Government of Canada | Gouvernement du Canada

Canada

Federal Contaminated Sites Action Plan (FCSAP)

Ecological Risk Assessment Guidance

Module 1: Toxicity Test Selection and Interpretation

March 2010

LIBRARY AND ARCHIVES CANADA CATALOGUING IN PUBLICATION

Federal Contaminated Sites Action Plan (FCSAP): Ecological Risk Assessment Guidance
Module 1: Toxicity Test Selection and Interpretation

Issued also in French under title:

Plan d'action pour les sites contaminés fédéraux (PASCF): Document d'orientation sur l'évaluation du risque écotoxicologique
Module 1: Sélection et interprétation des essais de toxicité

ISBN no. 978-1-100-22314-8

Cat. no. En14-92/1-2013E-PDF

DISCLAIMER

Her Majesty is not responsible for the accuracy or completeness of the information contained in the reproduced material. Her Majesty shall at all times be indemnified and held harmless against any and all claims whatsoever arising out of negligence or other fault in the use of the information contained in this publication or product.

The information in this document does not constitute legal advice and following this guidance will not necessarily ensure compliance with federal, provincial, or any other, regulatory requirement. In case of discrepancy between this information and any Acts of Parliament, most notably the *Canadian Environmental Protection Act, 1999* or the *Fisheries Act* or regulations made under these Acts, the Acts of Parliament and associated regulations take precedence. You should be advised that, notwithstanding any other regulatory or permitting requirements, any deposits, discharges and releases from your operations or activities must comply with all applicable federal Acts and regulations.

TABLE OF CONTENTS

1. BACKGROUND

The Federal Contaminated Sites Action Plan (FCSAP) was developed to support federal departments, agencies and consolidated crown corporations to reduce the risks to human health and the environment, as well as to reduce the financial liabilities associated with federal contaminated sites. Under FCSAP, ecological risk assessments (ERAs) are commonly used as a site management tool at federal contaminated sites. The FCSAP Ecological Risk Assessment Focus Group is developing guidance for ERA supplemental to the existing CCME guidance (1996a, 1997). The FCSAP ERA guidance consists of a comprehensive main ERA document and several specific technical guidance modules. This document is a technical guidance module on toxicity testing, which will provide the custodians of federal contaminated sites, FCSAP Expert Support[1], and project consultants, with nationally consistent advice on how to select and interpret toxicity tests as part of an ERA.

The intended audience for this module includes experienced risk assessment practitioners who are familiar with the role of toxicity testing in ecological risk assessments, and who would benefit from procedural guidance in the selection and interpretation of specific tests and endpoints. The guidance is intended neither as a prescriptive manual nor as a comprehensive listing of all technical details that may be relevant for a specific site. Moreover, it is essential that the practitioner not conduct test selection or interpretation in a vacuum, but rather understand and apply the principles and steps outlined in the main volume of the FCSAP guidance, particularly with respect to Problem Formulation (Section 2). In the course of conducting an ERA, it may also be valuable to solicit input from a professional consulting company with experience in this specialized type of work and/or a toxicology laboratory manager; each may be able to provide insights and information not covered in this module. As toxicity testing is a complex discipline area with significant implications for the risk assessment conclusions, clear communications among the toxicity laboratory, consultants, client, and regulators (including FCSAP Expert Support) with respect to scope, timelines, uncertainties, and logistical constraints are important.

1.1. Toxicity Testing in Ecological Risk Assessment

This document contains the technical guidance for selection and interpretation of tests within the FCSAP ERA framework. Application and interpretation of toxicity tests occurs in the Effects Assessment stage of an ERA (CCME, 1996, 1997) although consideration of toxicity testing issues should occur throughout the risk assessment process. For example, selection of specific toxicity tests as measurement endpoints occurs during the Problem Formulation stage, and requires consideration of advantages and uncertainties of potential test protocols, whereas consideration of uncertainties is often included in the Risk Characterization stage. Toxicity testing is commonly conducted in an ERA to support an evaluation of potential

[1] For the purpose of this module, "FCSAP Expert Support" is considered to include not only Environment Canada involvement, but also Department of Fisheries and Oceans and Health Canada, as appropriate. The degree of involvement from the respective parties should be evaluated during Problem Formulation.

environmental effects in a resident community, using extrapolation from the laboratory to the field.

Definition of Toxicity Test *– Toxicity tests are studies designed to determine whether exposure of test organisms (to contaminants in test media) causes an adverse effect (either lethal or sublethal) to those organisms. A toxicity test usually measures either: (a) the proportions of organisms affected (quantal); or (b) the degree of effect shown (graded or quantitative), following exposure to a specific test substance or test medium under controlled conditions. Tests may be categorized as acute or chronic in duration (Environment Canada, 1999):*

- *Acute testing occurs over a short period (minutes, hours, or a few days) in relation to the life span of the test organisms, for any discernable adverse effects (lethal or sublethal).*
- *Chronic testing occurs over a relatively long period of exposure, usually a substantial proportion of the life span of the organism (defined as 10% or more) and involving long term effects related to changes in metabolism, growth, reproduction, or ability to survive.*

The above definitions of acute and chronic testing relate to test duration only, and do not consider the importance or sensitivity of the life stage under evaluation. Other jurisdictions and investigators consider some short-term tests to be chronic tests based on endpoint type. Although Environment Canada provides a precise definition of chronic versus acute testing, in a risk assessment framework the duration of testing is not as important as the degree to which the test simulates an ecological process of interest. Some tests (*e.g.*, echinoid fertilization, bivalve larval development) have relatively short test durations, but represent exposures over critical life stages, and as such are often used as surrogates for potential long-term adverse responses.

The definition of toxicity testing provided above is broad, such that it applies to the full range of testing that may be contemplated under a FCSAP program. In a subset of cases, toxicity testing may be applied on a chemical-specific basis to determine the potential for adverse effects to a specific contaminant or contaminant group. Concentration-response relationships are of interest for Priority Substances List (PSL) and Domestic Substance List (DSL) evaluations performed under the *Canadian Environmental Protection Act*, but are also of interest for developing site-specific toxicity reference values for key constituents at contaminated sites. The definition of toxicity testing also uses the term "exposure" in the broad sense, such that the concept refers to both the magnitude of contamination in the test media or administered substance and also to the duration of exposure in the experiment.

This module emphasizes the selection of standardized toxicity tests that have widespread usage and that are commercially available. In addition to these tests, specialized and longer duration tests are under continuous development. For example, Vogt *et al.* (2007) extend the test protocol for *Chironomus riparius* to eleven generations to assess life-cycle parameters and genetic variability. For terrestrial (soil-based) testing, Campiche *et al.* (2009) describes the potential for extension of existing standard test protocols to incorporate multigenerational endpoints; the species include the springtail *Folsomia candida*, the white worm *Enchytraeus albidus*, the nematode *Caenorhabditis elegans* and the isopod *Porcellio scaber*. This module does not discourage the application of multigenerational tests or specialized test protocols, although the scientific benefits of such tests must be traded off against other test selection attributes considered in this module. Similarly, toxicity test protocols are under continual refinement and development,

such that additional tests may be added to the suite of tests specified in this module; these emerging tests may also be considered and evaluated using the selection criteria presented herein.

Toxicity testing in an ERA represents one (or several) lines of evidence in the weight-of-evidence (WOE) framework to detect possible effects and/or estimate risk to receptors of concern. Toxicity testing may also be used to develop site specific toxicity reference values (TRVs)[2]**.** Toxicity test technical issues should be considered when preparing the sampling and assessment plan for the effects assessment during the problem formulation stage, consistent with the national ERA framework (CCME 1996, 1997) and regional guidance for sediment assessment[3].

1.1.1. Toxicity Testing as a Line of Evidence in a Weight-of-Evidence Framework

In many cases, toxicity testing is conducted as one component of a weight-of-evidence (WOE) framework (Menzie *et al.*, 1996; Chapman *et al.*, 2002b; Reynoldson *et al.*, 2002; Grapentine *et al.*, 2002). An example of a WOE approach is the Sediment Quality Triad, which integrates information on bulk sediment toxicity, resident benthic macroinvertebrate communities, and comparisons of contaminant measurements to environmental quality guidelines (Chapman, 2000). Consideration of the other lines of evidence in a WOE framework may influence the selection of toxicity test endpoints. For example, relevance of using a specific organism is increased where these organisms are found to occur (*i.e.*, sediment toxicity tests using the genus *Chironomus* would be highly relevant to sites where chironomids are found in significant numbers). Conversely, it may be appropriate to test an organism that is either infrequently or not observed at a site, either based on presumed ecological or taxonomic relevance to species present, or as a test for effects to sensitive species that may have been already affected due to contamination. Test organisms and protocols are also dependent on the properties of the contaminants of concern found at the site; a site for which bioaccumulation and biomagnification processes are of concern is more likely to require toxicity tests that include a bioaccumulation option in the test protocol.

Different lines of evidence apply to different organisms and exposure pathways. To properly evaluate candidate tests, the practitioner should have a clear understanding of the study objectives, including consideration of:

- Degree of precision (repeatability) and resolution (characterization of spatial and temporal variance) required in measurement endpoints necessary to support effective management decisions for the site;

---

[2] A separate guidance module addresses toxicity reference values.

[3] The Canada-Ontario Decision-Making Framework for Assessment of Great Lakes Contaminated Sediment provides step-by-step science-based guidance for assessing risks posed by contaminated sediment. The framework is primarily concerned with risks to the environment but considers human health concerns associated with biomagnification of contaminants. It identifies all possible sediment assessment outcomes based on four lines of evidence (sediment chemistry, toxicity to benthic invertebrates, benthic community structure, and the potential for biomagnification) and provides specific direction on next steps in making sediment management decisions.

- Strength of the toxicity endpoint(s) in representing the ecological processes identified as assessment endpoints;
- Number and type of tests required to depict the variation in organism sensitivity and to characterize the applicable exposure pathways and ecologically relevant endpoints;
- Requirement (if any) for comparability with previous test history or legal/policy considerations (*e.g.*, rainbow trout testing for evaluating effluent toxicity); and
- Site-specific constraints on sample handling and analysis (*e.g.*, duration of tests, holding times, requirements for tiered testing, restrictions on sampling windows to accommodate seasonal organism availability, practical sample collection constraints, and access to sampling sites).

All of the above factors may influence the selection of test organisms/protocols and the interpretation of results, and all may vary on a project-specific basis. Thorough consideration of these factors prior to test selection improves the likelihood of obtaining useful data for risk characterization.

1.1.2. Toxicity Testing to Develop a Site-Specific TRV

Because toxicity tests integrate the cumulative effects of multiple potential stressors in a mixture, it is not always possible (or necessary) to discern the specific causative agent in an individual sample. However, at some sites, it is possible to develop a concentration-response relationship that quantifies the linkage between an exposure medium (overlying water, interstitial water, bulk sediment, or surface soil) and an effect measure of interest. An empirically-derived concentration-response relationship must be developed with care, particularly with respect to potential for spurious relationships, non-linear responses, and/or interaction responses among constituents.

When such relationships are defensibly derived, it is possible to calculate a threshold response level or TRV associated with an individual substance in that exposure medium. TRVs can be used to extrapolate from individual sampling stations (or areas) to broader spatial areas, provided that the underlying mechanism of toxicity is well understood and that the factors mediating such toxicity are controlled among sites. Typically, more evidence than simple statistical correlation is required to support a reliable site-specific toxicity reference value (TRV).

In developing site-specific TRVs, it is important to account for physical factors that can confound the relationship between a potential toxic agent and the response measure. For example, in soil toxicity testing, Natal-da-Luz *et al.* (2008) document how the avoidance response of soil invertebrates (including the earthworm *Eisenia andrei* and springtail *Folsomia candida*) can be influenced by the soil properties (*e.g.*, soil organic matter and texture) that affect behavior of the test species in the soil.

In cases for which TRV derivation is confounded by multiple potential toxicants, toxicity testing can be conducted using spiked soil or sediment samples, or alternatively toxicity identification evaluations (TIEs) may be performed to restrict the number of potentially toxic substances. Either of these approaches may be used to develop site-specific TRVs. For additional guidance on TRV derivation, please consult the separate FCSAP guidance on TRV development.

1.1.3. Other Roles of Toxicity Testing

In addition to TRV derivation and use as a single line of evidence, the following aspects of environmental risk may be addressed through toxicity testing:

- measure of contaminant bioavailability[4] (*i.e.,* if substances are not bioavailable, they will not be toxic, even if they are present at high levels that exceed environmental quality guidelines);
- measure of presence of unknown or unsuspected contaminants (*e.g.*, if toxicity is observed, and the suite of analytical results shows no contaminants of concern, additional unmeasured or poorly understood toxicants may be acting);
- interaction responses (effects of mixtures of two or more contaminants, whether antagonistic, additive, or synergistic);
- demonstration of risk reduction for remediation or disposal options;
- testing for regulatory compliance (*e.g.*, federal program for Disposal at Sea); and
- investigation of potential toxicity sources through toxicity identification evaluation (TIE).

These aspects of toxicity testing assist in the refinement of uncertainty in the effects assessment and risk characterization stages.

1.2. Scope of Module

The scope of this module is restricted to the selection of toxicity tests for aquatic and lower-trophic level organisms, including invertebrates (aquatic and terrestrial), amphibians, fish[5], and plants. The module does not provide guidance for testing of mammals or birds because the use of site-specific toxicity testing for these species is rarely conducted in ecological risk assessments[6].

If a practitioner wishes to consider tests on mammals or birds, they should consult a specialist in the assessment of toxicity to domestic animals or wild species maintained in a laboratory setting, such as the Center for Integrative Toxicology at Michigan State University or the National Wildlife Research Centre (Environment Canada's Wildlife Toxicology Division, Ottawa, Ontario). Shore and Rattner (2001) summarize avian/mammalian ecotoxicology methods and

---

[4] Short-term toxicity tests should not be used to evaluate bioavailability for substances that are persistent and biomagnifying. Furthermore, lack of toxicity is not necessarily evidence of lack of accumulation, as the organism may not be sensitive to the exposure levels administered.

[5] "Fish" as applied in this module means teleost fish (bony fishes), rather than the broader definition of fish as provided in the *Fisheries Act*, which includes a variety of other aquatic life, including aquatic mammals.

[6] Toxicity tests on birds and mammals are applied rarely, and are limited to higher-tier detailed risk assessments. Protocols for toxicity testing using birds and mammals are highly species-specific, may include laboratory and/or field exposures, have a variety of test durations, and incorporate a variety of endpoints including mortality, development, histopathology, biochemistry, genetic mutations, tissue bioaccumulation and biomagnification, and behavioural measures.

findings from numerous sources, including wild mammals that have been brought into the lab and tested using standard laboratory protocols, such as mink (*Mustela vison*) and voles (*Microtus* spp.). Animal welfare issues are of heightened importance when tests on birds or mammals are contemplated. The Canadian Council of Animal Care (CCAC) oversees the treatment of animals used for scientific purposes in Canadian universities, government and private laboratories. Canadian wild species accounted for only 4.4% of experimental animals in 2008 (CCAC, 2008). The Wildlife Toxicology Division leads Environment Canada's National Wildlife Toxicology Program, and is the main source of federal scientific knowledge and expertise with respect to the impacts of toxic substances on wildlife. The National Wildlife Toxicology Program focuses on migratory birds, with lesser attention to amphibians, reptiles, mammals, and plants. Accordingly, there is some Canadian guidance for avian toxicity testing (Mineau *et al.* 1994, 1996), but the majority of the program elements rely on the research and monitoring from other wildlife programs.

There is not a sufficient technical basis to prescribe specific tests or endpoints for a given pathway or ecosystem type on a Canada-wide basis, nor is it expected that such a degree of prescription will be possible or desirable in the future. However, there is sufficient understanding of the available tests to provide a framework for the decision process, and to provide a means of evaluating the advantages and disadvantages of each candidate test prior to selection. Accordingly, this guidance module emphasizes the following:

- Identification of candidate toxicity tests for combinations of pathways, receptors, and exposure media (**Table 1**, **Table 2**);
- Identification of test factors that must be considered during the selection process (mandatory considerations);
- Preliminary evaluation of factors commonly used to evaluate tests, for use in test selection and uncertainty assessment (**Table 3**); and
- Guidance for the interpretation of toxicity test results.

The guidance in this module is intended to help the risk assessor in the selection of appropriate toxicity test endpoints on a site-specific basis. However, it is not intended to replace consultation with FCSAP Expert Support, professional toxicologists, and/or testing laboratories prior to finalizing selection of toxicity tests. Prior to investment of significant resources, such consultation is recommended to ensure that the results generated are of value to the study. Use of the guidance presumes that the practitioner is familiar with applicable methods for sampling and handling of test media, such as Environment Canada (1994) and CCME (1993a,b).

Toxicity tests may be conducted using water, sediment or soil samples, or combinations of these environmental media (either through sampling of multiple site media, or through exposure to multiple media in the laboratory). The tests are normally conducted using field-collected samples; however, manipulation of exposure concentrations can be achieved through spiking or dilution (*e.g.*, to achieve a range of exposure concentrations for development of a site-specific TRV). In some cases, manipulation of the sample media is conducted through mixing (*e.g.*, elutriate, mixing/resuspension) or through physical/chemical alteration (*e.g.*, purging, treatment with chemical sorbents). This guidance module briefly discusses sample manipulations, but focuses on the default test methods most commonly applied in the laboratory. *In situ* methods (*e.g.*, field

mesocosm studies) are discussed under the specialized test heading (**Appendix A** – Section 3); however, details of the *in situ* protocols are not explored within this module.

In other jurisdictions, attempts have been made to quantify the merits and limitations of toxicity tests and to recommend standardized toxicity testing[7]. However, many risk assessors and ecotoxicologists maintain that test selection should always be conducted on a project-specific basis. For this reason, this module does not attempt to quantify the degree of test applicability, provide relative rankings among test types, or exclude any individual tests. Rather, this module emphasizes a careful and informed assessment of test attributes prior to selection. For transparency, the practitioner must document a defensible rationale for a selected test, including assessment of potential pitfalls, constraints, or advantages of the test, as identified in **Table 3**.

This module emphasizes laboratory toxicity test methods. *In situ* methods are also relevant tools and may provide advantages in providing useful information relative to some laboratory toxicity tests, depending on the study objectives. They are an option for increasing site-specific relevance and/or for decreasing the lab- to field-extrapolation uncertainties. **Appendix A** briefly discusses *in situ* methods; however, detailed evaluation of *in situ* methods is beyond the scope of this Guidance module. If *in situ* methods are considered, it is recommended that the project manager discuss appropriateness of tests and candidate test species with the testing laboratory and seek advice of qualified environmental toxicologists when appropriate, including advice from FCSAP Expert Support.

This guidance module addresses technical issues and does not provide detailed guidance regarding fulfillment of federal, provincial, or local regulatory requirements, which vary by jurisdiction.

2. GUIDANCE

This section summarizes guidance for toxicity test selection and interpretation. Details are provided in tabular summaries of toxicity tests (**Tables 2 and 3**); detailed guidance on

[7] Recent examples include the Southern California Coastal Water Research Project, which conducted a series of technical studies to provide a sound scientific foundation for the selection of methods for the sediment quality objectives (SQO) program (Bay *et al.*, 2007). These investigations led to a recommendation to emphasize five endpoints (*Eohaustorius estuarius* [amphipod] survival, *Leptocheirus plumulosus* [amphipod] survival, *Rhepoxynius abronius* [amphipod] survival, *Neanthes arenaceodentata* [polychaete] juvenile growth, and *Mytilus galloprovincialis* [mussel] embryo development) as preferred standardized tests for evaluating coastal sediment toxicity. The British Columbia Ministry of Environment also investigated the feasibility and utility of identifying a subset of the available toxicity tests that should be recommended for use in ecological risk assessment; this effort included a multi-stakeholder workshop (with participants from across North America) in September 2007 to evaluate test types, data interpretation, and weight-of-evidence evaluations. The workshop identified several test types that are useful, commonly applied, and that may serve as a preferred starting point for establishing a study design, subject to site- and media-specific constraints. However, consensus was not reached on a preferred toxicity test regime for all situations.

the test selection process is provided in **Appendix A**. A graphic showing the stepwise process is provided in **Figure 1**.

The main sections of this module are structured as follows:

- General procedures for test selection;
- Candidate test summary; and
- Test interpretation.

2.1. General Procedure for Test Selection

The process for selecting a toxicity test has the following steps (**Figure 1**):

1. **Problem Formulation Review** – Consider project-specific objectives and constraints, building upon the considerations explored in Section 2 of the main FCSAP ERA guidance document. Implementation of this module presumes that the receptors of concern, assessment endpoints, protection goals, and operable pathways have already been identified. Clear identification on these aspects (which are discussed in the Problem Formulation section of FCSAP ERA guidance) is essential to understanding what toxicity tests are needed and acceptable. The correspondence between ecological risk pathways, ecosystem types, and receptor types is summarized in **Table 1**; this table can be used to identify the suite of candidate test types potentially applicable to the ERA[8]. Where available, site data need to be examined in order to identify data gaps, to avoid redundant testing, and to maximize relevance to the ERA.
2. **Receptor Summary** – Choose applicable receptor group(s) and receptor type(s) for the risk assessment (see guidance module on Receptor Selection) and, if applicable, determine whether the site is considered freshwater[9] or saltwater (marine/estuarine). In brackish water, classification of freshwater versus saltwater may not be obvious; consultation with regulators, field testing of salinity regimes, and/or review of existing tide and salinity data may be required. Careful attention should be paid to organism salinity tolerances during test selection. The practitioner should provide a rationale for selection of species in the final report for the study in order to be clear on reasons why a particular species was selected over another.
3. **Identify Applicable Environment Canada Tests** – **Table 2** lists applicable toxicity testing protocols by ecosystems and organism types, including the available Environment Canada protocols. When available, consult the Environment Canada biological test

[8] Table 1 includes headers for terrestrial mammals and birds for sake of compatibility with the receptor selection guidance – however, identification of test protocols for these receptor types is beyond the scope of this module.

[9] CCME defines freshwater as water with salinity of < 0.5%. However, for the purposes of toxicity testing, the more pertinent considerations are the tolerances and ecological relevance of candidate organisms; therefore, freshwater species may be considered at salinities higher than 0.5%.

methods[10] for the test organisms of interest. These protocols have been reviewed by federal and provincial scientists, and they are accepted by federal and provincial regulators. Furthermore, Canadian Association for Laboratory Accreditation (CALA) accredited capability for these tests exists within the Canadian consulting laboratory community[11]. Accordingly, all other factors being equal, the Environment Canada test methods are preferred to those developed in other jurisdictions in order to provide consistency in practice and readily determinable quality assurance evaluations.

4. **Identify Alternate Available Tests** – Where an existing Environment Canada protocol is not available or not ideally suited to the endpoint of interest, consult **Table 2** and Section 2.2 of this module to identify alternate candidate tests and protocols. In some cases there may not be representation of the specific organism type (*e.g.*, periphyton). In these cases, the list of candidate tests may be expanded to include all entries within the "receptor group" (*e.g.*, Aquatic Primary Producer), which provides a broader level of taxonomic and/or ecological representation[12]. If test accreditation is considered to be a desired feature of alternate tests (based on discussion with FCSAP Expert Support), the practitioner should contact CALA for information on whether accredited capability for these tests exists within the Canadian consulting laboratory community.

5. **Screen for Appropriateness/Acceptability** – Evaluate the mandatory considerations identified in **Appendix A**. Eliminate from further consideration any tests that are incongruent with the study objectives identified in Step 1, or that do not meet the specified requirements (*e.g.*, inappropriate grain size, salinity, or total organic carbon content [TOC]). Critically evaluate the species under consideration with respect to their tolerances to the factors listed in Step 6(b), below. Note that a test does not strictly require a published testing protocol in order to be an acceptable test. If the test is relevant, is performed well and appropriately for the substance, organism, and endpoint of interest, it is acceptable provided that the advantages and disadvantages of the test have been thoroughly evaluated and documented. If testing is intended to be used to develop chemical-specific thresholds, consideration should be given to relevant program-specific guidance for data quality. For example, where the purpose of the study includes collection of data on a chemical (or category of chemicals) as part of a Screening Information Data Set (SIDS), it is prudent to apply data quality evaluation templates such as OECD (2001) to assess reliability, relevance, and adequacy of data generated by the candidate tests.

6. **Evaluate in Detail for Suitability** – For remaining tests, evaluate the three categories of test suitability, using the information supplied in **Table 3** and details in **Appendix A**, supported by additional literature review as warranted. Consultation with an environmental toxicology laboratory is also appropriate.

    a. Utility for Risk Assessment Purposes

---

[10] In addition to Table 2, a separate biological methods publications list is available online at: http://www.etc-cte.ec.gc.ca/organization/bmd/bmd_publist_e.html

[11] Information available online at: http://www.cala.ca/

[12] See glossary definition.

i. Availability of Toxicity Data (for assessing sensitivity to substances of concern)

ii. Relevance of Negative Control to Test Media (if reference media unavailable)

iii. Statistical Power (ability to detect responses)

iv. Availability of Multiple, Sublethal, and/or Chronic Test Endpoints

v. Geographic Relevance (temperate, tropical, or regional species)

vi. Tissue Production (for potential bioaccumulation assessment)

b. Organism Tolerance[13]

i. Ammonia Tolerance

ii. Sulphide Tolerance

iii. Substrate Tolerance (particle size, TOC)

iv. Salinity / Water Hardness Tolerance

c. Logistics and Planning Factors

i. Laboratory Handling of Organisms (transport, acclimation, culturing, maintenance)

ii. Organism Source (available and reliable)

iii. Seasonal Availability of Organisms (suitable condition for testing)

iv. Sample Volume Requirements (feasible to acquire and transport)

v. Availability of Standard Method (commercial availability)

vi. Test Cost (per sample for standard replication)

7. **Select Appropriate Test(s)** – Based on the factors evaluated in Step 6 and the study objectives identified in Step 1, identify relevant tests. Relevant tests must apply appropriate test organisms and allow the generation of appropriate measurement endpoints that are aligned with the assessment endpoints of the ERA. A relevance check[14] should be conducted to ensure that the selected tests remain compatible with the study goals. The practitioners should consult the FCSAP ERA Guidance on receptor selection as part of this verification to ensure that measurement endpoints, assessment endpoints, and selected receptors are congruent. In addition, consultation with FCSAP Expert

[13] The listed attributes emphasize tolerance to factors that commonly confound test interpretation. Organisms also vary in tolerance to additional factors, such as light duration/intensity, aeration/oxygen conditions (*e.g.*, dissolved oxygen concentration), pH value, temperature, *etc*. These factors are controlled for in protocols for organism acclimation and testing conditions and are less commonly used to distinguish among candidate tests.

[14] See glossary.

Support, toxicologists and testing laboratories should be conducted prior to finalizing selection of toxicity tests.

8. **Consider Test Modifications** – There are situations in which modifications to the standard test protocol may be warranted. Examples of common test modifications include a change in test duration (*e.g.*, increase in particle settling time prior to addition of larvae in bivalve larval development tests to reduce entrainment effects), a change in number of replicates (*e.g.*, increased replicates to improve statistical power for a sublethal endpoint), or a sample manipulation (*e.g.*, brine adjustment to obtain appropriate salinity ranges). When contemplating a test modification, the practitioner should evaluate the trade-offs among the factors of test relevance, reliability, and standardization. **Appendix A** provides further information on specialized toxicity testing. Where deviations from standard operating procedures are contemplated: (1) the rationale should be carefully documented in the ERA; (2) consultation with regulatory agencies and/or FCSAP Expert Support is required; and (3) such deviations should be implemented only where the benefits of the modifications are clearly warranted. If test modifications are deemed necessary, side-by-side testing with standardized tests should also be considered, depending on the study objectives.
9. **Consider Additional Tests** – The use of multiple toxicity tests in a battery approach is generally preferred to use of a single presumed "best" toxicity test endpoint. It is rare that all of the factors necessary to comprehensively evaluate test sensitivity or reliability are known in detail prior to testing. Depending on the goals of the study (determined in Step 1), it may be appropriate to repeat the process (*i.e.*, apply a test battery approach) until sufficient information is collected. Multiple tests can be applied in parallel, or as part of a tiered assessment.

2.2. Additional Suitability Considerations for Pore Water

Pore water testing can be conducted with any of the test protocols marked as "water column" under test media in **Table 3**. Pore water toxicity tests have been described as advantageous due to their 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, Nipper *et al.*, 2002). However, pore water testing requires careful consideration of the following suitability factors prior to selection:

- Sample generation – The pore water must be collected in sufficient volumes, including necessary refreshes.
- Extraction method – The method of pore water extraction, and consequent effects on sample alteration and bioavailability should be evaluated carefully (Anderson *et al.*, 2001). Pore water samples may also require centrifugation or filtering to limit the confounding influence of entrained particles.
- Sample processing –In addition to centrifugation/filtering, adjustment for low dissolved oxygen, pH, and/or sulphide concentrations can be provided in the laboratory, but comes with an associated risk of modifying the chemical speciation and bioavailability of sample constituents.

- Ecological relevance – Some authors have cautioned that pore water 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 pore water and whole-sediment toxicity, although limited, indicate that toxicity is greater in pore water samples but linked primarily to ammonia rather than site-specific contaminants of potential concern (COPCs) (Burgess *et al.*, 1993; Anderson *et al.*, 2001; McDonald, 2005). One disadvantage of pore water tests is that many standardized test protocols result in exposure to ammonia in test media that is greater than would be reflective to the *in situ* condition (due to natural flushing and dilution processes), particularly for organisms that cycle overlying water. As such, the confounding factor of high levels of background ammonia in uncontaminated field collected reference sediment can result in toxicity artefacts.

2.3. Candidate Test Summary

Table 3 summarizes approximately 75 of the most commonly applied toxicity tests in North America. Each test is evaluated with respect to the test suitability characteristics identified above. A guide explaining the codes is presented in **Appendix A** (see subsection on test suitability factors).

The test suitability summary provides preliminary indications of the strengths and limitations of each test, using a simple coding system. Where a constraint to a given test is identified, the practitioner may need to consult additional detailed literature on the topic, using references provided in **Appendix A** and **Appendix B** or other supporting information. This module does not comprehensively evaluate every technical issue that may emerge during testing; rather, it is intended to highlight the highest priority factors that influence test selection and reliability.

As an example of the coding system, the marine amphipod *Rhepoxynius abronius* is flagged in **Table 3** as being constrained in terms of salinity range and particle size tolerances. Further investigation of this factor reveals that *Rhepoxynius abronius* is inappropriate for testing with low to moderate salinity ($< 2.5\%$), that salinity adjustment is rarely attempted for this species, and that this species is unsuited to testing of sediments with a high percentage of fines. The practitioner would need to consider an alternate amphipod species, such as *Eohaustorius estuarius*, if site sediments were dominated by silts and clays.

An important aspect of test selection is the assessment of caveats that apply to individual test methods, as summarized in the notes at the end of **Table 3**. Whereas some tests will be excluded on the basis of strict protocol-specified thresholds for test validity, in other cases the decision will be less clear, and will require application of professional judgement. For example, several studies have demonstrated that sediment toxicity tests conducted with larvae of oysters and mussels can exhibit adverse effects (test artifacts) caused by settling suspended particulate material (Elphick *et al.*, 2004). Similarly, sensitivity of bivalve species to low ammonia concentrations in test sediments has been observed (McDonald, 2005). For some of these factors, protocol-specified thresholds are often unavailable or do not identify the full range of exposures over which such artefact responses may occur. **Table 3** is intended to guard against overlooking these factors by: (1) flagging potential sensitivity *a priori* to minimize the probability of obtaining suspect data; and (2) identifying potential confounding factors that may need to be considered in the *post hoc*

evaluation of data. In many cases, test modifications and/or inclusion of appropriate reference samples assist in appropriate interpretation of toxicity data for risk assessment purposes.

2.4. Interpretation of Toxicity Testing Results

Environment Canada has developed guidance documents for numerical interpretation of their biological test methods. Accordingly, practitioners are advised to consult the Environment Canada *Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology* (Environment Canada, 1999). This document is supported by the Environment Canada (2007) supplemental guidance on statistical methods for environmental toxicity tests. Also, two of the Environment Canada test methods provide additional guidance on data interpretation (Environment Canada, 1998b; Environment Canada, 2002). The above references provide the default procedures to be applied in the evaluation of individual toxicity test results. The remainder of this section outlines some additional considerations that apply to the use of toxicity test data within a risk assessment or weight-of-evidence (WOE) framework.

Prior to the interpretation of toxicity test results, the quality of data must be confirmed to the extent possible. This module presumes that data have been collected following established protocols and that Quality Assurance/Quality Control (QA/QC) procedures have been satisfactorily implemented in the collection of the data. The practitioner should confirm the above, thoroughly review the laboratory report, and confirm that no substantive protocol deviations have occurred that would influence test interpretation. General health of test organisms, negative control performance and reference toxicant results are also important considerations when evaluating the validity of test results. The practitioner should evaluate the documentation of field sample collection, handling, and transport, particularly where anomalous test results are observed.

2.4.1 General Considerations

Provision of guidance concerning toxicity test interpretation is challenging due to the diversity of test types, endpoints, and study objectives. However, the following general principles apply to the interpretation of all tests.

- Toxicity testing results should be evaluated in relation to the protection goals established for the site specific ecological risk assessment. Toxicity tests may provide one or more lines of evidence in the WOE framework for characterizing risk at the site. Decision points and criteria for acceptable or unacceptable risk should be identified prior to risk characterization. To this end, the practitioner should consult with the regulatory agencies and/or FCSAP Expert Support during the problem formulation stage.
- The interpretation of toxicity testing results relates to the objective of the toxicity testing. For example, the interpretation of a toxicity test conducted for purposes of a TIE may differ from the same test conducted as part of a WOE.

- Polarized decisions (*e.g.*, toxic or non-toxic) should be avoided, as they oversimplify the test results and result in loss of information[15].
- Interpretation should recognize the precision (or lack thereof) of the measurement endpoints. For example, amphipod survival rates of 89.6% and 90.4% are not distinguishable in terms of ecological relevance, given the precision of the test. Most toxicity test data should have no more than two significant figures presented in data summaries. Additional significant figures should only be provided if warranted by the precision of the experiment and where necessary to interpret the results.
- Interpretation should consider the suitability of the control or reference media against which site samples are compared. Relatively clean reference media of similar physical composition to site samples are preferred to laboratory negative controls (the latter are intended primarily for quality control purposes). The reference evaluation should consider not only the chemical contamination of the reference media but also physical and/or biological characteristics that may influence test performance. It is often beneficial to include media from reference or background sites in the toxicity testing regime.
- Where possible, the interpretation should not emphasize only statistical significance/confidence, or only effect size. Rather, to the extent possible, the interpretation should consider all aspects of statistical power (*e.g.*, effect size, sample size, variability of endpoint data, appropriate alpha [α] levels). Probability of both Type I and Type II errors (Allchin, 2001) should be considered. See Glossary under headings of *Statistical Power* and *Type I and Type II errors* for clarification of these concepts and default practice for assigning alpha [α] and beta [β] levels.
- When interpreting toxicity test results, attention should be paid to visual observations made during the test (as documented in laboratory notes, photographs, test media descriptions, *etc.*). Although qualitative, these observations can provide important insights for test interpretation and uncertainty assessment. For example, indications of microbial/microalgae growth in test vessels may signify an important confounding biological factor. Observations of turbidity in test vessels can provide indications of potential entrainment responses, such as may occur post-settling in some bivalve larval development tests. Observations of changing sediment colour or texture may also provide insight regarding chemical bioavailability processes (*e.g.*, oxidation of metals and presence of precipitates or flocculated material).

2.4.2 Test Interpretation Methods

This section summarizes some of the methods commonly used to identify toxicity test response thresholds. Environment Canada (1999, 2007) provides additional details for several of the methods discussed below. In some risk assessments, definition of numerical acceptable effect

[15] For some regulatory purposes, clear designation may be required where an existing decision framework exists. For example, the Disposal at Sea Program has prescribed decision points based on data from pass/fail tests. Details are found at http://www.ec.gc.ca/seadisposal/monitoring/index_e.htm#biological. This guidance is not intended to supplant those frameworks.

thresholds may not be required. This may occur in cases where the scope of the ERA is simply to summarize effects without any judgment about acceptability.

2.4.2.1 Point Estimation

The development of point estimation methods comes from standard ecotoxicity testing protocols (Rand, 1995). The inhibitory concentration (IC_X) is the concentration that results in a magnitude of effect of X% over a specified period of time, whereas the effect concentration (EC_X) is the concentration that results in a prescribed response in X percent of the test organisms over a specified period of time (see definition box below). To derive point estimates, data from fixed times of observation are characterized using one of several models (*e.g.*, logistic regression, or linear models based on probits or logits) (Rand, 1995). Although statistical details vary, all methods use some type of numerical interpolation to estimate the test concentration associated with a defined level of response (Zajdik, 2007). For quantitative tests, Environment Canada (2007) indicates that point estimates derived through regression are preferred to alternate approaches. Such estimates consider concentration-response relationships and are less susceptible to weaknesses of the hypothesis testing framework (Chapman *et al.*, 1996).

Point estimation procedures are conceptually attractive because they relate concentration (whether on a volume/volume basis or on a contaminant-specific basis) to a prescribed level of effect. Point estimation procedures (IC_X /EC_X) can address statistical power and significance using statistical confidence bands. Furthermore, point estimates apply best to tests conducted over a gradient of exposure (such as a dilution series). In sediment or soil testing, dilution series are rarely conducted, and due to the complex mixtures of contaminants present in field-collected samples, derivations of reliable contaminant-specific point estimates can be challenging. If contaminant-specific point estimates are required for soil or sediment, they can be derived empirically, through concentration-response assessment of multiple samples over a gradient, or experimentally, using spiked sediment toxicity tests. The latter entail an increased level of technical sophistication due to issues related to spiking of contaminants in a way that is representative of field bioavailability and not excessively disruptive of the physical structure of the sample matrix.

EC_X versus IC_X Definition *– For endpoints other than mortality, there is some confusion about the meaning of EC_X (Effect Concentration). A true EC_X typically applies to dichotomous variables and is the concentration at which the percentage of the test population demonstrating a specific response relative to controls over a specified time period is X – for example, for an EC_{20}, 20% of individuals tested may exhibit a specified level of reproductive impairment (e.g., 20% of bivalve larvae failed to develop normally). An EC_X can also be applied for mortality – for example 20% of the test population died; however, this is more commonly referred to as an LC_X (Lethal Concentration). In contrast, an IC_X (Inhibitory Concentration) is the concentration at which X% impairment occurs for a continuous response variable – for example, for an IC_{20}, the average individual organism in the test population would be expected to exhibit 20% reproductive impairment relative to control over a specified time period. Many guidance documents use the term EC_X loosely and provide examples that are either EC_X or IC_X.*

2.4.2.2 Use of NOAELs/LOAELs

Derivation of no-observed-adverse-effect and lowest-observed-adverse-effect levels (NOAELs, LOAELs) is based on strict statistical significance criteria (in a traditional hypothesis testing framework) applied to toxicity data collected at multiple exposure levels. Pairwise comparisons are conducted between treatment effects and the control (or reference) condition, and a binary determination of statistical significance is made. In most cases the level of significance (α) is set to 0.05, and the resulting probability of a Type II error ($1-\beta$; see Glossary) varies widely among test methods and samples. In addition, the value of a NOAEL or LOAEL is strongly dependent on the spacing of the treatments along an exposure gradient and the sample sizes used. Therefore, NOAEL/LOAEL assessments have received broad criticism (Chapman *et al.*, 1996), and are insufficient for evaluating toxicity data for risk assessment or regulatory purposes. Instead, application of NOAELs/LOAELs is recommended only as part of a more robust decision process. Environment Canada (2007) notes that use of the NOAEL/LOAEL procedure is decreasing in frequency and that the procedure is less desirable than point estimation.

2.4.2.3 Effect-Size Thresholds

The reduction threshold approach is a method of evaluating the significance of toxicity test results, using an effect-size based approach. As such, the use of effect-size threshold is a variant of the point estimation (IC_X /EC_X) approach (Section 2.4.2.1), and is based on a policy decision that a defined level of effect is acceptable. The magnitude of effect size (or endpoint reduction) deemed to be ecologically significant varies (by jurisdiction, land use, type of ecosystem, type of organism, *etc.*), but numerous investigators have identified 20% and 50% reductions to an individual-level survival, growth, or reproduction endpoint as being indicative of thresholds for environmental significance in laboratory tests (with moderate and large effect sizes, respectively). In practice, reduction thresholds apply most commonly to IC_X determinations, and represent the average magnitude of reduction for a growth or reproduction endpoint. These values (20% and 50% reductions) may have a technical basis as well as a policy-administrative basis. The main challenge with their implementation is that the degree of environmental impairment associated with a fixed numerical effect-size varies depending on the endpoint type and the scale of the endpoint. For example, a 20% reduction in growth (IC_{20}) measured in a laboratory test for a single invertebrate species may be considered a suitable threshold for environmental significance, whereas a 20% reduction in population of a wide-ranging and socially or commercially valuable fish stock (EC_{20} for mortality) would not. The latter would be inconsistent with federal legislation (*i.e.*, Fisheries Act considerations) and would therefore not be acceptable.. For field studies and endpoints with high ecological and/or human value, thresholds lower than 20% may be adopted. The Federal Government has not established a general policy regarding acceptable effect sizes. At the time of publishing, protection goals and acceptable effects levels are determined on a site specific basis. Practioners should consult with FCSAP Expert Support prior to adopting any specific effect-size as a threshold for environmental significance.

In summary, using reduction thresholds can provide a suitable compromise between sensitivity, confidence and reliability. However, because effects at or above a prescribed level may not always be concordant with ecological significance, these categorizations should be considered carefully, and where possible, combined with statistical significance measures (see *Ordinal*

Approaches, below). Furthermore, more than one line of evidence should be used in management decisions, as part of a weight-of-evidence procedure. Expert Support and other regulatory authorities should be consulted where a weight-of-evidence approach is needed, especially if there are plans to divest the site, and taking into consideration current and future land uses.

In the absence of policy-driven or test-driven thresholds, risk assessors may work with site managers to develop site-specific thresholds. In the absence of other information, default provisional values have been suggested (see Section 2.4.2.4). The provisional values should be applied with caution, as they were developed based on assessment of select marine sediment toxicity tests, and may not apply to all situations.

2.4.2.4 Ordinal Approaches for Effect-Size Thresholds

Effect-size thresholds are not necessarily binary (nontoxic/toxic) in nature. The following example of an ordinal approach to interpreting results of toxicity tests is suggested by the Southern California Coastal Water Research Project (Bay *et al.* 2007) for aquatic tests. This approach is not intended as prescribed approach for interpretation, but rather provides an example of how continuous data can be categorized while considering both statistical significance and effect size. The ordinal approach "retains more information about the toxicity response and thus provides greater potential resolution when combining the toxicity data with other lines of evidence in a sediment quality triad approach" (Bay *et al.*, 2007). The following ordinal categories are established:

- Nontoxic – Response not substantially different from that expected in uncontaminated media and have optimum characteristics for the test species;
- Low toxicity – A response that is of relatively low magnitude; the response may not be greater than test variability;
- Moderate toxicity – High confidence that a statistically significant effect is present; and
- Strong toxicity – Highest confidence that a toxic effect is present and the magnitude of response is among the strongest effects observed for the test.

These categories differ from the previous approaches (Sections 2.4.2.1 through 2.4.2.3) in that the categories are assigned based on an assessment of both statistical significance and the absolute magnitude of the observed responses. The approach relies on comparison of the test result (*e.g.*, percent survival relative to reference) to three quantitative effect thresholds, corresponding to the upper bound of the response range for the Low Toxicity, Moderate Toxicity, and Strong Toxicity categories. Specifically, the categories consider the following decision rules:

- A sample is considered "nontoxic" if either of the following conditions apply: (1) the unadjusted response value is less than a "low threshold"; or (2) the reference-normalized response exceeds only the "low threshold" and is not statistically different from the control.
- A sample is considered to exhibit "low toxicity" if either of the following conditions apply: (1) the reference-normalized response exceeds only the "low threshold" and is statistically different from the control; or (2) the reference-normalized response exceeds the "moderate threshold" but is not statistically different from the control.

- A sample is considered to exhibit "moderate toxicity" if the reference-normalized response exceeds the "moderate threshold" and is statistically different from the control.
- A sample is considered to exhibit "strong toxicity" if the reference-normalized response exceeds the "high threshold", irrespective of statistical significance.

This framework, although shown here as an example only, is conceptually consistent with recent guidance prepared for the provincial environment ministries in BC[16] and Ontario[17]. Bay *et al.* (2007) developed numeric thresholds for selecting species using test-specific characteristics, such as test variability (minimum significant difference [MSD]) and distribution of the toxicity response data. A statistical criterion was also used in the classification scheme using Student *t* tests with $\alpha \leq 0.05$ criterion for probability of a Type I error (comparison of negative control to test sediments, assuming unequal variances[18]). Rationales for the thresholds applied (that may be used as guidance for future threshold development) include:

- Low Threshold – Bay *et al.* (2007) developed the threshold separating the Nontoxic and Low categories considering the lowest acceptable control response value for the given test, as established in the test protocols. The rationale was that any response that fell within the range expected of animals exposed to optimum sediment conditions (*i.e.*, controls) should indicate a nontoxic condition in the test sample.
- Moderate Threshold – The narrative intent of the Moderate Threshold was to distinguish between samples producing a small response of uncertain significance and larger responses representing a reliably significant difference relative to the control (Bay *et al.*, 2007). The numerical thresholds for specific tests were derived considering the MSD[19], which was specific to each test method . The moderate threshold was set equal to the 90th percentile of the MSDs for a given toxicity test method. As the mean and median values of the thresholds were close to 80% (*i.e.*, 20% reduction), this was adopted for provisional use in evaluating other toxicity endpoints. In addition, Efroymson and Suter (1999) and Pack (1993) have suggested that reductions in survival, growth, or reproduction of 20% or greater are indicative of significant effects to wildlife. Accordingly, the 20% reduction approach has been adopted as common risk assessment

---

[16] The BC guidance for Detailed Ecological Risk Assessment (SAB, 2008) does not provide prescriptive decision rules for categorizing responses, but advocates the simultaneous consideration of effect size, statistical significance, and other aspects of statistical power in the interpretation of endpoint responses.

[17] A similar approach is used by Environment Canada with biological data, using a comparison-to-reference approach using the Canadian Aquatic Biomonitoring Network (CABIN). The degree to which the invertebrate assemblage is similar (or dissimilar) to the predicted assemblage determines its classification on a gradient of perturbation (*e.g.*, unstressed to severely stressed) relative to the reference sites. Additional information is available online at: http://cabin.cciw.ca/Main/cabin_about.asp

[18] Student *t* tests can be conducted using an assumption of equal variances (pooled variance) or different variances; the latter was adopted to acknowledge expected systematic differences in variance between test and control sediments.

[19] The MSD is defined as the minimum difference between the control and sample mean response that is necessary to be statistically different at $p \leq 0.05$ level.

practice in several North American jurisdictions to identify potentially ecologically relevant responses.

- High Threshold – The narrative intent of the High Threshold was to discriminate samples producing a severe and highly significant effect from those samples producing lesser effects. To make this determination for specific tests, Bay *et al.* (2007) considered a combination of test variability and response distribution that corresponded to the category definition. Test variability was evaluated using the 99th percentile MSD value, whereas the response distribution component of the high threshold was based on the distribution of numerous toxic samples from California (75th percentile of test samples having a mean response that was significantly different from the control response). The mean of the two values (test variability and response distribution approaches) was used as the High Threshold Response. Bay (pers. comm., 2008) has also indicated that a high frequency of benthic community impacts is observed at >50% amphipod mortality, which supports the choice of 50% as indicative of a greater potential ecological response.

For all response thresholds, it is preferable to apply test-specific data. Although the Bay *et al.* (2007) study suggested that provisional values of 10% reduction (low), 20% reduction (moderate), and 50% reduction (high) are reasonable thresholds for categorization of sediment toxicity responses, these should be updated where possible with more rigorously derived thresholds. Caution should be exercised in applying these thresholds to test types not evaluated in the California study.

2.4.2.5 Other Approaches

Other approaches have been considered for the interpretation of toxicity test results, including the "reference envelope" method and the Minimum Significant Difference /Minimum Detectable Difference (MSD/MDD) method (SFF, 2007). These alternative approaches can, in some cases, have some advantages over the methods described above. The reason for not including them in the guidance at this time is they are often constrained by baseline data availability (*e.g.*, establishment of regional reference performance data, robust data sets required for calculating numerous endpoint-specific MSDs). They are also more difficult to relate to protection goals, such as a 20% reduction response threshold. Finally, it should be noted that in some cases there may be test-specific thresholds that have been established for particular tests in particular geographic regions (*e.g.*, Bay *et al.*, 2007).

3. FURTHER INFORMATION

Appendix A provides detailed guidance on test selection, including a discussion of test attributes that influence the site-specific applicability of a test. **Appendix B** lists references for toxicity test protocols as well as references providing technical guidance related to toxicity testing.

Figure 1: General Procedure for Toxicity Test Selection

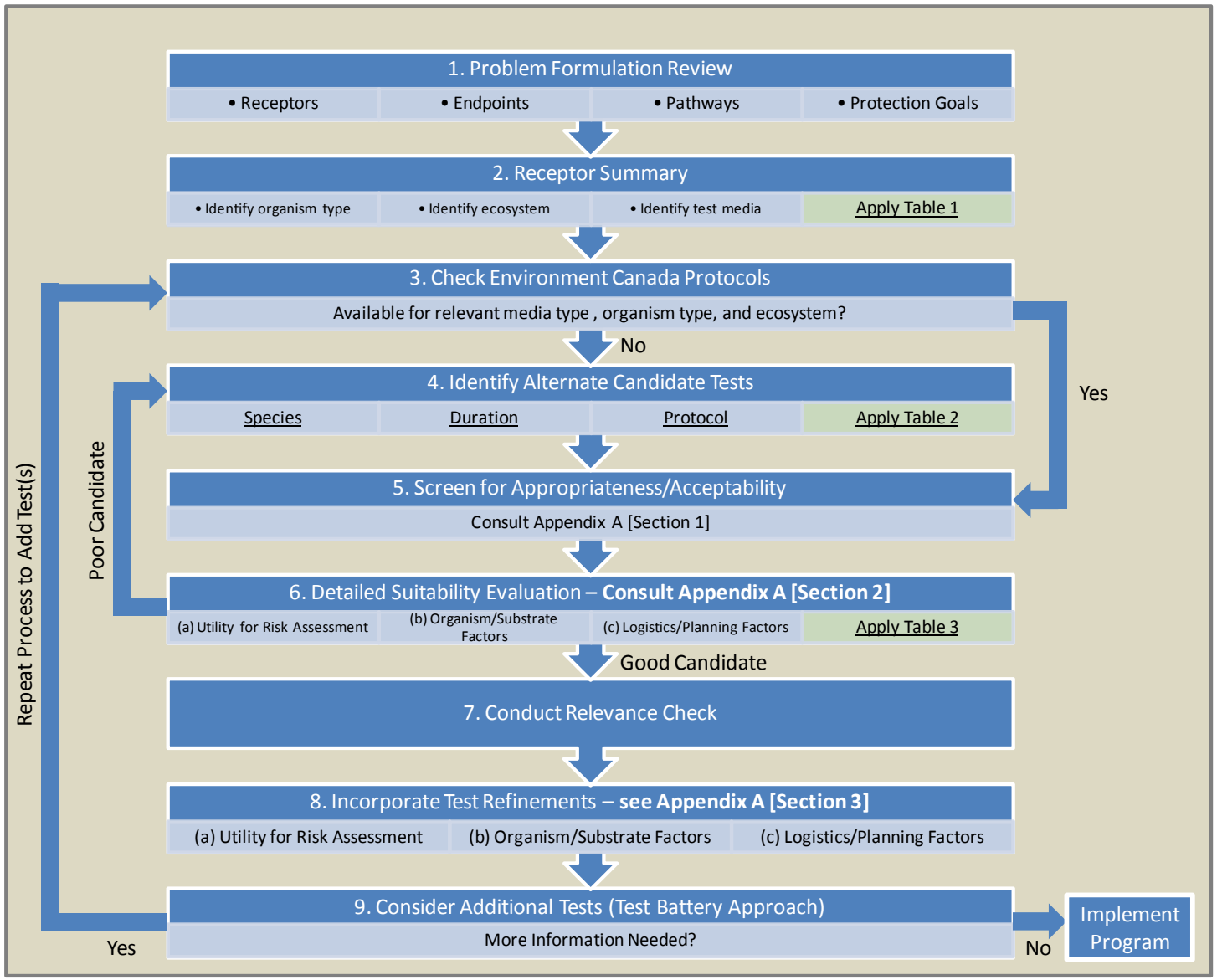

Table 1: Linkage between receptor types, applicable ecosystem types, and toxicity test media/pathways

Receptor Group	Organism Type	Organism Type ID	Applicable Ecosystem Types	Test Media	Pathways Simulated
Terrestrial Primary Producer	Moss / Grass / Shrub / Tree / Forb	PP-PLANT	Terrestrial – human-influenced land (all land uses), wildland (all types)	Soil	Translocation through roots from soil and porewater; direct contact between roots and soil.
Terrestrial Invertebrate	Ground-dwelling	INV-GROUND	Terrestrial – human-influenced land (all land uses), wildland (all types)	Soil	Direct contact with contaminated food, ingestion of contaminants through processing and feeding on soil and associated interstitial fluid.
	Aerial	INV-AERIAL	Terrestrial – human-influenced land (all land uses), wildland (all types)	Sediment*	*Protocols currently unavailable for soil-based toxicity tests for this pathway. Sediment-based test are available for early lifestages of emergent insects (*e.g.*, mayfly); this would reflect the pathway of uptake to larvae from contaminated sediment, pore water, and overlying water.
Terrestrial Mammal	Herbivorous	MAMMAL	Terrestrial – human-influenced land (all land uses), wildland (all types)	Diet	Mammalian tests are highly specialized and are beyond the scope of this module. Most laboratory testing with mammals is conducted via exposure to drinking water and/or contaminated food.
	Insectivorous	MAMMAL			
	Carnivorous	MAMMAL			
	Omnivorous	MAMMAL			
Terrestrial Bird	Herbivorous	BIRD	Terrestrial – human-influenced land (all land uses), wildland (all types)	Diet	Avian tests are highly specialized and are beyond the scope of this module. Most laboratory testing with birds is conducted via exposure to drinking water and/or contaminated food.
	Insectivorous	BIRD			
	Carnivorous	BIRD			
	Omnivorous	BIRD			
Aquatic Primary Producer	Phytoplankton	PP-PHYTO	Marine - deep (ocean, harbour); Freshwater – deep (lake)	Surface Water, Pore Water	Direct uptake through diffusion through cell membrane.
	Periphyton	PP-PERI	Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream)	N/A	Standard toxicty test methods not well developed for this receptor.
	Macrophyte	PP-MACRO	Marine – shallow/shoreline (rocky shore, tidal flat, estuary); Freshwater – shallow/shoreline (pond, marsh)	Surface Water, Pore Water, Sediment	Direct uptake through diffusion through cell membrane; absorption from sediment and pore water through roots.
Aquatic Pelagic Invertebrate	Zooplankton	PELAG-ZOO	Marine - deep (ocean, harbour); Freshwater – deep (lake)	Surface Water, Pore Water	Direct uptake through diffusion through cell membrane (unicellular); Respiration via gill membrane for planktonic crustaceans.

Table 1 (continued): Linkage between receptor types, applicable ecosystem types, and toxicity test media/pathways

Receptor Group	Organism Type	Organism Type ID	Applicable Ecosystem Types	Test Media	Pathways Simulated
	Others	PELAG-OTHER	Marine - deep (ocean, harbour); Freshwater – deep (lake)	Surface Water, Pore Water, Sediment	Direct uptake through diffusion, respiration and direct contact. Mysid tests can be conducted as whole sediment or water-only tests.
Aquatic Benthic Invertebrate	Epifauna; and	BENTHIC-EPI; and	Marine – shallow/shoreline (rocky shore, tidal flat, coastline, estuary); Marine – deep (ocean); Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream); Freshwater – deep (lake)	Sediment, Surface Water, Pore Water	Direct contact with substrate, absorption and respiration for all exposure media tested.
	Infauna	BENTHIC-INF		Sediment	Direct contact with substrate, absorption and respiration for all exposure media tested.
Fish	Benthivorous	FISH	Marine – shallow/shoreline (rocky shore, tidal flat, coastline, estuary); Marine – deep (ocean); Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream); Freshwater – deep (lake)	Surface Water, Pore Water	Standard toxicity tests evaluate the respiration and direct accumulation pathways from aqueous media. In nature, exposure also occurs through contact and incidental ingestion of contaminated sediments (especially for benthivores) and via dietary accumulation of prey, but these pathways are represented only in specialized test regimes.
	Planktivorous	FISH			
	Piscivorous	FISH			
Aquatic Mammal	Herbivorous	MAMMAL	Marine – shallow/shoreline (rocky shore, tidal flat, coastline, estuary); Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream)	Diet	Mammalian tests are highly specialized and are beyond the scope of this module. Most laboratory testing with mammals is conducted via exposure to drinking water and/or contaminated food.
	Piscivorous	MAMMAL			
	Omnivorous	MAMMAL			
Aquatic Bird	Herbivorous	BIRD	Marine – shallow/shoreline (rocky shore, tidal flat, coastline, estuary); Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream)	Diet	Avian tests are highly specialized and are beyond the scope of this module. Most laboratory testing with mammals is conducted via exposure to drinking water and/or contaminated food.
	Insectivorous	BIRD			
	Piscivorous	BIRD			
	Omnivorous	BIRD			
Amphibian	Carnivorous / Herbivorous / Omnivorous (lifestage and species dependent)	HERPTILE	Freshwater – shallow/shoreline (pond, marsh); Freshwater – flowing (river, stream); Terrestrial - wildland (for adult members of some species, such as wood frogs)	Sediment, Surface Water, Pore Water	Amphibian toxicity tests often emphasize accumulation/respiration from overlying water. However, specialized test protocol incorporate consideration of maternal transfer to eggs, exposure to sediments during larval and metamorphic stages.
Reptile	Omnivorous	HERPTILE	Terrestrial – wildland (grassland/prairie, forest, tundra, alpine)	N/A	Standard toxicty test methods not well developed for this receptor.

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
	FRESHWATER						
1	Frog - 96 h - Survival, Development and Growth	Frog	Amphibian	AMPHIB	*Xenopus laevis*	Acute	ASTM Method E1439-98 (ASTM, 2006)
2	Frog - 21 d - Survival, Development and Growth	Frog	Amphibian	AMPHIB	*Xenopus laevis*	Chronic	OECD (2008)
3	Amphibian - 10 d - Survival and Growth	Amphibian	Amphibian	AMPHIB	*Rana pipiens; R. clamitans; R. sylvatica; Bufo americanus*	Chronic	ASTM Method E2591-07 (ASTM, 2008)
4	Amphipod - 10, 14 d - Survival and Growth	Amphipod	Arthropod (Crustacea)	BENTHIC-EPI	*Hyalella azteca*	Acute	ASTM Method E1706-05 (ASTM, 2006); EPA/600/R-99/064 (USEPA, 2000); EPS 1/RM/33 (Environment Canada, 1997a)
5	Amphipod - 42 d - Survival, Growth and Reproduction	Amphipod	Arthropod (Crustacea)	BENTHIC-EPI	*Hyalella azteca*	Chronic	ASTM Method E1706-05 (ASTM, 2006); EPA/600/R-99/064 (USEPA, 2000)
6	Mayfly - 10 d - Survival	Mayfly	Arthropoda (Insecta)	BENTHIC-EPI	*Hexagenia limbata*	Acute	OMOE (1992)
7	Mayfly - 21 d - Survival, Growth and Moulting Frequency	Mayfly	Arthropoda (Insecta)	BENTHIC-EPI	*Hexagenia limbata*	Chronic	OMOE (1992); ASTM Method E1706-05 (ASTM, 2006)
8	Fatmucket Mussel - 48 - 96 h - Survival	Fatmucket Mussel	Mollusc (Bivalve)	BENTHIC-EPI	*Lampsilis siliquoidea*	Acute	ASTM Method E2455-06 (ASTM, 2008)
9	Fatmucket Mussel - 28 d - Survival and Growth	Fatmucket Mussel	Mollusc (Bivalve)	BENTHIC-EPI	*Lampsilis siliquoidea*	Chronic	Ingersoll et al. (2008; cf MOE, 2007)
10	Mussels - 10 d - Survival	Mussels	Mollusc (Bivalve)	BENTHIC-EPI	*Anodonta imbecillis*	Acute	USEPA/USACE (1998)
11	Cladoceran - 96 h - Survival, Growth and Development	Cladoceran	Arthropod (Crustacea)	BENTHIC-EPI	*Chydorus sphaericus*	Chronic	Dekker et al. (2006)
12	Oligochaete - 10 d - Survival	Oligochaete	Annelid (Oligochaeta)	BENTHIC-INF	*Pristina leidyi*	Acute	USEPA/USACE (1998)
13	Sludgeworm - 10 d - Survival	Sludgeworm	Annelid (Oligochaeta)	BENTHIC-INF	*Tubifex tubifex*	Acute	USEPA/USACE (1998)
14	Sludgeworm - 28 d - Survival and Reproduction	Sludgeworm	Annelid (Oligochaeta)	BENTHIC-INF	*Tubifex tubifex*	Chronic	ASTM Method E1706-05 (ASTM, 2006)
15	Amphipod - 28 d - Survival and Behaviour	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Diporeia sp.*	Chronic	ASTM Method E1706-05 (ASTM, 2006)

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
16	Midge - 10 d - Survival and Growth	Midge	Arthropoda (Insecta)	BENTHIC-INF	*Chironomus dilutus*	Chronic	ASTM Method E1706-05 (ASTM, 2006); EPA/600/R-99/064 (USEPA, 2000); EPS 1/RM/32 (Environment Canada, 1997b)
17	Midge - 30 d; 50 - 65 d - Survival, Growth and Adult Emergence, Reproduction	Midge	Arthropoda (Insecta)	BENTHIC-INF	*Chironomus riparius*	Chronic	ASTM Method E1706-05 (ASTM, 2006); EPA/600/R-99/064 (USEPA, 2000)
18	Blackworm - 10 d - Survival	Blackworm	Annelid (Oligochaeta)	BENTHIC-INF	*Lumbriculus variegatus*	Acute	USEPA/USACE (1998)
19	Blackworm - 28 d - Survival, Reproduction and Growth	Blackworm	Annelid (Oligochaeta)	BENTHIC-INF	*Lumbriculus variegatus*	Chronic	OECD Guideline 225 (OECD, 2007)
20	Fathead Minnow - 96 h - Survival	Fathead Minnow	Fish	FISH	*Pimephales promelas*	Acute	EPA-821-R-02-012 (USEPA, 2002a); USEPA/USACE (1998)
21	Fathead Minnow - 7 d - Survival and Growth	Fathead Minnow	Fish	FISH	*Pimephales promelas*	Acute	EPS 1/RM/22 (Environment Canada, 1992a); EPA-821-R-02-013 (USEPA, 2002b)
22	Rainbow Trout - 96 h - Survival	Rainbow Trout	Fish	FISH	*Oncorhynchus mykiss*	Acute	EPS 1/RM/9 (Environment Canada, 1990a); EPS 1/RM/13 (Environment Canada, 2000a); EPA-821-R-02-012 (USEPA, 2002a)
23	Rainbow Trout - 7 d - Embryo Viability	Rainbow Trout	Fish	FISH	*Oncorhynchus mykiss*	Acute	EPS 1/RM/28 (Environment Canada, 1998a)
24	Rainbow Trout - ~30 d - Alevin Viability, Hatching and Deformity	Rainbow Trout	Fish	FISH	*Oncorhynchus mykiss*	Chronic	EPS 1/RM/28 (Environment Canada, 1998a)
25	Rainbow Trout - ~70 d - Alevin Viability, Hatching and Deformity, Fry Survival and Behaviour	Rainbow Trout	Fish	FISH	*Oncorhynchus mykiss*	Chronic	EPS 1/RM/28 (Environment Canada, 1998a)
26	Waterflea - 7 d, 21 d - Survival and Reproduction	Waterflea	Arthropod (Crustacea)	PELAG-ZOO	*Daphnia magna*	Chronic	OECD Guideline 211 (OECD 2007); ASTM Method E1193-97; E1705-05 (ASTM, 2006)
27	Waterflea - 48 h - Survival	Waterflea	Arthropod (Crustacea)	PELAG-ZOO	*Ceriodaphnia dubia*	Acute	EPA-821-R-02-012 (USEPA, 2002a)
28	Waterflea - 7 d - Survival and Reproduction	Waterflea	Arthropod (Crustacea)	PELAG-ZOO	*Ceriodaphnia dubia*	Chronic	EPS 1/RM/21 (Environment Canada, 2007a); EPA-821-R-02-013 (USEPA, 2002b); ASTM Method E1706-05 (ASTM, 2006)

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
29	Waterflea - 48 h - Survival	Waterflea	Arthropod (Crustacea)	PELAG-ZOO	*Daphnia magna, Daphnia pulex*	Acute	EPS 1/RM/11 (Environment Canada, 1990b); EPS 1/RM/14 (Environment Canada, 2000b); EPA-821-R-02-012 (USEPA, 2002a)
30	Rotifer - 24 h - Survival	Rotifer	Rotifer	PELAG-ZOO	*Brachionus calyciflorus*	Acute	ASTM Method E1440-91 (ASTM, 2006)
31	Duckweed - 7 d - Growth	Duckweed	Aquatic Macrophyte	PP-MACRO	*Lemna minor*	Chronic	EPS 1/RM/37 (Environment Canada, 2007c)
32	Domestic Rice - 14 d - Growth, Chlorophyll Content	Domestic Rice	Emergent Macrophyte	PP-MACRO	*Oryza sativa*	Chronic	ASTM Method E1841-04 (ASTM, 2006)
33	Phytoplankton - 72 h - Cell Yield	Phytoplankton	Algae	PP-PHYTO	*Pseudokirchneriella subcapitata*	Chronic	EPS 1/RM/25 (Environment Canada,
	MARINE/ESTUARINE						
34	Mussels - 48 h - Larval Development and Survival	Mussels	Mollusc (Bivalve)	BENTHIC-EPI	*Mytilus sp.*	Chronic	EPA-600/R-95/136 (USEPA, 1995); ASTM Method E724-98 (ASTM, 2006); PSEP (1995); USEPA/USACE (1998)
35	Oyster - 48 h - Larval Development and Survival	Oyster	Mollusc (Bivalve)	BENTHIC-EPI	*Crassostrea gigas; C. virginica*	Chronic	EPA-600/R-95/136 (USEPA, 1995); ASTM Method E724-98 (ASTM, 2006); PSEP (1995); USEPA/USACE (1998)
36	Red Abalone - 48 h - Larval Development and Survival	Red Abalone	Mollusc (Univalve)	BENTHIC-EPI	*Haliotis rufescens*	Chronic	EPA-600/R-95/136 (USEPA, 1995)
37	Sea Urchin - 48 – 96 h - Larval Development and Survival	Sea Urchin	Echinoid	BENTHIC-EPI	*Strongylocentrotus droebachiensis; Arbacia punctulata*	Chronic	ASTM Method E1563-98 (ASTM, 2006); PSEP (1995); USEPA/USACE (1998)
38	Sea Urchin - 10:10 min; 20:20 min; 60:20 min - Fertilization	Sea Urchin	Echinoid	BENTHIC-EPI	*Strongylocentrotus purpuratus; Arbacia punctulata*	Chronic	EPS 1/RM/27 (Environment Canada, 1992b); EPA-600/R-95/136 (USEPA, 1995)
39	Sea Urchin - 10:10 min; 20:20 min; 60:20 min - Fertilization	Sea Urchin	Echinoid	BENTHIC-EPI	*Strongylocentrotus droebachiensis; Lytechinus pictus*	Chronic	EPS 1/RM/27 (Environment Canada, 1992b); EPA-600/R-95/136 (USEPA, 1995)
40	Clam - 7 d - Survival and Growth	Clam	Mollusc (Bivalve)	BENTHIC-INF	*Mulinia lateralis*	Chronic	Burgess and Morrison (1994)

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
41	Polychaete Worm - 20-28 d - Survival, Reproduction and Growth	Polychaete Worm	Annelid (Polychaeta)	BENTHIC-INF	*Capitella capitata*	Chronic	ASTM Method E1562-00 (ASTM, 2006)
42	Polychaete Worm - 10 d - Survival	Polychaete Worm	Annelid (Polychaeta)	BENTHIC-INF	*Neanthes arenaceodentata*	Acute	ASTM Method E1611-00 (ASTM, 2006)
43	Polychaete Worm - 20 d - Survival and Growth	Polychaete Worm	Annelid (Polychaeta)	BENTHIC-INF	*Neanthes arenaceodentata*	Chronic	PSEP (1995)
44	Polychaete Worm - 14 d - Survival and Growth	Polychaete Worm	Annelid (Polychaeta)	BENTHIC-INF	*Polydora cornuta*	Chronic	EPS 1/RM/41 (Environment Canada, 2001)
45	Amphipod - 10 d - Survival	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Eohaustorius estuarius;*	Acute	EPS 1/RM/26 (Environment Canada, 1992c); EPS 1/RM/35 (Environment Canada, 1998b); PSEP (1995)
46	Amphipod - 10 d - Survival	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Rhepoxynius abronius*	Acute	EPS 1/RM/26 (Environment Canada, 1992c); EPS 1/RM/35 (Environment Canada, 1998b); PSEP (1995);
47	Amphipod - 10 d - Survival	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Leptocheirus plumulosus*	Acute	PSEP (1995);
48	Amphipod - 10 d - Survival	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Ampelisca abdita*	Acute	PSEP (1995);
49	Amphipod - 10 d - Survival	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Eohaustorius washingtonianus; Foxiphalus xiximeus; Leptocheirus pinguis; Corophium volutator; Amphiporeia virginiana;*	Acute	EPS 1/RM/26 (Environment Canada, 1992c); EPS 1/RM/35 (Environment Canada, 1998b); PSEP (1995); EPA/600/R-94/025 (USEPA, 1994); EPA/600/R-01/020 (USEPA, 2001)
50	Amphipod - 28 d - Survival and Growth	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Grandidierella japonica*	Chronic	Nipper et al. (1989)
51	Amphipod - 28 d - Survival, Growth and Reproduction	Amphipod	Arthropod (Crustacea)	BENTHIC-INF	*Leptocheirus plumulosus*	Chronic	EPA/600/R-01/020 (USEPA, 2001)
52	Sand Dollar - 10:10 min; 20:20 min - Fertilization	Sand Dollar	Echinoid	BENTHIC-INF	*Dendraster excentricus*	Acute	EPS 1/RM/27 (Environment Canada, 1992b); EPA-600/R-95/136 (USEPA, 1995)
53	Silverside - 7 d - Survival and	Silverside	Fish	FISH	*Menidia beryllina*	Chronic	EPA-821-R-02-014 (USEPA, 2002c)

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
	Growth						
54	Sheepshead Minnow - 96 h - Survival	Sheepshead Minnow	Fish	FISH	*Cyprinodon variegatus*	Acute	EPA-821-R-02-012 (USEPA, 2002a)
55	Sheepshead Minnow - 7 d - Survival and Growth	Sheepshead Minnow	Fish	FISH	*Cyprinodon variegatus*	Chronic	EPA-821-R-02-014 (USEPA, 2002c)
56	Silverside - 96 h - Survival	Silverside	Fish	FISH	*Menidia beryllina, Menidia menidia, Menidia peninsulae*	Acute	EPA-821-R-02-012 (USEPA, 2002a)
57	Threespine Stickleback - 96 h - Survival	Threespine Stickleback	Fish	FISH	*Gasterosteus aculeatus*	Acute	EPS 1/RM/10 (Environment Canada, 1990c)
58	Topsmelt - 7 d - Survival and Growth	Topsmelt	Fish	FISH	*Atherinops affinis*	Acute	EPA-600/R-95/136 (USEPA, 1995)
59	Sanddab - 96 h - Survival	Sanddab	Fish	FISH	*Citharicthys stigmaeus*	Acute	USEPA/USACE (1998)
59	Mysid Shrimp - 7 d - Survival, Growth and Fecundity	Mysid Shrimp	Arthropod (Crustacea)	PELAG-OTHER	*Americamysis bahia*	Chronic	EPA-821-R-02-014 (USEPA, 2002c)
60	Mysid Shrimp - 2 d, 4 d, 10 d - Survival	Mysid Shrimp	Arthropod (Crustacea)	PELAG-OTHER	*Americamysis bahia; Holmesimysis costata; Neomysis americana*	Acute	EPA-821-R-02-012 (USEPA, 2002a) USEPA/USACE (1998)
61	Mysid Shrimp - 7 d - Survival and Growth	Mysid Shrimp	Arthropod (Crustacea)	PELAG-OTHER	*Holmesimysis costata*	Chronic	EPA-600/R-95/136 (USEPA, 1995)
62	Rotifer - 24 h - Survival	Rotifer	Rotifer	PELAG-ZOO	*Brachionus plicatilis*	Acute	ASTM Method E1440-91 (ASTM, 2006)
63	Giant Kelp - 48 h - Germination and Growth	Giant Kelp	Algae	PP-MACRO	*Macrocystis pyrifera*	Chronic	EPA-600/R-95/136 (USEPA, 1995)
64	Red Macroalgae - 48 h - Reproduction	Red Macroalgae	Algae	PP-MACRO	*Champia parvula*	Chronic	EPA-821-R-02-014 (USEPA, 2002c)
65	Diatom - 24 – 96 h - Growth	Diatom	Phytoplankton	PP-PHYTO	*Skeletonema costatum*	Chronic	ASTM Method E1218-04 (ASTM, 2006)
66	Bacterium - Microtox - 5, 15, 30 min - Light inhibition	Bacterium - Microtox	Bacterium		*Vibrio fischeri*	Acute	EPS 1/RM/24 (Environment Canada, 1992d); EPS 1/RM/42 (Environment Canada, 2002); PSEP (1995)
	TERRESTRIAL						
65	Roundworm - 24, 48 h - Survival	Roundworm	Nematoda	INV-	*Caenorhabditis*	Acute	WDOE (2004)

Table 2: List of toxicity tests, classification, and relevant reference protocols.

	Species - Duration - Endpoint(s)	Common Name	Taxonomy / Trophic Level	Organism Type ID	Scientific Name	Classification	Relevant Reference Protocols
				GROUND	*elegans*		
66	Roundworm - 96 h - Survival	Roundworm	Nematoda	INV-GROUND	*Panagrellus redivivus*	Acute	Samoiloff (1990)
67	Earthworm - 56, 63 d - Survival, Growth, and Reproduction	Earthworm	Annelid (Oligochaeta)	INV-GROUND	*Eisenia andrei*	Chronic	EPS 1/RM/43 (Environment Canada, 2004)
68	Earthworm - 48, 72 h - Avoidance	Earthworm	Annelid (Oligochaeta)	INV-GROUND	*Eisenia andrei, Eisenia fetida or Lumbricus terrestris*	Acute	EPS 1/RM/43 (Environment Canada, 2004)
69	Earthworm - 14 d - Survival	Earthworm	Annelid (Oligochaeta)	INV-GROUND	*Eisenia andrei, Eisenia fetida or Lumbricus terrestris*	Acute	EPS 1/RM/43 (Environment Canada, 2004)
70	Potworm - 14 – 42 d - Survival and Growth	Potworm	Annelid (Oligochaeta)	INV-GROUND	*Enchytraeus albidus*	Chronic	ASTM (2008)
71	Predatory Mite - 14 d - Survival and Reproduction	Predatory Mite	Arthropod (Arachnida)	INV-GROUND	*Hypoaspis (Geolaelaps) aculeifer*	Chronic	OECD Guideline 226 (OECD, 2008)
72	Springtail - 21, 28 d - Survival and Reproduction	Springtail	Arthropoda (Insecta)	INV-GROUND	*Orthonychiurus folsomi, Folsomia candida, F. fimetaria*	Chronic	EPS 1/RM/47 (Environment Canada, 2007d)
73	Terrestrial Plants (various) - 14 d, 21 d - Germination, Survival and Growth	Terrestrial Plants (various)	Plant	PP-PLANT	*Various*	Chronic	EPS 1/RM/45 (Environment Canada, 2004)

Table 2a: Instructions for the use of Table 2

Step	Instructions
1 - Choose Ecosystem Type	All tests are primarily organized according to broad ecosystem type, in the order of "Freshwater Aquatic", "Marine/Estuarine Aquatic", and "Terrestrial"; the relevant ecosystem type is determined in the Problem Formulation stage.
2 - Choose Receptor Type (Guild)	All tests are secondarily organized according to feeding guild (*i.e.*, broad organism type and functional group), with entries in the "Receptor Type" column sorted in alphabetical order; the relevant feeding guilds are determined during the Problem Formulation stage.
3 - Choose Test Organism	All organisms with a common ecosystem and feeding guild are grouped together in adjacent rows. For some tests, there is a single test protocol that applies to the organism. In other cases, there are multiple test protocols applicable to the row entry (*e.g.*, sediment tests for Hyalella may be conducted using Environment Canada, ASTM, or USEPA test protocols). Where the latter occurs, details are provided in the columns marked "Relevant Reference Protocols"

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
	FRESHWATER																								
1	Frog - 96 h - Survival, Development and Growth	AMPHIB	*Xenopus laevis*	F	X				+	+		+	-	+			NA	+	+	+	+	+	+	$$$	
2	Frog - 21 d - Survival, Development and Growth	AMPHIB	*Xenopus laevis*	F	X					+		++	-	+			NA	+		+	+	+	+	$$$$	
3	Amphibian - 10 d - Survival and Growth	AMPHIB	*Rana pipiens; R. clamitans; R. sylvatica; B. americanus*	F	X	X			+	+		+	++	+				+			-	+	-	$$$	Y
4	Amphipod - 10, 14 d - Survival and Growth	BENTHIC-EPI	*Hyalella azteca*	F		X			++	+	+	+	++	+	+		T	S	+	+	+	+	++	$$$	Y
5	Amphipod - 42 d - Survival, Growth and Reproduction	BENTHIC-EPI	*Hyalella azteca*	F		X			+	+	-	++	++	+	+		T	S	+	+	+	+	-	$$$$	Y
6	Mayfly - 10 d - Survival	BENTHIC-EPI	*Hexagenia limbata*	F		X			+	+		-	++	+		+	+	+			-	+	+	$$$	
7	Mayfly - 21 d - Survival, Growth and Moulting Frequency	BENTHIC-EPI	*Hexagenia limbata*	F		X				+		++	++	+			+	+			-	+	+	$$$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
8	Fatmucket Mussel - 48 - 96 h - Survival	BENTHIC-EPI	*Lampsilis siliquoidea*	F	X							-	+	+			NA			+	-	+	-	$$	
9	Fatmucket Mussel - 28 d - Survival and Growth	BENTHIC-EPI	*Lampsilis siliquoidea*	F		X						+	+	+						+	-	+	-	$$$$	
10	Mussels - 10 d - Survival	BENTHIC-EPI	*Anodonta imbecillis*	F		X						-		+								+	-	$$$	
11	Cladoceran - 96 h - Survival, Growth and Development	BENTHIC-EPI	*Chydorus sphaericus*	F		X			-			++	++	-	+			+			+	+	-	$	
12	Oligochaete - 10 d - Survival	BENTHIC-INF	*Pristina leidyi*	F		X						-		+								+	-	$$	
13	Sludgeworm - 10 d - Survival	BENTHIC-INF	*Tubifex tubifex*	F		X			+	+	+	-	++	+	+		+	+	+	+	+	+	+	$$	
14	Sludgeworm - 28 d - Survival and Reproduction	BENTHIC-INF	*Tubifex tubifex*	F		X				+		+	++	+	+		+	+	+	+	+	+	+	$$$$	
15	Amphipod - 28 d - Survival and Behaviour	BENTHIC-INF	*Diporeia sp.*	F		X			+			++	+	+	+							+	-	$$$$	
16	Midge - 10 d - Survival and Growth	BENTHIC-INF	*Chironomus dilutus*	F		X			++	+	+	+	++	+	+		T	+	+	+	+	+	++	$$$	Y
17	Midge - 30 d; 50 - 65 d - Survival, Growth and Adult	BENTHIC-INF	*Chironomus riparius*	F		X			+	+	+	++	++	+	+		T	+	+	+	+	+	+	$$$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
	Emergence, Reproduction																								
18	Blackworm - 10 d - Survival	BENTHIC-INF	*Lumbriculus variegatus*	F		X			+	+		-	++	+	+		+	+	+	+	+	+	+	$$	
19	Blackworm - 28 d - Survival, Reproduction and Growth	BENTHIC-INF	*Lumbriculus variegatus*	F		X				+		++	++	+	+		+	+	+	+	+	+	+	$$$$	Y
20	Fathead Minnow - 96 h - Survival	FISH	*Pimephales promelas*	F	X		X		++	+	+	-	+	-	+	+	NA	+	+	+	+	+	++	$$	
21	Fathead Minnow - 7 d - Survival and Growth	FISH	*Pimephales promelas*	F	X				++	+	+	+	+	-	+	+	NA	+	+	+	+	+	++	$$$	
22	Rainbow Trout - 96 h - Survival	FISH	*Oncorhynchus mykiss*	F	X				++	+	NA	-	++	+	+	+	NA	T	+	+	+	+	++	$	Y
23	Rainbow Trout - 7 d - Embryo Viability	FISH	*Oncorhynchus mykiss*	F	X				+	+	+	+	++	+	+		NA	+	-	+	-	-	++	$$$$	Y
24	Rainbow Trout - ~30 d - Alevin Viability, Hatching and Deformity	FISH	*Oncorhynchus mykiss*	F	X				+	+		++	++	+	+		NA	+	-	+	-	-	+	$$$$	Y
25	Rainbow Trout - ~70 d - Alevin Viability, Hatching and	FISH	*Oncorhynchus mykiss*	F	X					+		++	++	+	+		NA	+	-	+	-	-	+	$$$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
	Deformity, Fry Survival and Behaviour																								
26	Waterflea - 7 d, 21 d - Survival and Reproduction	PELAG-ZOO	*Daphnia magna*	F	X	X			++	+	+	+	++	-	+	+	NA	T	-	+	+	+	++	$$$	
27	Waterflea - 48 h - Survival	PELAG-ZOO	*Ceriodaphnia dubia*	F	X				++	-	+	-	++	-	+		NA	S	-	+	+	+	++	$	
28	Waterflea - 7 d - Survival and Reproduction	PELAG-ZOO	*Ceriodaphnia dubia*	F	X	X			++	+	+	+	++	-	+	+	NA	S	-	+	+	+	++	$$	Y
29	Waterflea - 48 h - Survival	PELAG-ZOO	*Daphnia magna, Daphnia pulex*	F	X				++	+	+	+	++	-	+	+	NA	T	-	+	+	+	++	$	
30	Rotifer - 24 h - Survival	PELAG-ZOO	*Brachionus calyciflorus*	F	X				+			-	++	-			NA		+	+	+	+	++	$	
31	Duckweed - 7 d - Growth	PP-MACRO	*Lemna minor*	F	X				+	-	+	+	++	+	+	+	NA	+	-	+	+	+	++	$$	Y
32	Domestic Rice - 14 d - Growth, Chlorophyll Content	PP-MACRO	*Oryza sativa*	F	X	X						+	++	+				+		+	+	+	-	$$$	
33	Phytoplankton - 72 h - Cell Yield	PP-PHYTO	*Pseudokirchneriella subcapitata*	F	X				+	-	+	+	++	-	+		NA	+	-	+	+	+	++	$$	Y

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
	MARINE/ESTUARINE																								
34	Mussels - 48 h - Larval Development and Survival	BENTHIC-EPI	*Mytilus* sp.	M	X	X	X		++	-	+	+	+	-	+	+	S	T	+	+	+	+	++	$$	Y
35	Oyster - 48 h - Larval Development and Survival	BENTHIC-EPI	*Crassostrea gigas; C. virginica*	M	X	X	X		+	-	+	+	+	-	+	+	S	T	+	+	-	+	++	$$	Y
36	Red Abalone - 48 h - Larval Development and Survival	BENTHIC-EPI	*Haliotis rufescens*	M	X					+	+	+	-	-	+		+			+		+	+	$$	
37	Sea Urchin - 48 – 96 h - Larval Development and Survival	BENTHIC-EPI	*Strongylocentrotus droebachiensis; Arbacia punctulata*	M	X	X	X		+	-	+	+	+	-	+	+	S	T	+	+	-	+	++	$$	Y
38	Sea Urchin - 10:10 min; 20:20 min; 60:20 min - Fertilization	BENTHIC-EPI	*Strongylocentrotus purpuratus; Arbacia punctulata*	M	X		X		++	+	+	+	+	-	-	+	+	S	+	+	-	+	++	$$	Y
39	Sea Urchin - 10:10 min; 20:20 min; 60:20 min - Fertilization	BENTHIC-EPI	*Strongylocentrotus droebachiensis;*	M	X		X		+	+	+	+	+	-	-	+	+	S	+	+	-	+	+	$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
			Lytechinus pictus																						
40	Clam - 7 d - Survival and Growth	BENTHIC-INF	*Mulinia lateralis*	M		X						+		+									-	$$	
41	Polychaete Worm - 20-28 d - Survival, Reproduction and Growth	BENTHIC-INF	*Capitella capitata*	M		X					+	++	++	+			S		+			+	-	$$$$	Y
42	Polychaete Worm - 10 d - Survival	BENTHIC-INF	*Neanthes arenaceodentata*	M		X				+	+	-	+	+	+		T	S	+	+	+	+	-	$$	Y
43	Polychaete Worm - 20 d - Survival and Growth	BENTHIC-INF	*Neanthes arenaceodentata*	M		X			+	+	+	+	+	+	+	+	T	S	+	+	+	+	++	$$$	Y
44	Polychaete Worm - 14 d - Survival and Growth	BENTHIC-INF	*Polydora cornuta*	M		X			-	+	+	+	+	+	+		+	+	+	+	-	+	-	$$$	Y
45	Amphipod - 10 d - Survival	BENTHIC-INF	*Eohaustorius estuarius*	M		X			++	+	+	+	+	-	+	+	T	T	+	+	+	+	++	$$	Y
46	Amphipod - 10 d - Survival	BENTHIC-INF	*Rhepoxynius abronius*	M		X			++	+	+	+	+	-	+	+	S	S	+	+	+	+	++	$$	Y
47	Amphipod - 10 d - Survival	BENTHIC-INF	*Leptocheirus plumulosus*	M		X			++	+	+	+	+	-	+	+	S	T	+	+	+	+	++	$$	Y

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
48	Amphipod - 10 d - Survival	BENTHIC-INF	*Ampelisca abdita*	M		X			++	+	+	+	+	-	+	+	S	S	+	+	+	+	++	$$	Y
49	Amphipod - 10 d - Survival	BENTHIC-INF	*Eohaustorius washingtonianus; Foxiphalus xiximeus; Leptocheirus pinguis; Corophium volutator; Amphiporeia virginiana*	M		X			-	+	+	+	+	-					+			+	+	$$	
50	Amphipod - 28 d - Survival and Growth	BENTHIC-INF	*Grandidierella japonica*	M		X				+				-	+			+	+			+	-	$$$$	
51	Amphipod - 28 d - Survival, Growth and Reproduction	BENTHIC-INF	*Leptocheirus plumulosus*	M		X			+	+	-	+	+	+	+		T	T	+	+	+	+	+	$$$$	Y
52	Sand Dollar - 10:10 min; 20:20 min - Fertilization	BENTHIC-INF	*Dendraster excentricus*	M	X		X		++	+	+	+	+	-	+		+	T	-	+	-	+	++	$$	Y
53	Silverside - 7 d - Survival and Growth	FISH	*Menidia beryllina*	M	X				++	+	+	+	+	-	+		NA	T	+	+	+	+	++	$$$	Y
54	Sheepshead Minnow - 96 h - Survival	FISH	*Cyprinodon variegatus*	M	X				++	+	+	+	-	-	+		NA	+	+	+	+	+	++	$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
55	Sheepshead Minnow - 7 d - Survival and Growth	FISH	*Cyprinodon variegatus*	M	X				++	+	+	+	-	-	+		NA	+	+	+	+	+	++	$$$	
56	Silverside - 96 h - Survival	FISH	*Menidia beryllina, Menidia menidia, Menidia peninsulae*	M	X				++	+	+	-	+	-	+		NA	T	+	+	+	+	++	$	Y
57	Threespine Stickleback - 96 h - Survival	FISH	*Gasterosteus aculeatus*	M	X				+	+	NA	-	++	+	+		NA	T	+	+	-	+	+	$	Y
58	Topsmelt - 7 d - Survival and Growth	FISH	*Atherinops affinis*	M	X				++	+	+	+	+	-	+		NA	T	+	+	+	+	++	$$$	
59	Sanddab - 96 h - Survival	FISH	*Citharicthys stigmaeus*	M	X		X		+	+	+	-	+	+			NA	+	+	+	+	-	+	$$	
60	Mysid Shrimp - 7 d - Survival, Growth and Fecundity	PELAG-OTHER	*Americamysis bahia*	M	X				+	+		++	+	-	+		NA	T	+	+	+	+	++	$$$	Y
61	Mysid Shrimp - 2 d, 4 d, 10 d - Survival	PELAG-OTHER	*Americamysis bahia; Holmesimysis costata; Neomysis americana*	M	X	X	X		++	+	+	-	+	-	+		NA	T	+	+	+	+	++	$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
62	Mysid Shrimp - 7 d - Survival and Growth	PELAG-OTHER	*Holmesimysis costata*	M	X					+		+	+	-	+		NA	T	+	+	+	+	++	$$$	
63	Rotifer - 24 h - Survival	PELAG-ZOO	*Brachionus plicatilis*	M	X				++			-	++	-			NA		+	+	+	+	++	$	
64	Giant Kelp - 48 h - Germination and Growth	PP-MACRO	*Macrocystis pyrifera*	M	X				+	-	+	+	+	-	+	-	S	+	-	+	+	+	+	$$$$	Y
65	Red Macroalgae - 48 h - Reproduction	PP-MACRO	*Champia parvula*	M	X					-	-	+	-	+			S	+	-	+	+	+	+	$$$$	Y
66	Diatom - 24 – 96 h - Growth	PP-PHYTO	*Skeletonema costatum*	M	X				+	-		+	++	NA			NA		-		+	+	-	$$	
67	Bacterium - Microtox - 5, 15, 30 min - Light inhibition		*Vibrio fischeri*		X	X	X		++	-	+	-	NA	NA	+	+	T	T	+	+	+	+	++	$	Y
	TERRESTRIAL																								
68	Roundworm - 24, 48 h - Survival	INV-GROUND	*Caenorhabditis elegans*	T				X				-	++	+	NA	NA			+			+	+	$	
69	Roundworm - 96 h - Survival	INV-GROUND	*Panagrellus redivivus*	T				X				-	++	+	NA	NA						+	-	$	
70	Earthworm - 56, 63 d - Survival, Growth, and Reproduction	INV-GROUND	*Eisenia andrei*	T				X		+		++	++	+	NA	NA		NA	+	+	+	+	+	$$$$	Y
71	Earthworm - 48, 72 h - Avoidance	INV-GROUND	*Eisenia andrei, Eisenia*	T				X		+		-	++	+	NA	NA		NA	+	+	+	+	+	$$	

Table 3: Summary of factors to consider in selection of toxicity tests for risk assessment purposes

				Media					Utility for Risk Assessment						Organism Tolerance				Logistics and Planning Factors						
Test ID	Species - Duration - Endpoint(s)	Receptor Type	Scientific Name	Environment	Water Column	Whole Sediment	Sediment Derivative	Soil	Availability of Toxicity Data	Relevance of Control	Statistical Power	Multiple / Chronic Endpoints Available?	Geographic Suitability	Tissue Production	Ammonia Tolerance	Sulphide Tolerance	Substrate Tolerance	Salinity / Water Hardness	Laboratory Handling	Organism Source	Seasonal Availability	Sample Volumes	Availability of Method	Cost per Sample	Caveats
			fetida or Lumbricus terrestris																						
72	Earthworm - 14 d - Survival	INV-GROUND	*Eisenia andrei, Eisenia fetida, Lumbricus terrestris*	T				X	+	+		+	++	+	NA	NA		NA	+	+	+	+	++	$$$	
73	Potworm - 14 – 42 d - Survival and Growth	INV-GROUND	*Enchytraeus albidus*	T				X		+		+	++	+	NA	NA			+			+	+	$$	
74	Predatory Mite - 14 d - Survival and Reproduction	INV-GROUND	*Hypoaspis (Geolaelaps) aculeifer*	T				X				+	+	+	NA	NA							-	$$$	
75	Springtail - 21, 28 d - Survival and Reproduction	INV-GROUND	*Orthonychiurus folsomi, Folsomia candida, F. fimetaria*	T				X	+	+		+	++	-	NA	NA			+	+	+	+	+	$$$	
76	Terrestrial Plants (various) - 14 d, 21 d - Germination, Survival and Growth	PP-PLANT	*Various*	T				X	++	+	+	++	++	+	NA	NA			+	+	+	+	++	$$	Y

Table 3a: Instructions for the use of Table 3

Step	Instructions
1 - Choose Ecosystem Type	All tests are primarily organized according to broad ecosystem type, in the order of "Freshwater Aquatic (F)", "Marine/Estuarine Aquatic (M)", and "Terrestrial (T)"; the relevant ecosystem type is determined in the Problem Formulation stage. The codes F, M, and T, are provided in the Column heading labelled "Environment".
2 - Choose Receptor Type (Guild)	All tests are secondarily organized according to feeding guild (*i.e.*, broad organism type and functional group), with entries in the "Receptor Type" column sorted in alphabetical order; the relevant feeding guilds are determined during the Problem Formulation stage.
3 - Choose Test Organism	All organisms with a common ecosystem and feeding guild are grouped together in adjacent rows. For some tests, there is a unique taxon and single test duration relevant to the test type. In other cases, there are multiple species or durations that apply to the general test protocol. Where the latter occurs, details are provided in the columns marked "Scientific Name" and "Species - Duration - Endpoints"
4 - Interpret Factor Scoring	For the chosen test species/protocol, the assessment of each factor is coded using symbols (++, +, -, blank) that summarize the significance of each factor in test selection. Interpretation of these symbols is provided below.
++	Significance: The test is considered highly advantageous for this attribute, or abundant information is readily available to evaluate the sensitivity of the test performance to the attribute;
+	Significance: Information is considered somewhat advantageous for this attribute, or information is available to evaluate the sensitivity of the test performance to the attribute;
Blank	Significance: The attribute is unknown with respect to test reliability and/or utility;
-	Significance: Attribute is considered to be a constraint to the test reliability and/or utility; practitioners should proceed with caution and determine whether the limitation applies to their situation, and assess the margin for error.
5 - Consider Scoring Rationale	For more information of the decision rules used to assign the factor scoring codes, consult the decision rules appended below this table [Notes - Part 2]. The practitioner should consult the primary literature regarding these attributes in order to inform decisions of test selection, suitability, and constraints.
6 - Note Applicable Caveats	For the chosen test species/protocol, an entry of "Y" (yes) under the rightmost (Caveats) column indicates that additional information is available for the specified test. Consult the [Notes - Part 3] material below, using the number of the test (leftmost column) to reference the relevant information.

Table 3b: Rating Systems for Factors of Considerations listed in Table 3 (more information in Appendix A)

Factor for Consideration	Rating System (see Appendix A for more information)
Availability of Toxicity Data	(++) if data are known to be available for most COPCs; (+) if data are known to be available for some COPCs; (-) if data are known to be absent for most COPCs; (blank) if unknown.
Relevance of Control	(+) toxicity test uses a field-collected natural negative control or suitable surrogate; (-) toxicity test does not use a field-collected natural negative control, the negative control uses a different medium, or the composition of the negative control could affect interpretation of results; (blank) if unknown.
Statistical Power	(+) toxicity test is known to have relatively low inter-replicate variability; (-) toxicity test is known to have relatively high inter-replicate variability; (blank) if unknown. Tests without replicates (e.g., 96-h rainbow trout LC_{50}) are scored as not applicable (NA) in Table 3.
Multiple / Chronic Endpoints Available?	(++) test includes multiple, long-term chronic endpoints; (+) test includes at least one chronic, long-term endpoint or a surrogate chronic endpoint (*e.g.*, early life stage); (-) test does not include a chronic or surrogate chronic endpoint.
Geographic Suitability	(++) test species is aligned with resident organisms found at most locations in Canada; (+) test species is aligned with resident organisms for some locations in Canada; (-) test species is not well-aligned with receptor selection for Canadian risk assessments.
Tissue Production	(+) test species likely has sufficient mass to permit tissue chemistry analyses; (-) test species does not have sufficient mass to permit tissue chemistry.
Ammonia Tolerance	(+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.
Sulphide Tolerance	(+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.
Substrate Tolerance	refers to Grainsize tolerance : (+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.
Salinity / Water Hardness	(+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.

Table 3b (continued): Rating Systems for Factors of Considerations listed in Table 3 (more information in Appendix A)

Factor for Consideration	Rating System (see Appendix A for more information)
Laboratory Handling	(+) known to be relatively tolerant to handling/culturing stress; (-) known to be relatively sensitive to handling/culturing stress; (blank) if unknown.
Organism Source	(+) reliable sources of organisms are available, whether from in-house or commercial cultures, or appropriate field collection locations; (blank) if unknown.
Seasonal Availability	(+) generally available on a year-round (or nearly year-round) basis; (-) seasonal constraints are known; (blank) if unknown.
Sample Volumes	(+) sample volumes do not generally place a significant constraint on sampling programs; (-) sample volumes can present a significant constraint on sampling programs; (blank) if unknown.
Availability of Method	(++) test is routinely offered by most laboratories; (+) test may have limited availability or require specialized effort to complete; (-) test has little or no availability.
Cost	($) costs are typically less than $500/sample; ($$) costs are $500 - 1000 per sample; ($$$) costs are $1000 - 1750/sample; ($$$$) costs typically exceed $1750/sample.

Table 3c: Explanation of Test Caveats listed in Table 3

Test ID	Caveat or Description
3	In general, amphibian toxicity tests can be conducted with a variety of field-collected species (subject to collection permit restrictions). The test duration and endpoints can be modified to meet study-specific objectives.
4	*H. azteca* is relatively tolerant to salinity if properly acclimated. However, this species is relatively intolerant to water samples with major ion inbalances or extremely soft or hard water. Note that *H. azteca* is epibenthic and does not burrow into sediment.
5	The chronic *H. azteca* toxicity test is not widely offered because of high failure rates in negative controls. Note that H. azteca is epibenthic and does not burrow into sediment.
16	The survival and growth of *Chironomus* sp. can be reduced in samples with extremely low organic carbon content because the organism does not have enough material to build its cocoon.

Table 3c (continued): Explanation of Test Caveats listed in Table 3

Test ID	Caveat or Description
19	*L. variegatus* is also used in bioaccumulation testing.
22	Rainbow trout testing can be conducted on estuarine samples provided that organisms are properly acclimated and provided that salinity is maintained at or below 10 ppt (Environment Canada 1990a, 2000a). See test protocol for information. 96-h LC50 tests can also be conducted with other salmonids (e.g., brown trout, coho salmon, chinook salmon) depending on hatchery availability.
23	Newly-fertilized eggs are highly sensitive to physical shock until water hardening has been completed. Tests are conducted in the dark (or low-light) conditions. Embryos are only available for short periods of the year.
24	Larval fish deformity assessments require expertise to complete properly.
28	*C. dubia* is sensitive to pH, salinity and extremely hard or soft water. There is considerable information about the biological tolerances of this species in the protocol documents as well as the scientific literature.
31	This test can be modified to work with sediment derivatives (*e.g.*, elutriates) as well
33	Test media contains EDTA which can reduce the bioavailablity of selected metals. High turbidity in the sample can reduce photosynthesis and thus impact growth rates. This species is also referred to as *Selenastrum capricornutum* in older literature.
34	*Mytilus* sp. is sensitive to large amounts of suspended particulates that can smother developing organisms. The elutriate version of the test (PSEP, 1995) specifies a 4-hour settling time which is routinely modified to 24-h to reduce the effect of entrainment (based on consultation with Canadian regulators). This test uses a water-only negative control, and therefore, a reference sediment should be considered for the elutriate version of this test (in addition to the water only control). *Mytilus* sp. is also sensitive to ammonia concentrations.
35	*Crassostrea* sp. is sensitive to large amounts of suspended particulates that can smother developing organisms. This test uses a water-only negative control, and therefore, a reference sediment should be considered for the elutriate version of this test. *Crassostrea* sp. is relatively insensitive to ammonia concentrations.
37, 38	The larval development tests with sea urchins tend to be sensitive to large amounts of suspended particulates that can smother developing organisms. This test uses a water-only negative control, and therefore, a reference sediment should be considered for the elutriate version of this test. Sea urchins may be sensitive to high ammonia concentrations in porewater as well as salinity outside a range of 28 - 32 ppt. Literature indicates that the multi-day exposures to sea urchins are particularly senstive to ammonia toxicity.
41	*C. capitella* prefers samples with high organic carbon content
42, 43	*N. arenaceodentata* is only available from a highly inbred population maintained by a single North American supplier.
44	Commercial cultures for *P. cornuta* are not currently available.

Table 3c (continued): Explanation of Test Caveats listed in Table 3

Test ID	Caveat or Description
45	*E. estuarius* is tolerant to wide range of sample salinities and sediment grain size.
46	*R. abronius* prefers a coarse-grained sample (i.e., dominated by sands) and cannot tolerate a wide range of salinities.
47	*L. plumulosus* prefers samples without large amounts of coarse material but can tolerant to a wide range of salinities.
48	*A. abdita* prefers fine-grained samples with relatively high organic carbon contents. This species is not tolerant of a wide range of salinities.
51	High variability in the growth and reproduction endpoints have been reported for this test.
52	Sand dollar gametes degrade quickly, and therefore, experience is required to initiate testing efficiently.
53	This species is tolerant of a wide range of sample salinities.
56	This species is tolerant of a wide range of sample salinities. Inland silversides are temperate (testing done at 25C), and the species is not native to Canada, so relevance to Canadian sites should be evaluated.
57	Organisms may not be not available during late-spring / summer when adult stickleback move into freshwater for spawning. Species tolerates a very wide range of salinities.
60	Mysid shrimp are relatively tolerant of a wide range of sample salinities
64	The giant kelp test is sensitive to samples with high amounts of particulate material because the settling particles prevent the gametophytes from properly attaching to the bottom of the test container. High turbidity can interfere with the ability to count gametophytes under the microscope.
65	High variability in the toxicological endpoint has been reported, which may be associated with the high handling stress. Additionally, this test uses a species that is not present in Canadian waters.
67	Bacteria are rarely included as a receptor group in an ecological risk assessment, and therefore, this test may not be well-aligned with common assessment endpoints.
70	*E. foetida* is the most common earthworm test species. It is a non-native species that inhabits composts. *E. foetida* has colonized natural soils in proximity to urbanized areas but may have little ecological relevance for wildland applications elsewhere in Canada
76	Consider selection of negative control carefully (e.g., clean natural soil or artificial soils). Seeds are readily available from commercial suppliers. Test protocols include 15+ different species, including grasses, legumes and garden produce. Test is readily adaptable to other species.

List of Acronyms

AEL	Acceptable Effect Level
ASTM	American Society for Testing and Materials
CABIN	Canadian Aquatic Biomonitoring Network
CALA	Canadian Association for Laboratory Accreditation
CCAC	Canadian Council of Animal Care
CCME	Canadian Council of Ministers of the Environment
COPC	Contaminant of Potential Concern
DSL	Domestic Substances List
EC	Effect concentration
ERA	Ecological Risk Assessment
FCSAP	Federal Contaminated Sites Action Plan
IC	Inhibitory concentration
LC	Lethal concentration
LOAEL	Lowest-observed-adverse-effect level
MDD	Minimum detectable difference
MSD	Minimum significant difference
NA	Not applicable
NOAEL	No-observed-adverse-effect level
OECD	Organization for Economic Cooperation and Development
PAHs	Polycyclic aromatic hydrocarbons
PCBs	Polychlorinated biphenyls
PSL	Priority Substances List
QA/QC	Quality assurance/quality control
ROC	Receptor of Concern
S	Sensitive to parameter under evaluation
SAB	Science Advisory Board of British Columbia
SIDS	Screening Information Data Set
SQO	Sediment quality objective
T	Tolerant of parameter under evaluation
TBT	Tributyltin

TIE	Toxicity identification evaluation
TRV	Toxicity reference value
USEPA	United States Environmental Protection Agency
UV	Ultraviolet
WOE	Weight-of-Evidence

Glossary

Acceptable effect level – The magnitude of effects that would be acceptable for a specific measurement endpoint.

Acute toxicity – A discernable adverse effect (lethal or sublethal) induced in test organisms with a short period of exposure in relation to the life span of the test organism (defined as less than 10% of organism's life span by Environment Canada).

Antagonistic toxicity – A phenomenon in which the toxicity of a mixture is less than the toxicity of the most toxic chemical when present singly at the same concentration. This is distinct from "sub-additive toxicity" as defined below.

Assessment endpoint – An explicit expression of the environmental value to be protected. An assessment endpoint must include a receptor (or receptor group – *i.e.*, a 'thing' to be protected) and a specific property of that receptor. For example, if the receptor is a fish community, endpoint properties could include the number of species, the frequency of deformities, the trophic structure, *etc.*

Bioaccumulation – A process by which chemicals are taken up by organisms directly from water as well as through exposure through other routes, such as consumption of food and soil/sediment containing the chemicals.

Bioassay – A test in which living organisms are used to estimate the strength or potency of a material, usually a medical drug. Bioassay has also been used to describe environmental toxicity tests (usually in older literature), but "toxicity test" is now the recommended term.

Bioavailable – The fraction of the total chemical in the surrounding environment which can be taken up by organisms. The environment may include surface water, interstitial water, soil, sediment, suspended particles, and food items.

Biomagnification – A phenomenon observed as the result of bioaccumulation by which tissue concentrations increase as the chemical passes up through the food chain (*i.e.,* two or more trophic levels.

Chronic toxicity – A discernable adverse effect (lethal or sublethal) induced in test organisms during relatively long period of exposure, usually a substantial proportion of the life span of the organism (*i.e.,* defined as 10% or more of lifespan by Environment Canada). Chronic refers to the duration of exposure only; in some situation the definition has been confused with "sublethal". Chronic toxicity tests may have both lethal and sublethal endpoints. Some short-term (acute) tests that evaluate sensitive life stages are useful for making inferences about potential long-term responses, but are not chronic tests in the strict sense as defined by Environment Canada.

Confounding factor – Any modifying variable in an experimental design that is not controlled for and that is influencing the experimental results in a non-random manner.

Contaminants of Potential Concern – Contaminants that have been selected for evaluation in the ERA. The process used to select COPCs is not covered in this module. A contaminant is any undesirable agent, substance or material present in sediment, soil or water.

Ecological Representation – refers to the ecosystem function represented (*e.g.*, feeding guild, trophic status).

Ecological risk assessment – Ecological risk assessment is the process that evaluates the likelihood and/or magnitude of potential adverse ecological effects (current or future) as a result of exposure to one or more stressors (in the context of contaminated sites, the stressors are usually chemical). A risk cannot exist unless: (1) the stressor has an inherent ability to cause adverse effects, and (2) it is coincident with or in contact with an organism long enough and at sufficient intensity to elicit the identified adverse effect(s).

Elutriate – An aqueous solution obtained after adding water to a solid substance (*e.g.,* sediment, soil, tailings, drilling mud, dredge spoil), shaking the mixture and then centrifuging or filtering it and decanting the supernatant.

Exposure pathways – The routes of exposure from environmental media (soil, water, air and/or aquatic sediment) to the receptors of concern.

In situ *testing* – Refers to "on site", and used to distinguish work conducted "in the field" from work done in the laboratory. *In situ* toxicity testing involves the exposure of test organism to the contaminated media under field conditions. Exposure conditions are partly controlled through application of an experimental design, but environmental variables such as light, pH, temperature, *etc.* are not controlled in the same way as their laboratory counterparts.

Interstitial water – See *pore water*.

Measurement endpoint – A measurement endpoint is a parameter that measures or describes an effect on a test organism, or that measures or describes a change in an attribute of an assessment endpoint or its surrogate in response to a stressor to which it is exposed.

Negative controls – A field-collected or artificially prepared substrate (water, sediment or soil, depending on the test) of known physicochemical composition and consistent quality. The negative control must not contain concentrations of contaminants that affect the test organism in any way, and the physical characteristics should be within the tolerance thresholds of the organism. Negative controls provide a basis for interpreting toxicity data and are used to monitor the health of test organisms, the relative sensitivity of test organisms over time and the "performance" of laboratories.

Pore water – The water that occupies the spaces (interstices) between sediment particles.

Positive controls (also reference toxicant tests) – Reference toxicants are chemicals used to measure the sensitivity of the test organisms in order to establish confidence in the toxicity data obtained for field-collected samples. In most instances, the reference toxicant test involves a range of concentrations and calculation of a point-estimate value (*e.g.,* LC_{50}) that is compared to previous reference toxicant testing conducted by the same laboratory for the same organism.

Potentiation – A phenomenon (more-than-additive toxicity) in which the toxicity of a mixture of chemicals is greater than that which would be expected from a simple summation of the toxicities of the individual chemicals present in the mixture (*i.e.*, greater than expected toxicity when mixed). This is a somewhat different concept from synergistic toxicity (defined below).

Problem Formulation – The problem formulation is a planning and screening process that defines the feasibility, scope, and objectives for the risk assessment. This process includes examination of scientific data and data needs, regulatory issues, and site-specific factors.

Quality Assurance/Quality Control – Quality assurance refers to the management and technical practices designed to ensure an end product (in this case, toxicity tests) of known or reliable quality. Quality control refers to the techniques and procedures used to measure and assess data quality and the remedial actions to be taken when the data quality objectives are not met.

Receptor of Concern – Any non-human individual organism, species, population, community, habitat or ecosystem that is potentially exposed to contaminants of potential concern and that is considered in the ecological risk assessment.

Reference sample – Also called "background sample". A field collected sample of sediment, soil or water collected from a site thought to be relatively free of contamination and included in the toxicity testing program because of its geochemical similarity (*e.g.,* particle size, hardness, organic content) to the samples collected from the contaminated site. Reference samples are used to assist in the interpretation of toxicity data (in addition to the negative controls).

Relevance check – Entails a review at an overview level to make sure that choices made considering the detailed factors outlined in Steps 2 through 6 of the module remain consistent and reasonable when applied against the broader objective s outlined in Step 1. A relevance check is equivalent to the process conveyed by the idiomatic expression "seeing the forest through the trees" and guards against excessive reliance on reductionist approaches to scientific decision-making.

Statistical power – Loosely defined, statistical power refers to the probability of correctly concluding that there is a difference between the variables being tested. Formally, statistical power is the probability of rejecting the null hypothesis when it is in fact false and should be rejected. Power cannot be directly set by the investigator before doing a toxicity test but can be strengthened by adding more organisms, more replicates, *etc.*, and can be evaluated at the end of the test. In the context of formal hypothesis testing, statistical power represents (1-β), where β represents the probability of a Type II error (*i.e.*, probability of incorrectly concluding no difference when in fact a difference does exist). A common convention, including the default practice specified by the USEPA (2009), is to require power to exceed 0.8 in order to provide "reasonable test of a hypothesis". However, there is no minimum power for a result considered as part of a weight-of-evidence analysis, because power is in part a function of the effect size of interest and the acceptable level of uncertainty.

Sub-additive toxicity – A phenomenon in which the toxicity of a mixture of chemicals is less than that which would be expected from a simple summation of the toxicities of the individual chemicals present in the mixture. In other words, the mixture toxicity is greater than that of any single chemical in the mixture, but less than expected based on model prediction.

Sublethal toxicity – Refers to effects that are detrimental to the organism, but below the level that directly causes death (*e.g.,* growth, reproduction)

Synergistic toxicity – A phenomenon in which a synergist (a substance that is nontoxic singly but increases the toxicity of other toxicants) acts to enhance toxicity.

Toxicity identification evaluation (TIE) – Consists of side-by-side toxicity testing using manipulated and non-manipulated samples using a systematic sample pretreatment (*e.g.*, pH adjustment, filtration, aeration, or addition of binding agents) followed by tests for toxicity. 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.

Toxicity test – A study designed to determine whether exposure of test organisms to test media causes an adverse effect (either lethal or sublethal) to those organisms. A toxicity test usually measures either (a) the proportions of organisms affected (quantal) or (b) the degree of effect shown (graded or quantitative), after exposure to a specific test substance or test medium under controlled conditions.

Toxicity test battery – The use of multiple toxicity tests on the same samples. Data from these tests are often interpreted using weight-of-evidence approaches.

Toxicity test endpoints – Endpoints are the statistic that is estimated at the end of the test (*e.g.,* LC_{50}, EC_{20}), but can also refer to the variable being measured in the toxicity test (*e.g.,* survival, growth and reproduction).

Type I and Type II errors – Type I error (probability of which is commonly designated as alpha [α]) occurs when an investigator concludes there is a significant difference between samples when actually there is none. Type II errors (probability of which is commonly designated as beta [β]) occurs when an investigator concludes there is no significant difference when actually there is. Common target values for conventional hypothesis tests are 0.05 and 0.20, respectively (USEPA, 2009). However, a larger alpha [α] translates into a greater statistical power, and for power analysis, the alpha-level is often relaxed from the traditional 0.05 to 0.1.

Weight-of-Evidence – Weight of evidence is a process for integrating the results of different types of data into an overall conclusion. It involves a framework for considering of the strengths and weaknesses of each type of data, and the nature of uncertainty associated with each of them. Weight-of-evidence frameworks may be quantitative or qualitative, may involve the exercise of professional judgement, but must always be transparent and consistent.

Appendix A – Detailed Toxicity Test Selection Guidance

1. MANDATORY CONSIDERATIONS

1.1. Linkage to Study Objectives

The most important consideration in the selection of a toxicity test is the degree of alignment with the assessment endpoints and protection goals for the risk assessment. A specific test may score highly for a number of attributes, but its value is significantly diminished if it is not highly ecologically relevant. The problem formulation and conceptual site model should be used to identify the primary "drivers" for test selection; some common considerations include:

- Is the investigation linked to regulatory testing or specific Environment Canada programs that should be conducted according to the methods outlined in those programs?
- Does the exposure route (*i.e.*, environmental media being tested) reflect the pathway and organism life stage of relevance to the receptors of concern?
- Is the assessment screening-level or detailed (adherence to standardized protocols is emphasized more in the former)?
- What is the uncertainty associated with lab-to-field extrapolation, and how can it be minimized by choice of species and/or refinement of test procedures?
- Are the test endpoints appropriate to the nature of the contamination under investigation (*e.g.*, bioaccumulative substances such as polychlorinated biphenyls [PCBs] and tributyltin [TBT] may require longer test durations and may benefit from incorporation of a tissue bioaccumulation component)?
- How can standard toxicity endpoints be aligned with other lines of evidence (*e.g.*, toxicity identification evaluation [TIE], benthic community structure)?
- Is the testing tiered in nature? For example, short term toxicity tests can be used to refine a preliminary assessment of potential risk, such that resources are not wasted conducting long-term sublethal testing on acutely toxic test media.
- Does the investigator need to understand the cause of any measured toxicity (in which case a TIE may be needed)?

One area of consensus from a study by the Sustainable Fisheries Foundation (SFF, 2007) was that use of multiple toxicity tests in a battery approach is generally preferred to use of a single presumed "best" toxicity test endpoint. It is rare that all of the factors necessary to comprehensively evaluate test sensitivity or reliability are known in detail prior to testing. The test battery approach (Bay *et al.*, 2007):

- Provides a degree of insurance against unknown or unanticipated factors;
- Reduces the influence of spurious results from a test; and
- Increases the overall sensitivity of the testing program by using species with different patterns of contaminant sensitivity.

1.2. Organism Tolerance (Test Acceptability)

Toxicity tests have been developed for a specified range of environmental conditions. Many of these conditions can be controlled in the laboratory, and are stipulated in the test protocols. However, other conditions are sample-specific; this requires that the environmental media conform to tolerance ranges specified for the test. In many instances, it is possible to screen out tests from further consideration based on the following:

- Ammonia and Sulphide Tolerances – It is inappropriate to conduct toxicity testing on sediment samples that exceed the thresholds for pore water ammonia and sulphides, unless the purpose of the investigation is to evaluate toxicity of these substances. Purging of sediments prior to testing may be permissible to reduce these parameters to acceptable ranges (which are species-specific). Prior to consideration of purging (or any other substantive alteration of test media), discussion with and involvement of FCSAP Expert Support and Expert Support toxicologists is strongly advised.
- Substrate Tolerance – Environment Canada protocols for testing of amphipods in sediment stipulate that the approximate particle size distributions for the sampled sediments be known prior to testing. Testing must not be conducted on species outside their specified performance ranges, and testing at the margins of the range is not recommended. The percent content of fines in sediments also influences performance of sediment resuspension test endpoints (such as bivalve survival and normal development).
- Salinity Tolerance – Many test organisms are highly sensitive to ranges and variations in salinity. Although salinity can be influenced in the laboratory through the use of overlying water (including refreshes), toxicity testing should not be conducted outside specified tolerance ranges for salinity.
- Water Hardness Tolerance – Some freshwater test organisms (*e.g.*, some daphnids) cannot tolerate high hardness in test waters. Measured test responses may be due to water hardness rather than toxicants in the sample.
- pH Tolerance – sample media pH can have a direct effect on the test organism, or an indirect effect through mediation of the toxicity of COPCs (*e.g.*, pH-influenced toxicity of ammonia). The pH conditions are typically monitored throughout the duration of a test (routine monitoring of water quality) and there are protocol-stipulated test acceptability ranges. If a given test media is known to represent an extreme in pH, this should be incorporated in the test selection process.

In terms of pH tolerance, the investigator should not rely only on the selection factors (or caveats) listed in Table 3. Sample pH is a unique factor in that it may be considered either as a controlled factor in the test (*i.e.*, constrained in laboratory to be within range) or alternatively considered as a toxicological property of interest. Therefore, the context of pH levels should be considered for any proposed test, including:

- Comparison to protocol-stipulated ranges – where pH is out of range, many protocols recommend conducting side-by-side trials with adjusted and unadjusted pH, to aid in the discrimination of sources of toxicity;

- Environmental significance[20] – Where the receiving environment is naturally alkaline or acidic, the condition of a sample manipulated in the laboratory to a circumneutral condition may not reflect the ambient environmental conditions relevant to the assessment of potential harm. Conversely, manipulation may be required to mimic the toxic response profile expected in the field (i.e., discharge of groundwater or effluent to a well-flushed surface water body). The investigator must assess whether the toxicity potential is best assessed through of a manipulated or unmanipulated sample; side-by-side comparisons are often helpful in this regard;
- Influence on ammonia toxicity – pH is a modifying parameter for ammonia toxicity, and is used in the calculation of unionized ammonia concentrations;
- Mode of toxicity – pH may be of interest as a modifying factor for the ionization of other substances of interest, or may be toxic as an inherent sample property;
- Changes over time – pH changes may occur over the course of testing as a result of sample refreshes and/or chemical changes throughout the test.

Guidance is available on a test-specific basis with respect to tolerance ranges of candidate test organisms. The availability of data for each of the above factors is summarized in **Table 3** on a test-specific basis.

The assessment of test acceptability based on protocol-stipulated organism tolerance is only a preliminary step in the selection of a suitable test organism. The performance ranges specified by test protocols provide clear articulation of conditions that are unsuitable for test application, but may not provide guidance on conditions for which a species is only marginally suitable. For example, Environment Canada (1998b) specifies a tolerance ranges for several amphipod species that are overlapping, such that practitioners may have several candidate species for which the sample may be tested without violating the protocol. Other literature may be available to evaluate the suitability of test conditions close to the protocol thresholds. There are also test factors that may not have protocol-stipulated tolerances (*e.g.*, total organic carbon content) but for which information on organism sensitivity is highly relevant. For this reason, the organism tolerance is evaluated in two stages: (1) screening against protocol-stipulated ranges (Section 1.2) and (2) refined evaluation based on assessment of the literature for remaining candidate tests (Section 2.2).

1.3. Physicochemical conditions

In evaluating a toxicity test, the potential interactions of the test substance with the test medium and other chemicals in the medium must be considered. This consideration is partly addressed above, when considering the ammonia and sulphide tolerances of the organism. For example, Schubauer-Berigan *et al.* (1995) discuss the interaction between pH and ammonia toxicity to two

[20] Note that, aside from risk assessment objectives, compliance with laws and regulations concerning discharges is required. Usage of sample pH manipulations and consideration of receiving environment adjustments may not be permissible at some sites (e.g., where discharges are acutely lethal and subject to *Fisheries Act* enforcement)

common freshwater toxicity test species. However, there are additional factors that must be considered in the selection of an appropriate test matrix; these factors can influence the bioavailability of the substance. For example, chemical interactions between metals cations and sulphide (or chloride) anions may be important in term of sequestering the metals. Presence of dissolved organic matter in a test medium can also influence bioavailability and toxicity. The effect of sample filtration, changes in pH and redox condition, and other physical and chemical changes in the sample should be taken into account when considering candidate toxicity tests. Sample collection, transport, storage and manipulation before and during testing may also alter sample properties that influence contaminant bioavailability (*e.g.*, oxidation of an anoxic sample, mixture of micro-scale layering by sediments during homogenization). These considerations are linked to Section 1.1 above because physicochemical factors determine (in part) the suitability of a toxicity test for representing the exposure pathway of interest.

2. TEST SUITABILITY FACTORS

Once the mandatory considerations for screening candidate tests have been evaluated, there are usually several candidate toxicity tests remaining to choose from. To further assist in test selection, broad test characteristics that most commonly influence selection of a specific test have been identified (**Table 3**). These characteristics relate to test feasibility, practicality, performance, and cost, and were subdivided into three groups of related factors, including:

- Utility/relevance for risk assessment purposes;
- Organism tolerance information; and
- Logistical and planning considerations.

These factors are described in the following subsections, which serve as a detailed explanation for the simplified categorizations found in **Table 3**. Detailed explanations for symbols and test-specific caveats are provided at the end of **Table 3**. Generally, the following interpretations apply:

- Factor scoring of (++) – the test is considered highly advantageous for this attribute, or abundant information is readily available to evaluate the sensitivity of the test performance to the attribute;
- Factor scoring of (+) – information is considered somewhat advantageous for this attribute, or information is available to evaluate the sensitivity of the test performance to the attribute;
- Factor scoring of (blank) – the attribute is unknown with respect to test reliability and/or utility;
- Factor scoring of (-) – the attribute is considered to be a constraint to the test reliability and/or utility; practitioners should proceed with caution and determine whether the limitation applies to their situation, and assess the margin for error.

2.1. Utility for Risk Assessment

This section discusses the relevance of the test data to the risk-based hypotheses identified in the problem formulation, and details the attribute scoring presented in **Table 3**.

1.1.1. Availability of Toxicity Data

Literature-based toxicity (concentration-response) data for a test species (or closely related species) is advantageous for risk assessment because it can be used to screen for significant toxicants in a particular sample. Although it is not possible to predict the relative sensitivities of various test organisms with a high level of precision (particularly for unique contaminant mixtures), it is sometimes possible to identify test endpoints that are known to be relatively sensitive for certain classes of contaminants (*e.g.*, divalent metals, polycyclic aromatic hydrocarbons). For example, if a test contains elevated concentrations of nickel, it may be advantageous to select *Ceriodaphnia* rather than *Daphnia* as the test species if other factors permit, as the former has demonstrated greater sensitivity to nickel in standardized toxicity testing (Kszos *et al.*, 1992). Pastorok and Becker (1990) observed that common marine sediment toxicity tests differed in their statistical sensitivity and biological sensitivity to particular sediments contaminated with different kinds of chemicals (*e.g.*, polycyclic aromatic hydrocarbons versus metals). SFF (2007) notes that significant data gaps remain with respect to the relative sensitivities of species to groups of contaminants. Although USEPA is conducting evaluations using five chemical classes to evaluate sensitivity in water-only exposures, the sensitivity in contaminated sediment mixtures is often highly uncertain. For this reason, *a priori* assumptions regarding concentration-response relationships should be made carefully, and should be reserved for situations in which only one or two COPCs dominate the contaminant profile.

The choice of a test species on the basis on sensitivity to one or more COPCs is an important aspect of test selection. Moreover, the objective is not necessarily to select the most sensitive species, but rather to select the species that best reflects the desired sensitivity in light of the protection goals for the site and the receptors of concern. Where detailed information is lacking regarding the species sensitivity distribution for a suite of organisms, selection of the most sensitive organism is often adopted as a conservative approach. In other instances, where a candidate test organism is suspected to be more sensitive but has low ecological relevance (*e.g.*, low functional relevance to the receptor group of interest) it may not be the preferred choice.

Factor scoring: (++) if data are known to be available for most COPCs; (+) if data are known to be available for some COPCs; (-) if data are known to be absent for most COPCs; (blank) if unknown.

1.1.2. Relevance of Control

Negative controls provide evidence that laboratory test conditions have not resulted in unacceptable (confounded) test performance. Also, in the absence of study-specific reference samples, many risk assessments interpret toxicity data relative to negative controls. The degree of relevance of a negative control is highly test-specific, and many protocols provide several options. Therefore, the importance of this factor can vary depending on the laboratory even for

the same test protocol; hence, consultation with an environmental toxicologist at the candidate laboratory should be conducted.

Negative controls that consist of field-collected, natural media provide more robust data than negative controls prepared from artificial substrates (*e.g.*, silica sand, laboratory prepared or adjusted water). Negative controls conducted using a media different from the site samples (*e.g.*, a sediment test conducted with a water-only negative control) are least relevant; in these situations, the importance of obtaining appropriate reference samples is heightened.

Factor scoring: (+) toxicity test uses a field-collected natural negative control or suitable surrogate; (-) toxicity test does not use a field-collected natural negative control, the negative control uses a different medium, or the composition of the negative control could affect interpretation of results (*e.g.,* nutrient-supplemented algal culture medium); (blank) if unknown.

1.1.3. Statistical Power

The statistical power of a test, which is the ability to correctly detect a reduction in endpoint performance, is influenced by several factors, including:

- Magnitude of the reduction to be detected (*i.e.*, the decision criteria or effect size);
- Number of replicates available;
- Overall inter-replicate variability; and
- Acceptable Type I error rate.

For the purposes of a site-specific risk assessment, *a priori* evaluation of the statistical power of a test is rare because: (1) the number of replicates is typically dictated by the protocol; and (2) the decision criteria are often dictated by policy. Nevertheless, different tests have systematic differences in the magnitude of response that can reliably be detected. A toxicity test with a high degree of inherent biological variability may present a challenge for interpretation, and knowledge of this variability can used to evaluate and compare candidate tests.

Factor scoring: (+) toxicity test is known to have relatively low inter-replicate variability; (-) toxicity test is known to have relatively high inter-replicate variability; (blank) if unknown. Tests without replicates for a defined exposure level (*e.g.*, 96-h rainbow trout LC_{50}) are scored as not applicable (NA) in **Table 3**.

1.1.4. Availability of Chronic and Multiple Endpoints

Risk assessment guidance often emphasizes chronic (long-term) toxicity endpoints, particularly those with sublethal endpoints such as growth or reproduction, over acute toxicity tests. However, neither the sensitivity nor the utility of a test endpoint is directly proportional to test duration; Bay *et al.* (2007) observed that sublethal tests and lethal tests are complementary, as none of the test methods ranked consistently highest with respect to sensitivity or reliability. However, measurement of chronic and sublethal effects is often considered to align more closely with assessment endpoints, and tests with multiple endpoints provide additional information on the nature of the biological response.

Factor scoring: (++) test includes multiple, long-term chronic endpoints; (+) test includes at least one chronic, long-term endpoint or a surrogate chronic endpoint (*e.g.*, sensitive life stage); (-) test does not include a chronic or surrogate chronic endpoint.

1.1.5. Geographic Suitability

In a risk assessment, toxicity test species and endpoints are chosen as measurement endpoints to be used as surrogates for the assessment endpoint of interest (natural species, populations, and communities). Although individual toxicity test species are assumed to have relevance to broader organism groups, it is advantageous to minimize the degree of geographical extrapolation. For example, a cold-water marine macro-algae native to Canadian waters likely provides a better alignment with the "marine plant" receptor group than a tropical aquatic diatom.

Factor scoring: (++) test species is aligned with resident organisms found at most locations in Canada; (+) test species is aligned with resident organisms for some locations in Canada; (-) test species is not well-aligned with receptor selection for Canadian risk assessments.

1.1.6. Tissue Production

Direct measurement of contaminant concentrations in organisms exposed to site-specific media provides an additional line of evidence that may have value in a risk assessment context. Some test protocols provide or accommodate options for tissue collection synoptic with the toxicity exposures.

Factor scoring: (+) test species likely has sufficient mass to permit tissue chemistry analyses; (-) test species does not have sufficient mass to permit tissue chemistry. Note that modifications to the test may be necessary to permit tissue analyses (*e.g.*, number of replicates or loading densities; use of preservatives after termination).

2.2. Organism Tolerance

The factors included in this category relate to the biological characteristics of the candidate test organisms, specifically their tolerance to potential confounding factors. Most toxicity tests are sensitive to some type of non-contaminant effect (*e.g.*, grain size, hardness) that can have a confounding effect on test results. Samples with natural physiochemical variables that are outside the biological tolerance limits for the test species may exhibit apparent toxicity that is artefact (not attributable to the presence of COPCs). Although the influences of potentially confounding factors are still not entirely known for many tests, knowledge of factors that can affect a test and the approximate range where effects occur is needed for study design and data interpretation (Bay *et al.*, 2007). In evaluating organism tolerances, it is assumed that the practitioner will first evaluate the sample relative to protocol-stipulated acceptability ranges (Section 1.2), such that further evaluation of organism tolerance (described below) is conducted to assist in discriminating among multiple candidate tests.

Natural physiochemical parameters that are frequently considered in the selection of test species include ammonia, sulphides, grain size, organic carbon content and salinity. Other parameters such as alkalinity, pH, and major ion concentrations should be considered on a site- and species-specific basis. An extensive literature search is required to provide a definitive threshold for each

parameter and species and is beyond the scope of this Guidance Module. However, it was possible to characterize the general level of toxicological information available to assess these factors, and in some cases to identify species that are known to be sensitive to one or more of these factors. Protocol documents for each test can be consulted to provide a summary of the technical knowledge, and additional references from the literature are provided below.

1.1.7. Ammonia

Word *et al.* (2005) provide a useful summary of the confounding effects of nonpersistent chemical characteristics, including ammonia. Their review identified that toxic effects to marine life occur in the range of 3 to 100 mg/L ammonia (as nitrogen), with larval stages of organisms more sensitive than adults. However, ammonia is a highly test- and species-specific consideration, particularly as labile substances do not typically remain at constant concentrations throughout the duration of toxicity tests. Ammonia is not a typical consideration for soil toxicity tests.

References for aquatic testing include : Ankley *et al.* (1995; *H. azteca*); Kohn *et al.* (1994; marine amphipods); Dillon *et al.* (1993; *N. arenaceodentata*); McDonald (2005; *M. galloprovincialis*; oysters); Schubauer-Berigan *et al.* (1995; *L. variegatus*, *C. tentans*); Nebecker and Schuytema (2000; amphibians, *P. promelas*); Schubauer-Berigan and Ankley (1991; *C. dubia*, *L. variegatus*); Boardman *et al.* (2004; *M. menidia*, *C. variegatus*; *M. bahia*; *P. pugio*; *M. mercenaria*); Phillips *et al.* (2005; *H. rufescens*, *H. costata*; *A. affinis*, *M. galloprovincialis*; *M. pyrifera*) ; Borgmann (1994) ; Borgmann and Borgmann (1997).

Factor scoring: (+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.

1.1.8. Sulphides

Word *et al.* (2005) provides an introduction to the confounding effects of sulphides in sediments. For additional information, consult Knesovich *et al.* (1994; multiple species); Dillon *et al.* (1993; *N. arenaceodentata*); Broderius *et al.* (1977; *P. promelas*); Kuester *et al.* (2005; *D. magna*); Losso *et al.* (2007; *P. lividius*, *C. gigas*); Oseid and Smith (1975; *H. limbata*). Sulphide is not a typical consideration for soil toxicity tests.

Factor scoring: (+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.

1.1.9. Substrate (Particle Size and TOC)

Grain size distribution is important for selecting species for whole-sediment toxicity tests. Relevant literature includes: Environment Canada (1998b; marine and estuarine amphipods). Suedel and Rogers (1994; *H. azteca*, *C. tentans*), Sibley *et al.* (1998; *C tentans*); Ankley *et al.* (1994: *H. azteca*, *C. tentans*, *L. variegatus*); Ringwood *et al.* (1997; Microtox™); Tay *et al.* (1998; Microtox™). Samples with high clay or silt content can also be a factor in testing that involves suspended sediment for larval species (van den Hurk ;1994; *C. gigas*) or those species that require adequate light penetration such as giant kelp (Devinney and Voise, 1978). Sediment

that is enriched with woody debris or organically enriched by wastewater or other eutrophication sources may not be appropriate for some test species. Similarly, low organic carbon may be a factor in whole-sediment toxicity tests that involve species that feed during the test exposure. For example, Suedel and Rogers (1994) found that *C. tentans* had low survival in samples with organic contents of 0.9% or less, as the organisms were unable to find sufficient material to construct their larval cases. Relevant literature includes: Suedel and Rogers (1994; *H. azteca*, *C. tentans*); Ankley *et al.* (1994: *H. azteca*, *C. tentans*, *L. variegatus*). Word *et al.* (2005) provides a useful summary of confounding factors in sediment toxicity tests, including persistent chemical characteristics such as organic carbon quantity and quality.

Factor scoring: (+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.

1.1.10. Salinity and Hardness

Freshwater organisms often require hardness similar to that of natural systems; extremely soft or hard waters can cause organism stress. Water hardness should be considered in study designs through the inclusion of appropriate reference samples; adjustment of sample hardness is possible but may alter contaminant bioavailability. Similar issues exist for marine organisms and salinity. Most marine tests are conducted at a standard salinity of approximately 2.8%, and adjustment of sample salinity is often required. This can be accomplished by the addition of hypersaline brine prepared from seawater or the addition of sea salts. The choice of salt amendment varies depending on laboratory; both methods are acceptable provided that negative control performance can be maintained. Many organisms, especially those from estuarine environments, can be acclimated over time to different salinities depending on study requirements.

Factor scoring: (+) toxicological information is generally available; (-) toxicological information is known to be absent; (blank) if unknown; (S) species known to be sensitive to the parameter relative to other test species; (T) species known to be tolerant to a wide range of the parameter.

2.3. Logistical and Planning Factors

The factors included in this category relate to the practicalities of testing. As risk assessments are generally conducted under financial and timeline constraints, logistical and cost factors can significantly influence the feasibility of specific tests.

1.1.11. Tolerance to Laboratory Handling

Organism transport and laboratory handling can be a significant stressor for some species. Although test protocols require that organisms be acclimated to laboratory conditions prior to testing, and provide guidance on appropriate handling techniques, species that are highly sensitive to collection and handling stress may require extra effort to ensure high-quality data. Some organisms can be cultured in-house, which removes uncertainty attributable to transport. However, the maintenance of robust cultures of organisms is challenging for some species. Although negative control tests will identify major test artefacts attributable to stressed

organisms, the potential loss of critical test information or expense of retesting can be a significant limiting factor.

Factor scoring: (+) known to be relatively tolerant to handling/culturing stress; (-) known to be relatively sensitive to handling/culturing stress; (blank) if unknown.

1.1.12. Organism Source

Maintaining reliable cultures or obtaining field-collected organisms is one of the largest constraints on toxicity test selection. Toxicity tests using species without a reliable source require considerable effort and are not routinely included in ecological risk assessments despite possible advantages in terms of organism sensitivity or alignment with receptor groups. Also, it may be necessary to acquire the same strain of organism if comparative tests are to be conducted among treatments.

Factor scoring: (+) reliable sources of organisms are available, whether from in-house or commercial cultures, or appropriate field collection locations; (blank) if unknown.

1.1.13. Seasonal Availability

Not all organisms are available on a year-round basis. Organism availability can be a significant constraint in the design and implementation of timely risk assessment and toxicity monitoring programs. For some species (particularly shellfish, but also some fish species), the reproductive or spawning status of organisms is a significant limiting factor in the selection of a test species. Field-collected organisms that are normally available year-round may also periodically be affected by extreme weather conditions (*e.g.,* high temperatures, storms) at their collection sites that temporarily reduce availability.

Factor scoring: (+) generally available on a year-round (or nearly year-round) basis; (-) seasonal constraints are known; (blank) if unknown.

1.1.14. Sample Volumes

Sample collection and transport considerations will be influenced by the amount of sample required. The volume of exposure media required for each test varies substantially among tests. Some water-column tests include the collection of additional refresh samples during the test exposure. These volumes can be prohibitive, especially for fish tests. For example, a 7-d early life stage test with rainbow trout can require as much as 140 L of water in total (including refreshes), rendering it impractical for samples collected from seeps that have limited discharge rates. Sediment samples for whole-sediment testing tend to be less susceptible to these constraints. However, porewater sampling methods require volumes comparable to overlying water tests, and extraction efficiency from bulk sediment is low for many sediment types.

Factor scoring: (+) sample volumes do not generally place a significant constraint on sampling programs; (-) sample volumes can present a significant constraint on sampling programs; (blank) if unknown.

1.1.15. Availability of (Standard) Method

Toxicity testing can be conducted with virtually any species and toxicological endpoint as part of a research program. However, the weight (strength of confidence) assigned to a test endpoint is partly a function of the precedent for use in similar applications. This Guidance Module focuses on standard test species and endpoints that have been documented in widely-available North American protocol documents. Methods defined as "standard" have a protocol that has received the rigorous testing necessary to be published as an Environment Canada, USEPA, or ASTM method. Such standard tests are preferred based on control acceptability criteria and quality assurance standards.

Not all test species and endpoints described in protocol documents are readily available in commercial or government toxicology laboratories, and only a subset of tests is routinely offered by most facilities.

Factor scoring: (++) test is routinely offered by most laboratories; (+) test may have limited availability or require specialized effort to complete; (-) test has little or no availability.

1.1.16. Cost

Cost is a limiting factor in many sediment assessment studies, and costs for toxicity tests are generally proportional to the amount of labour required for monitoring test conditions, set-up and tear-down time, and the costs for obtaining/maintaining organism cultures. As costs vary according to laboratory, and laboratories may offer discounts for large numbers of samples, these cost assignments are intended to be relative measures rather than absolute cost indicators. Tests with lower per-sample costs are more amenable to the toxicity test battery approach, and application of appropriately sensitive tests that are also relatively inexpensive may facilitate improved spatial characterization in the study design. The increased cost of chronic testing must be considered relative to the degree of uncertainty reduction afforded by the test. California State Water Resources Control Board (2005) demonstrated that for survival and growth endpoints, the *Leptocheirus* 28-day toxicity test was actually equal to or less sensitive (on average) relative to the 10-d version of the test. In terms of toxicity testing, longer duration tests are not necessarily better, particularly if chronic testing results in reduced representation of feeding types and test species (due to cost implications).

Factor scoring: ($) costs are typically less than $500/sample; ($$) costs are $500 - 1000 per sample; ($$$) costs are $1000 - 1750/sample; ($$$$) costs typically exceed $1750/sample.

3. SPECIALIZED TESTING CONSIDERATIONS

3.1. Protocol Adaptation

Although some protocol requirements are considered inviolable, other aspects of test implementation are more flexible and provide an opportunity to customize to project specific needs. For example:

- Dilution Water – Toxicity testing can be conducted using site water for dilution instead of laboratory water, to more closely approximate site-specific factors.
- Field Culturing – If cultured organisms are hypothesized to be more sensitive than field organisms (*i.e.*, lack of adaptive response to natural background concentrations of non-anthropogenic substances), field culturing in a natural condition can be substituted.
- Sample Properties – Sample collection, transport, storage and manipulation before and during testing may alter sample properties that influence contaminant bioavailability (*e.g.*, oxidation of an anoxic sample, mixture of micro-scale layering by sediments during homogenization). In these cases, testing of intact sediment core samples may provide a more environmentally realistic representation of exposure.
- Brine Adjustment – Where the purpose of the assessment is to evaluate the potential impact of a source medium on the receiving environment, and where the salinity conditions differ between the source and receiving environment, adjustment using hypersaline brine or dry salts allows use of a test species representative of the receiving environment. Dual controls are required in these cases to gauge the potential confounding effect of brine adjustment.
- Constructed (artificial dilution) water) – Where appropriate natural dilution water is not available, it is possible to create test waters for which the ion composition can be manipulated to match the site and be free of contaminants (*e.g.*, Borgman, 1996). Reconstituted water can account for physicochemical differences between negative controls and test samples, but have associated uncertainties.Protocol Modification – Some tests are amenable to changes in test duration *(e.g.,* settling time in a resuspension test), endpoints *(e.g.,* optional reproduction endpoints for some tests, such as sediment testing with mysid shrimp), or exposure media *(e.g.,* inclusion of solids at the base of a water column test vessel).

Any modifications should be discussed with regulators and FCSAP Expert Support to ensure that the modifications and test results of these modified tests would be acceptable.

3.2. *In Situ* Testing

In situ tests simulate continuous exposure to the site media under actual environmental conditions such as temperature, salinity, nutrients, stream flow, and insolation (*e.g.*, UV light) and also account for time-varying stressors such as those associated with flood events or tidal flow. In *in situ* testing, direct evaluation of exposure pathways relevant to an organism is conducted (such as sediment, porewater, and overlying water to an epifaunal invertebrate). Therefore, data from *in situ* tests may provide a more realistic assessment of field responses relative to data from laboratory tests. Anderson *et al.* (2004) also note that the "common practice of sediment homogenization and sieving may also subject animals in laboratory exposures to artefacts of chemical disequilibria." In contrast, there are disadvantages of *in situ* tests, specifically related to the lack of control over the conditions under which an *in situ* test occurs. Fluctuations in test conditions caused by storms, drought, flooding, or atypical conditions can prove problematic, and logistics are more difficult in the field relative to a controlled laboratory setting. Historically, *in situ* testing has been applied more frequently in freshwater environments, and had focused on the

invertebrates *Hyalella azteca* and *Chironomus dilutus* (Chappie and Burton, 1997). However, marine/estuarine methods have progressed beyond caged mussel/fish studies and now include validated testing using amphipods such as *Eohaustorius* spp. (Anderson *et al.*, 2004)

Detailed evaluation of *in situ* methods is beyond the scope of this Guidance Module. In general, the desirability of *in situ* endpoints should be considered during the problem formulation phase, trading off the advantages and disadvantages of the approach. Species and endpoints can be chosen using a similar procedure to that identified for standardized laboratory tests.

3.3. Toxicity Identification Evaluation

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.

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 environmental quality guideline exceedances. TIEs address this problem by indicating the contaminant group(s) most likely responsible for the observed responses. A properly conducted TIE will increase the confidence of the study conclusion by using multiple lines of evidence (*i.e.*, multiple treatments showing consistent indications of potential cause-effect), thereby reducing the chance of a spurious result.

Considerations that influence the usage and type of TIE include:

- TIEs are typically conducted after (or concurrently with) standardized toxicity testing. Careful consideration of how to integrate sample collection for both a standard toxicity testing program and a TIE is required. For example, sufficient sample volumes need to be collected in advance if a synoptic TIE is contemplated.
- TIEs are most effective for samples that exhibit pronounced (as opposed to marginal) toxic responses.
- TIEs are iterative, with the results of one type of manipulation leading to other potential manipulations that should be examined. The scope of the TIE cannot often be predicted in advance, although there should be discussion regarding the desired level of identification (broad contaminant type, class, element, speciation). 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 may need to consider a broad range of potential contaminants, as 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. Alternatively, if soil/sediment TIE is conducted, the technical limitations/uncertainties of the method must be considered.

TIEs are not usually recommended for screening level or preliminary assessments, and are conducted at a minority of sites. Consequently, standardized protocols should be emphasized unless the site-specific investigation strategy (and associated problem formulation findings) suggest a need to adapt the program to account for potential TIE needs.

Appendix B – References for Toxicity Testing

1. TOXICITY TEST PROTOCOL REFERENCES

Allchin, D. 2001. Error Types. *Perspectives on Science* 9(1):38-58.

ASTM International. 2006. Standard Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians. Method E729-96 (re-approved 2002). In: *2006 Annual Book of ASTM Standards*, *Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard Guide for Behavioural Testing in Aquatic Toxicology. Method E1604-94 (re-approved 2002). In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard Test Method for Measuring the Toxicity of Sediment-Associated Contaminants with Estuarine and Marine Invertebrates. Method E1367– 03. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology,* Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard Guide for Conducting In-situ Field Bioassays with Caged Bivalves. Method E2122-02. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting static acute toxicity tests starting with embryos of four species of saltwater bivalve molluscs. Method E724-98 (re-approved 2004). In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting *Daphnia magna* life-cycle toxicity tests. Method E1193-97 (re-approved 2004). In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting static toxicity tests with microalgae. Method E1218-04. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting the frog embryo teratogenesis assay - *Xenopus*. Method E1439-98 (re-approved 2004). In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for acute toxicity test with the rotifer *Brachionus*. Method E1440-91 (re-approved 2004). In: *2006 Annual Book of ASTM Standards, Biological*

Effects and Environmental Fate; Biotechnology, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting static acute toxicity tests with echinoid embryos. Method E1563-98 (re-approved 2004). In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting sediment toxicity tests with polychaetous annelids. Method E1611-00. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard test methods for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. Method E1706-05. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting renewal phytotoxicity tests with freshwater emergent macrophytes. Method E1841-04. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting laboratory soil toxicity or bioaccumulation tests with the lumbricid earthworm *Eisenia fetida* and the enchytraeid potworm *Enchytraeus albidus*. Method E1676-04. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2006. Standard guide for conducting laboratory toxicity tests with freshwater mussels. Method E2455-06. In: *2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

ASTM International. 2008. Standard guide for conducting whole sediment toxicity tests with amphibians. Method E2591-07. In: *2008 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology*, Volume 11.06. ASTM International, West Conshohocken, PA.

Burgess, R.M. and G.E. Morrison. 1994. A short-exposure, sublethal, sediment toxicity test using the marine bivalve *Mulinia lateralis*: statistical design and comparative sensitivity. *Environ. Toxicol. Contam.* 13:571-580.

Dekker, T., G.D. Greve, T.L. Ter Laak, M.E. Boivin, B. Veuger, G. Gortzak, S. Dumfries, S.M.G. Lucker, M.H.S. Kraak, W. Admiraal and H.G. van der Geest. 2006. Development and application of a sediment toxicity test using the benthic cladoceran *Chydorus sphaericus*. Environ. Pollut. 140:231-238.

Dillon, T.M., D.W. Moore and A.B. Gibson. 1993. Development of a chronic sublethal bioassay for evaluating contaminated sediment with the marine polychaete worm *Nereis (Neanthes) arenaceodentata*. *Environ. Toxicol. Chem.* 12:589-605.

Environment Canada. 1990a. *Biological test method: acute lethality test using rainbow trout*. Environmental Protection Series, Report EPS 1/RM/9, July 1990. Environment Canada, Conservation and Protection, Ottawa, ON. Amended May 1996 and May 2007.

Environment Canada. 1990b. *Biological test method: acute lethality test using Daphnia spp*. Environmental Protection Series, Report EPS 1/RM/11, July 1990. Environment Canada, Conservation and Protection, Ottawa, ON. Amended May 1996.

Environment Canada. 1990c. *Biological test method: acute lethality test using threespine stickleback (Gasterosteus aculeatus)*. Environmental Protection Series, Report EPS 1/RM/10. July, 1990. Environment Canada, Conservation and Protection, Ottawa, ON. Amended March 2000.

Environment Canada. 1992a. *Biological test method: test of larval growth and survival using fathead minnows*. Environmental Protection Series, Report EPS 1/RM/22, February 1992. Environment Canada, Conservation and Protection, Ottawa, ON. Amended November 1997.

Environment Canada. 1992b. *Biological test method: fertilization assay using echinoids (sea urchins and sand dollars)*. Environmental Protection Series, Report EPS 1/RM/27, December 1992. Environment Canada, Conservation and Protection, Ottawa, ON. Amended November 1997.

Environment Canada. 1992c. *Biological test method: acute test for sediment toxicity using marine or estuarine amphipods*. Environmental Protection Series, Report EPS 1/RM/26, December 1992. Environment Canada, Conservation and Protection, Ottawa, ON. Amended October 1998.

Environment Canada. 1992d. *Biological test method: toxicity test using luminescent bacteria (Photobacterium phosphoreum)*. Environmental Protection Series, Report EPS 1/RM/24, November 1992. Environment Canada, Conservation and Protection, Ottawa, ON.

Environment Canada. 1997a. *Biological test method: test for survival and growth in sediment using the freshwater amphipod Hyalella azteca*. Environmental Protection Series, Report EPS 1/RM/33, December 1997. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 1997b. *Biological test method: test for survival and growth in sediment using the larvae of freshwater midges (Chironomus tentans or Chironomus riparius)*. Environmental Protection Series, Report EPS 1/RM/32, December 1997. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 1998a. *Biological test method: toxicity tests using early life stages of salmonid fish (rainbow trout)*. Environmental Protection Series, Report EPS 1/RM/28, Second edition - July 1998. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 1998b. *Biological test method: reference method for determining acute lethality of sediment to marine or estuarine amphipods*. Environmental Protection Series, Report EPS 1/RM/35, December 1998. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2000a. *Biological test method: reference method for determining acute lethality of effluents to rainbow trout*. Environmental Protection Series, Report EPS 1/RM/13, Second edition – December 2000. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON. Amended May 2007.

Environment Canada. 2000b. *Biological test method: reference method for determining acute lethality of effluents to Daphnia magna*. Environmental Protection Series, Report EPS 1/RM/14, Second edition – December 2000. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2001. *Biological test method: test for survival and growth in sediment using spionid polychaete worms (Polydora cornuta)*. Report EPS 1/RM/41, December 2001. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2002. *Biological test method: reference method for determining the toxicity of sediment using luminescent bacteria in a solid-phase test*. Report EPS 1/RM/42, April 2002. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

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

Environment Canada. 2005. *Biological test method: test for measuring emergence and growth of terrestrial plants exposed to contaminants in soil*. Report EPS 1/RM/45, February 2005. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2007a. *Biological test method: test of reproduction and survival using the cladoceran Ceriodaphnia dubia*. Environmental Protection Series, Report EPS 1/RM/21, Second edition - February 2007. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2007b. *Biological test method: growth inhibition test using a freshwater alga*. Environmental Protection Series, Report EPS 1/RM/25, Second edition – March 2007. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2007c. *Biological test method: test for measuring the inhibition of growth using the freshwater macrophyte, Lemna minor*. Report EPS 1/RM/37, Second edition – January 2007. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Environment Canada. 2007d. *Biological test method: test for measuring survival and reproduction of springtails exposed to contaminants in soil*. Report EPS 1/RM/47, September 2007. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Ingersoll, C.G., D.D. MacDonald, J.M. Besser, W.G. Brumbaugh, C.D. Ivey, N.E. Kemble, J.L. Kunz, T.W. May, N. Wang and D. Smorong. 2008. *Sediment chemistry, toxicity and*

bioaccumulation data report for the US Environmental Protection Agency – Department of the Interior sampling of metal-contaminated sediment in the Tri-state Mining District of Missouri, Oklahoma and Kansas. Report prepared for US Environmental Protection Agency and US Fish and Wildlife Service by MacDonald Environmental Sciences Ltd. And US Geological Survey, Columbia, MO.

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.

Nipper, M.G., D.J. Greenstein and S.M. Bay. 1989. Short- and long-term sediment toxicity test methods with the amphipod *Grandidierella japonica*. *Environ. Toxicol. Chem.* 8:1191-1200.

OECD (Organization for Economic Cooperation and Development). 2007. *OECD guidelines for the testing of chemicals: sediment-water Lumbriculus variegatus toxicity test using spiked sediment*. OECD Guideline 225. Organization for Economic Cooperation and Development, Paris, France. October 2007.

OECD. 2008. *OECD guidelines for the testing of chemicals: predatory mite (Hypoaspis (Geolaelaps) aculeifer) reproduction test in soil*. OECD Guideline 226. Organization for Economic Cooperation and Development, Paris, France. October 2008.

OECD. 2008. *OECD Guideline for the testing of chemicals: the amphibian metamorphosis assay*. Draft test guideline. June 25, 2008.

OMOE (Ontario Ministry of the Environment). 1992. *Laboratory sediment biological testing protocol.* Ontario Ministry of the Environment, PIBS 2067E. Water Resources Branch, Toronto, ON. August 1992.

PSEP (Puget Sound Estuary Program). 1995. *Recommended guidelines for conducting laboratory bioassays on Puget Sound sediments*. Prepared for US Environmental Protection Agency, Region 10, Seattle, WA. Final Report, July 1995. 85 pp.

Samoiloff, M. 1990. The nematode toxicity assay using *Panagrellus redivivus*. *Toxicity Assessment* 5:309-318.

USEPA (United States Environmental Protection Agency). 1994. *Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods*. US Environmental Protection Agency, Office of Research and Development, Narragansett, RI. EPA/600/R-94/025. June 1994.

USEPA. 1995. *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to west coast marine and estuarine organisms*. US Environmental Protection Agency, National Exposure Research Laboratory, Cincinnati, OH. EPA/600/R-95-136. August 1995.

USEPA. 2000. *Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates*. US Environmental Protection Agency, Office of Research and Development, Duluth, MN. EPA/600/R-99/064. March 2000.

USEPA. 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.

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

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

USEPA. 2002c. *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms*. 3rd edition. US Environmental Protection Agency, Office of Water, Washington, DC. EPA-821-R-02-014. October 2002.

USEPA. 2009. Statistical Primer. Authored by Leska Fore (Statistical Design, Seattle WA) under contract to US Environmental Protection Agency, Office of Water, Washington, DC. Last updated on 10 July 2009. Available at: http://www.epa.gov/bioiweb1/statprimer/power.html

USEPA/USACE (USEPA/US Army Corps of Engineers). 1998. *Evaluation of dredged material proposed for discharge in waters of the U.S. - testing manual*. Inland Testing Manual. EPA-823-B-98-004. US Environmental Protection Agency, Office of Marine and Estuarine Protection, Washington, DC and US Army Corps of Engineers, Washington, DC. February 1998.

WDOE (Washington State Department of Ecology). 1996a. *Earthworm bioassay protocol for soil toxicity screening. Publication No. 96-327*. Washington State Department of Ecology, Olympia, WA. June 1996. (http://www.ecy.wa.gov/pubs/96327.pdf)

WDOE. 1996b. *Early seedling growth protocol for soil toxicity screening*. Publication No. 96-324. Washington State Department of Ecology, Olympia, WA. June 1996. (http://www.ecy.wa.gov/pubs/96324.pdf)

WDOE. 2004. *Nematode bioassay protocol for soil toxicity screening*. Publication No. 04-09-044. Washington State Department of Ecology, Olympia, WA. August 2004. (http://www.ecy.wa.gov/pubs/0409044.pdf)

2. TECHNICAL GUIDANCE REFERENCES

Anderson B.S., J.W. Hunt, B.M. Phillips, S. Tudor, F. Fairey, J. Newman, M. Puckett, M. Stephenson, E.R. Long, and R.A.Tjeerdema. 1998. Comparison of marine sediment toxicity test protocols for the amphipod *Rhepoxynius abronius* and the polychaete worm *Nereis* (*Neanthes*) *arenaceodentata*. *Environmental Toxicology and Chemistry* 17:859-866.

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.

Anderson, B.S., J. W. Hunt, B.M. Phillips, P.A. Nicely, R.S. Tjeerdema and M. Martin. 2004. A Comparison of *In Situ* and Laboratory Toxicity Tests with the Estuarine Amphipod *Eohaustorius estuarius*. *Arch. Environ. Contam. Toxicol.* 46:52–60.

Anderson-Carnahan, L., S. Foster, M. Thomas, W. Korth, and K.H. Bowmer. 1995. Selection of a suitable cladoceran species for toxicity testing in turbid waters. *Austral Ecology* 20(1):28-33.

Ankley, G.T., D.A. Benoit, J.C. Balogh, T.B. Reynoldson, K.E. Day, and R.A. Hoke. 1994. Evaluation of potential confounding factors in sediment toxicity tests with three freshwater benthic invertebrates. *Environmental Toxicology and Chemistry* 13(4):627-635.

Ankley, G.T., M.K. Schubauer-Berigan, and P.D. Monson. 1995. Influence of pH and hardness on the toxicity of ammonia to the amphipod *Hyalella azteca*. *Can. J. Fish. Aquat. Sci.* 52:2078-2083.

API (American Petroleum Institute). 1994. *Evaluation of Sediment Toxicity Tests for Biomonitoring Programs*. API Publication 4608, American Petroleum Institute, January 1, 1994.

ASTM (American Society for Testing and Materials) International. 2006. *Standard Guide for Selection of Resident Species as Test Organisms for Aquatic and Sediment Toxicity Tests*. Method E1850 - 04. In: 2006 Annual Book of ASTM Standards, Biological Effects and Environmental Fate; Biotechnology, Volume 11.06. ASTM International, West Conshohocken, PA.

Bairda, D.J. and P.J. Van den Brink. 2007. Using biological traits to predict species sensitivity to toxic substances. *Ecotoxicology and Environmental Safety* 67(2):296-301.

Bay, S., D. Greenstein and D. Young. 2007. *Evaluation of Methods for Measuring Sediment Toxicity in California Bays and Estuaries*. Southern California Coastal Water Research Project, Costa Mesa, CA. March 2007. Technical Report 503. Available at: www.sccwrp.org

Boardman, G.D., S.M. Starbuck, D.B. Hudgins, X. Li, and D.D. Kuhn. 2004. Toxicity of ammonia to three marine fish and three marine invertebrates. *Environ. Toxicol.* 19:134-142.

Borgman, U. 1994. Chronic toxicity of ammonia to the amphipod *Hyalella azteca*; Importance of ammonium ion and water hardness. *Environmental Pollution* 86: 329-335.

Borgmann, U, 1996. Systematic analysis of aqueous ion requirements of *Hyalella azteca*: *Arch. Environ. Contam. Toxicol.* 30:356–363.

Borgmann, U. And A.I. Borgmann. 1997. Control of ammonia toxicity to *Hyalella azteca* by sodium, potassium and pH. *Envrionmental Pollution* 95(3):325-331.

Broderius, S.J., L.L. Smith Jr, and D.T. Lind. 1977. Relative toxicity of free cyanide and dissolved sulfide forms to the fathead minnow *Pimephales promelas*. *J. Fish. Res. Board Canada* 34:2323-2332.

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.

California State Water Resources Control Board. 2005. *Sediment Quality Objectives for California Enclosed Bays and Estuaries – Development of Toxicity Indicators*. Bay Protection and Toxic Cleanup Program, Presentation to Scientific Steering Committee Meeting, July 26, 2005.

Campiche, S. 2009. Alternatives to Classical Reproduction and Mortality Tests to Assess the Effects of Toxic Substances on Soil Organisms. Swiss Federal Institute of Aquatic Science and Technology and École Polytechnique Fédérale de Lausanne. May 2, 2009. Available at: http://www.soil.ch/doku/verein_jt09_campiche_soil_ecotoxicologie.pdf

Carr, R.S., M. Nipper, W.J. Adams, W.J. Berry, G.A. Burton Jr, K. Ho, D. MacDonald, R. Scroggins, and 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: Biological, Chemical, and Ecological Considerations with a Review of Methods and Applications, and Recommendations for Future Areas of Research; 18–22 March 2000; Pensacola, FL. Society of Environmental Toxicology and Chemistry (SETAC). Pensacola, FL. 38 pp. Available at: http://www.setac.org/htdocs/files/PWSummary.pdf

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.

CCAC (Canadian Council of Animal Care). 2008. Facts & Figures - CCAC Animal Use Survey. January 2004. Available at: www.ccac.ca

CCME (Canadian Council of Ministers of the Environment). 1993a. Guidance Manual on Sampling, Analysis and Data Management for Contaminated Sites, Volume I: Main Report. Report CCME EPC-NCS62E. CCME, Winnipeg. December 1993.

CCME. 1993b. Guidance Manual on Sampling, Analysis, and Data Management for Contaminated Sites, Volume II: Analytical Method Summaries. CCME, Winnipeg.

CCME. 1996. *A Framework for Ecological Risk Assessment: General Guidance*. Canadian Council of Ministers of the Environment. Winnipeg.

CCME. 1997. *A Framework for Ecological Risk Assessment: Technical Appendices*. Canadian Council of Ministers of the Environment. Winnipeg.

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

Chapman, P.M. 2000. The Sediment Quality Triad: then, now and tomorrow. *International Journal of Environment and Pollution*. 13:351-356.

Chapman, P.M., F. Wang, J.D. Germano, and 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, and G.S. Lawrence. 2002b. Weight-of-Evidence Issues and Frameworks for Sediment Quality (And Other) Assessments. *Human and Ecological Risk Assessment* 8(7):1489 – 1515.

Chappie, D.J. and G.A. Burton, Jr. 1997. Optimization of in situ bioassays with *Hyalella azteca* and *Chironomus tentans*. *Environ. Toxicol. Chem*. 16:559-564.

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

Canada Ontario Agreement (COA). 2008. *Canada-Ontario Decision-Making Framework for Assessment of Great Lakes Contaminated Sediment*. Prepared by Peter Chapman (Golder Associates Ltd.) with the COA Sediment Task Group on behalf of Environment Canada and the Ontario Ministry of the Environment under the Canada Ontario Agreement. Available at: http://www.on.ec.gc.ca/greatlakes/FE582241-DAC0-4B93-907E-5615A5901657/Canada-Ontario-Sediment-Framework-En.pdf

Devinny, J.S. and L.A. Volse. 1978. Effects of sediments on the development of *Macrocystis pyrifera* gametophytes. *Marine Biology* 48:343-348.

Dillon, T.M., D.W. Moore and A.B. Gibson. 1993. Development of a chronic sublethal bioassay for evaluating contaminated sediment with the marine polychaete worm *Nereis* (*Neanthes*) *arenaceodentata*. *Environmental Toxicology and Chemistry* 12:589-605.

Efroymson, R. and G.W. Suter II. 1999. Finding a niche for soil microbial toxicity tests in ecological risk assessment. *Human and Ecological Risk Assessment* 5(4):635-868.

Elphick, J., H. Bailey, K. Bergmann, and M. Liebl. 2004. Modifications to the sediment bivalve larval development test to reduce testing artifacts associated with suspended particulates. Presented at Fourth SETAC World Congress, 25th Annual Meeting in North America, 14-18 November 2004. AMEC Earth and Environmental.

Environment Canada. 1994. *Guidance Document on Collection and Preparation of Sediments for Physicochemical Characterization and Biological Testing*. Environmental Protection Series, Report EPS 1/RM/29, December 1994. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

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. 2007. *Guidance Document on Statistical Methods for Environmental Toxicity Tests*. Environmental Protection Series, Report EPS 1/RM/46, March 2005, with June 2007 amendments. Environment Canada, Method Development and Application Section, Environmental Technology Centre, Ottawa, ON.

Grapentine, L., J. Anderson, D. Boyd; G.A. Burton, C. DeBarros, G. Johnson, C. Marvin, D. Milani, S. Painter, T. Pascoe, T. Reynoldson, L. Richman, K. Solomon, and P.M. Chapman. 2002. A Decision Making Framework for Sediment Assessment Developed for the Great Lakes. *Human and Ecological Risk Assessment* 8(7): 1641 – 1655.

Jop, K.M., A.M. Askew, and 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.

Knezovich, J.P., J.A. Jelinski, S.L. Anderson and D.J. Steichen. 1994. Sulfide tolerance of four marine bioassay test species used to evaluate sediment and porewater toxicity. *Bull. Environ. Contam. Toxicol.* 57:450-457.

Kohn, N.P., J.Q. Word, D.K. Niyogi, L.T. Ross, T. Dillon, and D.W. Moore. 1994. Acute Toxicity of Ammonia to Four Species of Marine Amphipod. *Marine Environmental Research* 38:1-15.

Kszos, L. A., A. J. Stewart, and P. A. Taylor. 1992. An evaluation of nickel toxicity to *Ceriodaphnia dubia* and *Daphnia magna* in a contaminated stream and in laboratory tests. *Environ Toxicol. Chem.* 11:1001-1012.

Kuester E., D. Falk, and R. Altenburger. 2005. Effects of hydrogen sulfide to *Vibrio fischeri*, *Scenedesmus vaculatus* and *Daphnia magna*. *Environmental Toxicology and Chemistry* 24:2621-2629.

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

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.

Losso, C., A.A. Novelli, M. Picone, D. Marchetto, C. Pantani, P.F. Ghetti, and A.V. Ghirandini. 2007. Potential role of sulfide and ammonia as confounding factors in elutriate toxicity bioassays with early life stages of sea urchins and bivalves. *Ecotoxicology and Environmental Safety* 66:252-257.

McDonald, B.G. 2005. Comparison of porewater and elutriate bivalve larval development toxicity testing in a sediment quality triad framework. *Ecotoxicology and Environmental Safety* 62:383-390.

McDonald, B.G. and P.A. Haynes. 2001. Silica sand as an artificial control sediment in a 20-day *Neanthes arenaceodentata* toxicity test. *Environ. Toxicol.* 16:172-176.

Menzie, C.A., M. Hope Henning, J.J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts

weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risks. *Human and Ecological Risk Assessment*: 2(2):277-304.

Mineau, P., R. Balcomb, R. Bennett, S. Dobson, M. Fry, M. Jaber, A. Leopold, R. Munk, B. Ringer, A. Rispin, L. Sileo, R. Solecki, and H. Thompson. 1996. *Testing for effects on reproduction. In Report of the SETAC/OECD Workshop on Avian Toxicity Testing*. Inter-Organizational Programme for the Sound Management of Chemicals, OECD Environmental Health and Safety Publications, Series on Testing and Assessment. No. 5, 44-62.

Mineau, P., D.C. Boersma, and B. Collins. 1994. An analysis of avian reproduction studies submitted for pesticide registration. *Ecotoxicology and Environmental Safety* 29:304-329.

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.

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

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

Natal-da-Luz, T, J. Römbke, and J.P. Sousa. 2008. Avoidance tests in site-specific risk assessment--influence of soil properties on the avoidance response of *Collembola* and earthworms. *Environmental Toxicology and Chemistry* 27(5):1112-1117.

Nebeker, A.V. and G.S. Schuytema. 2000. Effects of ammonium sulfate on growth of larval Northwestern salamanders, red-legged and Pacific treefrog tadpoles, and juvenile fathead minnows. *Bulletin of Environmental Contamination and Toxicology* 64(2): 271-278.

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.

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.

OECD (Organization of Economic Cooperation and Development). 2001. Guidance for Determing the Quality of Data for SIDS Dossiers (Reliablity, Relevance and Adequacy). Organization for Economic Cooperation and Development, Paris, France. September 18, 2001. Available at: http://www.oecd.org/dataoecd/62/7/1900131.pdf

Oseid, D.M. and L.L. Smith, Jr. 1975. Long term effects of hydrogen sulphide on *Hexagenia limbata* (Ephemeroptera). *Environmental Entomology* 4(1):15-18.

Pack, S. 1993. *A Review of Statistical Data Analysis and Experimental Design in OECD Aquatic Toxicology Test Guidelines*. Organization of Economic Cooperation and Development. Paris, France.

Pastorok, R.A., and D.S. Becker. 1990. Comparative sensitivity of sediment toxicity bioassays at three Superfund sites in Puget Sound. In: *Aquatic Toxicology and Risk Assessment*, W.G. Landis and W.H. van der Schalie, eds. American Society for Testing and Materials, Philadelphia.

Patyna, P.J., R.A. Davib, T.F. Parkertonb, R.P. Brownb and K.R. Coopera. 1999. A proposed multigeneration protocol for Japanese medaka (*Oryzias latipes*) to evaluate effects of endocrine disruptors. *Sci. Tot. Environ.* 233(1-3):211-220.

Phillips, B.M., P.A. Nicely, J.W. Hunt, B.S. Anderson, R.S. Tjeerdema, and F.H. Palmer. 2005. Tolerance of five west coast marine toxicity test organisms to ammonia. *Bull. Environ. Contam. Toxicol.* 75:23-27.

Rand, G.M. 1995. *Fundamentals of Aquatic Toxicology: Effects, Environmental Fate, and Risk Assessment.* Taylor & Francis. 1125 pp.

Ren, S. and P.D. Frymier PD. 2003. Use of multidimensional scaling in the selection of wastewater toxicity test battery components. *Water Res.* 37(7):1655-61.

Reynoldson, T.B., S.P. Thompson, and D. Milani. 2002. Integrating Multiple Toxicological Endpoints in a Decision-Making Framework for Contaminated Sediments. *Human and Ecological Risk Assessment* 8(7):1569 – 1584.

Ringwood, A.H., M.E. DeLorenzo, P.E. Ross, and A.F. Holland. 1997. Interpretation of Microtox Solid-phase Toxicity Tests: the Effects of Sediment Composition. *Environ. Toxicol. Chem.* 16(6):1135-1140.

SAIC (Science Applications International Corporation). 1993. *Refinements of Current PSDDA Bioassays: Final Report Summary*. Prepared for the US Environmental Protection Agency, Cincinnati, OH by SAIC, Bothell, WA.

Schubauer-Berigan MK, Monson PD, West CW, Ankley GT. 1995. Influence of pH on the toxicity of ammonia to *Chironomus tentans* and *Lumbriculus variegatus*. *Environmental Toxicology and Chemistry* 14:713-717.

Schubauer-Berigan, M.K. and G.T. Ankley. 1991. The contribution of ammonia metals and nonpolar organic compounds to the toxicity of sediment interstitial water from an Illinois River tributary, USA. *Environmental Toxicology and Chemistry* 10(7):925-940.

Science Advisory Board (SAB) for Contaminated Sites in British Columbia. 2008. Guidance for Detailed Ecological Risk Assessments (DERA) in British Columbia. Submitted to the Science Advisory Board by Golder Associates. March 31, 2006. British Columbia, Canada.

Shellenberger, T.E. 1978. A Multi-Generation Toxicity Evaluation of p,p'-DDT and Dieldrin with Japanese Quail: I. Effects on Growth and Reproduction. *Drug and Chemical Toxicology* 1(2):137-146.

Shore, R.F., and B.A. Rattner (Eds.). 2001. *Ecotoxicology of Wild Mammals*. John Wiley & Sons, Ltd., Chichester, United Kingdom, 730 pp. ISBN 0-471-97429-3.

Sibley, P.K., D.A. Benoit, and G.T. Ankley. 1997. Life cycle and behavioural assessments of the influence of substrate particle size on *Chironomus tentans* (Diptera: Chironomidae) in laboratory assays. *Hydrobiologia* 361(8):1-9.

Suedel, B.C., and J.H. Rodgers, Jr. 1994. Responses of *Hyalella azteca* and *Chironomus tentans* to particle size distributions and organic matter content of formulated and natural sediments. *Environmental Toxicology and Chemistry* 13:1639-1648.

SSF (Sustainable Fisheries Foundation). 2007. Workshop to Support the Development of Guidance on the Assessment of Contaminated Sediments in British Columbia - Workshop Summary Report. Prepared for B.C. Ministry of Environment, Land Remediation Section, Victoria, B.C. December, 2007 by Sustainable Fisheries Foundation, Nanaimo, B.C.

Tay, K.L., K.G. Doe, A.J. MacDonald, and K. Lee. 1998. The Influence of Particle Size, Ammonia, and Sulfide on Toxicity of Dredged Materials for Ocean Disposal, p. 559-574, Chapter 39 in: *Microscale Testing in Aquatic Toxicology -- Advances, Techniques, and Practice*, P.G. Wells, K. Lee, and C. Blaise (eds.), CRC Press, New York, NY, USA. 679 pp.

USEPA (US Environmental Protection Agency). 1991. Method for Aquatic Toxicity Identification Evaluations, Phase I Toxicity Characterization Procedures. EPA/600/6-91/003. Office of Research and Development, Washington DC. February 1991. Available at: http://www.epa.gov/npdes/pubs/owm0330.pdf

USEPA. 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.

USEPA. 1993. *Methods for Aquatic Toxicity Identification Evaluations: Phase II Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity*. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN. 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*. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN. 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.

USEPA. 2001. *Clarifications Regarding Toxicity Reduction and Identification Evaluations in the National Pollutant Discharge Elimination System Program*. U.S. Environmental Protection Agency, Office of Wastewater Management, Office of Regulatory Enforcement, Washington, DC. March 27, 2001. Available at: http://www.epa.gov/npdes/pubs/owmfinaltretie.pdf

USEPA. 2007. *Sediment Toxicity Identification Evaluation (TIE) Phases I, II, and III Guidance Document*. Edited by K.T. Ho and R.M. Burgess (U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Atlantic Ecology Division, Narragansett, Rhode Island) and D.R. Mount, T.J. Norberg-King, and J.R. Hockett (U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, Duluth, MN). EPA/600/R-07/080. September 2007. Available at:

http://www.epa.gov/nheerl/publications/files/Sediment%20TIE%20Guidance%20Document.pdf

Van Den Hurk, P. 1994. Effect of natural sediment properties on test results in bioassay with oyster larvae (*Crassostrea gigas*) on sediment elutriates. *Journal of Aquatic Ecosystem Health* 3(3):185-191.

Vandenbergh, G.F., D. Adriaens, T. Verslycke and C.R. Janssen. 2003. Effects of 17α-ethinylestradiol on sexual development of the amphipod *Hyalella azteca*. *Ecotox. Environ. Saf.* 54(2):216-222.

Vogt, C. C. Nowak, J.B. Diogo, M. Oetken, K. Schwenk and J. Oehlmann. 2007. Multi-generation studies with Chironomus riparius – Effects of low tributyltin concentrations on life history parameters and genetic diversity. *Chemosphere* 67(11):2192-2200.

Wenning, R.J. (Ed.). 2005. *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments: Proceedings from the Pellston Workshop on Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments*, 18-22 August 2002, Fairmont, Montana, USA. Society of Environmental Toxicology and Chemistry. Published by SETAC, 2005.

Word, J.Q., W.W. Gardiner, and D.W. Moore. 2005. Influence of Confounding Factors on SQGs and their Application to Marine and Estuarine Sediment Evaluation. Chapter 16 in: Wenning, R.J. (Ed.). 2005. *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments: Proceedings from the Pellston Workshop on Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments*, 18-22 August 2002, Fairmont, Montana, USA. Society of Environmental Toxicology and Chemistry. Published by SETAC, 2005.

Zajdik, B. 2007. *Statistical Software Feasibility Study and Improvement Project*. Zajdlik & Associates Inc. Prepared for: Biological Methods Division, Environment Canada and Members of the Statistical Working Group.

Acknowledgments

Environment Canada would like to thank the many reviewers in government and private sector in Canada and the United States who provided valuable comments as part of the public peer review process.

www.ec.gc.ca

Additional information can be obtained at:

Environment Canada
Inquiry Centre
10 Wellington Street, 23rd Floor
Gatineau QC K1A 0H3
Telephone: 1-800-668-6767 (in Canada only) or 819-997-2800
Fax: 819-994-1412
TTY: 819-994-0736
Email: enviroinfo@ec.gc.ca