note that decisions regarding how to organize different DERA tools are highly study- and site-specific, and therefore, these examples are provided for illustrative purposes only.

- A potential arrangement of DERA tools used to assess risks to avian and mammalian wildlife in the uplands (wildland) ecosystem is provided in Figure 6.

- A potential arrangement of DERA tools used to assess risks to aquatic receptors in the streams and rivers ecosystem is provided in Figure 7.

Key Issues for the DERA Practitioner:

- Has an appropriate DERA tool (or tools) been selected for each measurement endpoint documented in the problem formulation? Are the selected DERA tools specific to the relevant exposure pathway, appropriately ecologically relevant for the desired level of uncertainty, and adequately reliable for the objectives of the risk assessment?

- Has the relationships among different DERA tools been established to the degree needed for the problem formulation? Can I document a tiering or iterative strategy for how additional DERA tools could fit in the overall plan for this risk assessment if refinement of the risk estimates becomes necessary?

2.9.3 Prepare a sampling and analysis plan

A sampling and analysis plan (SAP) should be prepared prior to implementing any data collection activities. SAPs can be combined with the problem formulation or prepared as a stand alone document. SAPs should provide information about:

- Proposed study design (i.e., a rationale for number and location of samples) for each risk assessment tool, including consideration for how the data will be interpreted;

- Data collection activities needed to implement each risk assessment tool, including sampling, analytical or test methodologies to be followed. Shipping, transport and storage requirements are usually included.

- Quality assurance/quality control measures for each data collection activity are described and data quality objectives are specified; and

- Health and safety considerations are frequently included.

The level of detail in the sampling and analysis plan will vary depending on the complexity and nature of the risk assessment as well as the requirements of the client.

2.9.4 Review by Interested Parties

Review of the PF and SAP by the client and/or other interested parties may be appropriate depending on the outcome of Section 2.2.2. As a practical consideration, input on a PF is facilitated when a SAP is included, since the SAP provides details on proposed sample locations and the specific risk assessment tools.

3.0 EXPOSURE ASSESSMENT

This section focuses on central themes when selecting, applying and interpreting DERA tools within the exposure assessment phase of the risk assessment. Readers should also refer to specific DERA tools (described in Appendices I - III for direct measurement, modeling and interpretative tools, respectively) for additional information. A discussion of the synthesis tools is provided in the Risk Characterization section (Section 5.0) of this document.

This section has the following central themes:

- Section 3.1: Selecting an Appropriate Measurement of Dose;
- Section 3.2: Direct Measurement versus Modeling; and
- Section 3.3: Ecosystem-Specific Issues.

3.1 Selecting an Appropriate Measurement of Dose

Most screening-level ERAs focus on external dose, as quantified by the total contaminant concentrations in soil, water or sediment. However, DERAs should consider how abiotic factors influence the true external dose to which organisms are potentially exposed. External doses consist of two separate fractions[19] depending on the temporal scale involved (Semple et al. 2004)—the bioavailability of contaminants in soil and sediment typically decreases with aging as molecules of a COPC slowly move into locations within the environmental matrix that cannot be accessed by organisms (Alexander 2000).

Differentiation of the external dose fractions has numerous implications. In general, exposure assessment tools that measure the bioavailable fraction are preferred to those that only measure the total COPC concentration. Consideration of the degree to which the bioaccessible fraction can become bioavailable as a result of temporal or other changes is also important—for example, increased knowledge regarding sorption of organic compounds to soot carbon in sediment has implications for risk assessment methodologies such as the use of equilibrium partitioning and biota-sediment accumulation factors (Cornelissen et al. 2005).

DERAs should also consider how biotic factors influence the true internal dose[20] to which an organism is exposed (and reacts to). Differentiation of the bioabsorbed and bioreactive fractions also has implications for DERA. Tools that consider the bioabsorbed fraction (e.g., relative bioavailability factors for soil) or bioreactive fractions (e.g.,

[19] **Bioavailable:** The fraction of the total contaminant concentrations that is immediately available for uptake by organisms. **Bioaccessible:** The fraction of the total contaminant that may be available to an organism. This fraction includes the portion of the total that is currently bioavailable, plus the portions that may become bioavailable over time.

[20] **Bioabsorbed:** The fraction of the total contaminant concentration that is actually taken up by an organism (i.e., passes across the gill, integument or gut). The bioabsorbed fraction is not necessarily the same as the ingested dose, since a significant fraction of some contaminants may be excreted from the organism. **Bioreactive:** The fraction of the total contaminant concentration that is actually able to cause toxicity (i.e., the bioabsorbed fraction minus the fraction that is depurated, internally sequestered, or used by the organism for its own needs).

physiologically based pharmacokinetic [PBPK] models; organ-specific tissue residue guidelines) provide increased ecological realism, and are an area of ongoing research.

Key Issues for the DERA Practitioner:

- Operational definition of dose in terms of internal (or, ingested, as appropriate) versus external is adequate for most DERA applications; however, the ecological realism is enhanced when doses are considered in terms of the bioaccessible, bioavailable, bioabsorbed and bioreactive fractions. This latter approach (with selection of appropriate risk assessment tools) is recommended if justified by the desired level of information needed to support site management.

- Selection of the appropriate exposure dose is strongly influenced by the availability of applicable and appropriate effects data. Units and types of measurements need to be consistent between the exposure and effects assessment phases.

3.2 Direct Measurement versus Modeling

Environmental fate and transport models are often utilized in the exposure assessment; there is a broad range of model types of varying complexity available. Models include strictly abiotic models of contaminant transport (e.g., groundwater plume modeling) to biotic models (e.g., uptake models ranging from simple bioaccumulation factors to complex food web models).

3.2.1 Advantages and limitations

Direct measurement and modeling have different advantages and limitations in a DERA, as follows.

Advantages for direct measurement: Direct measurement of exposure to COPCs through chemical analyses is generally considered to be more reliable and credible than predicting COPC concentrations through modeling.

Limitations of direct measurement: Collecting sufficient exposure chemistry data may require considerable project resources, depending on the size of the area under investigation, number of COPCs and number of exposure pathways requiring sampling. Destructive sampling may be inappropriate, especially when the exposure assessment requires sampling of biological tissues. Direct measurement only provides a "snapshot" of the potential exposure at the time of sampling—seasonal or other trends are not captured unless sampling is repeated.

Advantages of models: Models can be used for interpolation (i.e., to fill in spatial, temporal or taxonomic gaps if the measured data are insufficient) or for extrapolation (i.e., once validated, models can be used to explore hypothetical scenarios regarding site

management or to assess the effects of changes in environmental conditions). Models can be used to gain a better understanding of the relative importance of different exposure pathways and the influence of factors that limit bioavailability (thus reducing the overall exposure). Models also facilitate a quantitative evaluation of the uncertainty in the exposure assessment that is more sophisticated than simply measuring the standard deviation or other summary statistics based on measured data.

Limitations of models: Models are limited in that the accuracy of a model's predictions is unknown until the model is validated against site-specific (measured) data. A substantial amount of data are required to parameterize some models (e.g., physical properties such as water volume and flow, sediment or soil organic carbon content for abiotic compartments, as well as biological properties such as lipid contents and feeding relationships for major species). Although some generic fate and exposure models are available, a certain level of expertise is required to determine if the generic model is appropriate for use and if not, to construct a site-specific model. ***Generic models should not be used unless they are deemed appropriate for the site*** since structural errors in a model may result in unrealistic estimates of exposure concentration. Less complex models (e.g., ORNL uptake models) are less sensitive to structural issues; however, concerns regarding their accuracy for a given site should still be considered.

3.2.2 Deterministic versus probabilistic models

Models can be either deterministic or probabilistic.

Deterministic models are advantageous because they: a) are relatively simple to implement and interpret (i.e., the model generates a single value only), and b) require less data (compared to probabilistic models). However, deterministic models ignore variability in parameterization by focusing on single values (e.g., mean, 95% upper limit). Selecting conservative estimates for these single values, by definition, implies an inherent bias whereas the model is automatically overprotective for a large fraction of the model's domain[21], and automatically underprotective for a smaller fraction of the model's domain. Deterministic models also ignore uncertainty in the parameterization by focusing on single values. The uncertainty analysis is therefore limited to qualitative statements about each individual parameter rather than a quantitative estimate of the total uncertainty in the model itself.

Probabilistic models are advantageous because they explicitly consider the variability and uncertainty in the distribution of each parameter; as a result, risk estimates are also provided as a distribution. As a result, risk estimates can be expressed in terms of a range or as mean with confidence intervals rather than a single value. Distinguishing between variability and uncertainty is a key issue. Both variability and uncertainty produce

[21] Domain refers to what is being modeled: receptors, changes over time or space, and so on.

statistical distributions of values, but those distributions are interpreted differently. For example, multiple water samples are collected, analyzed for a given COPC, and the results expressed as a statistical distribution. If the differences among individual measurements reflect spatial or temporal variability in the concentration, then the distribution reflects variability. If the differences among individual measurements reflect measurement error (imprecision in the analytical technique), then the distribution reflects uncertainty in the true value. In many cases, elements of both variability and uncertainty are present in the data; however, if the measurement error can be estimated separately based on laboratory replicates, then the remaining differences in the data can be attributed to variability alone. Information on correlations between different parameters is also needed to avoid unrealistic amplification of the risk estimate bounds.

USEPA (1997a) argues that "probabilistic analysis techniques such as Monte Carlo analysis, given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments"; however, not every assessment requires or warrants a quantitative characterization of variability and uncertainty. Deterministic models should be implemented first to determine if a probabilistic model would contribute to the site management objectives. USEPA (1997b) argues probabilistic approaches are unnecessary when: screening-level (or deterministic) risk estimates generated using conservative methods are clearly below levels of concern or the costs for site remediation are low (i.e., the financial implications of remediating areas that in fact do not present risks are acceptable to the client). However, probabilistic approaches should be considered if:

- Screening-level (or deterministic) risk estimates generated using conservative methods are clearly above levels of concern.

- It is necessary to quantify the uncertainty associated with point estimates of exposure, or it is necessary to prioritize different risk estimates for site management purposes (since rankings have little meaning if each risk estimate has varying levels of uncertainty and variability).

- The costs for site remediation are high (i.e., the financial implications of remediating areas that in fact do not present risks are not acceptable to the client).

Key Issues for the DERA Practitioner:

- Will having a bounded confidence interval for risk have any influence on risk estimates (and therefore, management decisions)?

- Are the necessary data available (i.e., estimates of variability or uncertainty for all important parameters; information on correlations among parameters)?

- Probabilistic approaches should be used when it is necessary to rank risk estimates or quantify the uncertainty associated with the risk estimates.

- If a probabilistic model is used, it may be necessary to obtain input from regulatory agencies regarding an acceptable probability for a defined level of predicted adverse effects.

3.2.3 Use of Modeling in DERA

We recommend that direct measurement form the basis of the exposure assessment in the majority of DERAs, supplemented by models under some circumstances, as described below. Direct measurement is strongly recommended for measuring COPC concentrations in different exposure media when those COPC concentrations are subsequently used as the basis for other models[22]. Models should be used to supplement direct measurement only in appropriate situations (e.g., site is relatively large and models can be used to infer COPC distributions with an adequate certainty; many media or species need to be sampled but only some are available; temporal variability needs to be considered).

All models must be validated to the extent possible, and the uncertainty assessment must include consideration of how minor variations in model parameterization impact the results (i.e., a sensitivity analyses). Validation and sensitivity analyses should be considered for all models irrespective of their complexity (even for simple bioaccumulation factors[23]), but are especially important for those pathways that contribute significantly to risk estimates. Lack of validation and sensitivity analyses was noted as an error in many risk assessments reviewed by USEPA (1993a). Probabilistic approaches provide an improved consideration of uncertainty, and should be included in the DERA framework wherever appropriate. Increasing use of probabilistic approaches for DERA are anticipated by USEPA as the science for this issue advances (Dearfield et al. 2005).

An iterative approach (i.e., use deterministic models at first with an increasing number of parameters converted to a probabilistic approach as needed relative to risk management goals) is recommended. Probabilistic approaches should quantify the uncertainty and

[22] For example, it is inappropriate to model groundwater concentrations based on measured soil COPC concentrations if those groundwater data are subsequently used to predict COPC concentrations in sea urchins. Risk assessors should avoid linking models wherever possible due to the compounding uncertainties involved.

[23] "Universal" bioaccumulation models (i.e., based on analyses of data from multiple sites) such as those presented by Efroymson et al. (2001) typically quantify the uncertainty in the model (e.g., BAF ± standard deviation) that should be considered in the sensitivity analyses. A reality check of the models against site-specific data is recommended wherever possible.

variability in as many parameters as possible (or at least, the parameters with the greatest impact on risk estimates); however, partial probabilistic models are acceptable, provided that discussion of the fact that the bounds on the risk estimates do not represent the total uncertainty/variability in the model is provided.

Key Issues for the DERA Practitioner:

- Has direct measurement been used to the extent possible?
- If models (probabilistic or deterministic) were used, were they validated against site-specific data?
- Has a sensitivity analyses of the model been included in the discussion of uncertainty? If probabilistic models are used, how do risk estimates reflect site-specific uncertainty and variability?

3.3 Ecosystem-Specific Issues for Consideration

3.3.1 Deep Aquatic Ecosystem

Provincial guidance regarding the design and implementation of sediment quality assessments is provided in BCMOE (2005) which should be reviewed in terms of its applicability within the objectives of the site-specific DERA. In addition, the following issues are presented for consideration for the exposure assessment for deep aquatic DERAs:

- Selecting analytes for sediment DERAs;
- Addressing subsurface sediment;
- Sampling design for sediment quality assessments; and
- Incorporating porewater chemistry data.

3.3.1.1 Selecting analytes for sediment DERAs

Chemistry samples need to be subjected to a broad range of analyses beyond site-specific COPCs. Data for multiple potential confounding factors will be required to properly interpret any subsequent effects data (e.g., toxicity testing, benthic community structure), including percent organic carbon, particle size distribution, as well as ammonia and sulphide concentrations in porewater. Acid volatile sulphide and simultaneously extractable metals (AVS-SEM) measurements provide information about the potential bioavailability of selected divalent metals. Additionally, the risk assessor should consider the contribution of other COPCs beyond those attributed to the specific site. Sediment assessments for urbanized harbours should consider the significant role that harbour- and

basin-wide non-point sources play in influencing sediment quality. Nearby point-sources (e.g., stormwater or combined sewer outfalls) should also be considered in the context of sediment transport patterns and COPCs selected accordingly (e.g., TBT should be measured if a former shipyard is nearby; pesticides should be measured if stormwater outfalls are in the area).

3.3.1.2 Subsurface versus surficial exposure pathways

Sediment deposition and burial lead to a gradual reduction in the bioavailability of most COPCs (and thus reduce exposure) over time. The DERA must explicitly consider if exposure to subsurface conditions will occur. Examples where exposure to subsurface conditions may occur include dredging, construction (e.g., installation of new pilings), general slope stability, propeller scour, 100 year storms, or floods. Exposure pathways involving subsurface sediments may be excluded from consideration in the DERA provided that they cannot be exposed under a reasonable likely scenario (Chapman and Anderson 2005). Inclusion of subsurface sediment is appropriate if the risk assessor cannot reasonably exclude future exposure scenarios.

3.3.1.3 Sampling design for sediment quality assessment

Gradient-based sampling designs are useful in order to assess the potential influence of other contaminant point-sources; a "near-field/far-field" approach is useful when assessing the potential influence of harbour-wide conditions. It may be necessary to tier the chemical analyses to minimize potential costs: consider having a rush analyses for the broader suite of potential COPCs on a subset of samples in order to determine analyte selection for the remaining majority of samples. Holding times are often an issue in this tiering approach. Consider having the analytical laboratory extract all samples on delivery since extracts for organic analytes can be held longer than the original sediment sample. Reconnaissance sampling (e.g., limited surface and core sampling in advance of the actual DERA sampling) may also be appropriate depending on the amount and quality of data available in the problem formulation.

3.3.1.4 Sediment porewater chemistry

Information on COPCs concentrations in porewater may be relevant. However, *ex situ* porewater collection methods (e.g., centrifugation; vacuum extraction) results in inevitable alteration of the speciation and bioavailability of the sample; *in situ* collection methods (e.g., peepers; solid phase extraction) result in limited sample volumes or require specialized analytical techniques (Chapman et al. 2002a). Consideration of the relative importance of the porewater exposure route within the context of the combination of ROPCs/COPCs selected in the problem formulation is recommended, since many benthic taxa are primarily exposed to surface water rather than porewater (e.g., epibenthic

amphipods inhabit sediment surfaces; clams extend siphons; some tube-dwelling organisms irrigate their tubes with surface water). Measurement of porewater COPCs as a surrogate for whole-sediment exposures is not recommended; however, such measurements are valuable in those instances where the porewater exposure route is of explicit interest (e.g., consideration of equilibrium partitioning of compounds from sediment particles; flux of porewater out of sediment).

3.3.2 Shoreline Ecosystem

Potential issues for consideration for the exposure assessment for shoreline DERAs include:

- Implications of variable geochemical conditions; and
- Implications of variable hydrological conditions (e.g., groundwater plumes).

3.3.2.1 Geochemical considerations

Exposure pathways in the shoreline ecosystem involve considerable alterations in contaminant biogeochemistry. For example, geochemical changes as COPCs discharge to aquatic receiving environments from groundwater have implications in terms of using groundwater chemistry data as a measure of exposure. Changes in redox potential, for example, influence the mobility and toxicity of different metals as they transition from groundwater to seepage zones to the receiving water body. Risk assessors should consider these changes in geochemistry, and consider sampling techniques that more closely approximate conditions at the point of discharge (e.g., mini piezometers in the shoreline; use of subsurface seepage samplers). Risk assessment tools that consider COPC geochemistry may not be applicable under all circumstances (e.g., AVS-SEM does not apply to oxygenated sediment; estuaries have unique and variable geochemistry that impact speciation and biotic ligand models).

3.3.2.2 Hydrological (groundwater plume) considerations

Expert advice from hydrogeologists regarding contaminant flow pathways (i.e., groundwater plumes) is recommended in order to select appropriate sampling locations for groundwater exposure assessments. For the majority of sites, the hydrogeological investigations conducted for site characterization purposes are sufficient; however groundwater plume models provide useful information regarding the likely exposure concentrations at various locations (thus indicating potential sample locations) within a groundwater plume. It may be necessary to implement additional hydrogeological studies if the site has considerable temporal or spatial variability. For example, groundwater discharges from shallow aquifers in an estuarine environment tend to be relatively complex, and thus require detailed examination to justify sample placement (e.g.,

Westbrook et al. 2005). Alternatively, if detailed hydrogeological investigations and/or groundwater plumes are not available, a "picket fence" (i.e., a row of samples along the shoreline) sampling approach is recommended in order to maximize the chances of intercepting the actual exposure pathway. Repeated sampling over time will likely be necessary to capture natural variations in groundwater flow patterns.

3.3.3 Upland Wildlands Ecosystem

A potential issue for consideration for the exposure assessment for uplands (wildland) is the appropriate level of detail in food chain models. Food chain models are frequently used to estimate the total exposure received by wildlife ROPCs through a combination of food, water and incidental soil ingestion. Food, water and soil ingestion rates for specific ROPCs are usually based on allometric scaling equations and assumptions regarding ROPC body weight. Other model parameters needed for calculating COPC exposure includes ROPC-specific dietary preferences as well as percent moisture data for each dietary item. Food chain models need to balance the use of modeled (e.g., allometric scaling formulae) and site-specific measured data for each parameter. In general, DERA food chain models should:

- Include more dietary items than would be normally assessed in a model constructed for screening-level purposes. For example, it is appropriate to divide the soil invertebrate dietary item into foliar, soil-dwelling and litter-dwelling invertebrates since differential COPC accumulation within the food chain of the soil invertebrate community is likely (e.g., Roth 1993). Differential accumulation of COPCs by different plant species (e.g., Torres and Johnson 2001) also means that the plant community should be subdivided into different functional groups such as grasses, forbs, shrubs and trees, and potentially, varying tissue types such as leaves, shoots and berries.

- Include more direct measurement of COPC concentrations in dietary items instead of using literature-based bioaccumulation factors or uptake models;

- Include more site-specific ROPCs that more closely mirror the selected measurement endpoints rather than default ROPCs which may not be as appropriate; and

- Utilize site-specific dietary preferences that reflect the relative abundance of dietary items actually available in the site of interest. For example, including earthworms as a dietary item in the food chain model is only meaningful if the site contains earthworms and ROPCs that consume earthworms.

- Utilize a metabolic-based model to estimate COPC dose where appropriate. Daily ingestion rates (kg food per day) are expressed in terms of daily required energy (calories per day), and the energy content of various dietary items is estimated (or

measured). These models are more complex and require additional data, but provide a more realistic representation of a receptor's feeding behaviour at a given site.[24]

3.3.4 Rivers and Streams Ecosystem

Specific issues for consideration in DERAs for rivers and streams were not identified. A generic consideration, however, is the fact that streams and rivers are highly dynamic, and therefore, it is appropriate to consider the potential influence of water flow and temperature on the exposure assessment. Additionally, see guidance in Section 3.3.2.2 if groundwater discharges to rivers and streams are being assessed.

3.3.5 Upland Human-Use Ecosystem

A potential issue for consideration for the exposure assessment for upland (human-use) is the appropriate level of detail in food chain models. As described in Section 3.3.1 food chain models are frequently used for estimating COPC exposure for wildlife ROPCs. Items identified for consideration in Section 3.3.1 are equally applicable for food chain models for uplands (wildlands) land use; however, the following additional items are applicable for food chains that model COPC exposure for the uplands (human use) ecosystem:

- Effect of human modifications to the environment that alter bioavailability must be considered (e.g., type, depth, and permanence of cover materials that isolate receptors from exposure). A permanent and impermeable barrier means that COPC uptake by dietary items from those particular areas of soil is negligible. Soil caps of clean material have a varying ability to block COPC transmittal depending on depth, quality of soil relative to the underlying material, and the future species assemblage[25].

- ROPCs for food chain models need to reflect the overall habitat quality and quantity. ROPCs should be tolerant of the level of human presence at the site, and included only if they utilize the area for feeding.

- Habitat range factors assume that the ROPC moves equally through all parts of a contiguous habitat range. Habitat range factors are not appropriate if habitat is highly fragmented, or adjacent areas contain habitat of relatively low quality that would limit the ability of the ROPC to move and feed equally in all areas. In some instances, the

[24] A metabolic-based ingestion model is described in USEPA (1993b) and elsewhere in the literature. The complexity of the model can be increased to reflect temporal changes (e.g., an organism's energy requirements vary depending on growth and reproductive status as well as season) and site-specific ecology (i.e., feeding behaviours tend to maximize the energetic return per unit effort by focusing on abundant food items with high energy contents).

[25] For example, a 0.5 meter soil cap is likely sufficient to block the accumulation of COPCs from the underlying material by grass, but may not be sufficient to block the accumulation by large shrubs.

site in question may contain higher habitat quality than its surroundings which would suggest ROPCs will preferentially feed in the area under investigation.

- Exposure data (e.g., soil chemistry) need to be specifically targeted to the areas included in the food chain model.

4.0 EFFECTS ASSESSMENT

This section focuses on central themes to select, apply and interpret DERA tools within the effects assessment phase of the risk assessment. Readers should also refer as needed to details for specific DERA tools (provided in Appendices I - III for direct measurement, modeling and interpretative tools, respectively). A discussion of the synthesis tools is provided in the Risk Characterization section (Section 5.0) of this document.

This section has the following central themes:

- Section 4.1: Ecologically Relevant versus Statistically Significant;
- Section 4.2: Using Literature-Based versus Site-Specific Data;
- Section 4.3: Using Toxicity Testing in a DERA;
- Section 4.4: Deriving Toxicity Reference Values for Food Chain Models;
- Section 4.5: Site Observations and Field Surveys; and
- Section 4.6: Ecosystem Specific Issues for Consideration.

4.1 Ecologically Relevant versus Statistically Significant

Effects data can be interpreted within the DERA based on ecological relevance and statistical significance.

BCMELP (1997) specifies that the permissible level of effects (i.e., what is considered ecologically relevant from a policy point of view) for measurement endpoints involving toxicity tests is the EC_{20} for aquatic ROPCs at all land uses and a variable EC_x for avian, mammalian, plant and soil invertebrate ROPCs as a function of land use[26]. This default guidance should be applied where appropriate, provided that the permissible level of effects makes sense in light of the selected measurement endpoint. For example, it is not permissible to have a 20% reduction in the survival of anadromous salmon (due to federal policy), or have a 50% reduction in the reproduction of a rare mammal species in the vicinity of a commercial operation (due to federal and provincial policy). No specific guidance on what constitutes a permissible level of effect exists for other types of measurement endpoints (e.g., site surveys). Discussion with the appropriate regulatory agencies is recommended to establish the level of effort required (which is strongly influenced by the permissible level of effects).

Reliance on statistical significance alone is equally problematic within the DERA framework since different lines of evidence have varying tendencies towards Type I and Type II errors. Risk assessors should consider statistical power without ignoring the actual magnitude of the observed effects. Test protocols should be consulted with respect to statistical considerations for toxicity testing; however, there is also general agreement

[26] Industrial or commercial land uses, EC_{50}; residential land uses, EC_{20}; urban park or agricultural land uses, EC_{10}

that EC_x approaches are preferred to NOECs and LOECs (both approaches are frequently reported as per test protocols) for DERA purposes since a no observed effect concentration is not equivalent to a no effect concentration (van der Hoeven 1997). A formal study design with respect to statistical consideration is necessary for most measurement endpoints that do not involve toxicity testing since regulatory protocols for site surveys (e.g., study design, replication, and desired power) are not available.[27]

Key issues for DERA practitioners:

- Interpretation of effects data in the DERA framework requires simultaneous consideration of statistical significance and ecological relevance. Reliance on one approach to the exclusion of the other should be avoided.

- Ecological relevance and policy-based decisions about what constitutes an acceptable level of effect is not the same thing. A reality check on the implications of the observed effects in light of their implications for each measurement endpoint is required.

- The statistical power of each measurement endpoint is an important consideration. Although formal power analyses are not always required, the tendency for an assessment endpoint towards false positive and false negative results should be considered in the uncertainty analysis.

4.2 Using Literature-Based versus Site-Specific Data

Literature-based toxicity data are frequently used to set threshold concentrations for use in a site-specific DERA. Examples of threshold concentrations include:

- Deriving an effects-based water, sediment, or soil quality guideline;
- Deriving toxicity reference values for ROPCs;
- Deriving an effects-based tissue residue guideline; and
- Deriving bioaccumulation factors or uptake models.

The first application (deriving an effect-based water, sediment or soil quality guideline) should only be used in conjunction with other risk assessment tools in a DERA since they involve "double-counting" of environmental concentrations (e.g., surface water data are considered a measure of exposure, as well as a measure of effect when compared to the threshold concentrations). Effects-based guidelines are primarily useful for identifying areas with the highest hazard potential to target other risk assessment tools appropriately.

[27] Formal consideration of statistical power may very lead to a decision that statistical significance is not a desired outcome of the study design; however, a clear statement to this effect is necessary so that the transparency of the risk assessment process is maintained.

Several guiding principles are proposed to facilitate the appropriate use of literature-based toxicity data in the DERA process for the remaining three applications:

4.2.1 Level of effort in literature search

The quality of the literature search dictates the reliability of the resulting threshold concentrations. Literature searches must be comprehensive if literature-based toxicity data are used in a DERA. A description of the nature of the literature search should be provided (e.g., list the search engines used or compendiums consulted; provide date ranges; provide the number of studies identified and retrieved; list the key words used). The following considerations for the design of literature searches are provided:

- Older toxicity data (i.e., pre-1990) are frequently relevant, but are less represented in electronic search engines since older articles tend to be listed only by the keywords selected by the author (newer articles tend to include full abstracts in the keywords). An electronic search engine provided by a single journal publisher is not adequate.

- Keywords should be kept as broad as possible since their use is highly inconsistent. For example, a search of "zinc" and "aquatic" and "toxicity" will miss many relevant papers because "aquatic" is not consistently utilized as a keyword.

- Original papers must be retrieved wherever possible[28]. Risk assessors should not rely on toxicity data reported by others (especially in online compendia) since these compendia do not necessarily provide adequate context for evaluating the quality of the study design or considering confounding factors. Transcriptional errors are also potentially present. Using compilations prepared by other jurisdictions or published in the peer-reviewed literature is acceptable, provided that the risk assessor reviews the methods involved to determine their adequacy relative to the considerations listed above. A reality check (i.e., a brief literature search) is recommended to determine if: a) the compilation is adequately comprehensive, and b) additional relevant toxicity data published since the compilation.

- Citation lists in relevant journal articles should be reviewed to identify other relevant papers which may not be captured through other aspects of the literature search.

[28] Most post-secondary institutions contain hard copy or electronic versions of the majority of relevant journals and have multi-institutional sharing agreements in place to access less-common journals. Alternatively, the Natural Research Council offers a fee-based documental retrieval system (http://cisti-icist.nrc-cnrc.gc.ca/docdel/) that can deliver journal articles electronically.

Key Issues for the DERA Practitioner:

- Original literature should be retrieved and reviewed wherever possible. Uncertainty associated with not reviewing the original literature must be documented.

4.2.2 Literature data review

All literature data retrieved must be reviewed in terms of its quality and relevance. Guidelines for reviewing toxicological data are provided in documents such as CCME (1999) and USEPA (2005a), but in general, guidelines can be divided into three categories:

Literature exclusion criteria: Factors that would immediately result in the paper being rejected for use, usually due to the fact that the toxicological investigation was conducted for reasons that are inconsistent with the DERA. USEPA (2005a) list exclusion criteria for evaluating toxicological data for deriving soil standards that include: study conducted to test biological toxins, drugs, or sewage; study used *in vitro* (e.g., cell lines, tissue cultures) methods rather than whole organisms; testing involved a mixture of chemicals[29]; data developed using QSAR or modeled results rather than measured data; data are not from a primary source; test duration not reported.

Study acceptance criteria: Study design should be evaluated further for literature not excluded. USEPA (2005a) suggest the following criteria for deriving soil standards: chemical form and concentration are reported; test medium was a natural or artificial soil; pH reported and within range of 4 – 8.5; organic content reported and less than 10%; study includes at least one control treatment with at least two additional test treatments; study reports ecologically relevant endpoints such as reproduction, population, growth or plant physiology. Note that some USEPA (2005a) criteria (e.g., soil pH is between 4 and 8.5) may not be applicable for all DERA applications; the objective is to match the available literature to site conditions to the extent possible using a transparent study evaluation method.

Study quality criteria: Studies that pass the literature exclusion and study acceptance criteria need to be reviewed in greater detail to ascertain their quality. USEPA (2005a) assigns a score of 0, 1 or 2 to each of the following nine quality criteria, and rejects any study that does not score 10 or greater. Potential factors for consideration include:

- Testing was done under conditions of high (or, for the purposes of DERA, appropriate) bioavailability;

[29] Note that the argument about excluding data for toxicity of contaminant mixture is based on the assumption that the mixtures tested in the study are not necessarily applicable to the site in question. If the mixture toxicity data are in fact applicable, then these data should be considered.

- Experimental designs were documented and appropriate;
- Concentrations of test substances in soil were reported;
- Control responses were acceptable;
- Chronic or life-cycle tests were used;
- Contaminant dosing procedure was reported and was appropriate;
- Dose-response relationship reported or can be established from available data;
- Statistical tests used and level of significance were described; and
- Origin of test organisms was described.

Key Issues for the DERA Practitioner:

- Literature data should be evaluated for relevance and quality using a consistent and transparent system.

4.2.3 Derivation methods

Derivation methods for establishing threshold values using laboratory-based toxicity data tend to utilize one of the following general approaches:

- **Single Toxicity Data Value** — Threshold values are based on the selection of a single data value (or the geometric mean of multiple data values; usually the lowest value[s] available), followed by application of a safety factor. Different toxicological measurements are used, depending on the application (e.g., NOEC, LOEC, EC_{25}, and LC50). Note also that the entire dose-response curve can be compared to the COPC exposure to improve the estimate of potential risks (i.e., apply the available ROPC-specific toxicity data in an increasingly probabilistic manner. See Section 3.2.2).

- **Species Sensitivity Distributions (SSDs) —** SSDs emphasize protection at the community level rather than traditional methods that emphasize protection of individual species (Posthuma et al. 2002). The basic premise of a SSD is that a "safe" concentration for the community at large can be extrapolated based on the distribution of toxicity data for the individual species that make up the community. In this respect, SSDs are fundamentally different from the common practice of dividing the lowest toxicity data point by a safety factor, and are superior since the SSD relies on the entire data distribution, not just the lowest data value.

4.2.4 Dealing with uncertainty in literature-based toxicity data

Incorporating literature-based toxicity data into the DERA process introduces considerable uncertainty if not done appropriately:

- Do a reality check of the methods, data quantity and data quality used to generate the literature-based threshold value. Applying additional uncertainty factors to compensate for poor quality or less relevant data is incorrect (e.g., deriving an avian toxicity reference value based on mammalian toxicity data and an extra uncertainty factor is not recommended).

- Reduce uncertainty to the extent possible by considering how factors that influence COPC bioavailability vary between the laboratory exposures and the actual field exposures that are the subject of the DERA. The ability to address these factors is influenced by the level of effort expended on the literature review.

- Apply uncertainty factors sparingly. Default values of 10 are typically applied for each area of uncertainty (intra-to-interspecies, acute-to-chronic, NOEC-to-LOEC, laboratory-to-field, and so on) resulting in an overall safety factor ranging from 10 to 10,000. Uncertainty factors are frequently misapplied—their original purpose was to compensate for sparse data sets, not to facilitate an extreme application of the Precautionary Principle that requires the use of an infinitely large (and thus overprotective) safety factor (Chapman et al. 1998). If uncertainty factors are necessary, they should be based on the available data instead of simply assuming a default value of 10^{30}. Note that situations where multiple default uncertainty factors are necessary suggest that the available data were not entirely relevant to the objectives of the DERA (see above).

- Use ECx-based data instead of NOEC and LOECs wherever possible. NOECs and LOECs are driven by the selection of test concentrations, and do not necessarily reflect an acceptable level of effects (Chapman et al. 1996). SSDs are preferred (provided that adequate data are available) over the single data point approaches; however, note that input from regulators regarding an acceptable percentage of species to be impacted will likely be required if a SSD approach is adopted.

ROPCs, measurement endpoints and risk hypotheses may need to be reexamined in light of whether or not sufficient toxicity data of appropriate quality are available. For example, a DERA conducted with brown trout (*Salmo trutta*) as a ROPC may wish to redefine the ROPC as "cold water salmonid fish" in order to incorporate rainbow trout toxicity data. If sufficient data are still not available, the risk assessor may wish to drop

[30] Default safety factors are often applied initially, and replaced only if risks are found to be unacceptable. This iterative refinement can be part of the tiering strategy (Section 2.9.2)

the use of literature-based threshold values and focus project resources on direct measurement of adverse effects or other alternate measurement endpoints instead.

Key Issues for the DERA Practitioner:

- Uncertainty factors are primarily intended to compensate for sparse data sets. Comprehensive literature searches should be used to determine if the data set are truly sparse, or simply difficult to assemble.

- Default and multiple uncertainty factors should be avoided where possible.

4.3 Using Toxicity Testing in a DERA

Several "big picture" issues regarding the appropriate use of toxicity data in the DERA framework are discussed below in greater detail; however, the reader should also refer to the modules outlining the advantages and disadvantages of different types of toxicity testing provided in Appendix I.

4.3.1 Which toxicity test(s) should be selected?

The number and types of toxicity tests selected are entirely dependent on the different routes of exposure and ROPCs being evaluated, and therefore, specific guidance for or against particular toxicity tests would be inappropriate[31]. The DERA must provide a detailed rationale for the selected toxicity tests, including consideration of toxicity-modifying factors such as grain size, pH, organic carbon content, and soil moisture, as well as confounding factors such as ammonia and sulphides. Linkage of the selected toxicity tests to the ROPCs and exposure pathways is necessary.

4.3.2 How much toxicity data are needed?

For most DERAs involving toxicity testing, a battery of toxicity tests (usually ranging from three to five tests) to reflect different trophic levels or major taxonomic groups is recommended. Several scenarios involving commonly-available toxicity tests are provided below for illustrative purposes:

- Potential toxicity tests to evaluate groundwater quality discharging to a marine rocky shoreline: 7-d giant kelp germination and growth; 48-h bivalve larval development; 10-min echinoderm fertilization; 7-d larval fish survival and growth.

[31] Mammalian and avian toxicity testing is exceptionally rare in DERA, and therefore, all further discussion regarding toxicity tests is focused on soil, sediment or water toxicity testing.

- Potential toxicity tests to evaluate marine sediment quality: 10-d amphipod survival; 48-h bivalve larval development (on sediment elutriate); 20-d polychaete survival and growth; 28-d amphipod survival, growth and reproduction. Porewater toxicity testing may also be appropriate, depending on the goals of the investigation (see Section 4.6.1).

- Potential toxicity tests to evaluate freshwater surface water quality: 7-d cladoceran survival and reproduction; 7-d larval fish survival and growth; 72-h algal growth; 7-d aquatic macrophyte growth; 7-d fish embryo development.

- Potential toxicity tests to evaluate soil quality: 7-d seed germination, growth and root elongation; earthworm survival and growth (various durations); 28-d collembolan reproduction; 42-d enchytraeid reproduction test.

Arguments that "we tested the most sensitive species" are often made to support testing a single species only; however, this argument is rarely valid unless a battery of various toxicity tests was previously conducted for the site. Without such a battery of site-specific data, the argument for single species testing would require that: a) a single COPC per exposure pathway is being evaluated (i.e., no mixtures of COPCs); b) literature-based toxicity data were available for multiple test organisms; c) the data from the literature were derived under similar test conditions as the site in question (e.g., consistent water hardness, grain size, organic carbon concentrations and so on); and, d) the most sensitive species to that particular COPC was also used for the site-specific toxicity testing. This scenario is extremely unlikely to occur.

4.3.3 What constitutes a chronic toxicity test?

The DERA should emphasize chronic toxicity data over acute toxicity data; however, the terms "acute" and "chronic" are not consistently defined or applied, in part, due the use of the terms to describe effect as well as duration. A review of definitions used by selected jurisdictions is provided below.

Environment Canada toxicity test methods: Environment Canada (1999) defines **acute** as within a short period (seconds, minutes, hours, or a few days) in relation to the life span of the test organisms, for any discernable adverse effects (lethal or sublethal). Conversely, chronic is defined as occurring during a relatively long period of exposure, usually a substantial proportion of the life span of the organism (such as 10% or more) and involving long term effects related to changes in metabolism, growth, reproduction, or ability to survive.

USEPA toxicity test methods: USEPA (2002a,b,c) published test methods for measuring toxicity of effluents and receiving waters. Acute test methods were those

designed to provide information on lethality (e.g., LC50) associated with 24-h to 96-h exposures. Short-term chronic test methods were developed for freshwater and marine/estuarine species with test durations ranging from <2 h to 9 days, or using embryo or larval life stages that are generally considered to be the most sensitive life stages.

Provincial water quality guideline derivation: BCMOE (1995) classified toxicity data as either acute toxicity, which refers to the results of short-term tests with toxicity endpoints that occur within 96 hours of exposure (e.g., less than or equal to a 96-h LC_{50}), or chronic toxicity which refers to tests with lethal or sublethal endpoints that exceed 96 hours of exposure duration. However, BCMOE (1995) notes that the normal longevity of the animal tested must be considered in this decision. For example, 96 hours is a relatively short time in the life cycle of most fish, whereas it may constitute most or all of the life cycle of some invertebrates or lower life forms.

USEPA water quality criteria derivation: USEPA (1985) classifies toxicity data as being "acute" or "chronic". Tests for daphnids, other cladocerans, or midges were deemed acute if the duration was near 48 h and the endpoint reported is either an EC_{50} for immobility or an LC_{50} for lethality. For embryos and larvae of crustaceans, molluscs and echinoderms (i.e., barnacles, clams, mussels, oysters, scallops, sea urchins, lobsters, crabs, shrimp, abalone), the test is considered acute if the duration is 48 to 96 h and the endpoint is an EC_{50} for incomplete shell development plus mortality. For all other freshwater or marine animal species, and older life stages, the test is considered acute if its duration is 96 h and the endpoint is an EC_{50} based on a combination of loss of equilibrium, immobility and lethality. Tests with single-celled organisms are not considered to be acute tests, even if the test duration is ≤96 h. Chronic toxicity data are defined as coming from life-cycle tests, except that partial life-cycle tests or early life-stage tests may be used for some fish species.

A summary of commonly-available toxicity tests, along with rationale for its designation as acute or chronic for the purposes of DERA is provided in Table 3.

Key Issues for the DERA Practitioner:

- Tests are defined as chronic only if the test duration represents a significant fraction (i.e., greater than 10%) of an organism's life cycle. These data are preferred for DERA purposes.

- Tests with a duration of less than 10% of the organism's life cycle, but measuring a sensitive stage of the life cycle should be properly described as a surrogate for chronic toxicity. These data are acceptable for DERA purposes, but the uncertainty associated with their use as a surrogate for chronic exposures should be noted.

- Tests with a duration of less than 10% of the organism's life cycle and not measuring a sensitive stage of the life cycle should be described as acute. Acute tests are valuable for screening purposes, but on their own, should not form the basis for concluding that effects are negligible in the DERA framework.

- The terms "lethal" and "sublethal" should also be used to describe the type of effect or endpoint being measured. Sublethal endpoints must be included, since lethality, on its own, should not form the basis for concluding that effects are negligible in the DERA framework.

4.3.4 Improving extrapolation from the lab to the field

Laboratory-based toxicity testing is advantageous for DERA purposes since it facilitates a standardized, quantifiable measure of adverse effects of field-collected samples to individual ROPCs. However, toxicity testing is also limited by the fact that its application requires an inherent extrapolation from the laboratory to the field. This extrapolation represents a source of uncertainty which cannot be avoided. However, the uncertainty can be reduced through the application of additional risk assessment "tools". The underlying intent of most tools is to conduct laboratory toxicity testing that more closely approximates site-specific environmental factors that influence COPC bioavailability, and biological factors that influence potential acclimation and adaptation to COPCs. Potential techniques (not a comphrensive list) are listed below:

Use of Site Water: Toxicity testing using site water for dilution instead of laboratory water will more closely approximate site-specific factors that influence bioavailability (e.g., pH, hardness; dissolved organic carbon concentration; major ion concentrations). This approach is similar to water-effect ratio testing (Jop et al. 1995)

Test Organism Acclimation: Several common metal COPCs are also essential elements, and therefore organisms used in toxicity testing could have increased sensitivity to these metals if they were cultured and/or acclimated in media with low metals concentrations. For example, Muyssen and Janssen (2002a,b) and Muyssen et al. (2002) found that culturing test animals (specifically, *Ceriodaphnia dubia* and *Daphnia magna*) in media deficient in zinc resulted in laboratory populations that were unnaturally sensitive to those same metals during toxicity tests.

Test Organism Adaptation: The ubiquitous nature of metals in the environment often leads to naturally-elevated background levels (i.e., in proximity to ore bodies); organisms have also evolved adaptive mechanisms to thrive in those areas. The ability of organisms to adapt to high concentrations of metals is not currently integrated or even considered in existing regulatory frameworks (Janssen et al. 2000). Adaptation of test organisms to these natural background concentrations of non-anthropogenic substances such as metals

and PAHs should be considered where appropriate (e.g., use field collected organisms for toxicity testing, especially for DERAs for metalliferous areas.

Key Issues for the DERA Practitioner:

- Toxicity test methods can be extensively modified, if and as appropriate, to reflect site-specific issues regarding COPC bioavailability, acclimation and adaptation.

4.4 Deriving Toxicity Reference Values for Food Chain Models

The selection of the toxicity reference value represents the effects assessment phase for wildlife food chain models. Selecting a TRV that appropriately balances conservatism and ecological realism is an essential step for the appropriate application of food chain models in DERA (Tannenbaum et al. 2003; Tannenbaum 2005). Understanding the inherent uncertainty in TRV derivation is also necessary.

4.4.1 Level of effort

TRVs proposed by Sample et al. (1996) were intended for screening-level ERA only, and therefore, the risk assessor should consider if refinement of the existing TRV derivation or development of a site-specific TRV is appropriate. Limitations of the Sample et al. (1996) TRV derivation approach (with recommendations for refinement) are discussed further in McDonald and Wilcockson (2003). A substantial literature search effort is required to derive site-specific TRVs, except for those few instances where existing compendia of toxicity data are adequate and appropriate for deriving TRVs (e.g., the ECO-SSL documents produced by USEPA).

4.4.2 Appropriate toxicological endpoints

Existing provincial risk assessment guidance (BCMELP 1997) states that acceptable toxicological endpoints include reproduction, growth, lethality, and tumour formation or other gross deformities in embryos and young, while subcellular responses (e.g., enzyme activity, DNA breakage, haematological parameters) are not suitable for risk assessment purposes. Subcellular responses may in fact be appropriate for DERA applications at certain sites, provided that the science is adequately well-developed to demonstrate how the subcellular response has resulted in an unacceptable adverse effect. TRVs for DERA purposes should be based on chronic data where possible since the use of acute toxicity data requires multiple uncertainty factors (UF) that may result in unrealistic risk estimates.

4.4.3 Permissible level of effects

BCMELP (1997) specifies the following permissible levels of effects for avian and mammalian ROPCs as a function of land use:

- Industrial or commercial land uses: EC_{50};
- Residential land uses: EC_{20}; and
- Urban park or agricultural land uses: EC_{10}.

A rationale for these policy decisions is provided in the "Tier I ERA Policy Decision Summary" (BCMELP 2000). As a last resort, BCMELP (1997) recommends that LOAEL values should be used without additional uncertainty factors for all land uses. TRVs for DERA should incorporate an EC_x-based approach wherever possible, which often requires that risk assessors retrieve and reanalyze the original mammalian and avian toxicological literature used as the basis for the TRV. Graphical interpolation of the EC_{50}, EC_{20} and EC_{10} values will be necessary for the majority of studies since historical investigations are unlikely to report data for individual replicates; additionally, the available statistical power in the original study may not support calculation of the EC_{10} or EC_{20} thresholds.

If data to support an ECx-based TRV are not available, we recommend that LOAEL-based TRVs should be used for common wildlife species (which recognizes the policy inherent in the use of an EC_x value that it is permissible to impact a wildlife population to a limited degree). However, for rare, threatened or endangered species, a NOAEL-based TRV should be used (since these species are protected at the individual organism level rather than at the population-level). Note that a policy decision with respect to what constitutes a permissible level of effects for rare, threatened or endangered species is not currently available.

4.4.4 Uncertainty factors

McDonald and Wilcockson (2003) noted that TRV derivation involved the use of multiple default UFs. For example, Sample et al. (1996) used an UF of 10 to extrapolate from subchronic (test duration less than one year) to chronic (test duration greater than one year), and a second UF of 10 to extrapolate from LOAEL-based TRVs to NOAEL-based TRVs. Existing provincial risk assessment guidance (BCMELP 1997) suggests that a LOAEL-based TRV that was derived from a feeding study measuring reproductive, growth, lethality or deformity endpoints would not require any additional uncertainty factors[32]; however, a UF of 10 should be applied if the laboratory and wildlife species "are not so closely related".

4.4.5 Allometric scaling

Mammals: Allometric scaling relies on an assumption that toxicity is dependent on body weight; this assumption is based on the underlying relationship between body weight and metabolic rate. In general, smaller organisms have higher metabolic rates as a function of body weight, which influences other toxicokinetic variables linked to metabolic rate (e.g., blood flow, renal clearance, respiration rate; metabolic half-life) (Bachmann et al. 1996; Kirman et al. 2003; Savage et al. 2004). The relationship between field metabolic rate and body weight has been well-documented, and is found to approximate a value of ¾ (Nagy et al. 1999; Savage et al. 2004). Other authors have argued that the scaling factor is closer to ⅔ (e.g., Dodds et al. 2001); however, these relationships were based on basal metabolic rates and earlier assumptions that metabolism was a function of surface area instead of body weight (Savage et al. 2004).

Allometric scaling (specifically, a scaling factor of ¾) was able to explain a substantial fraction of the variation in acute mammalian toxicity data sets (Goddard and Krewski 1992; Travis and White 1988), which is not surprising since toxicity is also dependent on toxicokinetic variables. Sample et al. (1996) converted TRVs expressed in terms of mg COPC/kg body weight/day) using the scaling factor of ¾. Sample and Arenal (1999) subsequently reexamined the ability of default scaling factors (e.g., 1, 0.75 and 0.66) to explain variations in acute toxicity data. Sample and Arenal (1999) calculated a mean scaling factor of 0.94 ± 0.03 (range: -0.15 to 1.69) for mammalian species based on a broader variety of compounds than previously examined; however, the majority of compound-specific scaling factors were not statistically different than any of the existing default scaling factors (0.66, 0.75 or 1). Sample and Arenal (1999) concluded that default scaling factors were appropriate for drug compounds (e.g., the data originally used to evaluate rodent-to-human scaling factors), but might not be applicable for all classes of compounds. However, Kirman et al. (2003) used physiologically-based pharmacokinetic

[32] BCMELP (1997) notes that preference should be given to studies with a duration of "weeks to months"; however, this is not an explicit requirement.

(PBPK) modeling to demonstrate that the ¾ scaling factor was applicable over a broad range of compounds other than drugs[33].

Birds: There are limited and contradictory data regarding the selection of an appropriate scaling factor for avian species. Nagy et al. (1999) calculated a scaling factor of 0.681 based on an analysis of field metabolic rates for 95 bird species. However, a single scaling factor for all birds may not be appropriate, given the likely differences in energy requirements for various avian species (e.g., passerine versus non-passerine). An examination of acute avian toxicity data (pesticides) failed to support the use of the ¾ or ⅔ scaling factors used for mammalian toxicity data; in fact, a scaling factor greater than 1 was proposed (scaling factor of 1.2; Mineau et al. 1996). Sample and Arenal (1999) found a mean scaling factor of 1.19 ± 0.05 (range: 1.16 to 3.09) was determined for avian species, which was consistent with the 1.2 scaling factor proposed by Mineau et al. (1996). No scaling factors were used for avian TRVs derived by Sample et al. (1996).

4.4.6 Recommendations for TRV Derivation

Existing provincial risk assessment guidance (BCMELP 1997) recommends against allometric scaling of TRVs, and instead, suggests that an uncertainty factor of 10 should be used to derive TRVs for "not so closely related" species. BCMELP (1997) suggests that uncertainty factors should not be used for closely-related species (for example, all rodents are considered to be closely-related, as are all waterfowl). Specific guidance for what constitutes closely-related species is not available.

The implications of allometric scaling versus uncertainty factors for deriving wildlife TRVs is a topic of ongoing scientific debate[34] for which a clear consensus has not yet been developed. Both approaches address the same underlying issue: toxicity data are rarely available for relevant wildlife species, and therefore, risk assessors are forced to rely on data for common laboratory species (e.g., mouse, rat, quail, chicken) which are often smaller than the wildlife receptors of potential concern. Further debate to develop a policy on this issue is recommended; however, the following information is provided for consideration:

- Compound specific allometric scaling factors are available for some compounds (Sample and Arenal, 1999), and are superior to using a generic scaling factor.

[33] Compounds tested by PBPK modeling by Kirnan et al. (2003) included benzene, carbon tetrachloride, chloroform, ethanol, ethylene oxide, methylene chloride, methylmercury, tetrachloroethene and vinyl chloride.

[34] Risk assessment guidance for the use of allometric scaling includes documents from: Oak Ridge National Laboratory (Sample et al., 1996); Great Lakes Water Quality Initiative (USEPA 1995), as well as the Total Risk Integrated Methodology model (USEPA 2005b). Guidance documents recommending against the use of allometric scaling include the ECO-SSL approach (USEPA 2005a) and BCMELP (1997).

- Physiological differences between different taxonomic groups are often cited as a major argument against the use of allometric scaling factors. Toxicity data should be from species with similar gastrointestinal physiology wherever possible; allometric scaling should not be used to extrapolate between distant taxonomic groups. Extrapolation between mammals, birds and amphibians should be avoided.

- Default uncertainty factors of 10 have minimal scientific basis; the problem is greatly compounded when multiple default uncertainty factors are applied.

Key Issues for the DERA Practitioner:

- Applying multiple, default UFs of 10 to the TRV derivation is inappropriate for DERA purposes since the purpose of the DERA is to emphasize ecological realism. Multiple UFs may be appropriate for a screening-level ERA, provided that they are not applied as a substitute for conducting an appropriate literature search.

- Risk assessors need to consider the overall uncertainty in the TRV derivation process, and either: a) apply a single UF (preferably not a default value); or b) use allometric scaling (preferably, a compound-specific factor). A discussion of the uncertainty in the TRV derivation process should be provided, irrespective of the option selected.

- Scientific consensus on uncertainty factors versus allometric scaling is not available at this time. An uncertainty factor approach provides a more conservative TRV, but does not necessarily improve the certainty in the risk estimate. An allometric scaling approach may provide a less conservative TRV (for wildlife species larger than the laboratory test species), but again, does not necessarily improve the certainty of the risk estimate.

4.5 Site Observations and Field Surveys

A site observation method is currently incorporated in the Tier-1 risk assessment guidance which is intended to "determine if plants and animals actually occur on site and whether or not these plants and animals show any obvious signs of toxicity". However, site observations regarding the presence/absence of specific plants or animals are more appropriate to the problem formulation phase of the DERA. Qualitative assessments of whether signs of toxicity are present (based on a question-based checklist) are not appropriate for DERA purposes. It would be an error, for example, to attribute bare patches of ground to phytotoxicity without consideration of soil type and level of disturbances (e.g., trampling, soil compaction). The remaining questions listed for the site observation method described in the available Tier-1 guidance manual focus primarily on habitat quality related issues which are also more appropriate to the problem formulation phase (i.e., they are not measuring effects related to site COPCs).

However, field surveys for measuring the potential magnitude of effects associated with COPCs are highly relevant to the objectives of the DERA. For example, measures of plant community characteristics can add substantially to the understanding of impacts to the plant community. Potential measures of effects include: biomass, dominant species, presence of sensitive species, structural stage, percent cover, and other biophysical characteristics such as soil type or moisture holding capacity. In general, field studies provide a level of ecological realism not readily attainable in laboratory studies, but multiple stressors frequently make it difficult to identify a particular stressor as the cause of observed ecological effects (USEPA 1993a). Consequently, field surveys for DERA purposes need to consider the following:

- Study designs need to be appropriate to achieve the desired statistical power, both in terms of sample locations (e.g., stratified or random sampling) as well as sample number. Assistance from a statistician is recommended. Note that for small sites, statistical power considerations may be less of an issue since the sampling program effectively samples all portions of the site.

- Field surveys for DERA purposes often involve comparison between impacted and reference locations. Selection of appropriate reference locations requires considerable project resources (i.e., it is necessary to document that the sites are consistent in all respects with the exception of the contamination). A reference envelope approach (i.e., the use of multiple reference locations to define the range of acceptable conditions) rather than basing the comparison on a single reference location is encouraged. Gradient designs are also beneficial, especially in those instances where obvious reference locations are not evident.

- The data to be collected from the site survey must reflect the assessment and measurement endpoints, risk hypotheses and decision criteria established in the problem formulation.

Key Issues for the DERA Practitioner:

- Field surveys need to be designed to address statistical power, the use of reference sites, and should be conducted by experienced biologists/ecologists (preferably with regional expertise).

- Qualitative site surveys are not appropriate for a DERA, except for those few instances where a more robust survey cannot be conducted. In effect, the qualitative survey represents a professional judgment; the ability to make a credible professional judgment regarding site effects (or lack thereof) based on a qualitative survey would be largely dependent on the level of effort involved and the credentials of the personnel involved.

4.6 Ecosystem-Specific Issues for Consideration

4.6.1 Deep Aquatic Ecosystem

A potential issue for consideration for the effects assessment for deep aquatic DERAs is the use of porewater toxicity testing. Porewater toxicity tests have been described as advantageous due to the tests' increased sensitivity to chemical contaminants, overall ecological realism and their ability to avoid confounding factors (e.g., grain size) common to whole-sediment toxicity tests) (Carr et al., 2001; Carr and Nipper, 2003). The increased sensitivity has described as follows:

- Porewater toxicity testing provides "an indication of potential sublethal effects which could otherwise not be analyzed" (Nipper et al., 2002); and

- "Porewater toxicity testing may be an order of magnitude more sensitive than whole-sediment toxicity testing, which allows for further investigation for those sediments that may be causing more complex changes to the benthic community." (Carr et al., 2001)

Other authors have cautioned that porewater toxicity testing has many inherent liabilities that may limit its utility for routine sediment quality investigations (e.g., Chapman et al., 2002a). Side-by-side comparisons of porewater and whole-sediment toxicity, although limited, indicate that toxicity is greater in porewater samples but linked primarily to ammonia rather than site-specific COPCs (Burgess et al., 1993; Anderson et al., 2001; McDonald, 2005). Ho et al. (2002) suggested that the increased influence of ammonia (relative to whole-sediment toxicity testing) may be an artifact of the test system (i.e., ammonia is water soluble, and therefore more likely to result in over-exposure in a porewater sample).

Key Issues for the DERA Practitioner:

- Porewater toxicity testing for DERA should: a) evaluate the potential role of ammonia; and b) collect data for porewater COPC concentrations as a measure of exposure that is relevant to the measure of effect.

4.6.2 Shoreline Ecosystem

Potential issues for consideration for the effects assessment for shoreline DERAs are:

- Phototoxicity; and
- Groundwater plumes

4.6.2.1 Phototoxicity

Phototoxicity should be considered when designing DERAs for shoreline ecosystems impacted by known phototoxic compounds such as PAHs. PAHs (and other compounds), once accumulated into biota, have the ability to absorb ultraviolet light (UV) energy. These photoactivated compounds can damage cellular membranes, resulting in biological impairment and death. Severe PAH phototoxicity has been demonstrated to multiple taxa, primarily using laboratory-based exposure systems. However, "the unanswered question…into the phototoxicity of contaminated sediment [and water] is whether phototoxicity is of ecological relevance or merely an interesting laboratory artifact" (Boese et al., 1999). Diamond and Mount (1998) noted that the risk from PAH phototoxicity depends on the "likelihood [for organisms accumulating PAH that can be photoactivated] of receiving activating solar radiation", and therefore, quantifying the UV exposure is equally as important as quantifying the PAH dose. The traditional practice of evaluating phototoxicity using laboratory-based toxicity tests has minimal ecological realism for the following reasons (McDonald and Chapman, 2002):

- UV doses in laboratory experiments are generally maximized by the use of environmentally unrealistic light sources (e.g., inappropriate photoperiods, wavelength distribution, and intensity);

- Attenuation of light in the water-column is minimized due to lower amounts of humic acid, dissolved organic carbon, and total suspended solids which absorb or block UV transmittal; and

- Laboratory exposure systems also prevent test organisms from utilizing behavioural adaptations, such as the utilization of refugia, which minimize the internal UV dose; laboratory-cultured organisms also lack resistance and/or tolerance mechanisms that may be present in natural populations.

Key Issues for the DERA Practitioner:

- Laboratory-based toxicity tests are not recommended for investigating the potential effects associated with phototoxicity unless steps are taken to improve the ecological realism of issues such as UV doses, light attenuation and refugia in the toxicity testing. Laboratory-based toxicity testing without these modifications grossly overestimates effects.

- Incorporation of *in situ* toxicity testing, or additional field based risk assessment tools (e.g., benthic community measurements; recolonization experiments) in the DERA is recommended.

4.6.2.2 Groundwater Plumes

Effects assessments for groundwater plumes frequently involve aquatic toxicity testing. The following modifications for toxicity tests designed for effluents and surface water should be considered if applied to groundwater samples:

- The objective of the DERA is to characterize effects at the point of discharge to the receiving environment; however, groundwater samples are normally collected from upland sites located at a distance from the receiving environment. The dilution series for the toxicity test should reflect the range of likely groundwater concentrations at the point of discharge, as determined by site-specific groundwater modeling (rather than assuming that a 10-fold attenuation exists).

- Regardless of the dilution series selected, the test should always include the maximum possible concentration. For freshwater sites, the maximum test concentration will be 100% groundwater. For marine sites, the maximum test concentration will vary from approximately 70 – 100%, depending on the amount of hypersaline brine needed to adjust the groundwater salinity to the surface water salinity.

- Toxicity testing requires that samples be well-oxygenated and have pH values that are capable of supporting aquatic life (typically pH 6.5 – 8.5). Sample manipulations to achieve the necessary test conditions may also alter contaminant bioavailability, and thus, represent a source of uncertainty in the toxicity data.

Site surveys of the groundwater discharge areas may also provide useful information regarding potential effects (e.g., measure the diversity and abundance of organisms in the discharge pathway). Note that soft-bottom benthic community surveys are not appropriate as a measure of effect for groundwater discharges since the benthic community reflects exposure to sediment-associated contaminants, not groundwater.

4.6.3 Upland Wildlands Ecosystem

Potential issues for effects assessments in uplands wildlands DERAs are:

- Consideration of the spatial scales of the available exposure data relative to the spatial scale of different terrestrial ROPCs; and

- Consideration of the potential for indirect effects.

4.6.3.1 Spatial Scale of Exposure Data

Mammalian and avian ROPCs: Spatial considerations are important with respect to the exposure data considered in a food chain model, including:

- The definition of a single "reasonable worst-case" soil concentration. Integration of multiple individual soil values is appropriate since the ROPCs are mobile (and assumed to move equally around all portions of the site). This assumption is used to simplify the food chain model but can be replaced by a habitat-weighted food chain model if needed (i.e., the ROPC preferentially spends more time in certain areas based on habitat quality considerations).

- Food chain models incorporate highly conservative soil concentrations (i.e., 95% upper confidence limits of the mean; 90th percentile)[35]. This conservative assumption is based on existing policy rather than science and is the primary factor mitigating against the use of other additional uncertainty factors in a food chain model (see Section 4.4.5).

In the event that the spatial coverage of the available soil data is adequate, alternatives to the use of a single reasonable worst-case soil concentration are recommended. Examples include the curve model described in BCMELP (1997), or the construction of a spatially explicit food chain model incorporating GIS software.

Soil Invertebrate and Plant ROPCs: Unlike birds and mammals, soil invertebrates and plants are immobile or have low vagility, and therefore, the use of site-wide "reasonable worst-case" concentrations for the exposure or effects assessment is inappropriate. If toxicity testing is used, risks to soil invertebrates and plants should be determined for individual soil samples collected on an appropriate scale (e.g., an individual soil sample is only representative of a very small area[36]). Compositing of multiple soil samples for the effects assessment is therefore problematic. Selection of specific locations to appropriately represent the range and mixture of contaminants present at the site (a decision typically based on the available site characterization data) is critical. Consequently, the spatial scale of effects data for soil invertebrate and plant ROPCs tends to be greater (i.e., more sampling per unit area) than for effects data for mammalian and avian ROPCs.

[35] Existing risk assessment guidance (BCMELP 2000) requires use of the lower of the 95% UCLM or the maximum COPC concentration.

[36] No specific guidance on what constitutes the appropriate area is available; however, it likely ranges from 1 to 25 square meters, depending on the heterogeneity of the contaminant and the potential range of exposure for the specific ROPCs (e.g., considering root networks; movement of soil invertebrates, etc).

4.6.3.2 Indirect Effects

Indirect effects occur when a toxicant-related effect on one species causes an indirect effect on a second species due to altered ecological interactions such as predation, competition or resource availability. Examples of indirect effects include:

- Indirect effects on a passerine bird population may occur as a result of changes in habitat availability (e.g., a soil COPC is phytotoxic, which in turn reduces forest cover and thereby changes the habitat).

- Indirect effects on an insectivorous small mammal population may occur since soil invertebrates may avoid areas with elevated soil COPC concentrations, thereby altering food availability.

- Indirect effects on fish populations occur as a result of a change in the zooplankton community (thus reducing food availability) associated with elevated water COPC concentrations.

Preston (2002) argued that single-species toxicity testing does not necessarily capture the complexity of the potential effect on an ecosystem-level effect. Numerous risk assessment tools are intended to compensate (at least partially) for this limitation, including: a) the use of a battery of toxicity tests; b) mesocosm toxicity testing; c) species sensitivity distributions; and d) integrating the results of field surveys with toxicity data using a weight-of-evidence approach. Increased consideration of the complexity of a site's ecology and the multitude of factors that drive an ecosystem's response to a chemical stress is recommended to address the implications of indirect effects; however, to date, minimal guidance on how to incorporate indirect effects into a risk assessment framework is available.

4.6.4 Rivers and Streams Ecosystem

No ecosystem-specific implications for the effects assessment are currently identified. See guidance in Section 4.6.2.2 if groundwater discharges to rivers and streams are being assessed.

4.6.5 Upland Human-Use Ecosystem

No ecosystem-specific implications for the effects assessment are currently identified. See Section 4.4.3 for discussion on permissible level of effects with respect to different potential land uses for this ecosystem type.

5.0 RISK CHARACTERIZATION

Risk characterization is the process of estimating the magnitude (and where possible, the probability) of adverse ecological impacts based on the information obtained from the exposure and effects assessments. Risk characterization provides the discussion of the "strengths, limitations and uncertainties arising from the data and models used to provide conclusions" (CCME, 1996) and accomplishes the following objectives:

- Risk characterization demonstrates how the results from multiple tools are integrated into a conclusion for each individual line of evidence, and how the conclusions from multiple lines of evidence are integrated into an overall conclusion regarding ecological risks. This integration is necessary to maintain transparency of the risk assessment process.

- Risk characterization requires that conclusions are presented in a clear and unambiguous manner (i.e., conclusions are stated in plain-language). The tendency for technical reports to obscure conclusions using jargon should be replaced by clear statements of what was estimated (and how). Emphasis on clarity in the risk characterization is necessary so that the DERA can be used by site managers in their decision making process.

- Risk characterization also requires that the uncertainty in the conclusions be discussed—again, the goal is provide site managers with information needed for site planning purposes.

This section focuses on the following three tools that are typically used in the risk characterization phase:

- Hazard quotients;
- Multivariate statistical analyses; and
- Weight-of-evidence approaches.

Additionally, guidance is provided regarding the application of best professional judgment, the appropriate terminology used to narratively describe risk estimates, and the role of uncertainty analysis in the DERA process.

5.1 Quotient Methods

Hazard quotients (HQs) are widely used in DERA due the prevalence of literature-based toxicity data and food chain models (Section 4.4). However, HQs measure hazard (as the name implies) rather than the classical definition of "risk" (i.e., they do not contain information about the probability that an adverse effect will occur). HQs are also subject to the following considerations:

- Quotient methods are only as reliable as the values in the numerator and denominator (with associated uncertainty).

- Quotient methods assume that both the numerator and denominator exist in all locations and all occasions (when, in fact, environmental concentrations are variable). The use of point estimates for the numerator and denominator mask the underlying uncertainty and variability in the data.

- HQs are not proportional to the magnitude of "risk". Although a very large HQ demonstrates a greater "risk" than a HQ slightly greater than 1, it is not true that minor changes in the HQ provide a meaningful differentiation (Ritter et al., 2002).

- The number of significant figures in the HQ should reflect the lowest number of significant figures in the numerator or the denominator. Although the ratio has an unlimited number of decimals, inclusion of excessive decimals implies a level of certainty that is not actually present. Most HQs can be rounded to the nearest whole number.

Key Issues for the DERA Practitioner:

- Hazard quotients do not provide a measure of risk. Hazard quotient approaches, if used as a line of evidence in a DERA, should be supplemented by other methods that provide more information about the magnitude and/or probability of adverse effects.

5.2 Multivariate Statistical Analyses

Multivariate statistical analysis refers to any of various statistical methods for analyzing more than two variables simultaneously. Assessing effects at a community or ecosystem levels usually involves measuring a large number of abiotic and biotic variables. Assessing each variable individually or with many pairwise bivariate analyses can be cumbersome, difficult to interpret, and cannot detect patterns that emerge from the interactions of variables. Multivariate techniques can be used to draw overall patterns from a large set of variables. Multivariate techniques can also be invaluable in displaying these patterns and communicating them to a non-technical audience.

There are three broad types of applications for multivariate techniques: ordination (data reduction), classification (clustering and discrimination), and canonical ordination (investigating relationships between sets of variables).[37] Appendix III, Section 6 provides an overview of the common multivariate statistical approaches and identifies potential pitfalls. See Sparks et al. (1999) for more information on specific techniques as they have

[37] Bayesian approaches provide alternative methods for statistical analyses that explicitly incorporate uncertainty. Specific guidance for application of Bayesian approaches in a DERA is not available at this time. Risk assessors should consult a statistician for further information.

been applied to risk assessment; additionally, a statistician with experience in biological or ecological investigations should be consulted as needed[38]. Note that the application of specific statistical techniques is subject to ongoing research, and therefore, the techniques listed below are meant only to illustrate the range of likely approaches. Selection of different statistical techniques will be study- and data set-specific.

5.3 Weight-of-Evidence (WOE) Assessment

All pollutants are contaminants, but not all contaminants are pollutants[39] because substances introduced into the environment may be more or less bioavailable to organisms depending on their chemical form, modifying factors in the environment, the environmental compartment they occupy, and the reactions (behavioural and physiological) of exposed biota (Chapman et al., 2003). Accordingly, determining when contamination has resulted in pollution requires not only chemical but also biological measurements (i.e., both exposure and effects assessment).

Since there are no perfect tools for determining pollution (e.g., we cannot measure all possible contaminants, run all possible tests, or determine the health of all organisms), risk assessments require that the results from multiple tools be integrated into a single conclusion regarding the likelihood and magnitude of ecological risks[40]. This integration is normally accomplished using a WOE assessment framework (Chapman et al. 2002b) that evaluates possible ecological risks based on appropriate, multiple lines of evidence (LOE). Although concurrent measurement and simultaneous consideration of multiple LOE are common, WOE-type approaches using a more linear approach are also available (e.g., sequential analysis of lines of evidence [SALES]; Hull and Swanson 2006). The manner to which the WOE incorporates different LOE (i.e., in sequence or simultaneously) is dependent on the study design (see Section 2.9.2).

WOE can be applied to any DERA for any environmental media, although the majority of WOE assessments to date address sediment quality issues. This is largely due to the evolution of the Sediment Quality Triad (e.g., Long and Chapman 1985; Chapman et al. 1997; Chapman 1990, 1996). Examples of WOE frameworks are provided for sediment (and other media) in Chapman et al. (2002b) and Chapman and McDonald (2005).

[38] For those risk practitioners without access to a statistician, Simon Fraser University offers statistical consulting services (www.stat.sfu.ca). Other college/university statistics departments, or other consulting firms may also offer similar services.

[39] Contamination refers to substances present where they would not normally occur, or at concentrations above natural background. Pollution refers to contamination that causes adverse biological effects in the natural environment.

[40] See Section 5.5 for a discussion of narrative descriptors of risk.

Examples of other WOE frameworks for non-sediment related assessments include:

- Johnston et al. (2002): WOE for an estuarine site which included LOE focused on pelagic fish, epibenthos, benthos, eelgrass, salt marshes and waterfowl.

- Sample and Suter (1999): WOE for piscivorous wildlife in a large river-reservoir system based on a literature-based food chain model, biomonitoring and field observations.

- Lowell et al. (2000): WOE for aquatic insects in large river systems based on a combination of field surveys, streamside artificial mesocosms, stable isotope analyses and bioindicators.

- Menzie et al. (1996): broad guidance for the construction of WOE assessments for DERAs.

5.3.1 Guiding principles

Guiding principles for all WOE assessments (irrespective of the environmental media under investigation) include:

- Lines of evidence incorporated in the WOE should include both: a) laboratory studies with individual organisms[41] and, b) field measurements of resident populations (Chapman and Hollert 2006). These different LOE provide complementary information that strengthens the ability of the WOE to make proper conclusions. Laboratory-based LOE provide the ability to measure contaminant-related effects under standardized conditions which reduce the influence of other non-contaminant related stressors, while field-based LOE capture information about adverse effects under realistic exposure conditions. The number and complexity of different risk assessment tools within each broad category of LOE can be (and should be) tiered.

- If the WOE indicates that adverse effects are present based on consideration of the laboratory- or field-based LOE, the risk assessor should consider implementing additional LOE to evaluate causation (e.g., *in situ* measurements of toxicity to assess differences between the laboratory and the field potentially related to tolerant field populations; measurements of contaminant body residues in organisms related to effects thresholds; chemical manipulations combined with laboratory toxicity measurements [TIE]). These causality investigations are often useful for resolving

[41] Inclusion of both laboratory and field studies is not possible in every instance (e.g., a WOE assessment of mammalian and avian ROPCs would not likely involve laboratory-based toxicity studies).

potential disagreements between different LOE. Criteria for evaluating the causality in other LOE can also be established (for example, see Lowell et al. 2000)[42].

- It is necessary to establish an *a priori* framework (to the extent possible) for integrating different LOE. The *a priori* framework should be agreed to by appropriate interested parties and should include a description of how the magnitude of response observed in each LOE and the concurrence among multiple LOE will be evaluated in terms of arriving at a risk estimate.[43] The use of an *a priori* framework means the data are fit to an agreed-upon decision-making framework, rather than the framework being fit to the data. This approach also matches the basic scientific paradigm of developing testable hypotheses prior to experimentation.

- WOE is not a static methodology. Its greatest strength is its flexibility in terms of the inclusion of different LOE to reflect the latest scientific knowledge and practices. The best available science should be used in applying any WOE assessment. Design and implementation of a WOE assessment reflects the experience of the scientists involved. Thus, WOE assessments also require suitable state-of-the-art expertise in the various disciplines comprising a particular assessment.

5.3.2 How to weigh different lines of evidence

WOE assessments need to be applied within the context of common sense; they should not be applied inflexibly. Critical to the WOE process are three factors: the weight assigned to each LOE; the magnitude of response observed in each LOE; and concurrence among multiple LOE (Menzie et al. 1996). The weight assigned to different LOE is determined as follows (Chapman and Anderson 2005):

- Chemistry data should not be used alone for decision-making except for "simple contamination where adverse biological effects are likely…when the costs of further investigation outweigh the costs of remediation, and there is agreement to act instead of conducting further investigations" (Wenning and Ingersoll 2002; Wenning et al. 2005).

[42] Lowell et al. (2000) established *a priori* causal criteria for evaluating different LOE in a WOE for northern rivers. Criteria included: spatial and temporal correlation; plausible explanation linking stressor and effect; experimental verification of stressor cause-effect relationship under controlled conditions; strength of the correlation, specificity of the effect to the COPC, evidence of COPC exposure in the body of the ROPC; consistency of association across other studies within the region and in analogous studies in other regions. Other examples of causality criteria are summarized in Lowell et al. (2000) and elsewhere in the literature.

[43] The complexity of the *a priori* framework is project dependent. For example, a terrestrial ERA might include upwards of 20 or more measurement endpoints (e.g., toxicity testing on multiple species in addition to different plant and soil invertebrate community metrics), which would make it difficult to establish the precise weighting of each different endpoint. However, it should still be possible to establish *a priori* what would constitute an unacceptable effect for each measurement endpoint, and to lay out general guidelines for how different types of data would be integrated.

- Greater weight must be applied to biological (effects) data than to exposure data.

- Within the effects data, LOE (e.g., laboratory toxicity tests, models) that contradict the results of properly conducted field surveys with appropriate power to detect changes "are clearly incorrect" (Suter 1996b) to the extent that those toxicity or model LOE are not indicative of adverse biological effects in the field. Conversely, data from field studies without appropriate statistical power should not be ignored, but rather, weighed appropriately in the WOE (along with toxicity, model or other LOE) depending on its strengths and limitations.

5.3.3 Numerical versus non-numerical ratings in WOE

The symbology of the WOE can vary from assessment to assessment. Numerical ratings for each measurement endpoint (or LOE) were proposed by Menzie et al. (1996), based on a set of eleven attributes scored between 1 and 5 based on *a priori* narrative criteria[44] (similar to causal criteria established by Lowell et al. 2000). The relative weight of each attribute was established on a scale between 0.0 and 1.0 based on a survey of 10 experienced risk assessors. The WOE was based on the sum of the (quality score x relative weight) scores. WOE frameworks proposed by Chapman and coauthors (e.g., Chapman et al. 2002b) used non-numerical rating systems (e.g., "○", "⊙", "●"). The specific symbols used in the WOE are not relevant.

Numerical ratings should only be used if the risk assessor can make meaningful differentiations between varying magnitudes of effect within a LOE as well as the relative weight between different LOE. Relative weighting systems such as those used by Menzie et al. (1996) are suitable; however, Menzie et al. (1996) emphasized that weighting systems must reflect collective professional judgment to minimize the influence of bias. When these conditions cannot be satisfied, numerical ratings are not recommended since they: a) likely reflect arbitrary and subjective differentiations; and b) imply a level of precision that is not actually present (i.e., a score of 5 is worse than a score of 6 when in fact the uncertainty in the LOE means that both scores are functionally equivalent with respect to management planning). Non-numerical rating systems are recommended under these circumstances.

[44] Attributes were: strength of association between measurement endpoint and assessment endpoint; site-specificity; stressor specificity; quality of data; availability of an objective measure for judging harm; sensitivity of the measurement endpoint to detect change; spatial representativeness; temporal representativeness; ability for the endpoint to be expressed quantitatively; correlation of stressor to response; and use of a standard method.

5.3.4 Using WOE in DERA

In summary, WOE assessments must be: a) objective, b) transparent, and c) scientifically rigorous as appropriate to the level of certainty needed for site management purposes. No specific framework is proposed (since the framework should be study- and site-specific), provided that it meets these criteria. WOE assessments provide the best means for risk characterization of environmental stressors (not restricted to just chemical contamination). They can be designed to address site-specific considerations as well as both localized and regional risks. And, because their findings can be made readily understandable to interested parties, they provide not only the best possible data for decision-making, but also a very high likelihood that the risk characterization will be accepted and used in subsequent risk management determinations.

5.4 Incorporating Professional Judgment

Professional judgment plays a major role in the DERA framework. Selection of COPCs, ROPCs and exposure pathways requires a degree of professional judgement. The construction of measurement endpoints and risk hypotheses (as well as selecting the risk assessment tools to test the hypotheses) also requires professional judgment based on education and experience, as does the interpretation of effects data (and if applicable, integration within a WOE framework). Risk hypotheses are accepted when the evidence in favour of the hypotheses is considered sufficient, and rejected when the evidence is not in favour of the hypotheses (or deemed insufficient). Some data will have decision points established by regulatory policy (i.e., a 20% reduction or greater in aquatic toxicity endpoints is considered unacceptable according to BCMELP 1997), while others can utilize statistically significant differences[45]. However, some data always require the risk assessor to judge whether they are evidence of an adverse effect or not.

Professional judgment is essential to risk assessment since the goal of the risk assessment is not limited to identifying those substances that are *scientifically proven* to be harmful, but also those substances for which there is scientific evidence that they *may* be harmful (Wandall 2004). Wandall (2004) argued that proper application of professional judgment in risk assessment required that (a) risk assessors are aware of what values[46] they are (often implicitly) relying on, (b) the values are justifiable, and (c) transparency is ensured. This requirement for transparency is the foundation of properly applied

[45] Note that decisions regarding statistical significance that rely on $p > 0.05$ also involve professional judgment, albeit to a lesser degree (since this judgment is a widely-accepted convention).

[46] Values refer to the attributes of "doing good science" rather than consideration of political or socioeconomic factors. Examples of values that scientists apply when creating and testing hypothesis include: ability of the hypothesis to explain the available data; simplicity of the hypothesis itself, fidelity of the hypothesis with other established facts; and whether a conservative burden of proof has been met (Wandall 2004).

professional judgment, and translates into the following guiding principles for applying judgment in the DERA framework:

- Risk assessors should determine if alternate or additional tools would provide data less reliant on professional judgment. Arguments against implementing these additional tools based on their cost or time required are not, on their own, sufficient to justify using professional judgment alone when alternate methods are available.

- All assumptions and decisions must be backed up with a rationale, especially for those instances where education and training (i.e., no citations are available) were used as the basis for the professional judgment.

- Declarative statements such as "the risk assessment found no evidence of impacts" should reflect where professional judgment was applied (e.g., "Our risk assessment, which was based on our professional judgment of XYZ data, found no evidence of impact").

5.5 Narrative Descriptors of Risk

Ideally, risk estimates should be a quantitative statement which includes a probability (e.g., "there is a 20% chance of 50% mortality"). However, few tools support the estimation of probability, and therefore, conclusions are usually presented as a qualitative statement (e.g., "there is a high chance of mortality occurring") (USEPA 1992). Provincial risk assessment guidance (BCMELP 1997) requires only that risk assessors provide an opinion regarding their results generated with respect to confidence, uncertainty and significance of impacts (a statement about probability is not required), As a result, hazard quotients are frequently used as a line of evidence in risk assessments despite the fact that they are not truly an estimate of risk (see Section 5.1). The emphasis on narrative descriptors of risk leads to situations where terms vary widely in application between different risk assessors.

The following operational guidance for the use of different narrative descriptors is provided:

- **Negligible risks:** Implies that adverse effects, based on the totality of available data, are not present, and that the risk assessor has high confidence that adverse effects will not be present in the future. This term should only be used in situations where multiple lines of evidence demonstrate a lack of adverse effects, and where each line of evidence (or the overall risk estimate) has relatively low uncertainty. Risk management or remediation is not necessary. Hazard quotients, if used as a line of evidence, tend to be less than one.

- **Low risks:** Implies that adverse effects are not present based on the totality of data available, but is different from the term "negligible" in that "low" risks is more

appropriate for situations where the conclusion is based on the balance of probabilities. Balance of probabilities suggests that adverse effects are unlikely to be present, although some data may indicate limited adverse effects, or the uncertainty is such that one cannot definitively exclude potential adverse effects in the future. Risk management or remediation is not necessary. Hazard quotients, if used as a line of evidence, tend to be less than one.

- **Moderate (or intermediate) risks:** Implies that some degree of adverse effects are likely, based on the totality of data available. Hazard quotients, if used as a line of evidence, are between 1 and 100 (BCMELP 1997), although some risk assessors may opt to subdivide this category into "moderate" for hazard quotients between 1 and 10, and "moderate-high" for hazard quotients greater than 10. Risk estimates suggest that risk management or remediation is necessary, unless further refinement of the risk estimate is conducted. Alternate term: "potential risks".

- **High (or severe) risks:** Implies that adverse effects are likely (and of relatively high magnitude) based on the totality of data. Hazard quotients, if used as a line of evidence, are greater than 100 (BCMELP 1997). Risk estimates suggest that risk management or remediation is necessary, and that this conclusion is unlikely to change even if further refinement of the risk estimate is conducted. Alternate term: "probable risks".

5.6 Uncertainty Assessment

Risk assessment involves estimation, extrapolation, and the use of models and assumptions which generate uncertainty in risk estimates. The following sources of uncertainty are identified within the context of providing operational guidance regarding uncertainty analyses for detailed ERAs.

- **Parameter uncertainty:** refers to missing or ambiguous data resulting from inadequate sampling, analytical errors, or lack of site-specific data. Note that parameter uncertainty is not the same as parameter variability; the variability in a data set can be characterized and evaluated, but it cannot be reduced.

- **Structural (or model) uncertainty:** refers to gaps in understanding or scientific theory on which models are based[47], although models can be improved as they incorporate more precise and site-specific physical, chemical, and ecological information. Inappropriate application of generic models results in structural uncertainty.

[47] "All models are wrong; some models are useful." George E.P. Box, statistician

5.6.1 Assessing uncertainty in risk estimates

Risk assessment exists to support sound management decisions. The uncertainty analysis is intended to make the risk assessment process more transparent by acknowledging and, to the extent possible, quantifying the uncertainty in the risk estimate. An incomplete uncertainty assessment is problematic since it contributes to a false sense of confidence regarding both the accuracy and the precision of the risk estimate. Identifying sources (and where possible, the magnitude) of uncertainty accomplishes two objectives: a) it helps decision-makers determine whether additional information should be obtained prior to making a decision, and b) provides a qualitative context for each particular risk estimate. Uncertainty analyses for DERAs should incorporate the following considerations:

1. **Identify and characterize sources of uncertainty**. Describe what is known and what is not known. Are we dealing with something that is unknowable, or about which we are totally ignorant? What would it take to reduce the uncertainties? Some uncertainties can be reduced and some cannot.

2. **Quantify uncertainty in the risk estimate.** Quantitative uncertainty analysis (e.g., with probabilistic methods such as Monte Carlo simulations and probability bounds analysis[48]) allows the assessor to see where further study is needed or where decisions can be made in the presence of uncertainty. In general, quantifying the uncertainty for for models or other highly quantitative risk assessment tools involves the following steps:

    - List all uncertain parameters (include additional parameters if necessary to represent uncertainty in model structure), and determine the maximum range of potential values for each uncertain parameter.

    - Determine a probability distribution for values occurring within this range. Consider correlations among parameters (e.g., if a maximum value is likely for one parameter, then what would be the likely values for other correlated parameters?). The objective of this step is to avoid having all parameters set to a maximum if such a scenario is ecologically irrelevant or otherwise impossible.

    - Propagate the uncertainty in the model parameters to produce a probability distribution of model predictions, and prepare quantitative statements of

[48] Monte Carlo methods are appropriate when input distributions are known precisely; however, they may not adequately represent the effects of uncertainty about how to parameterize variability in the input distributions. Probability bounds analysis is a tool for separating variability and uncertainty to obtain bounds on the result that explicitly account for uncertainty about the input distributions. As in Monte Carlo analysis, the overall slopes of the bounds indicate how much variability exists in the system. The distance between the bounds, on the other hand, is an indication of the uncertainty that exists due to lack of knowledge (i.e., incertitude).

uncertainty in terms of a confidence interval for the risk estimate that reflects the range of parameters used to calculate the risk estimate.

- Rank the parameters contributing most to uncertainty in the model prediction by performing a sensitivity analysis.

3. **Describe uncertainty in the risk estimate.** A quantitative approach to uncertainty analyses is preferred; however, it may not be possible (or appropriate) in all instances. A qualitative approach that follows the same logic as described above is recommended for those lines of evidence that do not lend themselves to the quantitative method described above. The qualitative approach involves a narrative description of: a) which lines of evidence were used in the risk estimate; b) how the results from individual areas or samples were integrated into an overall site-wide risk estimate; c) how different lines of evidence relate to one another; d) what the risk estimates could be if the worst-case values from individual areas or samples were used; and e) which of the lines of evidence had the greatest influence on the risk estimate. Note that some lines of evidence would also have a separate quantitative uncertainty analyses (e.g., statistical power of a toxicity test can be examined) in addition to the qualitative uncertainty analysis of the overall risk estimate.

5.6.2 When to refine risk estimates

The qualitative uncertainty assessment influences the appropriate degree of precaution with respect to the need to conduct additional investigations to reduce uncertainty. A matrix based on varying levels of estimated risk and uncertainty (based on Persons and Hopley 1999) is proposed:

	Low Magnitude of Risk	**High Magnitude of Risk**
Low Uncertainty in Risk Estimate	Low Precaution	Medium Precaution
High Uncertainty in Risk Estimate	Medium Precaution	High Precaution

Refinement of risk estimates for the “high” category of precaution is recommended; the “medium” category of precaution may also indicate a need to reduce uncertainty as necessary to support management actions. This refinement may involve one or more of the following strategies:

- Reduce parameter uncertainty by gathering additional data. Supplemental data collection should be targeted to deal with the underlying cause of the parameter uncertainty (e.g., address spatial coverage, improve analytical detection limits).

- Reduce structural (model) uncertainty by adopting a more appropriate model (or increasing the sophistication of the existing model). Risk assessment should be an iterative process where new data may require reassessment of previous approaches or conclusions. This iterative process allows risk assessment to be a dynamic process well suited to ecological study, and does not represent a failure of the initial risk estimate.

- Provide risk managers with multiple risk scenarios for consideration as a series of risk estimates with different assumptions and descriptions of uncertainty.

Several other strategies are often employed; however, they do not directly reduce parameter or model uncertainty. For example:

- Professional judgment is often used to fill in gaps in model structure. This may reduce uncertainty, but it may not, and there is likely no way to know. Conservative assumptions are often used as part this strategy; although it does not reduce uncertainty, it ensures that the majority of the uncertainty errs on the side of caution. The challenge in using conservative assumptions lies in balancing conservatism and ecological realism relative to site management needs.

- Increase the number and types of lines of evidence considered in a weight of evidence approach. This strategy does not reduce the uncertainty in any single line of evidence, but does reduce overall uncertainty in the conclusions of the risk assessment since the limitations of one line of evidence are frequently balanced by the strengths of another.

5.7 Linking Risk Assessment with Risk Management

The risk assessment should continue only to the point that an informed risk management decision can be made. Dialogue between the risk assessor and risk manager throughout the risk assessment process regarding how data collected will be relevant for management decisions is useful. Decisions regarding implementation of different risk assessment tools should be reviewed in terms of how iterative refinement of risk estimates supports the risk management goals for the site as established during the problem formulation (see Section 2.2.1). Risk managers "use the results of the risk assessment, along with information on technical feasibility, and social, economic and political concerns to reach a decision" (CCME 1996). Separation of risk management and risk assessment may be less clear when clients request input from risk assessors regarding appropriate (or likely to be acceptable) management actions; in those cases, any advice or conclusions regarding risk management should be distinguished from the risk estimates.

6.0 REPORT LIMITATIONS

This report was prepared for the exclusive use of the Science Advisory Board (SAB) and is intended to provide a professional opinion related to guidance for detailed ecological risk assessment (DERA). Any use that a third party may make of this report, or any reliance on or decisions made based on it, are the responsibility of the third parties. We disclaim responsibility for consequential financial effects on site management, or requirements for follow-up actions and costs.

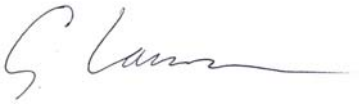
The services performed as described in this report were conducted in a manner consistent with the level of care and skill normally exercised by other members of the science professions currently practicing under similar conditions, subject to the time limits and financial and physical constraints applicable to the services. This report provides professional opinion and, therefore, no warranty is expressed, implied, or made as to the conclusions, advice and recommendations offered in this report. This report does not provide a legal opinion regarding compliance with applicable laws or regulations.

7.0 CLOSURE

This document was prepared as a collaborative effort of multiple Golder risk assessment practitioners. Modules provided in Appendices I – III were prepared by Adrian deBruyn, Blair McDonald, Cathy McPherson, Trish Miller, Christine Thomas, Barbara Wernick, and John Wilcockson. The main document was prepared by the undersigned.

GOLDER ASSOCIATES LTD.

Blair McDonald, M.E.T., R.P. Bio.
Environmental Scientist



Gary Lawrence, M.R.M., R.P. Bio.
Senior Environmental Scientist

Adrian M.H. deBruyn, Ph.D.
Environmental Scientist

Peter M. Chapman, Ph.D., R.P.Bio.
Senior Environmental Scientist/Associate

BGM/GSL/AMHD/PMC/lba
05-1421-040
O:\DATA\FINAL\2005\1421\05-1421-040\FINAL RPT-0330_06 SAB-DERA GUIDANCE MANUAL.DOC

8.0 REFERENCES

Alexander, M. 2000. Aging, bioavailability and overestimation of risk from environmental pollutants. Environ. Sci. Technol. 34: 4259-4265.

Anderson, B.S., J.W. Hunt, B.M. Phillips, R. Fairey, H.M. Puckett, M. Stephenson, K. Taberski, J. Newman, R.S. Tjeerdema. 2001. Influence of sample manipulation on contaminant flux and toxicity at the sediment-water interface. Mar. Environ. Res. 51: 198-211.

Bachmann, K., D. Pardoe, D. White. 1996. Scaling basic toxicokinetic parameters from rat to man. Environ. Health. Perspect. 104: 400-407.

BCMELP (British Columbia Ministry of Environment, Lands and Parks). 2000. Tier I ecological risk assessment policy decision summary. Available online: http://www.env.gov.bc.ca/epd/epdpa/contam_sites/standards_criteria/standards/tier1policy.html

BCMELP (British Columbia Ministry of Environment, Lands and Parks). 1997. Checklist for the Tier-1 Ecological Risk Assessment of Contaminated Sites in British Columbia. Prepared for British Columbia Ministry of Environment, Victoria. Institute of Environmental Toxicology and Chemistry, Western Washington University, Bellingham, WA; Ecological Planning & Toxicology Inc., Corvalis, OR; EVS Environment Consultants, North Vancouver, B.C. March 31, 1997.

BCMOE (British Columbia Ministry of Environment). 1995. Derivation of water quality criteria to protect aquatic life in British Columbia. British Columbia Ministry of Environment, Water Quality Branch, Victoria, BC.

BCMOE. 2005. Technical guidance on contaminated sites #19: assessing and managing contaminated sediment. British Columbia Ministry of Environment, Victoria, BC. Available online: http://www.env.gov.bc.ca/epd/epdpa/contam_sites/guidance/technical.

Boese, B.L., R.J. Ozretich, J.O. Lamberson, R.C. Swartz, F.A. Cole, J. Pelletier, J. Jones. 1999. Toxicity and phototoxicity of mixtures of highly lipophilic PAH compounds in marine sediment: can the sum PAH model be extrapolated? Arch. Environ. Contam. Toxicol. 36: 270-280.

Burgess, R.M., K.A. Schweitzer, R.A. McKinney, D.K. Phelps. 1993. Contaminated marine sediments: water column and interstitial toxic effects. Environ. Toxicol. Chem. 12: 127-138.

CCME (Canadian Council of Ministers of the Environment). 1996. A framework for ecological risk assessment: general guidance. National Contaminated Sites Remediation Program, Canadian Council of Ministers of the Environment, Winnipeg, MB.

CCME. 1999. A protocol for the derivation of water quality guidelines for the protection of aquatic life. Prepared by the Task Force on Water Quality Guidelines. Originally published in April 1991 as Appendix IX of Canadian Water Quality Guidelines (CCREM, 1987). Updated and reprinted with minor revisions and editorial changes.

Carr, R.S. and M. Nipper (eds). 2003. Porewater Toxicity Testing: Biological, Chemical, and Ecological Considerations. SETAC Press, Pensacola, FL. 346 pp.

Carr, R.S., M. Nipper, W.J. Adams, W.J. Berry, G.A. Burton Jr., K. Ho, D. MacDonald, R. Scroggins, P.V. Winger. 2001. Summary of a SETAC technical workshop: Porewater toxicity testing: biological, chemical, and ecological considerations with a review of methods and applications, and recommendations for future areas of research. Summary of the SETAC Workshop on Porewater Toxicity Testing, March 18-22, 2000, Pensacola, FL, USA.

Chapman, P.M. 1990. The Sediment Quality Triad approach to determining pollution-induced degradation. Sci. Tot. Environ. 97/98: 815-825.

Chapman, P.M. 1996. Presentation and interpretation of Sediment Quality Triad data. Ecotoxicology 5: 327-339.

Chapman, P.M., B.G. McDonald. 2005. Risk assessment using the Sediment Quality Triad. Chapter 10, pp. 305-330, In: Blaise C and Férard J-F (eds.), Small-Scale Freshwater Toxicity Test Investigations, Volume 2: Hazard Assessment Schemes. Kluwer Academic Press, Netherlands.

Chapman, P.M., J. Anderson. 2005. A decision-making framework for sediment contamination. Integr. Environ. Assess.Manage. 1: 163-173.

Chapman, P.M., H. Hollert. 2006. Should the Sediment Quality Triad become a tetrad, a pentad, or possibly even a hexad? J. Soil Sed. 6: 4-8..

Chapman, P.M., R.S. Cardwell, P.F. Chapman. 1996. A warning: NOECs are inappropriate for regulatory use. Environ. Toxicol. Chem. 15: 77-79.

Chapman, P.M., B. Anderson, S. Carr, V. Engle, R. Green, J. Hameedi, M. Harmon, P. Haverland, J. Hyland, C. Ingersoll, E. Long, J. Rodgers Jr., M. Salazar, P.K. Sibley, P.J. Smith, R.C. Swartz, B. Thompson, H. Windom. 1997. General guidelines for using the Sediment Quality Triad. Mar. Pollut. Bull. 34: 368-372.

Chapman, P.M., A. Fairbrother, D. Brown. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environ. Contam. Toxicol. 17: 99-108.

Chapman, P.M., F. Wang, J.D. Germano, G. Batley. 2002a. Pore water testing and analysis: the good, the bad and the ugly. Mar. Pollut. Bull. 44: 359-366.

Chapman, P.M., B.G. McDonald, G.S. Lawrence. 2002b. Weight of evidence frameworks for sediment quality and other assessments. Human Ecol. Risk Assess. 8: 1489-1515.

Chapman, P.M., F. Wang, C.R. Janssen, R.R. Goulet, C.N. Kamunde. 2003. Conducting ecological risk assessments of inorganic metals and metalloids: current status. Human Ecol. Risk Assess. 9: 641-697.

Cornelissen, G., O. Gustafsson, T.D. Bucheli, M.T.O. Jonker, P.C.M. van Noort. 2005. Extensive sorption of organic compounds to black carbon, coal, and kerogen in sediments and soils: Mechanisms and consequences for distribution, bioaccumulation and degradation. Environ. Sci. Technol. 39: 6881-6895.

Dearfield, K.L., E.S. Bender, M. Kravitz, R. Wentsel, M.W. Slimak, W.H. Farland, P. Gilman. 2005. Ecological risk assessment issues identified during the US Environmental Protection Agency's examination of risk assessment practices. Integ. Environ. Assess. Manag. 1: 73-76.

Diamond, S., D. Mount. 1998. Evaluating the role of photo-activated PAH toxicity in risk assessment. SETAC News 18: 17-18.

Dodds, P.S., D.H. Rothman, J.S. Weitz. 2001. Re-examination of the "3/4 law" of metabolism. J. Theor. Biol. 209: 9-27.

Efroymson, R.A., B.E. Sample, G.W. Suter II. 2001. Uptake of inorganic chemicals for soil by plant leaves: regressions of field data. Environ. Toxicol. Chem. 20: 2561-2571.

Eisler, R. 2000. Handbook of chemical risk assessment: health hazards to humans, plants and animals. Lewis Publishers, Boca Raton, FL, USA. 1829 pages (3 volumes).

Environment Canada. 1999. Guidance document on application and interpretation of single-species tests in environmental toxicology. Environmental Protection Series, Report EPS 1/RM/34, December 1999. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring. Environment Canada, National Environmental Effects Monitoring Office, Gatineau, QC. Available online: http://www.ec.gc.ca/EEM/English/MetalMining/

Goddard, M.J., D. Krewski. 1992. Interspecies extrapolation of toxicity data. Risk Anal. 12: 315-317.

Ho, K.T., R.M. Burgess, M.C. Pelletier, J.R. Serbst, S.A. Ryba, M.G. Cantwell, A. Kuhn, P. Raczelowski. 2002. An overview of toxicant identification in sediments and dredged material. Mar. Pollut. Bull. 44: 286-293.

Hull, R.N. and S. Swanson. 2006. Sequential analysis of lines of evidence – an advanced weight-of-evidence approach for ecological risk assessment. Integ. Environ. Assess. Manage. In press.

Jop, K.M., A.M. Askew, R.B. Foster. 1995. Development of a water-effect ratio for copper, cadmium, and lead for the Great Works River in Maine using *Ceriodaphnia dubia* and *Salvelinus fontinalis*. Bull. Environ. Contam. Toxicol. 54: 29-35.

Johnson, R.K. W.R. Munns Jr, P. L. Tyler, P. Marajh-Whittemore, K. Finkelstein, K. Munney, F.T. Short, A. Melville, S.P. Hahn. 2002. Weighing the evidence of ecological risk from chemical contamination in the estuarine environment adjacent to the Portsmouth naval shipyard, Kittery, Maine, USA. Environ. Toxicol. Chem. 21: 182-194.

Kirman, C.R., L.M. Sweeney, M.E. Meek, M.L. Gargas. 2003. Assessing the dose-dependancy of allometric scaling performance using physiologically based pharmacokinetic modeling. Reg. Toxicol. Pharmacol. 38: 345-367.

Long, E.R., P.M. Chapman. 1985. A sediment quality triad: measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Mar. Pollut. Bull. 16: 405-415.

Lowell, R.B., J.M. Culp, M.G. Dube. 2000. A weight-of-evidence assessment for northern river risk assessments: integrating the effects of multiple stressors. Environ. Toxicol. Chem. 19: 1182-1190.

Macfarlane, M.W., D.D. MacDonald, C.G. Ingersoll. 2003. Criteria for managing contaminated sediments in British Columbia: technical appendix. Environmental Management Branch, BC Ministry of Water, Land and Air Protection, Victoria, BC.

McDonald, B.G. 2005. Comparison of porewater and elutriate toxicity testing in a sediment quality triad framework. Ecotoxicol. Environ. Saf. 62: 383-390.

McDonald, B.G., P.M. Chapman. 2002. PAH phototoxicity – An ecologically irrelevant phenomenon? Mar. Poll. Bull. 44: 1321-1326.

McDonald, B.G., J.B. Wilcockson. 2003. Improving the use of toxicity reference values in wildlife food chain modeling and ecological risk assessment. Human Ecol. Risk Assess. 9: 1585-1594.

Menzie, C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: a weight-of-evidence approach for evaluating ecological risks. Hum. Ecol. Risk Assess. 2: 277-304.

Mineau, P., B.T. Collins, A. Baril. 1996. On the use of scaling factors to improve interspecies extrapolation of acute toxicity in birds. Regul. Toxicol. Pharmacol. 24: 24-29.

Muyssen, B.T.A., C.R. Janssen. 2002a. Tolerance and acclimation to zinc of *Ceriodaphnia dubia*. Environ. Poll. 117: 301-306.

Muyssen, B.T.A., C.R. Janssen. 2002b. Accumulation and regulation of zinc in *Daphnia magna:* links with homeostasis and toxicity. Arch. Environ. Contam. Toxicol. 43: 492-496.

Muyssen, B.T.A., C.R. Janssen, B.T.A. Bossuyt. 2002. Tolerance and acclimation to zinc of field-collected *Daphnia magna* populations. Aquat. Toxicol. 56: 69-72.

Nagy, K.A., I.A. Girard, T.K. Brown. 1999. Energetics of free-ranging mammals, reptiles and birds. Annu. Rev. Nutr. 19: 247-277.

Nipper, M., R.S. Carr, J.M. Biedenback, R.L. Hooten, K. Miller. 2002. Toxicological and chemical assessment of ordnance compounds in marine sediments and porewaters. Mar. Pollut. Bull. 44: 789-806.

Pearsons, T.N. and C.W. Hopley. 1999. A practical approach for assessing ecological risks associated with fish stocking programs. Fisheries 24: 16-23.

Posthuma, L., T.P. Traas, G.W. Suter II. 2002. General introduction to species sensitivity distributions. In: Species Sensitivity Distributions. L. Posthuma, G.W. Suter II and T.P. Trass (eds). Lewis Publishers, New York, NY, USA. pp 3-11.

Preston, B.L. 2002. Indirect effects in aquatic ecotoxicology: implications for ecological risk assessment. Environ. Manage. 29: 311-323.

Ritter, L., K. Solomon, P. Sibley, K. Hall, P. Keen, G. Mattu, B. Linton. 2002. Sources, pathways and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton inquiry. J. Toxicol. Environ. Health A. 65: 1-142.

Roth, M. 1993. Investigations on lead in the soil invertebrates of a forest ecosystem. Pedobiologia 37: 270-279.

SAB (Science Advisory Board). 2005. *Report on Screening Level Risk Assessment (SLRA Level 1 and SLRA Level 2)*. Submitted by the Science Advisory Board for Contaminated Sites in British Columbia to the British Columbia Ministry of the Environment, Victoria, B.C. August 2005.

Sample, B.E., C.A. Arenal. 1999. Allometric models for interspecies extrapolation of wildlife toxicity data. Bull. Environ. Contam. Toxicol. 62: 653-663.

Sample, B.E., G.W. Suter II. 1999. Ecological risk assessment in a large river-reservoir: 4. Piscivorous wildlife. Environ. Toxicol. Chem. 18: 610-620.

Sample, B., D. Opresko, G. Suter II. 1996. Toxicological benchmarks for wildlife: 1996 revision. Oak Ridge National Laboratory, Department of Energy. Oak Ridge, TN. ES/ER/TM-86/R3.

Savage, V.M., J.F. Gillooly, W.H. Woodruff, G.B. West, A.P. Allen, B.J. Enquist, J.H. Brown. 2004. The predominance of quarter-power scaling in biology. Funct. Ecol. 18: 257-282.

Semple, K.T., K.J. Doick, K.C. Jones, B. Burauel, A. Craven, H. Harms. 2004. Defining bioavailability and bioaccessibility of contaminated soil and sediment is complicated. Environ. Sci. Technol. 38: 228A-231A.

Sparks, T.H, W.A. Scott, R.T. Clarke. 1999. Traditional multivariate techniques: potential for use in ecotoxicology. Environ. Toxicol. Chem.18: 128-37.

Suter, G.W. II. 1996a. Guide for developing conceptual models for ecological risk assessments. ES/ER/TM-186. Office of Environmental Management, US Department of Energy, Oak Ridge, TN, USA.

Suter, G.W. II. 1996b. Risk characterization for ecological risk assessment of contaminated sites. ES/ER/TM-20. Office of Environmental Management, US Department of Energy, Oak Ridge, TN, USA.

Suter, G.W. II., R.A. Efroymson, B.E. Sample, D.S. Jones. 2000. Ecological risk assessment for contaminated sites. Lewis Publishers, Boca Raton, FL. 438 p.

Tannenbaum, L.V. 2005. A critical assessment of the ecological risk assessment process: a review of misapplied concepts. Integr. Environ. Assess. Manag. 1: 66-72.

Tannenbaum, L.V., M.S. Johnson and M. Bazar. 2003. Application of the hazard quotient method in remedial decisions: a comparison of human and ecological risk assessment. Human Ecol. Risk Assess. 9: 387-401.

Torres, K.C., M.L. Johnson. 2001. Bioaccumulation of metals in plants, arthropods and mice at a seasonal wetland. Environ. Toxicol. Chem. 20: 2617-2626.

Travis, C.C., R.K. White. 1988. Interspecific scaling of toxicity data. Risk Anal. 8: 199-125.

USEPA (United States Environmental Protection Agency).. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. Office of Water Regulations and Standards, US Environmental Protection Agency, Washington, DC. 98 pp.

USEPA. 1992. Framework for ecological risk assessment. Office of Research and Development, US Environmental Protection Agency, Washington, DC. EPA/630/R-92/001.

USEPA. 1993a. A review of ecological assessment case studies from a risk assessment perspective. Risk Assessment Forum, US Environmental Protection Agency, Washington, DC. EPA/630/R-92/005.

USEPA. 1993b. Wildlife exposure factors handbook. Office of Research and Development, US Environmental Protection Agency, Washington, DC. EPA/600/R-93/187.

USEPA. 1995. Great Lakes water quality initiative: technical support document for wildlife criteria. Office of Water, US Environmental Protection Agency, Washington, DC. EPA/820/B-95/009.

USEPA. 2002a. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms, 5th edition. EPA-821-R-02-012. US Environmental Protection Agency, Office of Water, Washington, DC. October 2002. 266 pp.

USEPA. 2002b. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms, 4rd edition. EPA-821-R-02-013. US Environmental Protection Agency, Office of Water, Washington, DC. October 2002. 335 pp.

USEPA. 2002c. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms, 2nd edition. EPA-821-R-02-014. US Environmental Protection Agency, Office of Water, Washington, DC. October 2002. 266 pp.

USEPA. 2005. Guidance for developing ecological soil screening levels. Office of Solid Waste and Emergency Response. OSWER Directive 9285.7-55. Available online: http://www.epa.gov/ecotox/ecossl/.

USEPA. 2005b. Fate, exposure and risk analysis (FERA). http://www.epa.gov/ttn/fera/.

Van der Hoeven, N. 1997. How to measure no effect. Part III: Statistical aspects of NOEC, ECX and NEC estimates. Environmetrics 8: 255-261.

Wandall, B. 2004. Values in science and risk assessment. Toxicol. Lett. 152: 265-272.

Wenning, R.J., C.G. Ingersoll (eds). 2002. Use of sediment quality guidelines and related tools for the assessment of contaminated sediments. Executive Summary Booklet of a SETAC Pellston Workshop. SETAC Press, Pensacola, FL, USA.

Wenning, R., C. Ingersoll, G. Batley, M. Moore. (eds.). 2005. Use of Sediment Quality Guidelines (SQGs) and Related Tools for the Assessment of Contaminated Sediments. SETAC Press, Pensacola, FL, USA.

Westbrook, S.J., J.L. Rayner, G.B. Davis, T.P. Clement, P.L. Bjerg, S.J. Fisher. 2005. Interaction between shallow groundwater, saline surface water and contaminant discharge at a seasonally and tidally forced estuarine boundary. J. Hydrol. 302: 255-269.

TABLES

Table 1: Levels of Biological Organization for Selecting Receptors of Potential Concern for Generic Ecosystem Types Considered in a Detailed ERA

GENERIC ECOSYSTEMS	RECEPTOR GROUP	
	HIGHER LEVEL OF RESOLUTION	LOWER LEVEL OF RESOLUTION
Aquatic **1. Deep Aquatic** **2. Shoreline** **3. Streams & Rivers**	Primary producers	Phytoplankton and periphyton Aquatic macrophytes
	Water-column invertebrates	Zooplankton Invertebrate planktivores
	Benthic community	Epibenthic invertebrates Invertebrate filter-feeders Benthic infauna
	Fish	Detrivorous fish Planktivorous fish Piscivorous fish
	Mammals	Piscivorous mammals
	Waterfowl	Piscivorous birds Benthivorous brids Detrivorous birds
	Amphibians	Amphibians
Terrestrial **4. Uplands (Wildlands)** **5. Uplands (Human Use)**	Microbes	Microbes
	Invertebrate	Litter-dwelling invertebrates Soil-dwelling invertebrates Arboreal invertebrates
	Plants	Mosses Grasses Shrubs Trees
	Small mammals	Small mammal ground insectivores Small mammal arboreal insectivores Small mammal omnivores Small mammal herbivores Small mammal carnivores
	Small birds	Avian ground insectivores Avian arboreal insectivores Avian omnivores Avian herbivores
	Large mammals	Large mammal herbivores Large mammal omnivores
	Carnivores	Raptors Carnivorous mammals
	Reptiles	Reptiles
	Amphibians	Amphibians

Multiple ecosystem types and transitional subtypes may exist within the boundaries of a single site, which may influence ROPC selection.

Table 2: Existing Guidance for Selecting Mammalian and Avian ROPCs based on Land Use Considerations (BCMELP 1997)

	INDUSTRIAL	COMMERCIAL	RESIDENTIAL	URBAN PARK	AGRICULTURAL
Large mammals (deer, elk, bear, coyotes, fox, skunk, raccoon)	Excluded	Excluded	Excluded	Included	Included
Large rodents (rabbits, beaver)	Excluded from urban areas	Excluded from urban areas	Excluded	Included	Included
Mustelids	Excluded	Aquatic mustelids may be included	Aquatic mustelids may be included	Aquatic mustelids may be included	Aquatic mustelids may be included
Small rodents (mice, vole)	Included	Included	Included	Included	Included
Bats	Excluded	Excluded	Included	Included	Included
Shorebirds, wading birds, waterfowl, seabirds	Excluded	Excluded	Included	Included	Included
Raptors	Include if threatened or endangered	Include if threatened or endangered	Include if threatened or endangered	Included	Included
Galliforms (e.g., pheasant, quail)	Excluded from urban areas	Excluded from urban areas	Excluded	Included	Included
Cavity-dwelling birds that eat foliar invertebrates	Excluded	Excluded	Included	Included if trees are present	Included if trees are present
Hummingbirds	Excluded	Excluded	Excluded	Excluded	Excluded

Table 3: Classification of Common Toxicity Tests for the Purpose of DERA

Test	Classification	Regulatory Agency	Rationale
Water Toxicity Tests			
48-h cladoceran (*Daphnia* sp.) survival	Acute	Environment Canada USEPA	Described as an "an acute test with the additional endpoint of immobility" by Environment Canada (EC 1RM11). Also, described as "acute" by USEPA (EPA-821-R-02-012).
48-h cladoceran (*Ceriodaphnia dubia*) survival	Acute	USEPA	Described as "acute" by USEPA (EPA-821-R-02-012).
48 or 96-h mysid survival (various species)	Acute	USEPA	Described as "acute" by USEPA (EPA-821-R-02-012).
96-h fish survival (various species)	Acute	Environment Canada USEPA	Described as "acute" by Environment Canada (EC 1RM9, EC 1RM10) and USEPA (EPA-821-R-02-012).
48-h bivalve larval development (various species)	Chronic Surrogate	USEPA	Described as a "estimate of chronic toxicity" by USEPA (EPA/600/R-95/136)
48-h echinoid larval development (various species)	Chronic Surrogate	USEPA	Described as a "estimate of chronic toxicity" by USEPA (EPA/600/R-95/136)
7-d cladoceran (*Ceriodaphnia dubia*) survival and reproduction	Chronic	Environment Canada USEPA	Described as "chronic" by Environment Canada. Also notes: "for tests with cladocerans, chronic is typically defined as continuing until three broods are produced." The document also refers to *Daphnia* sp. tests requiring 14 or 21 days duration as chronic (EC 1RM21). Described as "chronic" by USEPA (EPA/821/R-02/013).
20-min echinoid fertilization (various species)	Chronic Surrogate	Environment Canada	Described as "sublethal" by Environment Canada. Also notes an acute test for echinoids would have a duration of "a few days for echinoids, which generally have a life span of 4 – 8 years for sea urchins." However, the document also notes: "The fertilization assay is a sensitive sublethal test. The fertilization assay is not a chronic test, however, because of its very short duration relative to the life spans of the species (some years). The fertilization assay described in this report is not intended to replace chronic toxicity tests using echinoids, because it might not estimate the effects of longer exposures. However, this test can be expected to yield results closer to such chronic tests than would conventional lethality tests with marine or freshwater species" (EC 1RM27).The methodology used by USEPA (EPA-821-R-02-014) is comparable to Environment Canada. This test is described as an estimate of chronic toxicity

Table 3: Classification of Common Toxicity Tests for the Purpose of DERA (cont'd)

Test	Classification	Regulatory Agency	Rationale
7-d fish larval survival and growth (various species)	Chronic Surrogate	Environment Canada USEPA	Described as "sublethal" by Environment Canada. Also notes the test "is not of long enough duration relative to the life span of the fish, and is therefore not a chronic test". However, the document also notes: "The seven-day test is sensitive, however, because larval fish are usually among the most vulnerable stages of the entire life cycle. In general, the seven-day test could be expected to estimate the toxicity in a 30-day exposure of early life-stages of fathead minnows closely in some cases, and within a factor of 2 in other cases, but it might sometimes under-predict by an order of magnitude". The 7-d larval fish test" does not necessarily replace chronic toxicity tests, but comes much closer to results of such chronic tests than would a conventional lethality test with juvenile fish" (EC 1 RM22). Described as an estimate of chronic toxicity by the USEPA (EPA-821-R-02-014 and EPA-600-R-95/136)
7-d fish early life-stage survival (various species)	Chronic Surrogate	Environment Canada USEPA	7-d embryo (E) test described as an "acute" test, while the embryo-alevin (EA) test and embryo-alevin-fry (EAF) tests are referred to as "longer" tests. Also notes that "Because of the long life span of salmonids, early life-stage tests do not measure chronic toxicity, although the intent of this test is to estimate approximately, what such sublethal chronic toxicity might be." Also: "Results from full and partial life-cycle tests with several fish species and a variety of chemicals indicate that the early development stages (i.e., embryo, larval, and early juvenile) can be equally or more sensitive to aquatic contaminants than the adults" (EC 1RM28). Described as an estimate of chronic toxicity by USEPA.
9-d fish embryo-larval survival and teratogenicity (various species)	Chronic Surrogate	USEPA	Described as a "chronic estimate" by USEPA (EPA-821-R-02-014). Refer to 7-d fish early life stage for explanation.
72- or 96-h phytoplankton (*Selenastrum capricornutum*) growth inhibition	Chronic	Environment Canada USEPA	Defined as "chronic" by Environment Canada. Also notes that algae are exposed "over several generations" (EC 1RM25). Described as "chronic" by USEPA (EPA-821-R-02-013).
7-d duckweed (*Lemna* sp) growth inhibition	Chronic Surrogate	Environment Canada	Not defined as "acute" or "chronic" by Environment Canada (EC 1RM37).
48-h giant kelp (*Macrocystis pyrifera*) germination and growth	Chronic Surrogate	USEPA	Described as "chronic estimate" by USEPA (EPA-600/R-95/136)

Table 3: Classification of Common Toxicity Tests for the Purpose of DERA (cont'd)

Test	Classification	Regulatory Agency	Rationale
Sediment Toxicity Tests			
48-h bivalve larval development (various species)	Acute	PSEP	Not defined as "acute" or "chronic" by PSEP (PSEP 1995).
48-h echinoderm embryo growth and survival (various species)	Acute	PSEP	Not defined as "acute" or "chronic" by PSEP (PSEP 1995).
10-d amphipod survival (various species)	Acute	Environment Canada USEPA PSEP	Described as "acute" by Environment Canada. Also notes "amphipod species used for this test are known or presumed to have annual life cycles, so a 10-d test would represent an acute exposure" (EPS1/RM/26). Described as "acute" by USEPA. Note that this test could be considered chronic for two species (*A. abdita* and *L. plumulosus*) because of their relatively short life cycles. Reburial of surviving amphipods is an additional measurement that can be used as an endpoint (EPA/600/R-94/025).
10-d chironomid (*Chironomus* sp) survival and growth	Chronic	Environment Canada USEPA	Described as "chronic" by Environment Canada. In the laboratory, the life span for *Chironomus tentans* is approximately five to six weeks, so a 10-d test exposure would represent at least 10% of the organism's life span (EC 1RM32). Described as a "short term" test by USEPA (EPA/600/R-99/064). However, based on the life cycle of *C. tentans* as mentioned above this test has been classified as a chronic test
10 or 14-d amphipod (*Hyalella azteca*) survival and growth	Chronic	Environment Canada USEPA	Not defined as "acute" or "chronic" by Environment Canada (EPS1/RM/41). Described as "short term" by USEPA (EPA/600/R-99/064). However, life cycle of *H. azteca* is complete in 5 weeks, which represents more than 10% of the organism's life span, leading to a classification as a chronic test
14-d polychaete (*Polydora cornuta*) growth and survival	Chronic	Environment Canada	Described as "chronic" by Environment Canada. Under laboratory conditions the life cycle of the test organism, *P. cornuta*, can be completed in approximately 28 days (EC 1RM41).
20-d (*Neanthes* sp) polychaete survival and growth	Chronic	PSEP	Not defined as "acute" or "chronic" by PSEP although PSEP does note the life cycle of *Neanthes* is completed in 3-4 months. Test is considered chronic since a 20 day test duration is greater than 10% of the life cycle.
23-d chironomid (*Chironomus* sp) emergence	Chronic	USEPA	Described as "chronic" by the USEPA (EPA 600/R-99/064).
28-d amphipod (*Leptocheirus plumulosus*) survival, growth, and reproduction	Chronic	USEPA	Described as "chronic" by USEPA. The life cycle of *L. plumulosus* is complete in 4 weeks. A 28 day test exposure covers more than 10% of the life cycle and is therefore considered a chronic test (EPA/600/R-01/020).

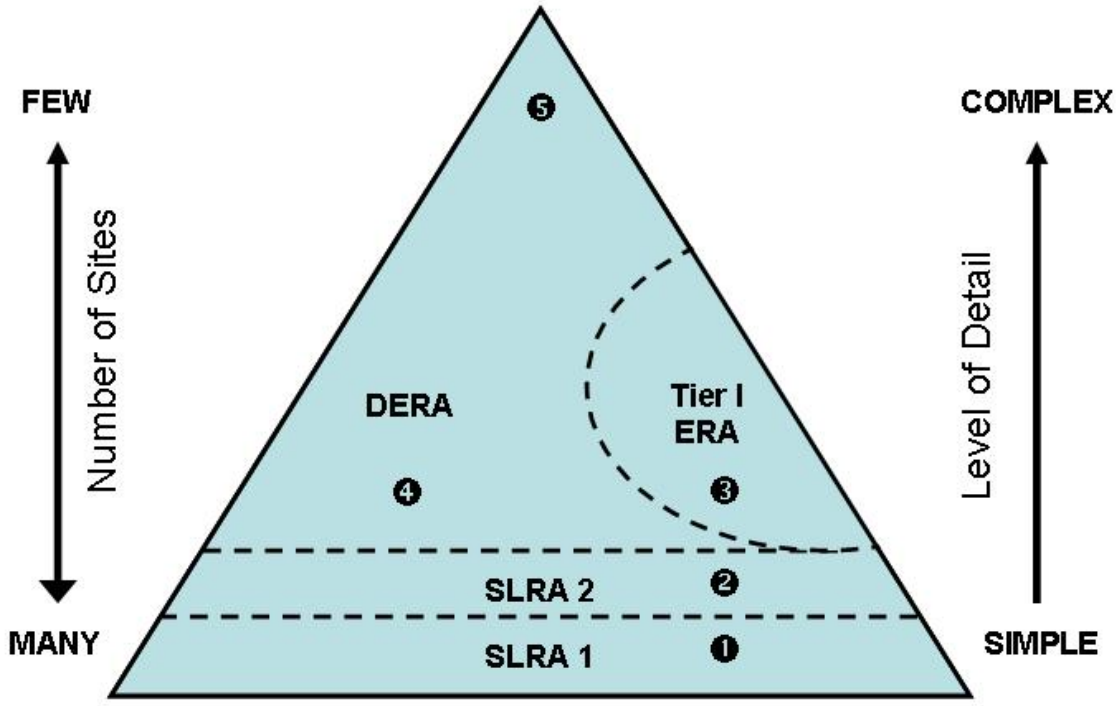
Table 3: Classification of Common Toxicity Tests for the Purpose of DERA (cont'd)

Test	Classification	Regulatory Agency	Rationale
42-d *Hyalella azteca* amphipod survival, growth, and reproduction	Chronic	USEPA	Described as a "long term" test by USEPA. Refer to 10-d *H. azteca* test for life cycle information (EPA/600/R-99/064).
Soil Toxicity Tests			
24- or 48-d nematode survival (various species)	Acute	Washington State Dept. of Ecology	Protocol available at: http://www.ecy.wa.gov/pubs/0409044.pdf
14-d seed germination (various species)	Acute	Washington State Dept. of Ecology	Protocol available at: http://www.ecy.wa.gov/pubs/96324.pdf
7-d earthworm survival	Acute	ASTM	Described as a lethal, short-term test by ASTM
14-d earthworm survival	Acute	Environment Canada	Described as an acute test (EPS 1/RM/43)
48 or 72-h earthworm avoidance	Acute	Environment Canada	Described as an acute test (EPS 1/RM/43)
56-d earthworm survival, growth and reproduction	Chronic Surrogate	Environment Canada	Described as "prolonged exposure" by Environment Canada (EPS 1/RM/43); decision to not describe the test as chronic is based on fact that test duration does not meet the criterion of >10% of an organism's life cycle (since earthworms can live for 4 -5 years). However, Environment Canada also notes that the intent of the test is to approximate a chronic exposure.
14- or 21-d seedling emergence and plant growth (various species)	Chronic Surrogate	Environment Canada (Draft)	Draft test methodology (June 2004) does not discuss the test's classification; however, the duration of the test is less than 10% of the lifespan of any of the twelve plant species described in the method (EPS 1/RM/45)
21 to 35-d collembolan (springtail) survival and reproduction (various species)	Chronic	Environment Canada (Draft)	Draft test methodology (August 2005) does not discuss the test's classification; however, the duration of the test is greater than 10% of the lifespan for at least one of the species (*Folsomia candida*; 28-d test duration versus 190 day life maximum life span) described in the method (EPS 1/RM/47)

Note: ASTM – American Society for Testing and Materials; PSEP – Puget Sound Estuary Program; USEPA – United States Environmental Protection Agency;

FIGURES

Figure 1: Anticipated Application of Different Provincial Guidance Manuals



Examples

1. SLRA Level 1 sufficient to screen out all exposure pathways; no triggers for DERA present
2. SLRA Level 2 sufficient to screen out all exposure pathways; no triggers for DERA present
3. DERA problem formulation completed. Tier I Guidance sufficient for complexity of site.
4. DERA problem formulation completed. DERA triggers are present for which Tier I guidance is not sufficient (e.g., presence of bioaccumulative compounds; contaminated sediment)
5. DERA problem formulation completed. Site is too complex for Tier I guidance.

Figure 2: Triggers for Conducting DERA (based on SAB 2005)

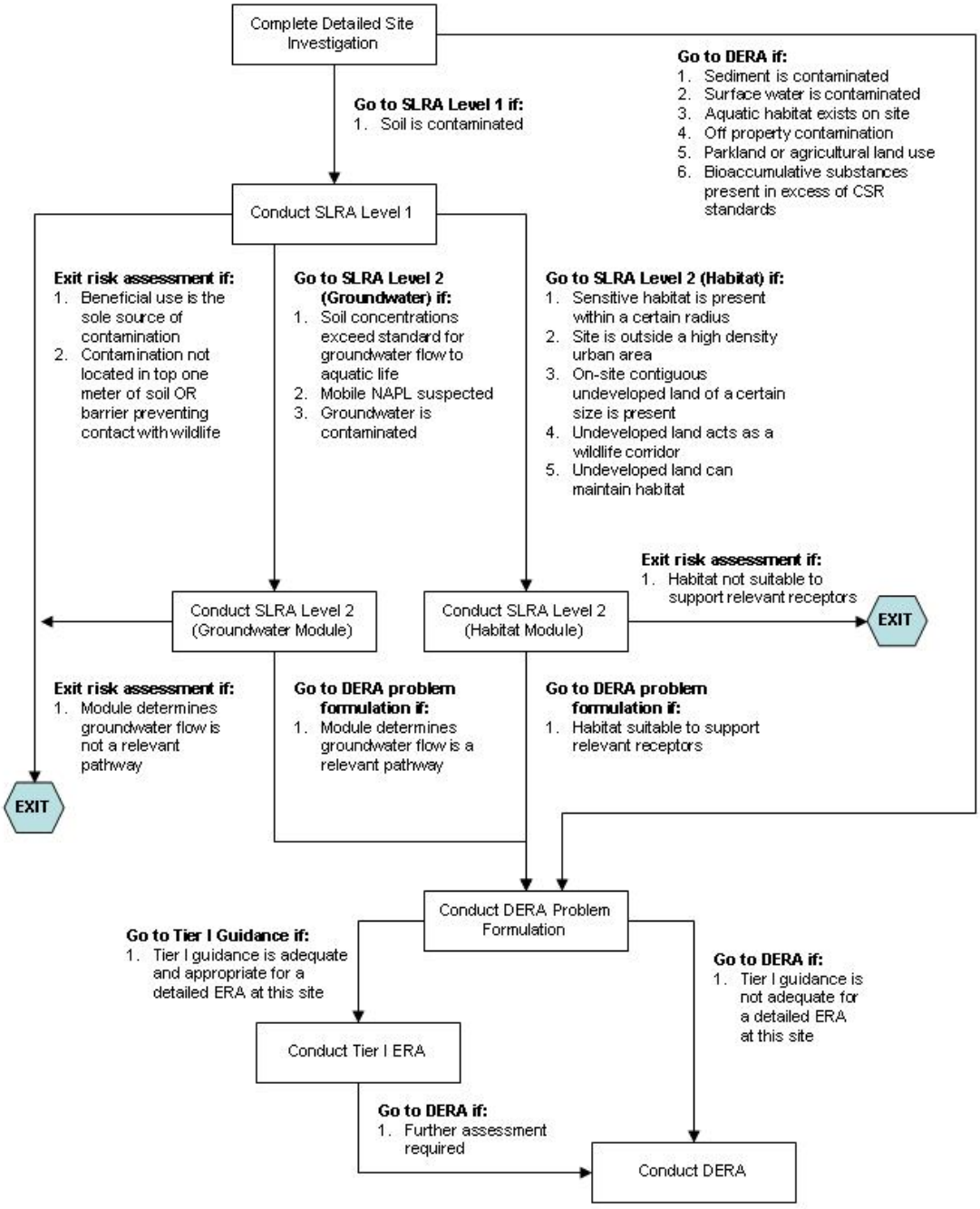


FIGURE 3: BROAD ECOSYSTEM TYPES IN BRITISH COLUMBIA

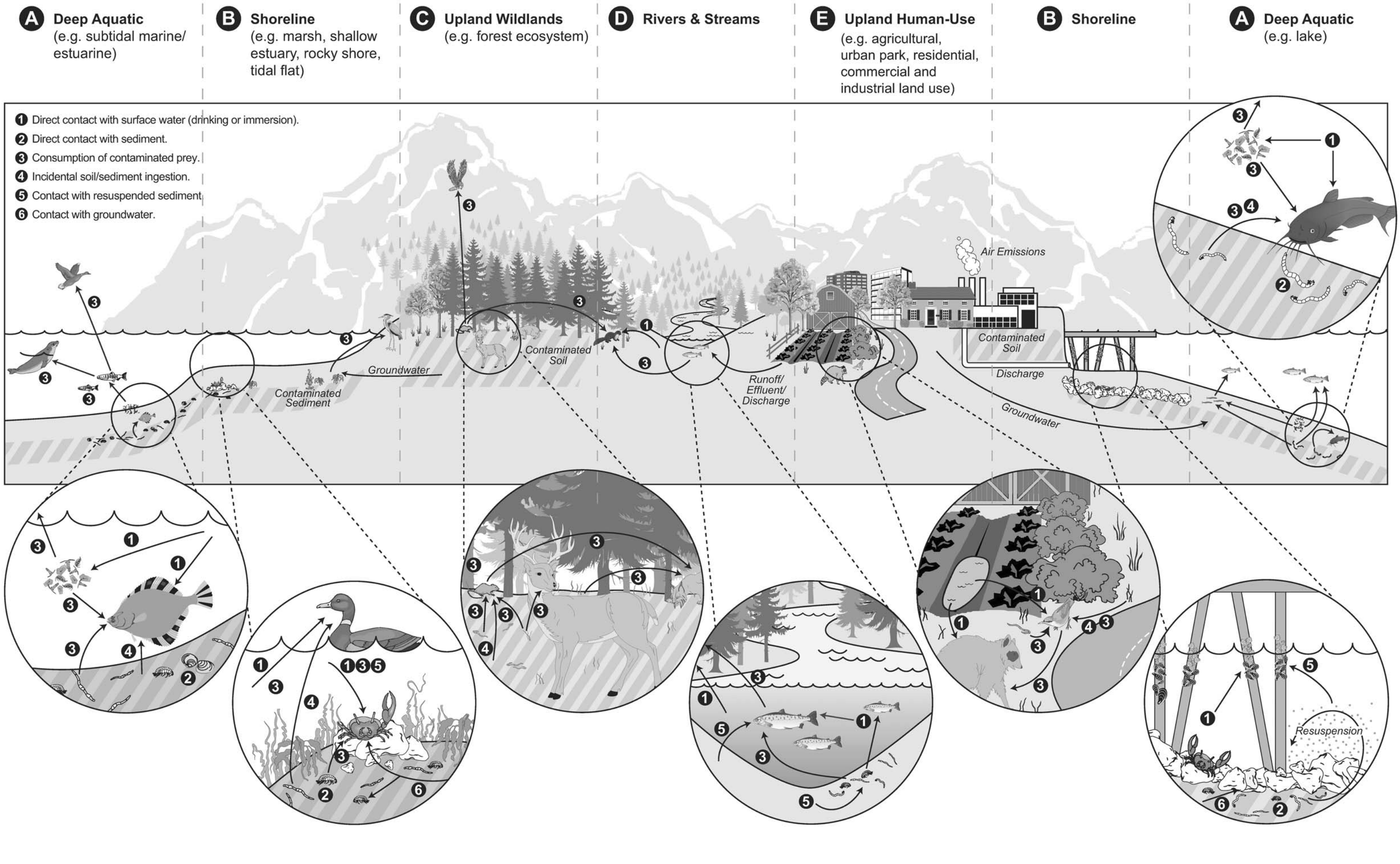


Figure 4: Example of a Box-Style Conceptual Model

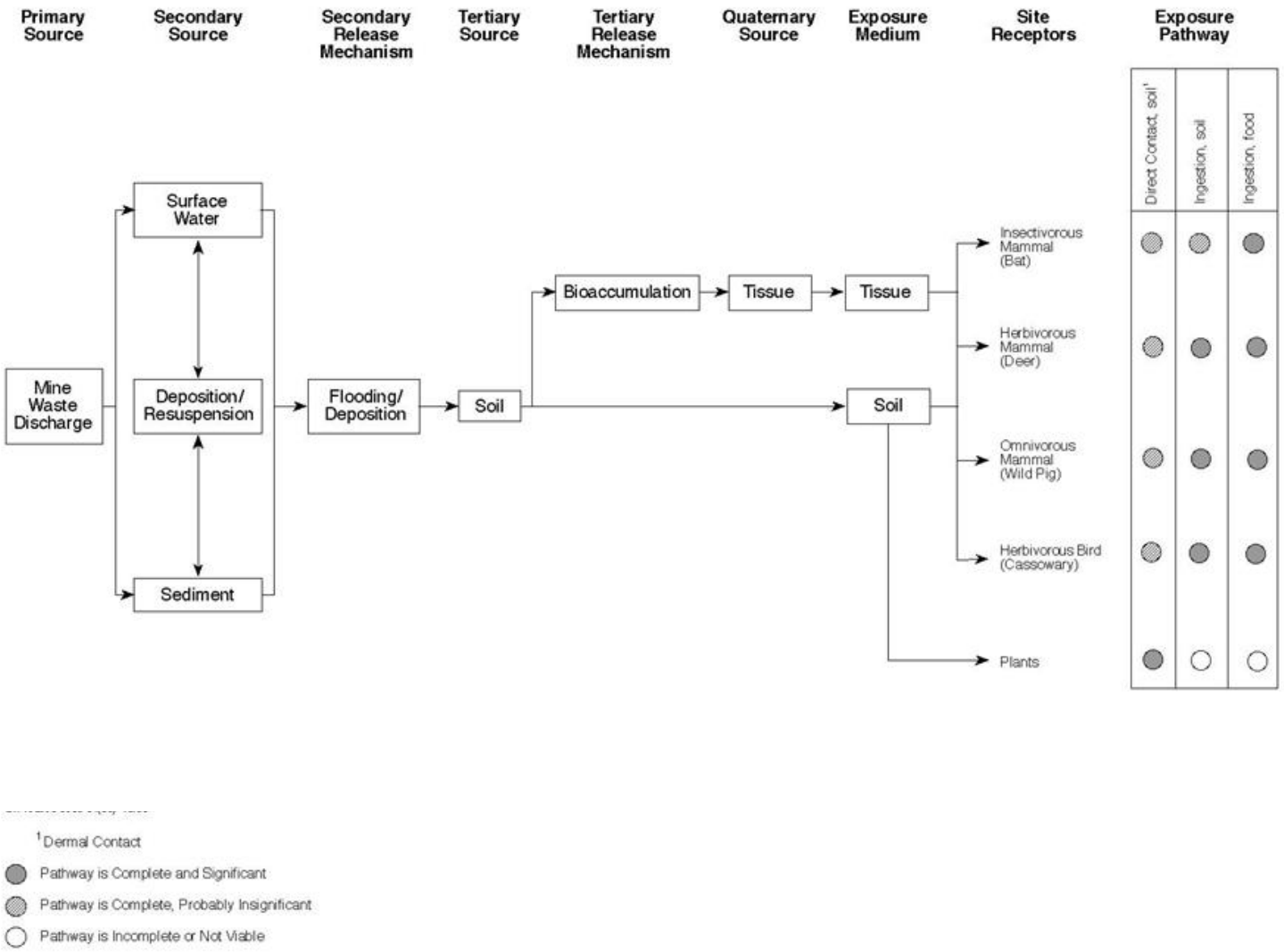


Figure 5: Example of Pictorial Conceptual Model

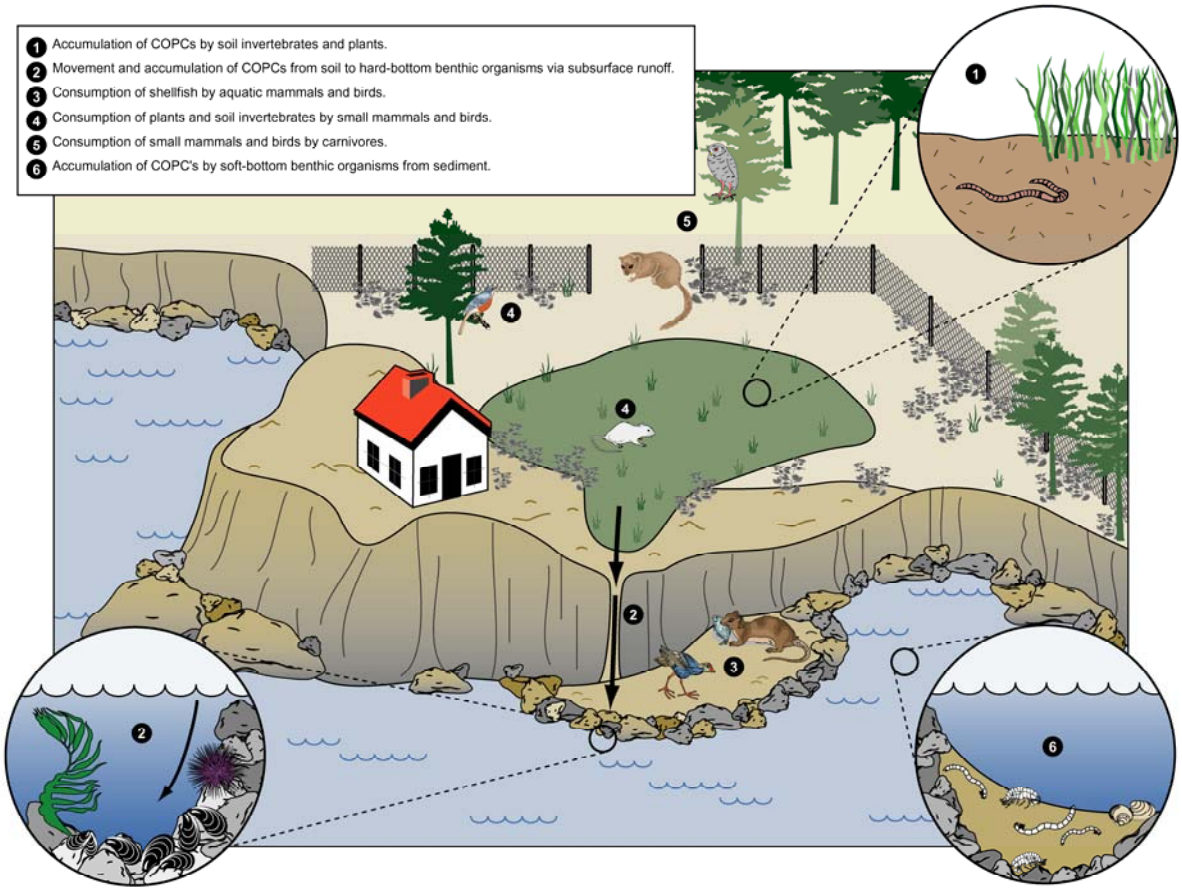


Figure 6: Illustrative Example of Tiering Risk Assessment Tools for Assessing Risks to Wildlife Receptors for the Wildlands Ecosystem Type

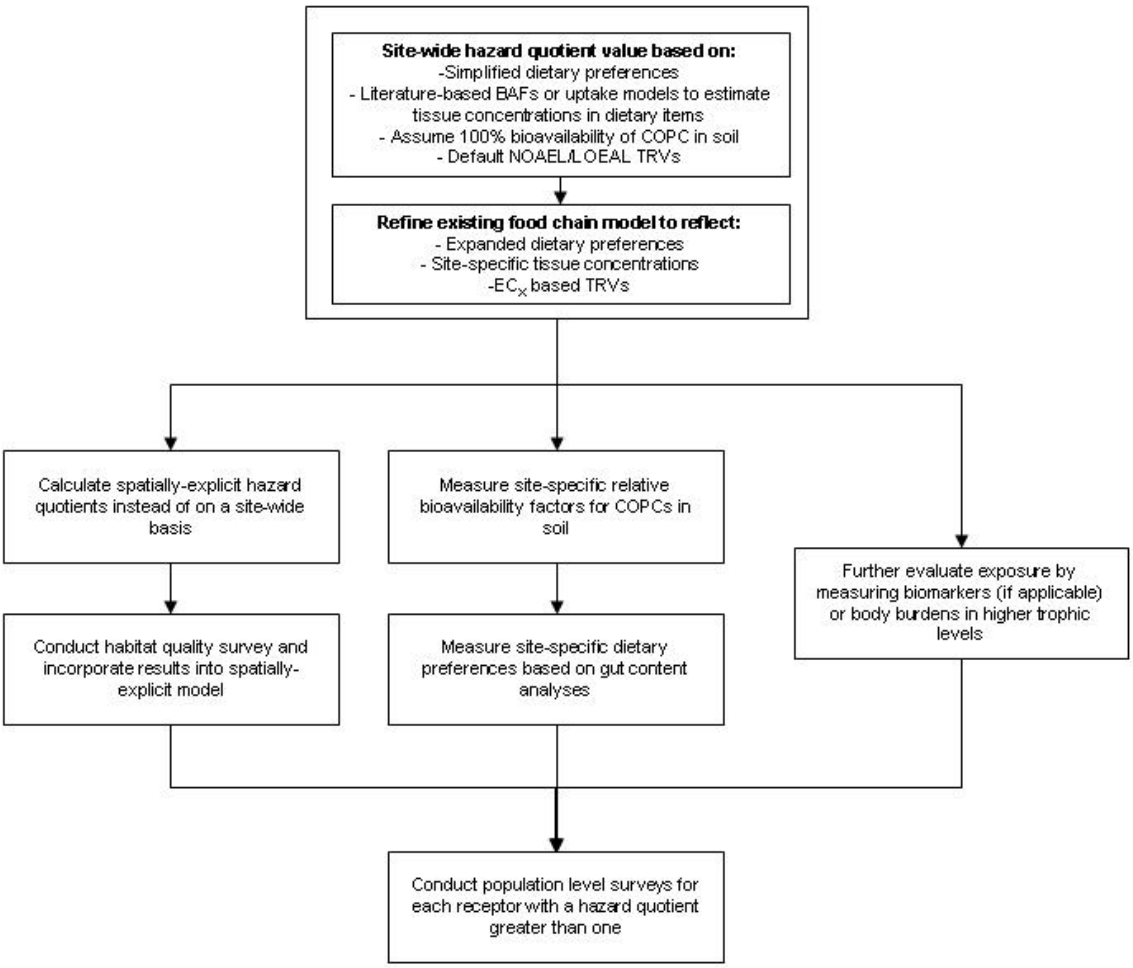
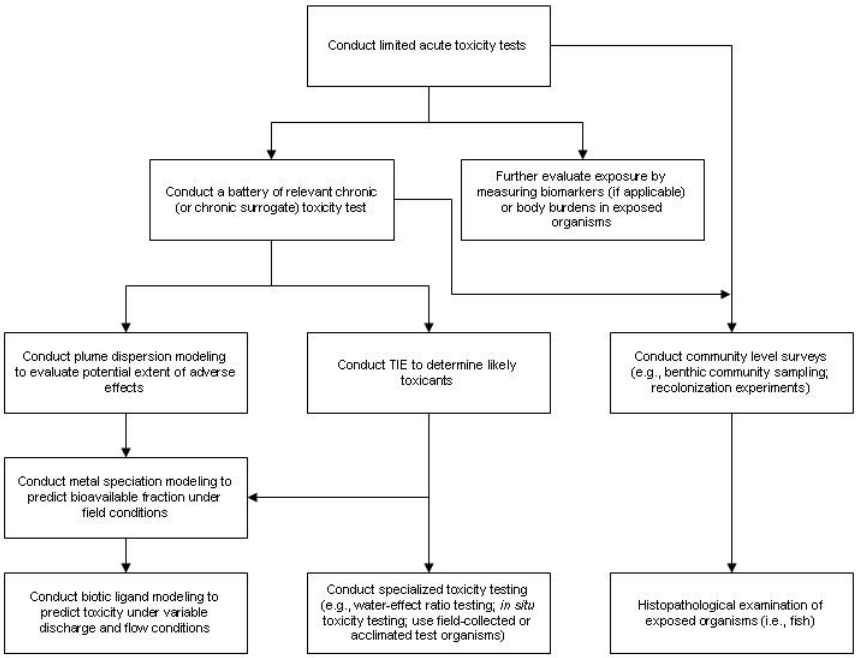


Figure 7: Illustrative Example of Tiering Risk Assessment Tools for Assessing Risks to Aquatic Receptors in the Stream and River Ecosystem Type



APPENDICES

APPENDIX I

DIRECT MEASUREMENT TOOLS

TABLE OF CONTENTS

DIRECT MEASUREMENT TOOL
#1: CHEMICAL ANALYSES OF SOIL, WATER AND SEDIMENT

What does this tool consist of? Measurement of the bulk concentration of contaminants in soil, sediment or water using analytical chemical techniques.

Which ecosystem(s) would this tool typically be applied in? All ecosystems.

How frequently is this tool used in a DERA? Common.

What are the benefits of using this tool in a DERA?

- Chemistry data provide direct measurement of COPC concentrations in the environmental media of concern.

- Many jurisdictions have published environmental quality criteria/guidelines against which bulk chemistry results can be screened to provide an initial list of COPCs (and an estimate of the magnitude of potential hazard).

- Remediation to numerical standards relies on COPC concentrations.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Bulk chemistry results do not provide a measure of bioavailability of COPCs or potential for effects, and therefore must be used in conjunction with other lines of evidence with the DERA framework.

- Numerous ancillary parameters need to be measured to facilitate an appropriate interpretation the data (e.g., *in situ* pH, hardness, TOC, AVS-SEM).

- The manifestation of biological effects may be influenced by the interaction of multiple COPCs (e.g., some parameters may be antagonistic and moderate the effects of another parameter, while other contaminants may be additive or synergistic), which cannot necessarily be predicted from bulk chemistry results.

- Not all COPCs can be analyzed with existing laboratory techniques, or if they can be measured, current laboratory techniques may not be able to detect environmental relevant concentrations (i.e., method detection limits may be above concentrations that may cause effects).

Where can I go for further information about this tool? The following guidance manuals are available online:

- Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring.

- Cavanagh, N., R.N. Nordin, L.W. Pommen and L.G. Swain. 1998. Guidelines for designing and implementing a water quality monitoring program in British Columbia. Field Test Edition. Ministry of Environment, Lands and Parks (now called the Ministry of Environment), Victoria, BC.

- Caux, P.Y., D.R.J. Moore, and D. MacDonald. 1997. Sampling strategy for turbidity, suspended and benthic sediments. Technical Appendix Addendum. Prepared for BC Ministry of Environment, Lands and Parks (now called Ministry of Environment) by Cadmus Group, Inc. and MacDonald Environmental Sciences Ltd.

Contact an accredited analytical chemistry lab for specific information about different analytical techniques.

DIRECT MEASUREMENT TOOL
#2: CHEMICAL ANALYSES OF TISSUES

What does this tool consist of? Measurement of the bulk concentration of contaminants in sampled tissues (e.g., whole benthic invertebrates; fish liver or fillet; soil invertebrates, plants) using analytical chemical techniques.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? This tool would commonly be used in a case in which COPCs are known to: a) bioaccumulate or biomagnify; and, b) be persistent. Direct measurement of COPCs in field collected tissues is also used when dietary ingestion is a relevant exposure pathway in the DERA.

What are the benefits of using this tool in a DERA? The presence of the COPC in tissues provides an indication that the organism has been exposed, or that it may expose higher trophic levels to the COPC. Measured tissue concentrations of COPCs can be used to quantify exposure of receptors if the effects profile is expressed in terms of internal or dietary concentration. Measured tissue concentrations can also be used as input to a food-chain or ecosystem model, or to validate a food chain or ecosystem model.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Bulk tissue concentrations do not provide a measure of effect, rather they provide a measure of exposure. Many species have the ability to metabolize or otherwise sequester some COPCs (e.g., copper is sequestered by metallothionein in fish; PAHs are metabolized and excreted in bile). See text of main document regarding selection of appropriate dose measurements.

- Ancillary measurements such as lipid and moisture content may be necessary to interpret the data, as environmental quality guidelines may be presented as "normalized" concentrations.

- The concentration of a given COPC in an organism may be affected by numerous factors such as: exposure route and duration; life stage and sex of the organism; physiological ability to detoxify and/or excrete the COPC; the condition of the exposed organism. Variability in the tissue chemistry data may be high.

Where can I go for further information about this tool? The following guidance manuals are available online:

- B.C. Ministry of Environment, Lands and Parks. 1997. Freshwater biological sampling manual. Prepared for the Resources Inventory Committee.

- B.C. Ministry of Environment, Lands and Parks. 1997. Fish collection methods and standards. Version 4.0. Prepared for the Resources Inventory Committee.

- Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring.

DIRECT MEASUREMENT TOOL
#3: CHEMICAL ANALYSES OF POREWATER

What does this tool consist of? Measurement of the bulk concentration of contaminants in porewater (freshwater or marine) using analytical chemical techniques.

Which ecosystem(s) would this tool typically be applied in? More commonly in Deep Aquatic and Shoreline ecosystems, but also applicable to Rivers and Streams.

How frequently is this tool used in a DERA? Occasionally used in aquatic ERAs.

What are the benefits of using this tool in a DERA?

- Measuring concentrations of COPCs in porewater provides direct information on the sediment associated contaminant fraction that is likely to be most available to some sediment dwelling organisms.

- Porewater testing can provide a complimentary line of evidence to bulk sediment chemistry data.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The choice of collection method should take into consideration the objectives of the sampling program. Porewater samples can be collected using *in situ* methods (e.g., peepers) or *ex situ* (e.g., centrifugation of bulk sediment), with advantages and disadvantages associated with each.

- There can be difficulty in collecting a sufficient volume of porewater for analytical testing, especially when low detection limits are required.

- Ancillary parameters often need to be measured to facilitate interpretation of the data (e.g., *in situ* pH, organic carbon, NH_3, H_2S).

- It is nearly impossible to avoid artifacts and chemical changes during sampling, extraction and storage (i.e., oxidation changes, etc.) of porewater samples.

- Porewater chemistry can also vary seasonally.

Where can I go for further information about this tool?

- Winger, P.V., P.J. Lasier, B.P. Jackson. 1998. The influence of extraction procedure on ion concentrations in sediment porewater. Arch. Environ. Contam. Toxicol. 35: 8-13.

- Carr R.S., Nipper M., Adams W.J., Berry W.J., Burton Jr. G.A., Ho K., MacDonald D., Scroggins R., Winger P.V. 2001. Summary of a SETAC Technical Workshop: Porewater toxicity testing: Biological, chemical and ecological considerations with a review of methods and applications, and recommendations for future areas of research. March 18 - 22, 2000, Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL. 38 p.

DIRECT MEASUREMENT TOOL
#4: SHORT-TERM/ACUTE TOXICITY TESTS

What does this tool consist of? Toxicity tests are studies specifically designed to determine whether exposure to a particular substance causes an adverse effect in a group of test organisms. These tests may be conducted on water, sediment or soil samples. Acute toxicity tests are defined as being of short duration, involving exposures ranging from minutes to a few days, relative to the lifespan of the test organism. Acute toxicity tests are defined as tests with duration of less than 10% of the lifespan of the test organism.

The endpoint most commonly measured in acute toxicity tests is lethality (or immobilization in the case of the cladoceran, *Daphnia* sp.). Some toxicity tests that use acute exposures also include measurement of sublethal endpoints (e.g., echinoid fertilization, trout embryo viability), and are sometimes used as surrogates for estimating chronic toxicity. Acute toxicity tests are usually conducted in a laboratory under controlled conditions, although they may also be conducted *in situ*. Tests may be conducted using either single-concentration or multi-concentration experimental designs, although sediments are generally tested without dilution. A negative (clean) control must always be tested concurrently, to assess natural background variability in the test population and determine test acceptability. Acute tests with daphnids or fish do not require replication, but other acute test methods do. Identical numbers of test organisms (of similar size/age) are exposed to the test material for a defined period of time under controlled laboratory conditions (i.e., temperature, light, water quality) and then the number of surviving organisms in each treatment is determined at the end of the test. Responses in the test treatments are compared to the negative control. If a multi-concentration test was performed, then an LC*p* (concentration estimated to be lethal to percentage "*p*" of the test population) can be calculated.

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic, Shoreline, Rivers and Streams, Uplands (Wildlands) and Uplands (Human-Use). Acute toxicity tests are applicable for all five ecosystem types. Standardized test methods are available for water (which may include groundwater, effluent, leachate, or receiving water), sediment and soil test species.

How frequently is this tool used in a DERA? Acute toxicity tests are commonly used in DERAs.

What are the benefits of using this tool in a DERA?

- Acute toxicity tests provide direct measurements of potential adverse effects to aquatic or terrestrial receptors of concern—information that cannot be obtained from chemistry measurements alone.

- Acute toxicity tests conducted in the laboratory are designed to be performed under controlled environmental conditions, so that the only variable under investigation should be the test material. The use of standard test methods and test species facilitates repeatability and reproducibility, and allows for comparison of data generated by different laboratories for different sites.

- Toxicity tests are useful for evaluating the effects of mixtures of contaminants (including contaminants not measured using analytical chemistry) and also provide an indication of the potential contaminant bioavailability under influence of modifying factors such as water hardness, organic carbon content or particle size.

- Toxicity tests can be used to predict potential adverse effects to receptors of concern; this differs from other biological assessment tools such as benthic community structure, which can only show whether than alteration has already occurred.

- Toxicity tests can be useful for identifying whether alterations to biological communities are due to contaminant exposure or some other stressor when interpreted as part of a weight-of-evidence assessment.

- Acute toxicity tests are particularly useful in DERAs as a screening tool in a tiered testing approach. For example, there would be little benefit to subjecting samples with high acute lethality to further chronic toxicity testing to evaluate potential sublethal effects. Samples demonstrating high acute lethality might need to be submitted for Toxicity Identification Evaluation (TIE) testing if identifying the stressor(s) causing the toxicity was important for site remediation. Conversely, samples demonstrating little or no acute toxicity could be subjected to further evaluation, such as chronic toxicity testing, in the next tier of the DERA investigation.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Acute toxicity tests do not provide information about the specific stressor(s) causing the observed toxicity, unless further evaluation is conducted using TIE manipulations.

- Acute toxicity tests do not provide information about sublethal effects such as growth, reproduction or development. On their own, they may provide enough information to identify areas that are not suitable for risk assessment (i.e., high toxicity) but they do not provide enough information about potential effects that may be associated with in-place risk management.

- Toxicity tests performed under laboratory conditions may not totally reflect "real world" conditions. Sample collection, transport, storage and manipulation during testing may alter sample properties that influence contaminant bioavailability (e.g., oxidation of an anoxic sample).

- Toxicity tests are performed with a limited number of species. Linkage to the ROPCs and measurement endpoints in the DERA framework is necessary.

Where can I go for further information about this tool?

- ASTM International. 2004. Standard guide for conducting acute toxicity tests on test materials with fishes, macroinvertebrates, and amphibians. Method E729-96 (re-approved 2002). In: 2004 Annual Book of ASTM Standards, Water and Environmental Technology, Volume 11.05. ASTM International, West Conshohocken, PA.

- Landis, W.G. and M-H Yu. 2004. Introduction to Environmental Toxicology: 3rd Edition. Lewis Publishers, Boca Raton, FL. 328 pp.

- Toxicity test method protocols can be found electronically at:

Environment Canada:	http://www.etc-cte.ec.gc.ca/organization/spd_e.html
USEPA :	http://www.epa.gov/waterscience/WET/
	http://www.epa.gov/ost/library/sediment/

DIRECT MEASUREMENT TOOL
#5: LONG-TERM/CHRONIC TOXICITY TESTS

What does this tool consist of? Toxicity tests are studies specifically designed to determine whether exposure to a particular substance causes an adverse effect in a group of test organisms. These tests may be conducted on water, sediment or soil samples. Chronic toxicity tests are defined as being of relatively long duration, involving a substantial portion of the test organism's lifespan (10% or greater). Surrogates for chronic tests are also used (i.e., test has a duration that is less than 10% of the organism's life cycle but measures a sensitive life stage). In addition to the test duration, an important distinction between acute and chronic tests is the endpoints measured. While lethality can be measured in both acute and chronic toxicity tests, it is the measurement of sublethal endpoints such as growth, development or reproduction that is most important in chronic toxicity tests. Chronic toxicity tests are usually conducted in a laboratory under controlled conditions in a manner identical to acute toxicity tests although they may also be conducted *in situ*.

Which ecosystem(s) would this tool typically be applied in? Chronic toxicity tests are applicable for all five ecosystem types. Standardized test methods are available for water (which may include groundwater, effluent, leachate, or receiving water), sediment and soil test species. Test methods are more widely developed for water-column testing than for sediment or soil testing.

How frequently is this tool used in a DERA? Chronic toxicity tests are likely to be commonly used in DERAs, although chronic "surrogates" are likely to be used more frequently than chronic tests involving full life-cycle exposures.

What are the benefits of using this tool in a DERA?

- All benefits of acute toxicity testing are also applicable to chronic toxicity tests.

- Chronic (and chronic surrogate) toxicity tests are particularly useful in DRAs when they are used in conjunction with acute toxicity tests in a tiered testing framework. Chronic toxicity tests generally require greater expense and effort than acute tests, so using acute testing as a screening tool to identify those samples that warrant further assessment using chronic toxicity testing is beneficial for prioritizing and focusing available resources.

What are the common "pitfalls" or issues that should be considered when using this tool in a DRA?

- All pitfalls applicable to acute toxicity testing are also applicable to chronic toxicity testing.

Where can I go for further information about this tool?

- Toxicity test method protocols can be found electronically at:

  Environment Canada: http://www.etc-cte.ec.gc.ca/organization/spd_e.html
  USEPA : http://www.epa.gov/waterscience/WET/
  http://www.epa.gov/ost/library/sediment/

- USEPA (US Environmental Protection Agency). 2001. Methods for assessing the chronic toxicity of marine and estuarine sediment-associated contaminants with the amphipod *Leptocheirus plumulosus*. US Environmental Protection Agency, Office of Research and Development, Newport, OR. EPA/600/R-01/020. 104 pp.

DIRECT MEASUREMENT TOOL
#6: MULTI-GENERATIONAL TOXICITY TESTS

What does this tool consist of? Toxicity tests are studies specifically designed to determine whether exposure to a particular substance causes an adverse effect in a group of test organisms. These tests may be conducted on water, sediment or soil samples. Multi-generational toxicity tests are an extension of full or partial life-cycle chronic toxicity tests.

In chronic toxicity tests that measure reproduction (e.g., three-brood *Ceriodaphnia dubia* cladoceran test, 28-d *Leptocheirus plumulosus* amphipod test), the number of offspring produced by the test organism is commonly used as the reproduction endpoint. The number of offspring (F1 generation) produced is used to quantify reproductive effects in the parents (P generation), but says nothing about the quality or condition of the offspring themselves. In a multi-generational toxicity test, test organisms are exposed to the stressor(s) of concern for two full generations, from the egg stage of the P generation through to the production of juveniles of the F2 generation. The F1 and F2 generations are isolated and reared under the same exposure conditions that were used for the P generation. Each generation may be evaluated in terms of effects on survival, growth and hatching success; the P and F1 generations may also be evaluated in terms of endpoints such as time to maturity, sex ratios, fecundity, and development of secondary sex characteristics.

Which ecosystem(s) would this tool typically be applied in? Multi-generational chronic toxicity tests are applicable for all five ecosystem types.

How frequently is this tool used in a DERA? Multi-generational toxicity tests are rarely used in DERAs.

What are the benefits of using this tool in a DERA?

- Multi-generational toxicity tests may be useful for the assessment of contaminants of potential concern associated with adverse teratogenic or endocrine-disrupting effects on receptors of concern, which might not be apparent in the parent generation but would be manifested in the offspring of the first or subsequent generations.

- Multi-generational toxicity tests provide a rigourous measure of the potential for adverse chronic effects, because of their extended duration relative to the organism's lifespan.

What are the common "pitfalls" or issues that should be considered when using this tool in a DRA?

- Life history characteristics of the candidate test organism need to be considered, and may limit the number of suitable test species. Ideally, the test species should have a fairly short life cycle with an early onset of sexual maturity (to reduce the overall length of the exposure period) and consistently produce large numbers of offspring so that there will be sufficient numbers of test organisms available from the F1 and F2 generations.

- As the exposure time increases for any toxicity test, the chance of an unexpected event (e.g., equipment failure, reduced organism health) leading to a catastrophic loss of experimental data increases. Costs associated with multi-generational toxicity tests are likely to be high because of the increased degree of monitoring and need for measurement of test endpoints throughout the study.

- Multi-generational tests are not performed routinely and therefore the toxicology database is limited and there is greater uncertainty about the amount of variability that is expected to be associated with each endpoint. This may make interpretation of the test results more difficult.

Where can I go for further information about this tool?

- Lock, K. and C.R. Janssen. 2002. Multi-generation toxicity of zinc, cadmium, copper and lead to the potworm *Enchytraeus albidus*. Environ. Pollut. 117:89-92.

- Newsome, C.S. 1980. A multigeneration fish toxicity test as an aid in the hazard evaluation of aquatic pollutants. Ecotox. Environ. Saf. 4:362-369.

- Vandenbergh, G.F., D. Adriaens, T. Verslycke and C.R. Janssen. In press. Effects of 17α-ethinylestradiol on sexual development of the amphipod Hyalella azteca. Ecotox. Environ. Saf.

DIRECT MEASUREMENT TOOL #7: *IN SITU* TOXICITY TESTS

What does this tool consist of? *In situ* toxicity tests involve conducting the toxicity test in the field (i.e., at the location under investigation) rather than in the laboratory. These tests can be conducted to evaluate water or sediment toxicity, using techniques adapted from laboratory-based acute or chronic toxicity test methods. In situ exposures can also be designed to provide information on contaminant update and accumulation, similar to laboratory-based bioaccumulation tests.

Test organisms, of similar size/age and obtained from an uncontaminated location, are placed in screened enclosures that allow contact with the environmental compartment of interest. Concurrent placement of additional enclosures in uncontaminated reference locations (e.g., upstream of the study area, or in a separate waterbody) is also necessary to assessment natural background responses of the test organisms. The enclosures may be suspended in the water column or anchored to be in contact with the sediment surface. *In situ* toxicity tests conducted with "eyed" eggs of salmonid fish may involve burying incubation enclosures in gravel and monitoring development to assess mortality and hatching rate. The size of the enclosures depends on the size and type of test organism being used, and the screen size needs to be such that organisms cannot escape, but that water can flow through without the screen becoming fouled or clogged. At the end of the exposure period, surviving test organisms are recovered; if the experimental design includes assessment of sublethal endpoints (e.g., growth) or tissue chemistry analyses, that is also done using the surviving specimens from each treatment. Responses between exposure and reference treatments are compared.

Which ecosystem(s) would this tool typically be applied in? *In situ* toxicity tests are applicable for all five generic ecosystem types, but are most common in the aquatic ecosystems. Fish and bivalves have been used most often for *in situ* testing, but other invertebrates have also been used.

How frequently is this tool used in a DERA? *In situ* toxicity tests are likely to be used occasionally for DERAs.

What are the benefits of using this tool in a DERA?

- As with laboratory toxicity tests, *in situ* toxicity tests provide a direct measure of potential adverse effects of exposure on test organisms, which cannot be determined from chemistry measurements alone.

- The benefit of using *in situ* toxicity tests is that it allows direct exposure of test organisms to actual site conditions, and therefore eliminates the need for

extrapolation of laboratory-based toxicity testing results to field conditions. *In situ* exposure integrates the environmental variables to which organisms would normally be exposed at a given location (e.g., fluctuations in temperature, water flow, water quality, food supply) as well as factors that may affect the bioavailability of the contaminants of potential concern.

- *In situ* toxicity tests can be designed to use acute or chronic exposures, and to measure lethal and/or sublethal effects, provided that the test species chosen is able to tolerate the exposure without demonstrating adverse effects in the reference locations.

- Provided that stocks are available from an uncontaminated location, it may be possible to use native species for *in situ* testing, rather than surrogate species such as those that are typically used in standardized laboratory toxicity tests.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- *In situ* testing requires that approved transplant permits from applicable regulatory authorities be in place prior to conducting testing. Depending on the target test species, it is possible that permission to transplant test organisms in sensitive watersheds may not be granted. This may make it difficult to locate populations of naïve test organisms that have no previously been exposed to the stressor(s) of concern.

- While *in situ* toxicity tests represent more realistic exposure scenarios than then controlled conditions associated with laboratory experiments, there is a higher degree of variability associated with the field exposures and that can make interpretation of *in situ* test results more difficult. Depending on the exposure duration, fluctuations in temperature and food supply may affect the health of the test organisms and their physiological response to the stressor(s) of concern.

- There is a risk of test chambers being lost or damaged during the exposure period, as a result of adverse weather conditions (storms, high or low water flows), predation or theft, and therefore the loss of associated data. Logistics associated with inspection and monitoring of enclosures during the *in situ* exposure requires consideration of how the enclosures will be anchored, their accessibility, and how to inspect them without causing undue stress to the test organisms as a result of disturbance.

Where can I go for further information about this tool?

- ASTM International. 2004. Standard guide for conducting *in-situ* field bioassays with caged bivalves. Method E2122-02. In: 2004 Annual Book of ASTM Standards, Water and Environmental Technology, Volume 11.05. ASTM International, West Conshohocken, PA.

- BCMWLAP (British Columbia Ministry of Water, Land and Air Protection). 2003. British Columbia field sampling manual for continuous monitoring and the collection of air, air-emission, water, wastewater, soil, sediment, and biological samples. British Columbia Ministry of Water, Land and Air Protection, Water Air and Climate Change Branch, Victoria, BC. January 2003. 383 pp.

- Chappie, D.J. and G.A. Burton Jr. 2000. Application of aquatic and sediment toxicity testing *in situ*. Soil Sed. Contam. 9:219-245.

- Environment Canada. 1999. Guidance document on application and interpretation of single-species tests in environmental toxicology. Environmental Protection Series, Report EPS 1/RM/34, December 1999. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

DIRECT MEASUREMENT TOOL
#8: BEHAVIOURAL TOXICITY TESTS

What does this tool consist of? Toxicity tests are studies specifically designed to determine whether exposure to a particular substance causes an adverse effect in a group of test organisms. These tests may be conducted on water, sediment or soil samples. Behavioural toxicity tests can be used to measure sublethal responses, other than changes in growth or numbers of offspring produced, in receptors that are exposed to stressors of concern. Examples of behaviours that can be assessed include changes in locomotion, respiration, habitat selection, feeding, avoidance (of predators or contaminants), competition, and reproductive behaviour. Changes in behavioural responses are compared to controls to determine whether the observed change in outside the typical range of variability for that species-behaviour combination. These tests can involve short-term or long-term exposures.

Which ecosystem(s) would this tool typically be applied in? Behavioural toxicity tests are applicable for all five ecosystem types. There is a standardized avoidance test method for soil using earthworms.

How frequently is this tool used in a DERA? Behavioural toxicity tests are likely to be used rarely in DERAs, except that earthworm avoidance might be used occasionally.

What are the benefits of using this tool in a DERA?

- The 48-h acute avoidance test with earthworms has been found to be a useful tool for screening soil samples to be included for 56-d chronic toxicity tests.

- Behavioural toxicity tests provide an alternative mechanism for assessment of sublethal effects in receptors of concern.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The behavioural characteristics of the test organism must be understood well before they can be used as measures of sublethal responses. Criteria for defining measured responses also need to be unambiguous, so that subjective judgement of behaviour by observers is avoided.

- Behavioural toxicity test results can easily be influenced by test organism health, care and handling, testing conditions, and prior exposure or experience with the stressor prior to testing.

- Interpretation of behavioural toxicity test results in the context of ecological effects is complicated because of uncertainty as to whether the observed behavioural change is likely to impact on relevant endpoints such as survival, growth or reproduction. If the behavioural change is associated with a short-term exposure, organisms may be able to recover without any long-term effects.

Where can I go for further information about this tool?

- ASTM International. 2004. Standard guide for behavioural testing in aquatic toxicology. Method E1604-94 (re-approved 2002). In: 2004 Annual Book of ASTM Standards, Water and Environmental Technology, Volume 11.05. ASTM International, West Conshohocken, PA.

- Environment Canada. 2004. Biological test method: tests for toxicity of contaminated soil to earthworms (*Eisenia andrei*, *Eisenia fetida*, or *Lumbricus terrestris*). Environmental Protection Series, Report EPS 1/RM/43, June 2004. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

- Morgan, J.D., G.A. Vigers, D.M. Janz, A.P. Farrell and J. Manville. 1991. Acute avoidance reactions and behaviourial responses of juvenile rainbow trout to Garlon 4, Garlon 3A and Vision herbicides. Environ. Toxicol. Chem. 10:73-79.

DIRECT MEASUREMENT TOOL
#9: TOXICITY IDENTIFICATION EVALUATION (TIE)

What does this tool consist of? Toxicity identification evaluations (TIEs) consist of side-by-side toxicity testing using manipulated and non-manipulated samples. Manipulations (chemical or physical) are selected to target specific toxicants (or groups of toxicants) known or suspected to be present in a sample. Differences in the toxicity between the manipulated and non-manipulated samples support inferences about chemical compounds or sample-related factors that are contributing to the original toxicity.

Which ecosystem(s) would this tool typically be applied in? TIEs can be applied in all five ecosystem types. TIE procedures are relatively well-developed for aqueous samples (e.g., porewater, groundwater, overlying water, and effluent), somewhat less developed for whole sediments, and relatively limited for soils. Techniques for whole sediment TIEs have received increased attention in recent years.

How frequently is this tool used in a DERA? Rare for DERAs involving soils and sediment. Occasional for DERAs involving aqueous samples.

What are the benefits of using this tool in a DERA? TIEs directly evaluate cause-effect relationships. TIEs can be used to determine the relative influence of physical versus chemical-related effects. Assessing the relative contribution of different chemicals also improves the ability of the risk characterization to guide appropriate risk management planning. TIEs are particularly useful for identifying contributions of ancillary chemicals (e.g., ammonia, sulphide, dissolved oxygen) to observed toxicity. At many sites, effects are often incorrectly ascribed to contaminants (e.g., metals, PAHs) on the basis of sediment quality value exceedances. TIEs address this problem by indicating the contaminant group(s) most likely responsible for the observed responses.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- TIEs are typically conducted after (or concurrently with) other toxicity testing. Careful consideration of how to integrate sample collection for both toxicity testing and a TIE is required. For example, sufficient sample volumes need to be collected in advance if TIE is contemplated for a sediment quality Triad.

- TIEs are most effective when applied to samples that exhibit pronounced toxic responses. TIEs are less useful when toxicity is minor to moderate in magnitude.

- TIEs operate in an iterative fashion where the results of one type of manipulation lead to other potential manipulations that should be examined. The scope of the TIE cannot often be predicted in advance, although there should be consensus regarding the desired level of identification (i.e., do you need to know which divalent metal is causing the toxicity, or is it enough for site management purposes to know that a contaminant group is responsible? [e.g., divalent metals, non-polar organics]). The tiered approach, although cost-efficient, can be problematic in practical terms because site managers often require certainty in project cost and timelines at the beginning of a project.

- The TIE needs to consider a substantially broader range of potential contaminants and factors than would normally be measured to meet CSR requirements. Non-listed contaminants or physical factors may also be contributing to the toxicity.

- TIEs often require substantial professional judgment in interpreting the multiple lines of evidence. The physical and chemical manipulations of samples can cause complex interactions in the bioavailability of different sample constituents. For example, purging of sediments to reduce the influence of volatiles can have the side-effect of increasing the bioavailability of metals. The TIE investigator needs to be aware of the influence of different manipulations, and interpretation can be complex where multiple stressors of concern are present.

- TIEs are most easily conducted on aqueous samples, and for this reason, sediment assessments often apply TIEs to porewater extracted from sediments. The investigator needs to be aware of the physicochemical implications of processing sediments to obtain porewater, and understand the ecological relevance of porewater toxicity testing to the receptors of concern.

- TIEs are conducted on individual samples using individual test organisms, and as such, represent a “snapshot” of cause-and-effect relationships. Seasonal (and other) variations are not considered; additionally, the toxic mode of action may vary for different organisms. Interpretation of TIE results under these circumstances as indicative of the overall ecological effects at a site is problematic. TIEs may need to be repeated using multiple test organisms and samples.

Where can I go for further information about this tool? The following TIE guidance manuals are available online:

- USEPA. 1991. Method for Aquatic Toxicity Identification Evaluations, Phase I Toxicity Characterization Procedures. EPA/600/6-91/003.

- USEPA. 1993. Methods for Aquatic Toxicity Identification Evaluations: Phase II Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA 600/R-92-080

- USEPA. 1993. Methods for Aquatic Toxicity Identification Evaluations: Phase III Toxicity Confirmation Procedures for Samples Exhibiting Acute and Chronic Toxicity. EPA 600/R-92-081

- USEPA. 1996. Marine Toxicity Identification Evaluation (TIE), Phase I Guidance Document. U.S. EPA, ORD, EPA/600/R-95/054.

DIRECT MEASUREMENT TOOL #10: HISTOPATHOLOGY

What does this tool consist of? Histopathology involves microscopic examination of tissues (e.g., gonads; liver) for cellular damage (e.g., lesions); usually applied to fish. Frequently combined with biomarker approaches.

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic and Rivers & Streams

How frequently is this tool used in a DERA? Rare

What are the benefits of using this tool in a DERA?

- Provides information on adverse effects that may be occurring in individual organisms at exposure concentrations lower than those that result in adverse effects on growth or reproduction. Provides information regarding the "health" of organisms.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Cause and effect relationships may not be clear; organisms may also suffer from disease which cause similar histopathological alterations.

- Histopathology is highly specialized: sample collection, preparation and analysis require substantial expertise.

- Substantial numbers of samples may be required to achieve the necessary statistical power.

Where can I go for further information about this tool?

- AETE (Aquatic Effects Technology Evaluation. 1998. Technical Evaluation of Histopathology as an Environmental Monitoring Tool for the Mining Industry in Canada. Report 2.2.2. Available online: http://www.nrcan.gc.ca/mms/canmet-mtb/mmsl-lmsm/enviro/metals/aete.htm

- USEPA. 1987. Guidance for Conducting Fish Liver Histopathology Studies During 301(H) Monitoring. EPA 430/987/004.

DIRECT MEASUREMENT TOOL #11: DEFORMITY ASSESSMENTS

What does this tool consist of? Visual inspection of organisms (usually larval fish or amphibians) from either chronic toxicity testing or from field sampling. The frequency and magnitude of deformities (e.g., edema, ocular or skeletal malformation) are measured. Deformity assessments are usually limited to those DERAs involving compounds with a known tendency to cause deformity (e.g., PCBs or pesticides for amphibians; selenium for larval fish).

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic or Rivers and Streams.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA?

- Malformations may occur at concentrations that are lower than thresholds for reproductive or growth effects; it is potentially a more sensitive toxicological endpoint.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Toxicological testing needs to be designed with malformation endpoints in order to properly address statistical power considerations. Adding malformation endpoints to existing chronic toxicity testing protocols is not appropriate.

- Malformation is expressed through a complex mode of toxic action involving the interaction of the contaminant with various stages of organism development. Timing of the exposure may be a significant confounding factor.

- Not all malformations are equal in terms of their ecological relevance. To date, few studies have explored the ecological relevance of malformation to larval organisms in terms of their population level impacts.

Where can I go for further information about this tool?

- A literature search for the compound of interest is recommended.

DIRECT MEASUREMENT TOOL #12: STABLE ISOTOPE ANALYSES

What does this tool consist of? Ratios of the stable isotopes of carbon (^{13}C vs ^{12}C) and nitrogen (^{15}N vs ^{14}N) can be used to infer feeding relationships, which can in turn be used to model the trophic transfer of contaminants. Stable isotope ratios are expressed in "delta" units (written $\delta^{13}C$ and $\delta^{15}N$). The combination of $\delta^{13}C$ and $\delta^{15}N$ is sometimes referred to as an organism's "stable isotope signature". $\delta^{13}C$ reflects the carbon source at the base of an organism's food web (e.g., benthic algae vs. phytoplankton, or a mixture of the two). $\delta^{13}C$ typically changes very little between diet and consumer ("you are what you eat"). $\delta^{15}N$ reflects the organism's trophic level (TL), and tends to increase between diet and consumer. An organism's $\delta^{15}N$ must be interpreted relative to the $\delta^{15}N$ of the base of the food web, commonly by using clams or some other herbivore to provide a long-term average $\delta^{15}N$ for the basal resource:

$$TL = TL_{baseline} + (\delta^{15}N_{organism} - \delta^{15}N_{baseline})/3.4\ ‰$$

where TL_{clam} is the trophic level of the species used as a baseline (2.0 for clams or other herbivores, 1.0 for plants) and 3.4‰ is the average enrichment in $\delta^{15}N$ between diet and consumer (called "trophic fractionation").

Which ecosystem(s) would this tool typically be applied in? Stable isotope analysis can be applied to establish feeding links and diet compositions in the food web of any ecosystem type.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? Stable isotope analysis provides site-specific, time-integrated diet information for receptors, and can be invaluable in estimating the exposure of these receptors to COPCs via their diets. Trophic level, inferred from $\delta^{15}N$, is especially important for estimating exposure to chemicals that biomagnify. Stable isotope analysis requires only a small amount of material (typically < 1 mg) and is inexpensive (~$10-20/sample).

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Site-specific measurements are essential. There can be tremendous spatial variation in the stable isotope signatures of basal resources (phytoplankton, vascular plants, detritus), produced by local variation in nutrient sources, currents and mixing, growth rates, etc. This will produce spatial variation in stable isotope ratios of the animal species farther up the food chain. It is not possible to assume that the stable isotope signature of an organism is the same as that measured in other areas.

- Small and short-lived species can exhibit large temporal variation in stable isotopes, which can make it difficult to correctly interpret feeding relationships from a 'snapshot' study. This problem can be circumvented for phytoplankton by sampling large-bodied, long-lived herbivores (e.g., clams) and then inferring the mean phytoplankton signature from this. A similar approach can be used for zooplankton if strictly zooplanktivorous fish are available; otherwise, it is best to have repeated (e.g., seasonal) sampling to capture this temporal variability.

- There can be substantial variation in stable isotope signatures among tissues within an animal. For small animals that are consumed whole (e.g., insects), it is appropriate to use a whole-body analysis. For larger animals (e.g., fish), analysis is typically done on muscle tissue. If non-lethal sampling is desired, it is possible to use scales, hair or feathers, but it is then necessary to know how the stable isotope signature of this tissue relates to that of the animal's bulk muscle tissue (the edible part, and the main repository of nitrogen in animals).

Where can I go for further information about this tool?

- Cabana, G., and J. B. Rasmussen. 1994. Modeling food chain structure and contaminant bioaccumulation using stable nitrogen isotopes. Nature 372: 255-257.

- Peterson, B. and B. Fry. 1987. Stable isotopes in ecosystem studies. Annu. Rev. Ecol. Syst. 18: 293–320.

- Post, D.M. 2002. Using stable isotopes to estimate trophic position: models, methods, and assumptions. Ecology 83: 703-718.

- Vander Zanden, M. J. and J. B. Rasmussen. 1999. Primary consumer δ13C and δ15N and the trophic position of aquatic consumers. Ecology 80: 1395–1404.

DIRECT MEASUREMENT TOOL #13: BIOMARKERS

What does this tool consist of? Biomarkers are measurable biological parameters that change in response to xenobiotic exposure and other environmental or physiological stressors, and can be indices of toxicant exposure or effects. Examples: bile fluorescent aromatic compounds (FACs); liver enzyme induction (EROD, CYP1a); hematological parameters; steroid hormone levels. Often combined with histopathology.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? If the biomarkers are sufficiently specific and well characterized, they can provide meaningful data for the risk assessment process by providing an indication of the degree of exposure of humans or animals in natural populations to a specific xenobiotic or class of xenobiotics.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Most biomarkers are effective as indices of exposure, but adequate information is rarely available on the underlying dose-response curves. Biomarkers are rarely useful in providing information about effects.

- Biomarkers tend to measure changes in sub-organism parameters (e.g., biochemistry; enzyme activity) which do not necessarily translate into a relevant endpoint for DERA purposes (e.g., organism-level endpoint such as survival, growth, deformity and reproduction)

- The degree of a change in a biomarker parameter can be influenced by multiple endogenous (e.g., age) and exogenous (e.g., chemical exposures) factors. Many biomarkers respond to multiple COPCs or groups of COPCs (e.g., CYP1A responds to multiple types of chemicals), making it difficult to correlate the degree of change in the biomarker with COPC exposure. In generally, most biomarkers are not specific enough for DERA purposes.

- Caution is urged in an attempt to utilize biomarkers in the risk assessment process until more complete documentation is available on the specificity, sensitivity, and time course of changes, and on the impact of multiple exposures or the time of exposures (Chambers et al. 2004).

Where can I go for further information about this tool?

- McCarty, L.S., M. Power and K.R. Munkittrick. 2004. Bioindicators versus biomarkers in ecological risk assessment. Hum. Ecol. Risk Assess. 8: 159-164.

- Chambers, J.E.; J.S. Boone, R.L. Carr, H.W. Chambers and D.L. Straus. 2004. Biomarkers as predictors in health and ecological risk assessment. Hum. Ecol. Risk Assess. 8: 165-176.

- Fossi, M.C. 1994. Nondestructive biomarkers in ecotoxicology. Environ. Health Perspect. 102 (S12): 49-54.

DIRECT MEASUREMENT TOOL
#14: BENTHIC COMMUNITY SURVEYS

What does this tool consist of? Taxonomic identification and enumeration of benthic organisms collected using standardized sampling techniques. Diversity, abundance, multiple other indices can be calculated from this data.

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic and Rivers and Streams

How frequently is this tool used in a DERA? Common

What are the benefits of using this tool in a DERA? Provides a direct measurement of potential long-term toxicant-related effects under actual field conditions.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Benthic community data are sensitive to habitat alteration and other physical factors not related to the COPCs.

- Microscale variation in contaminant distribution can result in substantial variation in benthic community data. Synoptic sampling for chemistry and benthic community data is essential.

- Statistical power is often limited when only one benthic community replicates are collected from each station. Multiple replicates are recommended wherever possible.

Where can I go for further information about this tool?

- Environment Canada provides detailed guidance for the use of benthic community surveys for environmental effects monitoring programs for pulp and paper mills and metal mines. Available online: http://www.ec.gc.ca/eem/english/Publications/default.cfm

- D.M. Rosenberg, D.M., I.J. Davies, D.G. Cobb, and A.P. Wiens. Undated. Protocols for measuring biodiversity: benthic macroinvertebrates in freshwaters. Available online: http://www.eman-rese.ca/eman/ecotools/protocols/freshwater/benthics/intro.html

DIRECT MEASUREMENT TOOL
#15: OTHER POPULATION AND COMMUNITY SURVEYS

What does this tool consist of? Population and community surveys consist of a quantification of types and numbers of organisms.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Common for DERAs for aquatic ecosystems (i.e., deep aquatic; rivers & streams, shoreline). Rare for DERAs for terrestrial ecosystem (uplands wildlands; uplands human use).

What are the benefits of using this tool in a DERA? Direct measurement of *in situ* communities has high ecological relevance. Community surveys complement other lines of evidence (e.g., toxicity tests) since the community structure reflects the response of the ROPCs to environmentally-relevant COPC (e.g., bioavailability, as well as adaptation/acclimation are considered). Data from screening-level qualitative surveys assist in the selection of appropriate ROPCs, and also for determining the relationships between different ROPCs (e.g., significant feeding preferences).

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Community surveys should be considered in conjunction with other lines of evidence because natural variability can confound the interpretation of the survey data. For example:
    - *In situ* communities can be influenced by numerous abiotic habitat factors (e.g., variations in sediment grain size, soil quality, etc) as well as landscape- or watershed-level influences (e.g., habitat alteration from forest fire; logging, etc). It is difficult to establish cause-effect relationships between the community-level measurement and the specific COPCs under investigation.
    - Seasonal influences must be considered (e.g., seasonal patterns in food availability; site occupation; variation in sensitivity to COPCs due to life history stage).
    - Mobile species may also be exposed to stressors outside the contaminated site.

- Historical/baseline data for a given ecosystem are not always available. These data are important for establishing the bounds of natural variability both spatially and temporally.

- Considerable resources may be required to conduct surveys that provide robust enough data (e.g., large enough sample size to detect statistically significant differences; multiple sampling events to address seasonality of the community) for meaningful assessment of potential contaminant-related effects. Study designs for *in situ* community measurements require explicit consideration of statistical issues.

Where can I go for further information about this tool?

- Suter, G.W. 1996. Risk characterization for ecological risk assessment of contaminated sites. Prepared by Lockheed Martin Energy Systems, Inc., Oakridge, TN for U.S. Department of Energy. ES/ER/TM-200.

- B.C. Ministry of Environment, Lands and Parks. 1997. Freshwater biological sampling manual. Prepared for the Resources Inventory Committee.

- Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring.

APPENDIX II

MODELLING TOOLS

TABLE OF CONTENTS

MODELLING TOOL #1: LITERATURE-BASED BIOACCUMULATION/BIOCONCENTRATION FACTORS AND UPTAKE MODELS

What does this tool consist of? Internal concentrations of chemicals in organisms can be related to concentrations in their ambient environment. For most chemicals at relatively low (i.e., typically encountered) ambient concentrations, the ratio of internal to ambient concentrations ($C_{internal}/C_{ambient}$) is assumed to be independent of ambient concentration (i.e., constant). In these cases, chemical accumulation is expressed by a bioconcentration factor (BCF) or bioaccumulation factor (BAF). BCFs and BAFs are simply ratios of the animal's internal concentration (often lipid-normalized, for organics) to the concentration in the water (for aquatic organisms) or air (for terrestrial organisms).

- The BCF is intended to reflect the tendency of a chemical to accumulate in a species via passive diffusion, according to equilibrium partitioning. BCFs are measured in a laboratory, under conditions of water exposure only (i.e., no dietary exposure).

- The BAF is intended to reflect the tendency of a chemical to accumulate in a species via all routes, including passive diffusion from the environment and uptake from the diet. BAFs may be measured in the lab, but are more commonly measured in the field. The BSAF (biota-sediment accumulation factor) is a closely-related approach applied to sediment-associated species.

In some inorganic chemicals, the ratio $C_{internal}/C_{ambient}$ has been observed to vary with $C_{ambient}$. The form of this relationship, often called an "uptake model", is described in documents such as Sample et al. (1999).

Which ecosystem(s) would this tool typically be applied in? All ecosystems.

How frequently is this tool used in a DERA? Common in both aquatic and terrestrial ecosystems.

What are the benefits of using this tool in a DERA? Literature-based BCFs/BAFs and uptake models make it possible to estimate the tissue COPC concentrations for an organism based on data regarding COPC concentrations in its environmental media.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- BCFs may underestimate exposure for COPCs primarily absorbed from the diet (e.g., chemicals with low solubility in water). BCF may greatly underestimate exposure to biomagnifying COPCs. BAFs are always preferred to BCFs.

- Laboratory-derived BCFs are usually based on maximum chemical bioavailability (i.e., low dissolved organic carbon, highly digestible food, etc.), and may therefore overestimate exposure to COPCs under field conditions that reduce bioavailability. Field-derived BAFs may have the same limitation if there are differences in bioavailability between the system in which the BAF was measured and the system in which the DERA is being conducted (i.e., may over- or underestimate exposure, depending on which system has higher bioavailability).

- Laboratory-derived BCFs are not always measured over a long enough period for the animal to approach steady state, and may therefore underestimate the degree of bioaccumulation that will occur under real-world conditions.

- BCFs/BAFs are not available for all species. They should only be extrapolated between species that are very similar with respect to bioaccumulation, especially with respect to their ability to detoxify the chemical.

- BAFs measured in the field vary enormously from site to site due to differences in both the physicochemical environment as well as the interaction of organisms with their environment. The use of literature-based BAFs, BCFs, BSAFs, or uptake model contributes considerable uncertainty to the DERA.

Where can I go for further information about this tool?

- Torres KC, Johnson ML. 2001. Bioaccumulation of metals in plants, arthropods, and mice at a seasonal wetland. Environ. Toxicol. Chem. 20: 2617-2626.

- Torres KC, Johnson ML. 2001. Testing of metal bioaccumulation models with measured body burdens in mice. Environ. Toxicol. Chem. 20: 2627-2638.

- Sample, B. E., G. W. Suter II, J. J. Beauchamp, and R. A. Efroymson. 1999. Literature-derived bioaccumulation models for earthworms: development and validation. Environ. Toxicol. Chem. 18:2110-2120.

- Bechtel Jacobs Company. 1998. Empirical Models for the Uptake of Inorganic Chemicals from Soil by Plants. U. S. Department of Energy, Oak Ridge, TN. Available online at: http://www.esd.ornl.gov/programs/ecorisk/documents/bjcor-133.pdf

MODELLING TOOL
#2: SITE-SPECIFIC BIOACCUMULATION FACTORS OR UPTAKE MODELS

What does this tool consist of? Co-occurring samples of soil and soil invertebrates are collected from the site and analyzed for the contaminants of potential concern. Other combinations of environmental media can also be sampled (e.g., soil and plant tissue; sediment and benthic invertebrates).

- Bioaccumulation factors (BAFs) are determined for each pair of samples, and a summary of the range of BAFs calculated (e.g., mean, 95% upper confidence limit of the mean; 90^{th} percentile or maximum)

- Uptake models can be developed using regression analyses to fit an appropriate model (based on *p* and R^2 values) to the available co-occurring soil and tissue data (e.g., linear, exponential, power). Multivariate analyses can be used to improve the predictive ability of these uptake models if data are also available for soil parameters that influence contaminant bioavailability (e.g., soil pH; sediment AVS concentrations; organic carbon concentration).

The summary BAFs or uptake models are then used to predict tissue concentrations across the remainder of the site based on the available soil chemistry data.

Which ecosystem(s) would this tool typically be applied in? Uplands (Wildlands) for soil measurements; Deep Aquatic for sediment measurements

How frequently is this tool used in a DERA? Common for Uplands (Wildlands); occasional for Deep Aquatic.

What are the benefits of using this tool in a DERA? Construction of a site-specific BAF or uptake model is a compromise between the use of literature-based BAFs and uptake models (see Modelling Tool #1) versus collection of substantial numbers of tissue samples from the site, as follows:

- Literature-based BAFs and uptake models are developed for specific sites (or limited number of sites); their application represents a substantial source of uncertainty. This uncertainty can be as much as several orders of magnitude in terms of both over-predicting and under-predicting tissue concentrations. Developing a site-specific BAF or uptake model substantially reduces this uncertainty.

- For large sites, developing a site-specific BAF or uptake model is advantageous in that it reduces the sampling effort (and costs) that would be required to provide sufficient spatial coverage for tissue samples.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The sampling used to develop the site-specific BAF or uptake model needs to reflect the full range of contaminant concentrations across the site. Bioaccumulation of many contaminants is dependent on concentration—a sampling program that focuses on worst-case areas is not necessarily a conservative approach. Uptake models are frequently superior to BAFs since they facilitate consideration of the soil (or sediment) concentration in the resulting tissue predictions.

- Consider sampling for other ancillary parameters that influence contaminant bioavailability in soil to facilitate development of multivariate uptake models. Risk assessors should be familiar with these factors for the applicable combinations of receptors and contaminants-of-concern.

- An inherent assumption in the collection of co-occurring samples is that the tissue items collected are also relatively immobile and in direct contact with the environmental media of interest. This assumption is not completely true: earthworms and benthic invertebrates have a limited degree of mobility; root systems may extend over a considerable area. Multiple soil or sediment samples from within the potential area from which the organism is accumulating contaminants may be appropriate to evaluate this uncertainty. This approach is not suitable for highly mobile taxa that are not in continual contact (e.g., benthic fish/sediment; predatory beetles/soil) unless the soil/sediment sampling plan also reflects the increased area of potential exposure.

- The uncertainty in the site-specific BAF or uptake model is strongly influenced by sample size. Determination of a minimum site-specific sample size should consider of contaminant distribution, heterogeneity, seasonal effects, and size of the area of interest. A minimum sample size of 10 is recommended unless it can be demonstrated that a smaller sample size is appropriate. Note that minimum sample sizes of greater than 10 may be necessary depending on the factors above.

- Sampling must consider the confounding effect of soil particles in the tissue analyses. In most instances, the objective is to predict the bioaccumulation of contaminants within the tissue of the organism, and therefore, organisms should

be well-rinsed (and blotted dry) as well as depurated (if applicable) to reduce the influence of this confounding factor.

Where can I go for further information about this tool? Examples of one or more aspects of the issues discussed above can be found in the following peer-reviewed scientific literature:

- Efroymson RA, Sample BE, Suter GW II. 2001. Uptake of inorganic chemicals for soil by plant leaves: regressions of field data. Environ. Toxicol. Chem. 20: 2561-2571.

- Hunter BA, Johnson MS, Thompson DJ. 1987. Ecotoxicology of copper and cadmium in a contaminated grassland ecosystem. II. Invertebrates. J. Appl. Ecol. 24: 587-599.

- Torres KC, Johnson ML. 2001. Bioaccumulation of metals in plants, arthropods, and mice at a seasonal wetland. Environ. Toxicol. Chem. 20: 2617-2626.

- Torres KC, Johnson ML. 2001. Testing of metal bioaccumulation models with measured body burdens in mice. Environ. Toxicol. Chem. 20: 2627-2638.

MODELLING TOOL
#3: BIOMAGNIFICATION OR TROPHIC TRANSFER FACTORS

What does this tool consist of? The biomagnification factor (BMF) or trophic transfer factor (TTF) is the ratio of chemical concentration between a species and its diet. BMF usually refers to organic chemicals (usually lipid-normalized concentrations) and the TTF usually refers to metals. The food web magnification factor (FWMF) or trophic magnification factor (TMF; especially in Europe) is an expression of the average BMF across several trophic levels. All of these terms reflect the tendency of a substance to biomagnify, i.e., increase in concentration at higher trophic levels.

Measured BMFs for many substances in many types of organisms are available in the literature. BMFs for metals are usually near or below 1 since most metals and metalloids do not biomagnify (mercury and selenium are notable exceptions). BMFs for organic substances range from well below 1 (e.g., in poorly-absorbed or rapidly-metabolized chemicals) to values on the order of 10 (fish, invertebrates) to 100 or more (birds, mammals).

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Common.

What are the benefits of using this tool in a DERA? Biomagnification factors can be used to assess the exposure of receptors to COPCs in their diets. BMFs are often used in food web models to predict exposure throughout the food web.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- BMFs are taxon- (and sometimes species-) specific. It may be possible to generalize among similar species within a higher taxon (e.g., compile several BMFs for a chemical in fish, and use these to estimate the BMF for another fish species), but these cannot be used to estimate the BMF of another taxon (e.g., extrapolate from fish to a bird species). Different taxa have different gut absorption efficiencies, experience different degrees of gastrointestinal magnification, and have very different capacities to metabolize and excrete various chemicals. These differences produce very large differences among species in BMFs.

- The same taxa may occupy different levels in the food web at different locations, depending on the availability of prey items and competitive pressures. See Direct Measurement Tool # 12 for methods to determine site-specific food webs.

- BMFs are chemical-specific. In no case is it possible to use a measured BMF for one chemical to estimate the BMF for another chemical. BMFs are highly sensitive to the metabolizability of the chemical, and this is difficult to predict from chemical structure.

- Accurately estimating BMFs from models requires information on metabolizability, which is not often available.

Where can I go for further information about this tool?

- Kelly, B.C., McLachlan, M.S. and Gobas, F.A.P.C. 2004. Intestinal absorption and biomagnification of organic contaminants in fish, wildlife and humans. Environ. Toxicol. Chem. 23: 2324-2336.

- Gobas, F.A.P.C. and J.B. Wilcockson 1999. Mechanism of biomagnification in fish under laboratory and field conditions. Environ. Sci. Technol. 33: 133-141.

- Campbell L.M., A.T. Fisk, X. Wang, G. Köck and D.C.G .Muir. 2005. Evidence for biomagnification of rubidium in freshwater and marine food webs. C. J. of Fish. Aquat. Sci. 62:1161-1167.

MODELLING TOOL
#4: MASS-BALANCE BIOACCUMULATION MODELS

What does this tool consist of? These are mechanistic models used to estimate the bioaccumulation of chemicals in organisms. The basic form of the model is an individual-based chemical mass balance, balancing the sum of inputs (dietary uptake, respiratory absorption) against the sum of outputs (fecal egestion, respiratory elimination, metabolic transformation, growth dilution). These models are typically used to estimate steady-state concentrations (i.e., by assuming that internal chemical concentrations are not changing over time), but the approach can also be used in a time-dependent formulation. Taxon-specific mass-balance bioaccumulation models have been developed for many species and higher taxa, and recently some very general models have been developed that can be used for almost any animal species.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Occasional

What are the benefits of using this tool in a DERA? Mechanistic bioaccumulation models can be used to estimate internal concentrations of COPCs for any receptor. If it is not possible to directly measure internal concentrations (e.g., the species is protected or inaccessible), mechanistic models may be the best way to obtain estimates of exposure. These models can be made more or less site-specific, depending on how much site-specific information is available. These models have been validated in a wide variety of environments, and typically can predict internal concentrations with fairly good precision (often within a factor of 2-3).

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Substantial information may be required to parameterize this model. It is not always appropriate to use generic parameters; it is always best to have site-specific values.

- As with all mechanistic models, some direct chemical measurements are necessary to validate the model.

- Mechanistic bioaccumulation models typically assume that the organism is at steady state (unchanging internal concentration) and are more difficult to apply to situations where this is unlikely to be a good assumption, for example if ambient concentrations vary a lot over time (especially in the case of a spill) or if the receptor migrates between contaminated and uncontaminated areas.

Where can I go for further information about this tool?

- Arnot, J.A. and F.A. P. C. Gobas. 2004. A Food Web Bioaccumulation Model for Organic Chemicals in Aquatic Ecosystems. Environ. Toxicol. Chem. 23: 2343-2355.

- Connolly, J. P.; Pedersen, C. J. 1988. A thermodynamic-based evaluation of organic-chemical accumulation in aquatic organisms. Environ. Sci. Technol. 22: 99.

- deBruyn, A.M.H. and F.A.P.C. Gobas. 2005. A Bioenergetic Biomagnification Model for the Animal Kingdom. in press, Environ. Sci. Technol.

- Glaser, D.; Connolly, J. P. 2002. A model of p,p '-DDE and total PCB bioaccumulation in birds from the Southern California Bight. Continental Shelf Res. 22: 1079.

- Gobas, F.A.P.C. 1993. A Model for Predicting the Bioaccumulation of Hydrophobic Organic Chemicals in Aquatic Food-Webs: Application to Lake Ontario. Ecol. Modelling. 69: 1-17.

- Kelly, B.C. and F.A.P.C. Gobas. 2003. An Arctic Terrestrial Food-Chain Bioaccumulation Model for Persistent Organic Pollutants. Environ. Sci. Technol. 37: 2966-2974.

MODELLING TOOL
#5: FUGACITY FATE AND TRANSPORT MODELS

What does this tool consist of? Fugacity-based fate models are used to predict chemical concentrations in abiotic media (water, air, soil, etc.) and in biota in a specified environment. Fugacity is directly proportional to chemical concentration, but is normalized to the sorptive capacity of a particular medium. Fugacity is effectively a measure of the tendency of a chemical to migrate between media.

Which ecosystem(s) would this tool typically be applied in? Fugacity models can be used in any type of real or hypothetical ecosystem at any scale. Many fugacity models have been developed for individual bodies of water or watersheds, but the approach has also been applied at regional, continental and global scales. The fugacity approach has also been used in detailed models of bioaccumulation and the distribution of chemicals within an organism.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? Fugacity models can be used to predict chemical concentrations in any abiotic medium or type of organism. When direct measurements of some concentrations are available, these may be used to validate the model. Well-defined methods exist to estimate the necessary parameters (e.g., sorptive capacities). There are four levels of complexity in fugacity modeling, so it is possible to construct a very simple model (with few parameters) when this is appropriate, and to increase the level of complexity as necessary.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Fugacity models are most appropriate for neutral organic substances because it is straightforward to estimate the sorptive capacities of environmental media and biota for these chemicals. Some fugacity models have been developed for charged organic substances, but the fugacity approach is difficult to apply to inorganics.

- As with any model, the output is only reliable if the model is well-constructed and the parameters are accurate. The simpler fugacity models make many simplifying assumptions, and may not accurately reflect reality. The more complex fugacity models require a large number of parameters to describe the ecosystem, so there is greater potential for compounding errors and uncertainty.

- Most existing fugacity models come with default parameter sets, but these are not appropriate for all ecosystems. It is essential to evaluate all parameter choices with respect to the particular system being assessed.

Where can I go for further information about this tool?

- Mackay D. 2001. Multimedia Environmental Models: The Fugacity Approach - Second Edition, Lewis Publishers, Boca Raton, Florida.

- Woodfine D.G., M. MacLeod, D. Mackay and J.R. Brimacombe. 2001. Development of continental scale multimedia contaminant fate models: integrating GIS. Environ. Sci. & Pollut. Res. 8:164-172.

- Kelly, B.C. and F.A.P.C. Gobas. 2003. An arctic terrestrial food-chain bioaccumulation model for persistent organic pollutants. Environ. Sci. Technol. 37: 2966-2974.

A detailed introduction to fugacity-based multimedia fate models and a wide selection of downloadable models is available from the Canadian Environmental Modelling Centre at Trent University:

http://www.trentu.ca/cemc/CEMC200102.pdf
http://www.trentu.ca/cemc/new.html

MODELLING TOOL
#6: PHYSIOLOGICALLY-BASED PHARMACOKINETIC MODELS

What does this tool consist of? Physiologically-based pharmacokinetic (PBPK) models mechanistically predict the uptake and distribution of substances within an individual organism. PBPK models represent various parts of the body as interconnected "compartments", usually specifying at least three such compartments (e.g., blood, rapidly-perfused tissues and slowly-perfused tissues) and often specifying many more than three (e.g., a compartment for each major organ). Transfer among compartments is usually considered to be via blood, and is therefore a function of tissue-blood partition coefficients, the volume of the tissue, and the flux of blood through the tissue. Mathematically, PBPK models use differential equations to describe the chemical concentration in each compartment as a function of the concentrations in other compartments.

Which ecosystem(s) would this tool typically be applied in? PBPK models are usually applied to mammals, and could be used in any ecosystem in which mammals are a receptor of concern.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? PBPK models can provide detailed information on exposure of receptors to COPCs via all routes simultaneously (ingestion, inhalation/gill exchange/transdermal absorption). PBPK models can predict total uptake rates, internal whole-body concentrations, or concentrations in specific target organs, and can therefore be used with dose-response relationships (ecological effects profiles) based on any of these measures of exposure.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- PBPK models have been described as data-hungry, resource intensive, complex, time consuming, compound-specific and difficult to validate.

- PBPK models require so much detailed information on the physiology of the receptor and the physical-chemical properties of the chemical, that they typically can only be constructed for very well-known species, such as humans.

Where can I go for further information about this tool?

- Cahill, T., Cousins, I., and Mackay D. 2003. Development and application of a generalized physiologically based pharmacokinetic model for multiple environmental contaminants. Environ. Toxicol. Chem. 22: 26-34.

- Clark, L.H., Setzer, R.W. and Barton, H.A. (2004) Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. Risk Anal. 24: 1697-1718.

MODELLING TOOL
#7: SEM-AVS

What does this tool consist of? The bulk concentrations of metals in sediments are poor predictors of their bioavailability to aquatic organisms. A comparison of the molar concentrations of acid volatile sulphide (AVS) and simultaneously extractable metals (SEM) has been found to be a useful predictive tool for assessing the potential for divalent metals(e.g., cadmium, lead, zinc) to cause toxicity in sediments. If the ratio of SEM:AVS is less than 1 or SEM minus AVS is less than zero, then toxicity is not expected. If the ratio of SEM to AVS is greater than 1 or the difference is greater than 0, then benthic organisms may or may not be exposed to toxicity.

Which ecosystem(s) would this tool typically be applied in? SEM-AVS can applied in aquatic systems but is generally most relevant for anaerobic sediments where sulphides can accumulate (i.e., this tool is not very useful in highly oxidized environments).

How frequently is this tool used in a DERA? This tool is commonly used in a DERA of freshwater and marine sediments.

What are the benefits of using this tool in a DERA? AVS-SEM data provide information regarding the potential bioavailability of selected divalent metals and can therefore help assess the potential for effects if bulk sediment chemistry results exceed published sediment quality guidelines.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The ratio of SEM to AVS is not reliable at low AVS concentrations.
- AVS may be lost from a sample prior to analysis if handled improperly (e.g., the sample is not placed in a container immediately and without headspace), thereby resulting in an overestimation of potential for divalent metals to be bioavailable.
- A ratio of SEM to AVS of greater than one does not necessarily mean that the divalent metals present will cause toxicity as many additional factors control binding of metals to sediments (e.g., particulate organic carbon and iron and manganese oxyhydroxides).
- Fe(III) has been observed to oxidize acid-insoluble copper sulphide complexes and therefore increase SEM_{Cu} during the AVS-SEM extraction procedure, without a corresponding increase in AVS. Therefore, an artifact of the analysis may be an overestimation of the potential for copper to become bioavailable (i.e., an artificially higher SEM:AVS).

Where can I go for further information about this tool? Examples of one or more aspects of the issues discussed above can be found in the following peer-reviewed scientific literature:

- Allen, H.A., G. Fu, and B. Deng. 1993. Analysis of acid-volatile sulfide (AVS) and simultaneously extracted metals (SEM) for the estimation of potential toxicity in aquatic sediments. Environ. Toxicol. Chem. 12: 1441-1453.

- Carlson, A.R., G.L. Phipps and V.R. Mattson. 1991. The role of acid-volatile sulfide in determining cadmium bioavailability and toxicity in freshwater sediments. Environ. Toxicol. Chem. 10: 1309-1319.

- Chapman, P.M., F. Wang, C. Janssen, G. Persoone, and H.E. Allen. 1998. Ecotoxicology of metals in aquatic sediments: binding and release, bioavailability, risk assessment, and remediation. Can. J. Fish. Aquat. Sci. 55: 2221-2243.

- DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, A.R. Carlson, and G.T. Ankley. 1992. Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. Environ. Sci. Technol. 26:96-101.

- DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environ. Toxicol. Chem. 9: 1487-1502.

- Simpson, S.L., S.C. Apte, and G.E. Batley. 1998. Effect of short-term resuspension events on trace metal speciation in polluted anoxic sediments. Environ. Sci. Technol. 32: 620-625.

MODELLING TOOL
#8: ORGANIC CARBON AND LIPID NORMALIZATION

What does this tool consist of? Biota-sediment or biota-soil accumulation factors (BSAFs) for hydrophobic chemicals are most easily predicted and interpreted when the chemical concentrations in sediment/soil and biota are normalized to the organic carbon (OC) content of the sediment/soil and the lipid content of the organism:

$$C_{S,OC} = C_S / \phi_{OC} \text{ and } C_{Biota,L} = C_{Biota} / \phi_L$$

where C_S and C_{Biota} are the chemical concentrations in sediment/soil and biota (any units, as long as they are consistent), and Φ_S and Φ_L are the OC and lipid fractions (unitless) in sediment/soil and biota, respectively. When concentrations are normalized in this way, the BSAF is theoretically (assuming equilibrium partitioning) equal to the relative sorptive capacities of lipid and OC (usually estimated to be ~ 1.7), multiplied by the ratio of biota lipid to sediment/soil OC fractions (Φ_L / Φ_S).

Which ecosystem(s) would this tool typically be applied in? OC and lipid normalization can be applied in any system in which chemical concentrations in organisms might be predicted from concentrations in soil or sediment

How frequently is this tool used in a DERA? Common in all DERAs.

What are the benefits of using this tool in a DERA? OC and lipid normalization provides a simple method to estimate the exposure of soil- or sediment-associated receptors from measured or estimated concentrations of COPCs in soil or sediment.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Predicting BSAFs from OC and lipid-normalized concentrations assumes that the organism and the soil or sediment are at chemical equilibrium, and that all of the chemical in sediment/soil is bioavailable. Empirical studies suggest that this is often not true. The true BSAF may be higher than predicted if the chemical is biomagnified, or lower than predicted if the chemical is rapidly metabolized or if the chemical in sediment/soil has low bioavailability.

- Predicting BSAFs from OC- and lipid-normalized concentrations is only appropriate for neutral (nonionic) organic chemicals.

Where can I go for further information about this tool?

- Di Toro, D.M. and L.D. De Rosa. Equilibrium Partitioning and Organic Carbon Normalization. National Sediment Bioaccumulation Conference. Available online at http://www.epa.gov/waterscience/cs/ditoro.pdf.

- Di Toro, D.M., C. Zarba, D.J. Hansen, W. Berry, R.C. Swartz, C.E. Cowan, S.P. Pavlou, and H.E. Allen. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. Environ. Toxicol. Chem. 10: 1541-1583.

- Wong, C. S., P. D. Capel, and L. H. Nowell. 2001. National-scale, field-based evaluation of the biota-sediment accumulation factor model. Environ. Sci. Technol. 35:1709-1715.

- Parkerton, T. F., J. P. Connolly, R. V. Thomann, and C. G. Uchrin. 1993. Do aquatic effects or human health end points govern the development of sediment-quality criteria for nonionic organic chemicals? Environ. Toxicol. Chem. 12:507-523.

MODELLING TOOL
#9: BIOAVAILABILITY ASSESSMENT MODELS

What does this tool consist of? The degree to which dietary contaminants are available for uptake by a consumer (bioavailability or bioaccessibility) can be estimated *in vitro* by measuring the fraction of dietary contaminant that is solublized under conditions that mimic the consumer's gut. Approaches range from simply mimicking the pH of a consumer's gut, to including a full enzyme complement, to using real digestive fluid extracted from wild or cultured animals. The more elaborate approaches are sometimes called physiologically based extraction tests (PBETs). Bioavailability assessment models are mainly applied to soil- and sediment-feeding organisms, because bioavailability in soil and sediment is known to be highly variable among ecosystems.

Which ecosystem(s) would this tool typically be applied in? Any ecosystem in which soil- or sediment-feeding organisms are ROPCs.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? These methods are a quick and inexpensive way to improve ecological relevance in assessment of dietary exposure to soil- or sediment-associated COPCs.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

The digestive fluid extraction approach is probably not useful for compounds for which ingestion is likely to be a minor route of uptake (e.g., hydrophilic organic compounds) or those for which intestinal absorption rather than solubilization constrains uptake (e.g., chromium).

Where can I go for further information about this tool?

- Oomen AG, Hack A, Minekus M, Zeijdner E, Cornelis C, Schoeters G, Verstraete W, Van de Wiele T, Wragg J, Rompelberg CJ, Sips AJ, Van Wijnen JH. 2002. Comparison of five in vitro digestion models to study the bioaccessibility of soil contaminants. Environ. Sci. Technol. 36: 3326-3334.

- Weston, D. P., Millward, R. N., Mayer, L. M., Voparil, I., and Lotufo, G. R. 2002. Sediment extraction using deposit-feeder gut fluids: A potential rapid tool for assessing bioaccumulation potential of sediment-associated contaminants, ERDC/EL T R-02-18, U.S. Army Engineer Research and Development Center, Vicksburg, MS. Available online at: http://el.erdc.usace.army.mil/elpubs/pdf /trel02-18.pdf

MODELLING TOOL
#10: METAL SPECIATION MODELS

What does this tool consist of? Water chemistry parameters are used to calculate the freely dissolved ion fraction of a metal in aqueous solution. Metals in aqueous solutions form numerous chemical species in solution of which only a proportion are freely dissolved and are therefore considered bioavailable.

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic, Shoreline, and Rivers and Streams.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA?

- Metal speciation models provide an estimate of the bioavailable fraction of metals in an aqueous solution. Generally, only metals in the ionized form are considered to be bioavailable. Calculating the ionized form is superior to using total dissolved metal concentrations (which has often been used as a surrogate for the bioavailable portion) since the dissolved fraction contains a combination of metal ions, soluble complexes and small particles of insoluble precipitates.

- The estimate of the bioavailable fraction can help the risk assessor bound the exposure of metals to aquatic receptors for a given site. In many cases, the model is used to show that the actual exposure is much less than what the measured dissolved concentration of metal in solution would indicate.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Metal speciation models provide a measure of bioavailability, but do not provide any information on the interaction of the dissolved metal species and the site of action (or biotic ligand) on the receptor. Consequently, interpretation of a "low bioavailability" has to be done with care. Even if the bioavailable portion of a metal is low, the proportion of bioavailable metal interacting with the biotic ligand may be high. For this reason, biotic ligand models should used where available.

- There are several metal speciation models available, and most require some knowledge of chemical thermodynamics. These models are based on dissociation constants for each of the potential metal to ligand complexes. There are several

sources of these dissociation constants and scientific advancements results in the periodic modification of the dissociation constants.

- The models calculate metal speciation using different mathematic algorithms. One area where models diverge is in the description of interactions between metals and dissolved organic carbon. The interaction of metals with organic matter in water is highly complex and some models provide a more realistic description of this interaction than others. If organic binding is likely to account for a large proportion of the metal-ligand binding, then it is advisable to use a model which uses a more sophisticated approach to modeling this interaction. The Windermere Humic Aqueous Model (WHAM) is one example of a metal speciation model which provides a more sophisticated approach to modeling the metal to organic ligand binding.

- A detailed understanding of how water quality guidelines were derived for the metal in question. Specifically, it is necessary to consider how differences in water quality parameters (e.g., pH, hardness, organic carbon, major ions) in toxicity tests used to derive the criteria vary from the conditions in the field. Metal speciation models are useful for instances where the bioavailability of metals in the toxicity test upon which the criteria were based is high but the estimated bioavailability of a metal in the site water is low.

Where can I go for further information about this tool?

- Schecher, W.D. & D.C. McAvoy. 2003. MINEQL+: A Chemical Equilibrium Modeling System, Version 4.5 for Windows, User's Manual. Environmental Research Software, Hollowell, Maine.

- Tipping. E., 2005. Windermere Humic Aqueous Model (WHAM) - A Chemical Equilibrium Model And Computer Code For Waters, Sediments And Soils Incorporating A Discrete Site / Electrostatic Model of Ion-binding By Humic Substances. Centre for Ecology and Hydrology. http://www.ife.ac.uk/aquatic_processes/wham/

- Tipping, E. 1994. WHAM - A chemical equilibrium model and computer code for waters, sediments, and soils incorporating a discrete site/ electrostatic model of ion-binding by humic substances. Computers Geosciences 20:973-1023.

- USEPA. 2003a. 2003 Draft Update for Ambient Water Quality Criteria for Copper. U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Washington, DC, USA.

MODELLING TOOL
#11: BIOTIC LIGAND MODELS

What does this tool consist of? Biotic ligand models (BLMs) utilize ancillary water quality parameters (e.g., pH, hardness, dissolved organic carbon, major ions) and the measured dissolved concentration of the metal of interest to derive a site-specific water quality criteria. BLMs predict the concentration of a metal bound to biotic ligands, which are located on the respiratory surfaces of aquatic organisms and are considered the cellular "site of action". The concentration of metal bound to the biotic ligand is directly related to the metal-mediated acute effect.. The estimated concentration of a metal bound to the biotic ligand for a given site is compared to toxicity reference values obtained from laboratory-based toxicity testing..

BLMs incorporate thermodynamically based metal speciation models in order to estimate the bioavailability of dissolved metals in water. Unlike metal speciation models, BLMs take one step further by also estimating the competition for binding which occurs between the metal of interest and natural ions for the biotic ligand.

Which ecosystem(s) would this tool typically be applied in? Deep Aquatic, Shoreline, and Rivers and Streams.

How frequently is this tool used in a DERA? Rare, however becoming more common. The USEPA has recently provided a draft manual for deriving site specific water quality criteria based on a biotic ligand (BLM) approach.

What are the benefits of using this tool in a DERA?

BLMs provide improved estimates of dissolved metal concentrations unlikely to result in a deleterious effect. They are useful for reducing the uncertainty associated with using total concentrations for evaluating metal toxicity in freshwater ecosystems..

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- BLMs are relatively new and therefore the availability of calibrated, validated models is low. USEPA uses a BLM to derive site specific water quality criteria for copper, and anticipates developing BLMs for other metals in the future..

- The existing USEPA BLM model is based on a complex metal speciation model, but for ease of use, the number of required water quality parameters was reduced. There is no ability for the user to modify thermodynamic dissociation constants,

meaning that it is necessary to apply the USEPA default values for a substantial number of parameters rather than incorporate site-specific values.

- BLMs are largely based on the results of acute toxicity tests using a small number of freshwater aquatic organisms; they incorporate an acute-to-chronic ratio to extrapolate the model to chronic conditions. Research in the development of truly chronic BLMs as well as BLMs for marine organisms is ongoing.

Where can I go for further information about this tool?

- Di Toro, D.M., H.E. Allen, H.L. Bergman, J.S. Meyer, P.R. Paquin and R.C. Santore. 2001. A biotic ligand model of the acute toxicity of metals I. Technical basis. Environ. Toxicol. Chem. 20:2383-2396.

- Niyogi, S. and C. M. Wood. 2003. Effects of chronic waterborne and dietary metal exposures on gill metal-binding: Implications for the biotic ligand model. Human and Ecological Risk Assessment, 9:813-846.

- Paquin P.R., J.W. Gorsuch, S. Apte, G.E. Batley, K.C. Bowles, P.G.C. Campbell, C.G. Delos, D.M. Di Toro, R.L. Dwyer, F. Galvez, R.W. Gensemer, G.G. Goss, C. Hogstrand, C.R. Janssen, J.C. McGeer, R.B. Naddy, R.C. Playle, R.C. Santore, U. Schneider, W.A. Stubblefield, C.M. Wood and K.B. Wu. 2002. The biotic ligand model: a historical overview. Comp Biochem Physiol, Part C, 133: 3-35

- Santore, R.C., D.M. DiToro, P. R. Paquin, H.E. Allen and J.S. Meyer. 2001. Biotic ligand model of the acute toxicity of metals. 2. Application to acute copper toxicity in freshwater fish and Daphnia. Environ. Toxicol. Chem. 20: 2396-2402.

- USEPA. 2003. 2003 Draft Update for Ambient Water Quality Criteria for Copper. U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Washington, DC, USA.

MODELLING TOOL
#12: QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIPS (QSARs)

What does this tool consist of? QSAR models are mathematical equations that describe a relationship between the toxicity (or other properties) of chemicals and their measured physico-chemical properties or structures. A QSAR derived for some members of a family of chemicals can then be used to predict unmeasured values for other members of the same family. QSARs are often used to estimate bioaccumulation or toxicity of new industrial chemicals/pesticides for which bioaccumulation or toxicity testing has not been conducted.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Rare unless dealing with new or unusual chemicals.

What are the benefits of using this tool in a DERA? This tool is useful for new compounds about which little is known. QSARs provide a means to screen these chemicals, so that testing can be focused on chemicals that are most likely to be of ecological concern.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Most QSARs are statistical models (i.e., regression equations), not mechanistic models. They describe a statistical correspondence between structure and activity in the set of chemicals used to develop the model. In applying a QSAR to estimate properties of new chemicals, we are assuming that the correspondence will continue to hold. This may not always be true. If it is not true (i.e., if the new chemical is different in some unknown way), then the predictions may be completely inaccurate.

- Regression-based QSARs also describe the strength of the statistical relationship (r^2). This information should be used to put confidence limits on the estimated value, although this is rarely done.

Where can I go for further information about this tool?

- Gobas, F.A.P.C.; Kelly, B.C.; Arnot J.A. 2003. Quantitative Structure Activity Relationships for Predicting the Bioaccumulation of POPs in Terrestrial Food-Webs. QSAR & Combinatorial Science 22: 329 – 336.

- Arnot, J.A.; F.A.P.C. Gobas. 2003. A Generic QSAR for Assessing the Bioaccumulation Potential of Organic Chemicals in Aquatic Food Webs. QSAR & Combinatorial Science 22: 337 – 345.

- Cronin, M.T.D. and D.J. Livingstone, eds. 2004. Predicting Chemical Toxicity and Fate. CRC Press.

A relevant USEPA website is also available: http://www.epa.gov/oppt/newchems/tools/

APPENDIX III

INTERPRETATIVE TOOLS

TABLE OF CONTENTS

INTERPRETIVE TOOL #1: HAZARD QUOTIENTS

What does this tool consist of? A hazard quotient (HQ) is the ratio of a receptor's observed or predicted exposure to a toxicity reference value (TRV). TRVs can be narratively defined in several ways to reflect the desired assessment endpoint (i.e., the TRV can reflect an acceptable level of risk or reflect a threshold below which adverse effects are not believed to occur). TRVs are calculated in a number of ways, including: 1) selecting the lowest reported value; or 2) derived from a statistical distribution of reported values (analogous to a species sensitivity distribution; see Interpretative Tool #3). HQs are normally interpreted based on a binary decisions (i.e., potential risks are present if HQ > 1; risks considered negligible if HQ <1).

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Common.

What are the benefits of using this tool in a DERA?

- Hazard quotients require few data, are easy to calculate and easy to interpret.
- Toxicity reference values have been derived for many chemicals in many species.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Hazard quotients are only as good as the TRVs used to calculate them. Theoretically, the TRV is an exposure level at which the specified effects do not occur in the specified receptor. In practice, TRVs are often estimated from lab-based toxicity data, often in acute tests, and often for model species. The series of calculations, extrapolations, and approximations used to estimate the TRV from what data are available (e.g., uncertainty factors, acute-to-chronic ratios, interspecific extrapolation factors, etc.) is a potentially large source of error and must be assessed carefully.
- Hazard quotients do not say anything about the magnitude of possible effects. A larger HQ is presumably associated with more severe effects, but this is not quantifiable. As such, HQs are most useful in screening-level risk assessments to indicate when a more detailed assessment is warranted.
- Producing an appropriate compilation of data with which to derive a TRV may require some calculations and approximations. For example, TRVs are commonly

expressed in units of intake rates (daily dietary dose, mg/kg/d or similar). When available toxicity data are in terms of dietary concentrations, they must be converted to intake rates with an estimated or measured feeding rate. It may also be necessary to estimate the desired effects concentrations (e.g., NOAELs) from other reported values (e.g., LOAELs) using some form of uncertainty factor or application factor.

Where can I go for further information about this tool?

- Efroymson, R. A., M. E. Will, and G. W. Suter II. 1997. Toxicological Benchmarks for Contaminants of Potential Concern for Effects on Soil and Litter Invertebrates and Heterotrophic Processes: 1997 Revision. ES/ER/TM-126/R2, Oak Ridge National Laboratory, Environmental Sciences Division.

- McDonald, B.G., and J.B. Wilcockson. 2003. Improving the use of toxicity reference values in wildlife food chain modeling and ecological risk assessment. Human and Ecological Risk Assessment 9: 1–10.

INTERPRETIVE TOOL #2: EC_{20} APPROACH

What does this tool consist of? The EC_{20} approach is a method of evaluating the significance of any toxicity predicted or observed for a contaminant and a receptor in an ecological risk assessment. The approach is based on a policy decision that up to a 20% reduction in growth, reproduction success or other aquatic toxicity endpoint is acceptable. Other ECx values are used for other types of endpoints, including terrestrial toxicity tests

Which ecosystem(s) would this tool typically be applied in? This tool applies to: Deep Aquatic, Shoreline, Rivers and Streams, Uplands (Wildlands) and Uplands (Human-Use)>>

How frequently is this tool used in a DERA? This tool is frequently used in DERA although the percent effect range varies by land use.

What are the benefits of using this tool in a DERA? The EC_{20} (or ECx) provides a sliding scale for evaluating toxicity data and determining the significance within an ecological context. The ECx approach is preferred to using NOECs or LOECs as the basis for interpreting toxicity data.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The ECx will vary according to existing risk assessment policy, which is determined largely by land use. Risk assessors must consider this provincial policy in light of regulatory requirement for other overlapping jurisdictions.

- The EC_{20} approach is a convention for evaluating the significance of a toxic response from a biological perspective and should not be confused with statistical significance. Suter et al. (1995) recommended if using both types of significance criteria, any significant effects should be identified as either biologically significant (e.g. > 20% effect) or statistically significant (<5% chance of the difference from reference is due to chance).

- The ECx approach reflects a policy decision regarding a permissible level of effects, and should not be misinterpreted as biologically or ecologically relevant.

Where can I go for further information about this tool?

- Suter, G.W., B.E. Sample, D.S. Jones, T.L. Ashwood and J.M. Loar. 1995. Approach and Strategy for Performing Ecological Risk Assessments for the U.S.

Department of Energy's Oak Ridge Reservation: 1995 Revision. Prepared by the Environmental Restoration Risk Assessment Program, Lockheed Martin Energy Systems, Inc. Oak Ridge, Tennessee 37831. Prepared for the U.S. Department of Energy Office of Environmental Management under budget and reporting code EW 20.

INTERPRETIVE TOOL
#3: SPECIES SENSITIVITY DISTRIBUTIONS

What does this tool consist of? A species sensitivity distribution (SSD) is the probability distribution of some measure of toxicity of a certain chemical to a set of animal species. Single-species toxicity data (e.g., LC50s or NOECs) for many species are fit to a distribution such as the lognormal or log-logistic. From this distribution of species sensitivities, a hazardous concentration (HC*p*) is identified at which a certain percentage (*p*) of all species is assumed to be affected.

Which ecosystem(s) would this tool typically be applied in? An SSD can be derived for any ecosystem type for which sufficient toxicity data are available.

How frequently is this tool used in a DERA? Occasional?

What are the benefits of using this tool in a DERA? SSDs provide a way to combine toxicity data for many species, reducing the effect of uncertainty in individual toxicity measurements. Calculating an HC*p* then allows a simultaneous assessment of toxic effects in all potential receptors. As an added benefit, calculating an HC*p* forces an explicit recognition of the magnitude of effect being considered (the chosen endpoint of the single-species toxicity tests) and the percentage of affected species that is judged to be acceptable (*p*).

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- The HC*p* has all the limitations of the single-species toxicity data used to generate the SSD. If LC_{50}s are used, the HC*p* will estimate the concentration at which 50% lethality occurs in *p*% of species, which may not be an adequately protective level. If the LC_{50}s are highly uncertain or have low ecological relevance (e.g., due to unrealistic test conditions), the HC*p* will be similarly limited.

- Calculating an HC*p* explicitly recognizes that some fraction of species will be affected at any given concentration, but does not consider which species these are. If the species that fall into the affected *p* percentage are critically important to ecosystem function, the resulting ecological effects may be greater than predicted.

- This approach is based on the assumption that the toxicity data are from a random sample of species. In practice, the species for which data are available may not be representative of the real set of species in the ecosystem of interest. Whenever possible, the toxicity data used to generate the SSD should come from species actually present in the system under consideration, or very similar species. If the COPC has a particular target taxon (e.g., a herbicide), the SSD must include representatives of that taxon, or it may be advisable to construct a SSD for the target taxon and another for non-target taxa.

Where can I go for further information about this tool?

- U.S. Environmental Protection Agency. 1998. Guidelines for ecological risk assessment; notice. Fed Reg 63:26846–26924.

- Kooijman SALM. 1987. A safety factor for LC_{50} values allowing for differences in sensitivity among species. Water Res 21: 269–276.

- Wagner C, Løkke H. 1991. Estimation of ecotoxicological protection levels from NOEC toxicity data. Water Res 25:1237–1242.

- Aldenberg T, Slob W. 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. Ecotoxicol Environ Saf 25:48–63.

- Newman, M.C., D.R. Ownby, L.C.A. Mezin, D.C. Powell, T.R.L. Christensen, S.B. Lerberg, and B.-A. Anderson. 2000. Applying species sensitivity distributions in ecological risk assessment: Assumptions of distribution type and sufficient numbers of species. Environ. Toxicol. Chem. 19: 508-515.

- Posthuma L, Suter GW, Traas TP, eds. 2002. Species sensitivity distributions in ecotoxicology. Boca Raton: Lewis Publishers.

INTERPRETIVE TOOLS
#4: SUMMARY METRICS

What does this tool consist of? Summary metrics are numerical expressions of the characteristics of a biological community and are based on taxonomic data (i.e., species identification and abundance). They can include attributes such as: abundance (number of organisms), richness/diversity (number of taxa); presence/absence of sensitive taxa; ratios of indicator taxa (e.g., percent Ephemeroptera-Plecoptera-Trichoptera or EPT in benthic invertebrate communities); and, ratios of functional feeding groups. Numerous indices have also been developed as a means of describing biological communities (e.g., Schwartz Dominance Index; Bray-Curtis Index; Index of Biotic Integrity).

Which ecosystem(s) would this tool typically be applied in? Any ecosystem in which a biological community survey has been conducted.

How frequently is this tool used in a DERA? Common in DERA in which biological community surveys have been conducted.

What are the benefits of using this tool in a DERA? Summary metrics simplify complex taxonomy data so that the patterns and relationships that describe the structure of a biological community can be assessed. They can be used as measurement endpoints for assessment endpoints involving biological community structure and can be incorporated into statistical analyses of differences between exposure and reference sites and correlations with habitat variables (e.g., water depth, grainsize distribution) and measures of exposure to provide information about effects potentially related to contaminants of concern.

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- Because summary metrics by definition simplify complex data sets, a variety of metrics need to be used to assess the structure of a biological community to determine presence/magnitude of effects. Different metrics focus on different aspects of a community, for example, species richness versus evenness of distribution of individuals among the species identified.

- Summary metric such as richness and diversity indices do not necessarily provide an evaluation of the ecological function of the organisms in a biological community.

- Biological communities are highly variable, so biotic indices typically have low statistical power to detect ecological effects. Indices are best used as one component of a weight of evidence approach.

Where can I go for further information about this tool?

- Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring.

- Simon, T. P. and J. Lyons, 1995, Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems, In: W.S. Davis and T.P. Simon (eds). Biological Assessment and Criteria Tools for Water Resource Planning and Decision Making, Lewis Publishers, Boca Raton, Florida, pp. 245-262. Available online at: http://www.epa.gov/bioindicators/pdf/ba_cch16.pdf

- Angermeier, P. L., and J. R. Karr, 1986, An index of biotic integrity based on stream fish communities: considerations in sampling and interpretation, N. Am. J. Fish. Manage. 6: 418-429.

- Bryce, S.A., R.M. Hughes, and P.R. Kaufmann. 2002. Development of a bird integrity index: using bird assemblages as indicators of riparian condition. Environ. Manage. 30: 294-310.

- Lydy, M.J., Crawford, C.G., and Frey, J.W., 2000, A comparison of selected diversity, similarity, and biotic indices for detecting changes in benthic-invertebrate community structure and stream quality: Arch. Environ. Contam. Toxicol. 39: 469-479.

INTERPRETIVE TOOL
#5: MULTIVARIATE STATISTICAL ANALYSES

What does this tool consist of? Multivariate analysis refers to any of various statistical methods for analyzing more than two variables simultaneously. Assessing effects at the community and ecosystem levels usually involves measuring a large number of abiotic and biotic variables. Assessing each variable individually or with many pairwise bivariate analyses can be cumbersome and difficult to interpret. Multivariate techniques can be used to draw overall patterns from a large set of variables.

There are three broad types of applications for multivariate techniques: ordination, clustering/discrimination, and investigating relationships between sets of variables.

- Ordination techniques (e.g., principal components analysis) reduce a large set of variables into a smaller set of 'factors', each of which is a combination of variables that captures as much as possible of the information in the original variables: in this way, a multidimensional set of data can be reduced into a more interpretable form.

- Clustering/discrimination techniques identify natural groupings among sampling units (e.g., most-similar groups of sampling sites) and the parameters that contribute most to this similarity (e.g., abundances of certain species).

- Techniques such as canonical correspondence analysis identify the degree of covariance between sets of variables (e.g., concentrations of several chemicals vs. abundances of several species), as well as identifying the variables within each set that contribute most to this covariance.

Ordination, classification and canonical ordination techniques can be applied to any ecosystem, and are common in DERAs. Further information is provided below for each group of techniques. Note that Bayesian approaches are sometimes used in risk assessment; however, specialized training is required, and therefore, no guidance for its application in risk assessment is provided at this time.

Ordination

What is it? Ordination techniques reduce a large set of variables into a smaller set of "derived factors", each of which is a combination of variables that captures as much as possible of the information in the original variables. In this way, a multidimensional set of data can be reduced into a more interpretable form. Commonly used ordination techniques include principal components and factor analysis (PCA and FA),

correspondence analysis and detrended correspondence analysis (CA and DCA), and metric and nonmetric multidimensional scaling (MDS and NMDS).

How is it useful in risk assessment? Ordination is usually treated as an exploratory tool for generating hypotheses and directing further research. If the reduction in dimensionality is sufficient, the results can be plotted for a visual analysis of relationships among sites or among variables. In some cases, a derived factor is readily interpretable (e.g., as a gradient of contamination) and can be used as a composite variable in further analyses (e.g., as an explanatory variable in multiple regression). Examples of ordination in DERA include exploring overall trends in a collection of response variables (measurement endpoints) such as in a set of chemical analyses, taxonomic data, or any other set of appropriately-related variables measured at a number of sites.

Issues to watch out for: Data sets frequently have missing values, skewed or bimodal distributions (e.g., many zeroes for rare species), and categorical or semiquantitative values. Different techniques have different sensitivities to these common issues; however, all ordination techniques de-emphasize the importance of individual variables (e.g., a particularly sensitive receptor or high-priority COPC) and therefore may mask important information. Important information can also be masked when variables are subject to ordination techniques without consideration of how those variables relate to one another.

Classification

What is it? Clustering and discrimination techniques identify natural groupings among sampling units (e.g., most-similar groups of sampling sites) and the parameters that contribute most to this similarity (e.g., abundances of certain species or concentrations of certain chemicals). The most commonly used clustering techniques are *k*-means clustering and two-way indicator species analysis (TWINSPAN). The most commonly used techniques to discriminate among established groups are linear discriminant analysis (LDA), Hotelling's T^2, and multivariate analysis of variance (MANOVA).

How is it useful in risk assessment? Cluster analysis is useful as an exploratory tool to identify natural groupings of measured values in space or time, so that further analysis or remediation can be stratified and/or focused on 'hotspots' of exposure or effects. Cluster analysis produces a dendrogram (a tree diagram) where sites may be grouped at varying levels of similarity. Discrimination techniques can be used to identify the variables that are most strongly associated with an established grouping scheme, to detect statistically-significant multivarate differences among groups, and to derive "rules" for predicting to which group (e.g., impacted vs unimpacted) a new sample belongs.

Issues to watch out for: The results of cluster analysis may be sensitive to the particular technique used, e.g., the choice of distance measure (how similarity among cases is calculated). Clustering typically produces ambiguous and/or unstable results when

samples are arranged continuously along gradients. As with their univariate counterparts (Student's *t* and ANOVA), T^2 and MANOVA are sensitive to the assumptions of multivariate normality and constant within-group variances and covariances. Discriminant analysis is often applied to the same set of data for which the rules were derived (the 'training' set), but this gives a highly inflated estimate of the success with which the categorization rules will determine group membership for a new sample. A better approach is to use cross-validation (split-sample or leave-one-out) to test the categorization rules. As with ordination techniques, cluster analysis de-emphasizes the importance of individual variables and may miss important univariate trends.

Canonical ordination

What is it? Canonical ordination techniques explore the degree of covariance between two sets of variables, as well as identifying the variables within each set that contribute most to this covariance. Commonly used techniques include canonical correlation analysis, redundancy analysis (RDA), and canonical correspondence analysis (CCA).

How is it useful in risk assessment? Canonical ordination techniques can be used to explore the relationship between exposure and effects when one or both of these are multivariate. For example, the data may include a concentration by site matrix for several chemicals and abundance by site matrix for several species. A technique such as CCA will reveal the strength of the overall correspondence (among sites) of abundances (effects) with concentrations (exposure). It is also common to include other site characteristics in this type of analysis, to assess to what extent species' abundances are determined by habitat characteristics vs. chemical concentrations. In CCA, explanatory variables can be of many types (e.g., continuous, ratio scale, nominal) and do not need to meet distributional assumptions. Hypothesis testing is possible with CCA by means of a randomization test.

Issues to watch out for: As with regression, one cannot necessarily infer direct causation from canonical ordination techniques. In addition, the independent effects of highly correlated variables (e.g., covarying concentrations of several metals) are difficult to disentangle. The outcome of CCA, in particular, is highly dependent on the scaling of the explanatory variables; log transformation of explanatory (exposure and environmental) variables is probably most often appropriate. CCA focuses more on species composition than RDA (which focuses on *relative* abundance); thus, if you have a gradient along which *all* species are positively correlated, RDA will detect such a gradient while CCA will not.

Multivariate techniques can be used to draw general patterns from very complex sets of data. Each technique has associated methods for graphical representation of these general patterns, which can aid in conveying complex ideas to non-technical stakeholders. Multivariate techniques can be used to assess community-level ecological

effects, which have more ecological relevance than studies at lower levels of biological organization.

What are the common "pitfalls" or issues that should be considered when using multivariate statistical analyses in a DERA?

- Many multivariate techniques have no established method for determining the statistical significance of observed patterns, and are suitable only for exploratory data analysis.

- Like all statistical methods, there are assumptions that must be carefully assessed before applying multivariate methods. For example, most multivariate methods (except cluster analysis) assume multivariate normality. Most are sensitive to outliers.

- Application of multivariate techniques may require some modification of the field study design, including the appropriate level of replication, the endpoints to be measured, and the taxonomic resolution required.

Where can I go for further information?

- Sparks, T.H., Scott, W.A., Clarke, R.T. 1999. Traditional multivariate techniques: potential for use in ecotoxicology. Environmental Toxicology and Chemistry 18**:** 128-137. (and the rest of the Special Section in ET&C Volume 18)

- Fairbrother, A. and R.S. Bennett. 2000. Multivariate statistical applications for addressing multiple stresses in ecological risk assessments. Pages 69-115 in Ferenc, S.A. and J.A. Foran, editors. Multiple Stressors in Ecological Risk and Impact Assessment: Approaches in Risk Estimation. SETAC Press, Pensacola, FL.

- USEPA Statistical Primer: Multivariate Methods. Available online at: http://www.epa.gov/bioindicators/primer/multivariate.html

INTERPRETIVE TOOL
#6: PROBABILISTIC METHODS

What does this tool consist of? Probabilistic methods estimate the likelihood of adverse effects, and the probable magnitude of those effects, by incorporating statistical distributions for exposure and/or ecological effects profiles. If exposure concentrations have been measured, the distribution of observed values may be incorporated into the exposure profile to reflect either uncertainty or variability in exposure (but not both; see below). If exposure concentrations are being estimated from a model, simulation methods can be used to generate a distribution of predicted exposures from variability or uncertainty in model structure or input parameters.

The most commonly-used simulation method is Monte Carlo analysis, a technique that randomly generates values for all uncertain or variable parameters and calculates the resulting exposure; many (usually > 10,000) such simulation scenarios give the range of possible exposures, each with an associated probability. Probability bounds analysis is another simulation technique. Ecological effects profiles can also incorporate statistical distributions as the cumulative distribution function of effects (i.e., a dose-response curve) for a single species, or as a species sensitivity distribution for multiple species.

Which ecosystem(s) would this tool typically be applied in? Any.

How frequently is this tool used in a DERA? Rare.

What are the benefits of using this tool in a DERA? Probabilistic methods produce very informative risk characterization statements that can include both a probability of observing a particular effect and the probable magnitude of that effect (e.g., "a 90% likelihood of no more than 50% mortality").

What are the common "pitfalls" or issues that should be considered when using this tool in a DERA?

- One should avoid developing probability distributions that intermingle or try to represent both variability and uncertainty since single probability distribution must be interpretable either as an expression of variability (90% of the time, or in 90% of the population) or as an expression of uncertainty (with 90% confidence). An intermingling of these two interpretations would be meaningless.

- Simulation methods (e.g., Monte Carlo) typically assume that all parameters are independent, and that a particular randomly-chosen value for one parameter will have no influence on the most likely value for another parameter. In reality, many ecological parameters are highly correlated (e.g., feeding rate and growth rate of a

species, or feeding rates of several species that are all a function of temperature). There are ways to account for these correlations in simulations, but this requires additional information about the form of the correlation, which is rarely available.

- Simulations do not easily take into account uncertainty in the structure of the model, and will therefore always underestimate to some degree the true uncertainty in model output.

- Monte Carlo simulations require accurate estimates of the magnitude of variability or uncertainty in parameters, and require that you know the form of the distribution of these parameters (e.g., lognormal). These data are often unavailable.

Where can I go for further information about this tool?

- US Environmental Protection Agency. 2001. Risk Assessment Guidance for Superfund (RAGS), Volume III - Part A: Process for Conducting Probabilistic Risk Assessment. EPA 540-R-02-002, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C. Available online at http://www.epa.gov/superfund/programs/risk/rags3a/index.htm.

- 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.

- Haas, C.N. 1999. On modeling correlated random variables in risk assessment. Risk Anal. 19: 1205-1214.

- USEPA 1997a. Policy For Use Of Probabilistic Analysis In Risk Assessment. Available online at: http://www.epa.gov/osa/spc/pdfs/probpol.pdf

- USEPA. 1997b. Guiding Principles for Monte Carlo Analysis. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC, EPA/630/R-97/001, 1997.