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
PO Box 9342 Stn Prov Govt
Victoria, B.C. V8W 9M1

Report Prepared - December, 2007 - by:

Sustainable Fisheries Foundation
#24 - 4800 Island Highway North
Nanaimo, B.C. V9T 1W6

Sustainable Fisheries Foundation

Building Partnerships for the Future

Funding for this Workshop provided through a Grant from the British Columbia Ministry of Environment

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
PO Box 9342 Stn Prov Govt
Victoria, B.C. V8W 9M1

Report Prepared - December, 2007 - by:

Sustainable Fisheries Foundation
#24 - 4800 Island Highway North
Nanaimo, B.C. V9T 1W6

MESL-SFF-SEDWK-1107-V2

Table of Contents

List of Tables

List of Figures

List of Appendices

Acknowledgements

This document was prepared by Don MacDonald and Dawn Smorong, on behalf of the Sustainable Fisheries Foundation (SFF). The authors would like to take this opportunity to thank everyone who attended the workshop and shared their insights on the selection of toxicity tests, on the interpretation of toxicity test results, and on the integration of multiple lines-of-evidence in sediment risk assessment. The free exchange of ideas and willingness to work towards consensus greatly simplified the process of developing relevant recommendations. The SFF would also like to thank Remi Odense, Mike MacFarlane, and Glenn Harris (B.C. Ministry of Environment), Chris Ingersoll (U.S. Geological Survey), Beth Power (Azimuth Consulting), Steve Bay (Southern California Coastal Water Research Project), Dave Mount (U.S. Environmental Protection Agency), and Jay Field (National Oceanic and Atmospheric Administration) for their excellent input on the design of the workshop. Helpful and timely comments on this document were provided by Chris Ingersoll (U.S. Geological Survey), Beth Power (Azimuth Consulting), Dave Mount (U.S. Environmental Protection Agency), Doug Bright (UMA Engineering Ltd.), Gary Lawrence (Golder Associates), Jay Field (National Oceanic and Atmospheric Administration), Scott Carr (U.S. Geological Survey), and Deanna Lee (Environment Canada). Funding for this workshop was provided by the B.C. Ministry of the Environment, through a grant to the SFF.

Executive Summary

On September 24-26, 2007, the Sustainable Fisheries Foundation (on behalf of the B.C. Ministry of the Environment) convened a Workshop to Support the Development of Guidance on the Assessment of Contaminated Sediments in British Columbia. The workshop, which was convened in Victoria, B.C., brought together a total of 29 participants representing toxicologists, risk assessors, and regulators from British Columbia and elsewhere in North America. Workshop participants were challenged with the task of developing recommendations on:

- The selection of selecting whole-sediment and pore-water toxicity tests for evaluating risks to aquatic receptors associated with exposure to contaminated sediments;
- The interpretation of the whole-sediment and pore-water toxicity tests for evaluating risks to aquatic receptors associated with exposure to contaminated sediments; and,
- The integration of information on multiple endpoints and multiple lines-of-evidence (LOEs) to obtain a weight-of-evidence (WOE) for assessing risks to aquatic receptors associated with exposure to contaminated sediments.

The workshop included a series of plenary sessions to provide workshop participants with a common understanding of the issues that needed to be addressed and the experience that has been gained by various experts in the field. After each plenary session, a work group session was convened to provide participants with an opportunity to share information, discuss technical issues, and develop recommendations regarding the three key topics. To guide these discussions, a series of focus questions were developed and posed to the work group participants (see Appendix 1 for a list of focus question questions posed for each work group to discuss). Each of the work group sessions was facilitated to ensure that all participants had an opportunity to share their views and contribute to the development of the workshop recommendations (See Appendix 2 for a list of participants in each work group). Each group independently elected a rapporteur to capture the results and recommendations that emerged from the discussions. After each work group discussion, each work group presented these findings to all workshop participants and a short question and answer period was conducted. Consensus (i.e., general agreement) among the workshop participants was considered to have been reached if two or more of the work groups developed similar recommendations.

Relative to the selection of toxicity tests (Work Group Session 1), workshop participants recognized that a tiered-assessment framework is used to evaluate contaminated sediments in British Columbia. This tiered framework relies upon comparisons of whole-sediment chemistry data to the numerical sediment quality criteria in the first tier. In the second tier, whole-sediment toxicity testing can be used, in conjunction with information on other lines-of-evidence (LOEs), to assess risks to sediment-dwelling organisms associated with exposure to contaminated sediments. It was further recognized that data from toxicity identification evaluations, sediment-spiking studies, whole-sediment bioaccumulation tests, and/or benthic invertebrate community structure assessments could be used to further inform sediment management decisions at complicated sites. In general, it was generally recognized that the weight-of-evidence (WOE) considered should reflect the weight of the decision at sites with contaminated sediments in the province.

Workshop participants generally agreed that a suite of whole-sediment toxicity tests should be applied to assess contaminated sediments in British Columbia. The following suite of toxicity tests was recommended for assessing sediment quality conditions at all freshwater sites (termed core toxicity tests):

- 10-d whole-sediment toxicity tests with the midge, *Chironomus dilutus* or *C. riparius* (Endpoints - survival and growth);
- 10-d to 14-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoint - survival); and/or,
- 28-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoints - survival and growth);

The 28-d toxicity test with the amphipod, *Hyalella azteca*, was considered to be a better choice than the short-term exposures with amphipods under most circumstances. For marine and estuarine sites, the following suite of core toxicity tests was recommended:

- 10-d whole-sediment toxicity tests with the amphipod, *Eohaustorius estuarius* or *Rhepoxynius abonius* (Endpoint - survival); and,
- Sublethal toxicity tests, which should be selected based on characteristics of the site under investigation.

It was understood that, currently, such sublethal toxicity tests are largely limited to 20- to 28-d toxicity tests with polychaetes or 28-d toxicity tests with bivalve mollusks. However, sediment-water interface tests with bivalves or echinoderms appear promising and could be applied when standard methods become available.

Workshop participants recognized that a number of approaches have been developed to interpret the results of whole-sediment toxicity tests with benthic invertebrates, including:

- Comparison to Negative Control – In this approach, the responses of test organisms exposed to site sediments are compared statistically to the responses of test organisms exposed to negative control sediments. Treatments that have responses that are statistically significantly different from those observed in the control treatment(s) would be designated as toxic;
- Minimum Significant Difference (MSD) – In this approach, the results of power analyses conducted with data from multiple studies for a specific toxicity test can be used to identify the MSD or minimum detectable difference (MDD) from the control treatment. Treatments with response levels greater than the MSD would be designated as toxic;
- Reference Envelope – In this approach, sediment samples from a number of reference sites are collected and tested using the same toxicity tests that are applied to the site. The results of the toxicity testing conducted on these samples can be used to develop a reference envelope (i.e., normal range of responses of test organisms exposed to reference sediments, as defined by ASTM 2006b). Sediment samples with response levels that fall outside the normal range of responses would be designated as toxic; and,
- Multiple Category Approach – In this approach, sediment samples are classified into various groups (e.g., not toxic, low toxicity, moderate toxicity, or high toxicity), based on the magnitude of the observed response and the statistical significance that is determined.

All of the work groups recognized that the results of individual toxicity tests may be used within a weight-of evidence (WOE) framework for evaluating risks to the benthic invertebrate community associated with exposure to contaminated sediments. In this context, it is reasonable to utilize one or more of the first three approaches to designate sediment samples as toxic and not toxic. Workshop participants also generally agreed that such WOE evaluations require information on the magnitude of toxicity in addition to, or instead of, toxicity designation information. Hence, it was generally agreed that the information on the magnitude of the response be retained to support further analyses of the toxicity data (i.e., WOE evaluations). The multiple category approach was considered to be useful in this respect.

While WOE approaches can be defined in various ways, workshop participants generally agreed that a WOE approach is: *A tool or mechanism to improve understanding of, interpretation of, and inferences to be drawn from multiple LOEs to inform recommendations to be made by risk assessors to risk managers and site managers.* Such WOE assessments facilitate prioritization of concerns relative to the risks posed by contaminated sediments and improve the confidence that can be placed in decisions regarding the management of contaminated sediments. By integrating information from multiple LOEs to assess risks to ecological receptors, WOE assessments provide a basis for identifying key stressors at a site, determining if something needs to be done to manage contaminated sediments, and, if so, where such remedial activities should be focused. Workshop participants identified a number of approaches that could be used to integrate multiple LOEs to assess risks to benthic invertebrates, including:

- Best professional judgement approach;
- Tiered approaches;
- Decision matrix approach;
- Semi-quantitative approach; and,
- Fully quantitative approach.

While each of the work groups evaluated and/or ranked the candidate WOE approaches in terms of their applicability for assessing and managing contaminated sediments in British Columbia, it was generally agreed that BCMOE should not establish prescriptive guidance on the selection of WOE approaches. Rather, practitioners should be afforded the flexibility to select the WOE approach that is most appropriate for integrating the types of data and information that were collected at a site. In addition, workshop participants developed a series of guiding principles that should be used to identify the most appropriate methods for integrating multiple LOEs at sites with contaminated sediments in British Columbia. The results of the work group discusions on all three of the topics addressed during the workshop are summarized in this document and detailed in Appendix 3, 4, and 5.

1.0 Introduction

There are a substantial number of sites with contaminated sediments in British Columbia. While numerical sediment quality criteria are available to support assessment and management of sediments at these contaminated sites, responsible persons also have the option under the *Contaminated Sites Regulation* (CSR) of applying risk-based standards to support decisions on the management of contaminated sediments. Development of such risk-based standards necessitates the design and implementation of detailed ecological risk assessments (DERA), which are intended to provide a basis for evaluating risks to ecological receptors associated with exposure to contaminated sediments and identifying the concentrations of chemicals of potential concern (COPCs) that pose tolerable risks to ecological receptors.

The B.C. Ministry of the Environment (BCMOE) has not yet promulgated specific guidance on ecological risk assessment at sediment contaminated sites. While the Science Advisory Board (SAB) is currently in the process of developing guidance on DERA, the draft documentation does not provide practitioners with specific guidance on procedures for assessing risks to ecological receptors associated with exposure to contaminated sediments. As a result, risk assessors in the province are currently applying a broad range of approaches to acquire the data and information needed to complete these assessments. In reviewing the results of these assessments, Ministry staff have identified three key issues that need to be addressed in the near term to achieve greater consistency in the ecological risk assessments that are being submitted on behalf of responsible persons, as follows:

1. Guidance is required on the selection of whole-sediment and pore-water toxicity tests for evaluating risks to aquatic receptors associated with exposure to contaminated sediments.
2. Guidance is required on the interpretation of the results of the whole-sediment and pore-water toxicity tests for evaluating risks to aquatic receptors.
3. Guidance is required on the integration of data and information on multiple lines-of-evidence (LOEs) to obtain a weight-of-evidence (WOE) for

assessing risks to aquatic receptors associated with exposure to contaminated sediments.

To address these issues, a workshop was convened during September 24 to 26, 2007. The workshop was designed to engage regulators, practitioners, and other experts in focussed discussions and debates on these topics, leading to the development of recommendations that can be used to establish BCMOE guidance on these issues. As is the case with all guidance that is issued by the BCMOE, compliance with the recommendations that are presented in this report, or any Ministry guidance that emerges from them, is not mandatory. However, responsible persons and contaminated sites professionals should support any deviations from the recommendations/guidance with defensible and documented rationale.

The results of the workshop are summarized in Sections 3, 4, and 5 of this document; Section 2 is provided as background information. It should be noted that the workshop discussions were largely constrained to issues relating to the assessment of risks to sediment-dwelling organisms; issues related to other ecological receptors and human health were generally not discussed.

2.0 Framework for Assessing and Managing Sites with Contaminated Sites in British Columbia

The procedures for assessing and remediating contaminated sites that fall under provincial jurisdiction are specified in two components of the *Environmental Management Act*, including the CSR and the Special/Hazardous Waste Regulation. The site management process proposed under the CSR is intended to establish a framework for assessing and remediating contaminated sites in the province. The process consists of five main elements, including site identification and screening; site investigation and determination/decision; planning; remediation; and, monitoring and evaluation. However, every site need not proceed through each component of the process (MacDonald and Ingersoll 2003a; 2003b). It is important to note that sites with contaminated sediments are not necessarily defined by the legal boundaries of a property and more commonly are defined by the spatial extent of contamination originating from a property. The following summary provides an overview of the existing contaminated site management framework (Figure 1). More detailed information on the elements of this framework is included in the CSR and in a series of associated Fact Sheets that have been published by the Ministry (see http://www.env.gov.bc.ca/epdiv/).

An overview of the framework is presented here to provide context for the balance of the workshop summary report. This framework was not a product of the workshop, however.

2.1 Site Identification and Screening

The first step in the site assessment and management process involves screening the site under consideration. This step in the process is initiated through the preparation of a site profile. In most jurisdictions in British Columbia, site profiles must be submitted to the responsible government agency when an application for subdivision, zoning, development, demolition of a structure, or removal of soil is received by a local government or when ordered by a regional manager. Following its submission, the site profile is assessed by provincial or local government official and a determination is made regarding the need for further investigations at the site. No further action is required at sites that are considered not to be potentially contaminated.

2.2 Site Investigation and Determination

Information from the site profile or other sources may indicate that a site is potentially contaminated. In this situation, preliminary site investigations (PSIs) and/or detailed site investigations (DSIs) may be required to determine if the site is contaminated, as defined under the CSR. Initially, a Stage I PSI is conducted to determine the probability that a site is contaminated. This assessment is conducted using archival records, conducting site visits, and relying on knowledge of the historical activities that were conducted on site. Next, a Stage II PSI or a DSI is conducted to provide the additional information needed to confirm or refute the potential for site contamination, primarily by sampling and chemical analysis of environmental media. The results of the Stage II PSI and/or DSI are used to determine if a site is contaminated, as defined under the CSR. For sites with contaminated sediments, the measured concentrations of COPCs in whole sediments from the site are compared to the numerical sediment quality criteria (SedQC) to determine if the site is contaminated. The determination of whether a site is contaminated (as defined under the CSR) is generally made by the Director. Detailed guidance on the information requirements for PSIs and DSIs for sediment-contaminated sites is provided in MacDonald and Ingersoll (2003a; 2003b; 2003c).

2.3 Site Management Planning

Assuming the site poses unacceptable risks, the first priority in the planning stage of the site management process is to determine who is potentially responsible for the contamination and who is potentially liable for clean-up costs. In addition, the need for and relative priority for remediation is assessed at this stage of the process. Other important planning steps include activating the remediation process (either through a voluntary remediation agreement or a remediation order), developing a remediation action plan, and initiating the approvals process.

The development of sediment quality standards (SedQS) represents another important step in the site management planning process. The legislation provides for the use of two distinct approaches to the establishment of SedQS at sediment contaminated sites, including the criteria-based approach and the risk-based approach. Using the criteria-based approach, SedQS may be established by directly adopting the SedQC or by deriving site-specific SedQS (e.g., by adjusting the SedQC to reflect site-specific conditions, such as elevated background levels of one or more COPCs; see MacDonald and Ingersoll 2003a; 2003b; 2003c; and Ingersoll and MacDonald 2003 for more information on the derivation of site-specific SedQCs). By comparison, risk-based SedQSs can be established at risk levels that are less than or equal to those upon which the sediment quality criteria are based (i.e., a 20% probability of an EC_{20}

or greater for sensitive sites and a 50% probability of an EC_{20} or greater for typical sites). Such numerical or risk-based standards can be used for determining if remedial measures are required at the site and if they have been satisfactorily completed. In some cases, risks posed by contaminated sediments are simply designated as acceptable of unacceptable. Such designations are then used to make decisions regarding management of the site.

2.4 Remediation

The remediation step in the process covers all of the activities that are associated with cleaning-up or securing a contaminated site. The legislation defines two broad types of remediation, including removal of contaminated materials (so that they no longer remain at the site) and treatment of the contaminated materials on-site. The legislation also provides environmental quality standards that are used to determine when the cleanup is complete. Alternatively, risk-based procedures may be used to determine the level of contamination that can remain on-site. In such situations, additional institutional controls may have to be established to assure that specific uses of the site and nearby areas are not unacceptably impacted.

2.5 Monitoring and Evaluation

Following the implementation of remedial measures, confirmatory sampling and analysis is normally conducted to determine if they have reduced the level of contamination or risk to tolerable levels. A Certificate of Compliance (C of C) is issued if the numerical standards in the CSR have been satisfied or if risk-based standards and related assessment procedures have been appropriately applied. When the contamination is managed on-site, certain conditions must be met by the site manager. Such conditions are generally established to assure the protection of the environment and human health, or the notification of potentially affected parties (e.g., off-site parties, future site owners).

3.0 Workshop Goals and Objectives

The goal of the workshop was to develop guidance for assessing contaminated sediments in British Columbia. More specifically, guidance is required on:

1. The selection of toxicity tests;
2. The interpretation of toxicity tests; and,
3. The application of WOE assessments in risk characterization.

To complement these goals, the following workshop objectives were established:

- Organize a workshop in the fall of 2007;
- Encourage experts to share technical information and engage in a series of technical discussions on the selection of toxicity tests for assessing freshwater, estuarine, and marine sediments, on the interpretation of toxicity test results, and on the integration of toxicity test results with other LOEs for assessing risks to ecological receptors; and,
- Prepare a workshop summary report that summarizes the results of the technical discussions and documents the key recommendations emerging from the workshop.

This report provides a brief summary of the results of the workshop on the assessment of contaminated sediments in British Columbia.

4.0 Approach to Addressing the Key Issues Related to the Assessment and Management of Contaminated Sediments

A variety of approaches could be used to develop guidance on the selection of whole-sediment and/or pore-water toxicity tests, on the interpretation of toxicity test results, and on the integration of multiple LOEs to assess risks to ecological receptors associated with exposure to contaminated sediments. To expedite the development of such guidance, the Sustainable Fisheries Foundation, in partnership with the BCMOE, organized and convened an expert's workshop on the assessment of contaminated sediments in British Columbia, which involved:

- Identification of a group of internationally-recognized experts in the fields of sediment quality assessment, ecological risk assessment, and contaminated sites assessment and management;
- Providing workshop participants with access to salient background information on toxicity test selection, on toxicity test interpretation, on WOE evaluation, and on the framework for assessing and managing contaminated sites in British Columbia (i.e., through pre-workshop reading assignments and technical presentations during the workshop);
- Engaging workshop participants in focussed work group discussions regarding the three key topics addressed at the workshop. The focus questions that were posed to workshop participants were included in the workshop agenda (Appendix 1). A total of three work groups were established to discuss the topics identified for each of the three work group sessions. The individuals who were assigned to each of the work groups are identified in Table 1;
- Encouraging each of the work groups to share the results of their discussions during the plenary sessions that were convened following each work group session. Workshop participants were then encouraged to reflect on the results presented by the three work groups during these plenary sessions; and,
- Summarizing the results of the workshop in a brief report to BCMOE. In summarizing these results, agreement among two or three of the work groups on a specific topic was considered to provide general agreement from the workshop participants. Hence, the recommendations provided in Section 5 of this document are considered to represent general agreement by the workshop participants. In contrast, the recommendations provided in Section 6 of this document were prepared by the Sustainable Fisheries Foundation and may or may not represent general agreement by the workshop participants.

Workshop attendees were selected based on their knowledge and extensive experience in the fields of sediment quality assessment and ecological risk assessment. In addition, participants were selected to achieve balanced representation by the private sector and government agencies. A total of 30 individuals attended the workshop. A complete list of workshop attendees and their affiliations is provided in Appendix 2.

5.0 Results and Discussion

This workshop was designed to provide a basis for the developing recommendations to BCMOE on the assessment of contaminated sediments in British Columbia. More specifically, the workshop was intended to provide a basis for developing general agreement on recommendations for:

- Selection of whole-sediment and/or pore-water toxicity tests for evaluating risks to benthic invertebrates associated with exposure to contaminated sediments;
- Interpretation of the results of whole-sediment and/or pore-water toxicity tests for evaluating risks to benthic invertebrates associated with exposure to contaminated sediments; and,
- Integration of information on multiple LOEs for evaluating risks to benthic invertebrates associated with exposure to contaminated sediments.

To generate such recommendations, workshop participants were asked a series of focus questions on these topics within the individual work group sessions. The presentations that were made by the work group rapporteurs were reviewed to identify the generally agreed upon recommendations that emerged from the work group sessions. Agreement among two or three of the work groups on a specific topic was considered to represent general agreement among the workshop participants. The key recommendations that emerged from the workshop are highlighted in the following sections of this report.

5.1 Selection of Whole-Sediment and/or Pore-Water Toxicity Tests

The BCMOE has not yet provided guidance on the selection of whole-sediment or pore-water toxicity tests for assessing contaminated sediments in British Columbia (i.e., at sites that are designated as contaminated under the CSR). As a result, ecological risk assessors in the province have selected a variety of whole-sediment and/or pore-water toxicity tests to evaluate the toxicity of freshwater, estuarine, and marine sediments in British Columbia. While such flexibility has provided practitioners in the field with the latitude to attempt to match the suite of toxicity tests to the conditions that prevail at the sites under investigation, application of different toxicity tests at various sites makes it difficult to compare results across sites (i.e., to determine if similar levels of protection are afforded to ecological receptors at sites across the province). In addition, uncertainty regarding the relative sensitivity of various test species and endpoints makes it difficult to determine the level of protection afforded to benthic invertebrates by sediment quality standards derived using the results of the selected toxicity tests. For these reasons, it may be desirable to establish a core suite of toxicity tests that should be applied at freshwater or at marine and estuarine sites with contaminated sediments, regardless of the conditions that prevail at the site.

Workshop participants were asked a series of questions to focus their discussions regarding the selection of whole-sediment and/or pore-water toxicity tests for assessing contaminated sediments in British Columbia (see Workshop Agenda for a complete list of the Focus Questions). In considering these focus questions, workshop participants recognized that a tiered-assessment framework is used to evaluate contaminated sediments in British Columbia. This tiered framework relies upon comparisons of whole-sediment chemistry data to the numerical sediment quality criteria in the first tier. In the second tier, whole-sediment toxicity testing can be used, in conjunction with information on other LOEs, to assess risks to sediment-dwelling organisms (and/or other ecological receptor groups) associated with exposure to contaminated sediments. Matching whole-sediment chemistry and whole-sediment toxicity data are often used to develop concentration-response relationships and associated site-specific sediment quality standards. It was further recognized that data from toxicity identification evaluations, sediment-spiking studies, whole-sediment bioaccumulation tests, and/or benthic invertebrate community structure assessments can also be used to help inform sediment management decisions at complicated sites. These latter data could be generated in the second or third tier of an assessment. In general, it was generally recognized that the weight-of-evidence (WOE) considered should reflect the weight of the decision (i.e., risks to human health, risks to ecological receptors, clean-up costs) at sites with contaminated sediments in the province.

In response to the first focus question, workshop participants identified a wide range of whole-sediment and pore-water toxicity tests that are candidates for evaluating the toxicity of freshwater, estuarine, and marine sediments in British Columbia. A listing of the toxicity tests that could be used to assess risks to benthic invertebrates in freshwater ecosystems is presented in Table 2, while the candidate toxicity tests for assessing marine and estuarine sediments are listed in Table 3.

Workshop participants used a variety of approaches to evaluate the strengths and limitations of the candidate toxicity tests that were identified. However, it was generally agreed that selection criteria that have been established previously by the American Society for Testing and Materials, U.S. Environmental Protection Agency, and others are relevant to evaluating candidate toxicity tests (see MacDonald and Ingersoll 2003b; 2003c for more information). Some of the selection criteria that were identified by workshop participants included:

- Ecological relevance;
- Sensitivity to COPCs;
- Sensitivity to physical variables;
- Reproducibility;
- Availability of test organisms;
- Field validation, contact with exposure medium;
- Taxonomic diversity;
- Variability of response; and,

- Method standardization.

Each of the work groups applied these or similar selection criteria, either formally (i.e., using a matrix approach) or informally (i.e., using a professional judgement approach) to identify the strengths and limitations of each of the candidate toxicity tests. Appendix 3A, 3B, and 3C, respectively, provide detailed information on the strengths and limitations of the candidate toxicity tests that were identified by the three work groups.

By integrating the information on the strengths and limitations of the candidate toxicity tests, work group participants recommended a suite of toxicity tests that should be applied at sites with contaminated sediments in British Columbia (i.e., core toxicity tests).

5.1.1 Toxicity Tests for Assessing Freshwater Sediments

Workshop participants identified a diverse array of whole-sediment and pore-water toxicity tests that could be used to evaluate contaminated sediments at freshwater sites in British Columbia (Table 2). Among all three work groups, it was generally recognized that 10- to 14-d and/or 28-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca*, (Endpoints - survival or survival and growth) represent an essential element of the recommended suite of toxicity tests for freshwater sites and, potentially, low salinity estuarine sites (i.e., *Hyalella* has been used successfully in toxicity tests conducted at salinities of up to 10 ppt). While amphipod reproduction can also be measured in 42-d exposures, experience at the USGS-Columbia Environmental Research Laboratory indicates that this endpoint is variable and tends to provide little information beyond that provided by the other two endpoints. It was also generally agreed that the results of 10-d whole-sediment toxicity tests with midge, *Chironomus dilutus* or *C. riparius*, (Endpoints - survival and growth) provide important information for evaluating freshwater sediments.

Workshop participants also noted that toxicity testing with other species, evaluating non-lethal endpoints over longer durations of exposure, can provide relevant information for assessing contaminated sediments in the province. For example, 20-

to 60-d whole-sediment toxicity tests with midge, *C. dilutus*, (Endpoints - survival, growth, and reproduction) were considered to be potentially relevant at certain types of sites (e.g., sites with sediments contaminated with pesticides or PCBs; i.e., chemicals that are more likely to cause chronic effects on reproduction than acute toxicity). In addition, whole-sediment toxicity tests with oligochaetes (*Tubifex tubifex* or *Lumbriculus variegatus*), mayflies (*Hexagenia limbata*), other amphipod species (e.g., *Diporea* sp.), and cladocerans (e.g., *Ceriodaphnia dubia*, *Daphnia magna*, *Chydorus* sp.) have been used in numerous investigations. However, difficulties associated with the availability of test organisms and/or lack of sensitivity of these species in whole-sediment exposures can limit their utility. Freshwater mussels (e.g., *Lampsilis siliquoidea*) were also considered to provide relevant information for assessing freshwater sediments at certain types of sites, particularly those that are contaminated by metals. Neither solid-phase nor aqueous-phase toxicity tests with the bacterium, *Vibrio fisheri* (i.e., Microtox) were recommended for assessing contaminated sediments at freshwater sites, however.

Overall, workshop participants generally agreed that a suite of whole-sediment toxicity tests should be applied to assess contaminated sediments at freshwater sites in British Columbia. This recommendation reflects the fact that species sensitivities to COPCs vary and that no one species is uniquely sensitive to all COPCs. Such a suite of toxicity tests should include:

- 10-d whole-sediment toxicity tests with the midge, *Chironomus dilutus* or *C. riparius* (Endpoints - survival, growth, and biomass);
- 10- to 14-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoint - survival); and/or,
- 28-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoints - survival, growth, and biomass).

In selecting a suite of toxicity tests from the three most highly recommended toxicity tests, workshop participants generally agreed that the suite should include the 10-d whole-sediment toxicity test with midge and either the 10- to 14-d or the 28-d whole-sediment toxicity test with amphipods. Of the two amphipod toxicity tests, the 28-d exposure was generally considered to be more useful due to its greater sensitivity to COPCs.

5.1.2 Toxicity Tests for Assessing Marine and Estuarine Sediments

Workshop participants identified a diverse array of whole-sediment and pore-water toxicity tests that could be used to evaluate contaminated sediments at marine and estuarine sites in British Columbia (Table 3). Among all three work groups, it was generally recognized that 10-d whole-sediment toxicity tests with marine and estuarine amphipods represent an essential element of the recommended suite of toxicity tests for marine and estuarine sites. While *Eohaustorius estuarius* and *Rhepoxynius abronius* were the most highly recommended species for conducting such toxicity tests, all three work groups noted that toxicity testing could be conducted using other amphipod species, considering additional endpoints (i.e., survival, growth, emergence, reburial, and reproduction) and exposure durations (i.e., up to 28-d for *Leptocheirus plumulosus*). It was noted that *Ampelisca* is a tube-dwelling species and, hence, may receive less exposure to COPCs than other amphipod species (potentially making it less sensitive to sediment-associated COPCs). Toxicity testing with amphipods was recommended because they tend to be sensitive species and their responses are often correlated with responses of the benthic community in the field.

Workshop participants also noted that toxicity testing with other species, evaluating non-lethal endpoints over longer durations of exposure can provide relevant information of assessing contaminated sediments in the province. For example, some workshop participants indicated that 20- to 28-d whole-sediment toxicity tests with polychaetes (e.g., *Neanthes arenoceodentata*; Endpoints - survival and growth) can provide useful information for assessing risks to benthic invertebrates associated with exposure to contaminated sediments. In addition, 48- to 96-hr sediment-water interface toxicity tests with echinoderm (e.g., *Arbacia punctulata*) or bivalve mollusc larvae (e.g., *Mytilus edulis*; Endpoint - development) were identified as emerging toxicity tests that could provide broader taxonomic coverage by the recommended suite of toxicity tests and reduce uncertainties associated with the traditional use of these species and life stages (i.e., in pore-water exposures). However, the work group participants did not achieve general agreement on whether or not such additional toxicity tests should be included in the suite of core toxicity tests for assessing marine and estuarine sites. The need for standardization of the sediment-water interface toxicity testing protocols was identified as one of the current limitations associated with applying these tests on a routine basis. Therefore, the relevance of such toxicity tests should be evaluated on a case-by-case basis to determine if one or more of these ancillary tests should be used to assess contaminated sediments at a site.

Workshop participants also achieved general agreement on several of the issues related to toxicity test selection. First, it was generally agreed that pore-water toxicity tests and elutriate toxicity tests should not be included in the core suite of tests that are applied at marine and estuarine sites in the province. Some of the challenges associated with the use of pore-water toxicity tests that were identified by workshop participants included responsiveness to hydrogen sulfide and ammonia, and depletion of hydrophobic organics during the course of the test. However, pending refinements to the standard methods for and/or better adherence to guidance on pore-water testing procedures may help to mitigate these challenges in the future. Elutriate toxicity tests were considered to be more relevant for assessing the effects of open-water disposal of dredged materials than evaluating the toxicity of in-place sediments. It was further generally agreed that *Neanthes arenoceodentata* should be selected for use in longer-term whole-sediment toxicity tests, if polychaete survival and growth is to be evaluated at a site with contaminated sediments. Most of the workshop participants generally agreed that the 28-d whole-sediment toxicity test with the amphipod, *Leptocheirus plumulosus*, would not provide useful information on sub-lethal responses to contaminated sediment exposures due to the high degree of variability that is typically observed in the test results. Similarly, neither solid-phase not aqueous-phase toxicity tests with the bacterium, *Vibrio fisheri* (i.e., Microtox) were recommended for assessing contaminated sediments at marine or estuarine sites as these tests provide an indication of exposure to contaminants rather than specific measures of effects on benthic organisms. However, there was little agreement on whether or not specific toxicity tests could be recommended for use in evaluated sediments that are contaminated with specific classes of COPCs.

Overall, workshop participants generally agreed that a suite of whole-sediment toxicity tests should be applied to assess contaminated sediments at marine and estuarine sites in British Columbia. Such a suite of toxicity tests should include:

- 10-d whole-sediment toxicity tests with the amphipod, *Eohaustorius estuarius* or *Rhepoxynius abronius* (Endpoint - survival); and,
- Sublethal test, which should be selected based on site-specific characteristics. In the near-term, the selection of such sublethal tests is restricted by the lack of standardization of the sediment-water interface tests and the limitations associated with pore-water tests. Therefore, sublethal toxicity testing is largely limited to 20 to 28-d whole-sediment toxicity tests with polychaetes; however, 28-d tests can be conducted with bivalves by refining the bioaccumulation test to evaluate survival and growth.

5.1.3 Selection of Toxicity and Bioaccumulation Tests for Other Receptor Groups

Workshop participants recognized that, while benthic invertebrates represent essential elements of aquatic ecosystems, other ecological receptors can also be exposed to contaminated sediments (e.g., microbiota, aquatic plants, fish, amphibians, reptiles, birds, and mammals). However, insufficient time was available to develop specific recommendations for toxicity tests that could be used to evaluate risks to these other receptor groups associated with exposure to contaminated sediments. In addition, bioaccumulation was identified as an essential element of sediment quality investigations at sites with bioaccumulative COPCs. However, time limitations prevented workshop participants from making specific recommendations relative to bioaccumulation testing.

5.1.4 Supplemental Toxicity Testing

Workshop participants generally agreed that management decisions can be enhanced by data and information that enables risk assessors and risk managers to identify the stressors that are causing or substantially contributing to the adverse effects that are observed in laboratory and/or field studies. For this reason, it was suggested that a tiered toxicity-testing framework could be developed for assessing contaminated sediments. The toxicity tests recommended in Section 5.1.1 and 5.1.2 could be selected for application in the first tier of the framework. Additional tools, such as sediment-spiking studies, toxicity identification evaluations (TIEs), and other procedures for evaluating COPC bioavailability, could be used in the second tier of the framework to provide further information to support sediment management decisions.

5.2 Interpretation of the Results of Whole-Sediment and/or Pore-Water Toxicity Tests

In British Columbia, whole-sediment chemistry data are used at the majority of sites to determine if a site has contaminated sediments and to support decisions regarding the management of contaminated sediments (i.e., using the criteria-based approach). However, responsible parties can also apply a risk-based approach to the assessment and management of contaminated sediments in British Columbia. Such risk-based approaches to assessing and managing risks to benthic invertebrates associated with exposure to contaminated sediments typically rely, to a greater or lesser extent, on the results of whole-sediment and/or pore-water toxicity tests.

A number of approaches can be used to interpret the results of whole-sediment toxicity tests with benthic invertebrates. For example, the responses of test organisms exposed to site sediments can be compared statistically to the responses of test organisms exposed to control sediments. Treatments that have responses that are significantly different from those observed in the control treatment(s) can be designated a toxic. Alternatively, the results of power analyses conducted with data from multiple studies for a specific toxicity tests can be used to identify the minimum significant difference (MSD) or minimum detectable difference (MDD) from the control treatment. Treatments with response levels greater than the MSD can be designated as toxic. Another approach to toxicity designation involves collection and testing of sediment samples from a number of reference sites. The results of the toxicity testing conducted on these samples can be used to develop a reference envelope (i.e., normal range of responses of test organisms exposed to reference sediments, as defined by ASTM 2006a). Sediment samples with response levels that fall outside the normal range of responses (e.g., survival below the 2.5th percentile for the reference samples) can be designated as toxic (MacDonald *et al.* 2002). Finally, a multiple category approach can be used to classify sediment samples into various groups (e.g., not toxic, low toxicity, moderate toxicity, or high toxicity), based on the magnitude of the observed response and the statistical significance that is determined.

Existing BCMOE guidance (Ingersoll and MacDonald 2003) recommends a procedure for designating samples as toxic or not toxic based on statistical comparisons to the negative control treatment(s) and comparison to the results obtained for reference sediment samples (i.e., reference envelope approach; see Figure 2 for an overview of these procedures).

While BCMOE has issued guidance on the interpretation of whole-sediment and pore-water toxicity tests (Ingersoll and MacDonald 2003), workshop participants generally agreed that further guidance was needed to support assessments of risks to benthic invertebrates exposed to contaminated sediments in the province (see Appendix 4A, 4B, and 4C for summaries of the work group discussions on this topic). More specifically, workshop participants recognized that it is desirable to maintain consistency in the interpretation of toxicity tests to ensure that the narrative objectives for managing sites with contaminated sediments are being respected at all sites and to facilitate comparisons of risks to benthic invertebrates between sites. The sediment management objectives (SMOs) that have been established by BCWLAP (2004) include:

- At sites with sensitive habitats, the principal SMOs are to restore sediment quality conditions to a state that will facilitate restoration of productive and diverse benthic invertebrate communities in the near-term and to minimize the risks to organisms at high trophic levels in the food web; and,
- At sites with typical habitats, the principal SMOs are to restore sediment quality conditions to a state that will facilitate restoration of productive and

diverse benthic invertebrate communities in the longer-term and to minimize the risks to organisms at high trophic levels in the food web.

Using these narrative SMOs, BCWALP (2004) has established numerical criteria for assessing sediment quality conditions at sensitive and typical contaminated sites in the province (see Schedule 2 of BCWALP 2004 for a list of factors that need to be considered in the selection of SQC for use at a site). The numerical criteria for sensitive sites are intended to define concentrations of COPCs below which there is a relatively low probability of observing statistically significant adverse effects in standardized toxicity tests with sensitive species and life stages (i.e., the incidence of toxicity would be ≤20%). By comparison, the numerical criteria for typical sites are intended to define concentrations of COPCs below which there is a moderate probability of observing statistically significant adverse effects in standardized toxicity tests with sensitive species and life stages (i.e., the incidence of toxicity would be 20 to 50%).

With this background information in mind, workshop participants were asked to describe how the results of whole-sediment and/or pore-water toxicity tests should be interpreted to support decisions on the management of contaminated sediments in British Columbia. In considering this question, workshop participants recognized that the results of toxicity tests can be used for assessing toxicity on a per-sample or a per-site basis. Workshop participants further noted that the approach that is used to interpret toxicity test results could differ, depending on the scope of the evaluation that is being conducted (see Appendix 4A, 4B, and 4C for summaries of the work group discussions).

Overall, it was generally agreed that site-wide ecological risk assessments represent the most important application of whole-sediment or pore-water toxicity data. More specifically, it was generally agreed that the results of the toxicity testing program that is implemented at a site should support the development of site-specific, risk-based sediment quality standards (i.e., numerical or narrative). While numerical standards can be derived in various ways, the development and interpretation of concentration-response relationships represents a model that could be considered to facilitate discussions on the interpretation of toxicity test results. In this context, workshop participants generally agreed that designation of samples as toxic or not toxic is not required early in the site assessment process. Rather, the magnitude of effect data (e.g., percent survival or biomass) can be used directly in the development of concentration-response relationships for COPCs at the site. The magnitude of effect data can also be used to classify sediment samples into risk categories, without having to designate individual sediment samples groups as toxic or not toxic. This approach to the interpretation of whole-sediment toxicity data was considered to be desirable because no information is lost during the data interpretation process. Toxicity test results can also support the development of narrative standards by defining conditions that are considered to pose either tolerable or unacceptable risks to ecological receptors.

Workshop participants also recognized that interpretation of toxicity test results may necessitate designation of individual sediment samples as toxic or not toxic (e.g., hot spot identification, evaluation of the spatial extent of toxicity). In these cases, workshop participants generally agreed that a step-wise approach should be used to interpret the results of individual toxicity tests. The recommended steps in this process include:

- Conduct the selected toxicity tests in accordance with the standardized protocol and described in a project Quality Assurance Project Plan (QAPP);
- Evaluate the validity of the toxicity tests. The project data quality objectives, which are documented in the QAPP, should define the performance criteria for measurement data that should be used to evaluate toxicity test acceptability. At minimum, such performance criteria should define the acceptable range of negative control and positive control (i.e., reference toxicant) results. Evaluation of potential test interferences should also be conducted during this step in the process (e.g., comparison of ammonia and hydrogen sulfide levels to lowest observed effect concentrations);
- Compare the results obtained for each sediment sample to the negative control results. Sediment samples for which the measured response is significantly greater than that for the negative control and for which the

measured response is greater than the MSD (or MDD) would be designated as potentially toxic;

- Compare the results obtained for each sediment sample to the reference envelope (see ASTM 2006b and MacDonald *et al.* 2002 for descriptions of reference sediments and the reference envelope approach). Sediment samples for which the measured response is greater (i.e., more toxicity is measured) than the lower limit of responses for reference sediment samples would be designated as potentially toxic (e.g., if the reference envelope for amphipod survival in a 10-d whole-sediment toxicity test is 77 to 98%, then sediment samples for which amphipod survival is less than 77% would be designated as toxic). Comparison to reference sites is only appropriate if reference sites have been well characterized and satisfy criteria for negative controls (i.e., response in reference sediments should not be significantly different from that of negative controls); and,
- Sediment samples that are designated as toxic using both the reference envelope and control comparison approaches should be identified as those that pose the highest risks to the benthic invertebrate community. Sediment samples for which the response of the test organism falls within the reference envelope should not be designated as toxic and should be considered to pose the lowest risks to the benthic invertebrate community. The following contingency table illustrates application of this process (which was provided by Dave Mount, with modifications provided by Chris Ingersoll, to address questions regarding the process);

Toxicity Designation Matrix	**Reference Comparison - NS**	**Reference Comparison - SD**
Control Comparison - NS	Not Toxic	Potentially Toxic
Control Comparison - SD	Potentially Toxic	Toxic

NS = not significant; SD = significantly different.
For the potentially toxic designations, the resultant data should be further examined to determine final toxicity designations.

It is important to note that the above described approach to toxicity designation is generally consistent with the recommendations that were developed by the three work groups, but that none of the work groups articulated this framework at this level of detail. It is also important to note that this framework is consistent with existing BCMOE guidance on the interpretation of toxicity test results (Ingersoll and MacDonald 2003). Therefore, it is not unreasonable to adopt this step-wise approach to toxicity designation when sample-by-sample designations are required to support data analyses. The MSD approach or comparison to negative control approach could also be applied when insufficient information is available to develop a reference envelope.

All of the work groups recognized that the results of individual toxicity tests can be used within a WOE framework for evaluating risks to the benthic invertebrate community associated with exposure to contaminated sediments. Workshop participants generally agreed that such WOE evaluations require information on the magnitude of toxicity in addition to, or instead of, toxicity designation information. Hence, it was generally agreed that the information on the magnitude of the response be retained to support further analyses of the toxicity data (i.e., WOE evaluations). Such information on effect size, as well as test endpoint variability and various statistical considerations, can be useful in the data interpretation process.

5.3 Integration of Multiple Lines-of-Evidence to Assess Risks to the Benthic Invertebrate Community

Risks to the benthic invertebrate community associated with exposure to contaminated sediments can be evaluated using several indicators of sediment quality conditions. For example, data on whole-sediment chemistry, whole-sediment toxicity, pore-water chemistry, pore-water toxicity, benthic invertebrate community structure, and/or invertebrate-tissue chemistry are often used to evaluate risks to the benthic invertebrate community associated with exposure to contaminated sediments. For each of these indicators of sediment quality conditions, a variety of metrics (which are also referred to a measurement endpoints) are commonly used to characterize sediments at a site. Collectively, the data generated on the various measurement endpoints that are included within an ecological risk assessment for contaminated sediments provide multiple LOEs for assessing risks to the benthic invertebrate community. The challenge that faces ecological risk assessors is to integrate the information on multiple LOEs to create a weight-of-evidence for assessing risks to ecological receptors and support decisions on the management of contaminated sediments.

While WOE approaches can be defined in various ways, workshop participants generally agreed that a WOE approach is: ***A tool or mechanism to improve understanding of, interpretation of, and inferences to be drawn from multiple LOEs to inform recommendations to be made by risk assessors to risk managers and site managers.*** Such WOE assessments facilitate prioritization of concerns relative to the risks posed by contaminated sediments and improve the confidence that can be placed in decisions regarding the management of contaminated sediments. By integrating information from multiple LOEs to assess risks to ecological receptors, WOE assessments provide a basis for identifying key stressors at a site, determining if something needs to be done to manage contaminated sediments, and, if so, where such remedial activities should be focussed at the site.

Workshop participants identified a number of approaches that could be used to integrate multiple LOEs to assess risks to benthic invertebrates. The candidate WOE approaches identified included:

- Best professional judgement approach;
- Tiered approaches;
- Decision matrix approach (consistent with the approach used in California; Bay *et al.* 2007);
- Semi-quantitative approach (consistent with the approach used in the Calcasieu Estuary; MacDonald *et al.* 2002); and,
- Fully quantitative approach (e.g., Menzie *et al.* 1996).

Appendix 5A, 5B, and 5C summarize the results of work group discussions on the various WOE approaches. Based on the input provided by the three work groups, it is apparent that workshop participants generally agreed that each of the candidate WOE approaches have certain strengths and limitations that influence their use in specific applications. While each of the work groups evaluated and/or ranked the candidate WOE approach in terms of their applicability for assessing and managing contaminated sediments in British Columbia, it was generally agreed that BCMOE should not establish prescriptive guidance on the WOE approaches. Rather, practitioners should be afforded the flexibility of selecting the WOE approach that is most appropriate for integrating the types of data and information that were collected at a site.

Workshop participants generally recognized that the applicability of all of the candidate WOE approaches can be influenced by limitations on data quality and data quantity. For this reason, it was emphasized that all sediment risk assessments should be supported by detailed problem formulations. In addition to the other essential elements of problem formulations, well-developed conceptual site models are needed to provide clear linkages between sources and releases of COPCs, the fate and effects of COPCs, key exposure pathways, and the ecological receptors potentially at risk. Such conceptual site models provide a defensible basis for identifying the assessment endpoints that need to be evaluated and selecting the measurement endpoints that are to be used to evaluate the status of each assessment endpoint. Workshop participants also indicated that the data quality objectives (DQOs) process (USEPA 2006) provides a systematic basis for determining data requirements for conducting ecological risk assessments. In addition, the DQOs process also supports the establishment of performance criteria for measurement data (as documented in the project QAPP) that should be used to determine if adequate quantities of data, of sufficient quality, have been collected to support the study. Only when the investigators are satisfied that the project DQOs have been met should a WOE approach be selected and employed to integrate multiple LOEs.

While workshop participants emphasized the importance of retaining flexibility in the selection of WOE approaches, it was recognized that there is a need for some level of consistency across sites in terms of the procedures that are used to integrate multiple LOEs. Such consistency is needed to ensure that the results of sediment risk assessments are generally comparable across sites, such that benthic invertebrates are afforded with similar levels of protection throughout the province (i.e., consistency with the provincial SMOs is essential). For this reason, workshop participants generally agreed that BCMOE should adopt a series of written guiding principles that can and should be applied in the selection of appropriate WOE approaches for assessing contaminated sediments at freshwater, estuarine, and marine sites in British Columbia. Ministry staff and/or other reviewers of ecological risk assessment could, then, use the guiding principles to determine if the WOE approach that has been selected for a site is appropriate.

The results of the work group discussions convened during the workshop provide an excellent basis for recommending a series of guiding principles that can be considered by BCMOE. More specifically, workshop participants generally agreed that WOE approaches that are used to integrate multiple LOEs at sites with contaminated sediments in British Columbia should have the following characteristics (some of which have included in the draft DERA guidance prepared for the SAB; Golder 2007):

- ***Supportive of Management Decisions*** - WOE approaches must generate results that can be used directly by risk managers and/or site managers to make decisions on the management of contaminated sediments;
- ***Scientific Defensibility*** - WOE approaches must be demonstrated to provide a basis for understanding the study results in a manner that is consistent by the accumulated body of scientific knowledge on the effects of sediment-associated COPCs on ecological receptors. That is, WOE approaches should yield results that are science-based, derived from the available scientific studies, and do not take into consideration socio-economic or management factors;
- ***Consistency with Narrative Intent*** - WOE approaches must support the development of site-specific sediment quality standards that are consistent with the SMOs that have been established by BCMOE (see BCWALP 2004 for further information);
- ***Consistency with Uncertainty Assessment*** - WOE approaches must rely on the results of the uncertainty assessment that is conducted as part of a DERA. As such uncertainty analyses document the confidence that can be placed in individual LOEs, these results can and should be used to

establish weightings for the various LOEs explored in the investigation (i.e., greater weight should be assigned to LOEs in which greater confidence can be placed). Both variability and true uncertainty should be explicitly considered in the uncertainty analysis and associated placement of weights on various LOEs;

- ***Reproducibility*** - WOE approaches must by sufficiently well-documented and objective to ensure that an independent investigator can apply the procedures to the data for a site and generate results that are consistent with those obtained by the primary investigators;
- ***Transparency*** - WOE approaches must be described in language that is readily-accessible to everyone involved in the assessment and management of contaminated sediments. In addition, the methods that are applied, the assumptions that are made, and the decisions that are taken to support implementation of a WOE approach must be clearly documented; and,
- ***Reliability*** - WOE approaches should provide a basis for confirming that the resultant classifications of sediment samples into the various risk categories are reliable [For example, MacDonald *et al.* (2002) presented an approach for evaluating the reliability of the WOE approach that was used to classify sediment samples into low, moderate, and high risk categories in the Calcasieu Estuary].

Workshop participants generally agreed that risk assessments must provide risk managers with the information that they need to inform decisions regarding the management of contaminated sediments. More specifically, the results of the result assessment must support the identification of a relatively small range within which uncertainty is not quantifiable, but above which the risk is unacceptable (as defined in the text of the DERA) and below which the risk is acceptable (as defined by law and modified by the uncertainty analysis). Within this range, the risk assessor must make a recommendation, which is supported by the text of the DERA, that will inform the risk manager's decision on remedial action. WOE approaches that embody the seven guiding principles described above and that support risk management decisions are likely to provide an effective basis for integrating multiple

LOEs. Such WOEs are likely to be considered to be reasonable by BCMOE and other reviewers.

6.0 Additional Issues and Concerns Related to the Assessment and Management of Contaminated Sediments in British Columbia

This workshop was designed to bring together experts in the fields of sediment quality assessment, ecological risk assessment, and contaminated sites assessment to share technical information and engage in technical discussions regarding the assessment and management of contaminated sediments in British Columbia. Workshop participants responded to this challenge by sharing their knowledge, experience, and insights relative to the selection of whole-sediment and pore-water toxicity tests, interpretation of toxicity test results, and integration of multiple LOEs within a WOE framework. The specific advice and guidance offered by workshop participants relative to each of these topics is highlighted in Sections 5.1, 5.2, and 5.3 of this document, respectively. More detailed compilations of the input provided by workshop participants is presented in Appendices 3, 4, and 5.

Workshop participants made excellent progress in terms of resolving the key issues that were to be addressed during the meeting. However, several additional issues and concerns emerged during the workshop that would benefit from further examination by BCMOE. These issues and concerns were subsequently examined by the Sustainable Fisheries Foundation (SFF) and used to develop additional recommendations to BCMOE (i.e., the following are SFF recommendations and do not necessarily reflect the views of the other workshop participants), as highlighted below:

- Workshop participants generally agreed that there is a higher level of variability (uncertainty) in a number of the chronic toxicity tests (evaluating sub-lethal endpoints) than there is in many of the acute toxicity tests (e.g., with amphipods). As suggested by ASTM, further investigations to evaluate the factors causing or contributing to such variability (e.g., species sensitivity, substrate type, handling stress) could aid in the interpretation of test results. For this reason, it is recommended that further research be conducted to evaluate the utility of such tests and the potential for standardization of such tests. In the near term, such research should focus on the chronic toxicity test with *Leptocheirus plumulosus* and on sediment-water interface toxicity tests with echinoids and/or bivalves.

- Based on the information presented at the workshop, it appears that DERAs are currently being conducted at a variety of sites in British Columbia. However, limitations on the resources available for assessing contaminated sediments at certain sites (particularly smaller sites) have resulted in such DERAs being conducted with limited data and information. While difficult to quantify, it is virtually certain that the uncertainty associated with such DERAs is higher than would be associated with more robust investigations. This higher level of uncertainty could result in inappropriate decisions regarding the management of contaminated sediments at certain sites. To address this concern, it is recommended that BCMOE establish further guidance for determining if the criteria-based approach or the risk-based approach should be used at sites with contaminated sediments (i.e., guidance for determining when it is appropriate to use risk assessment at a contaminated sediment site). Such guidance could reflect the size of the site (e.g., only sites greater than one hectare in surface area of contaminated sediment would qualify), estimated volume of contaminated sediment (e.g., only sites with >10,000 m^3 of contaminated sediment would quality), and/or other relevant factors. Alternatively, minimum data requirements for conducting risk assessments should be established. In this way, practitioners can ensure that risk assessment is applied in a consistent manner at sites with contaminated sediments in the province;

- USEPA (1997; 1998) has established an eight-step process to guide the design and implementation of ecological risk assessments (ERA) at hazardous waste sites. This process has proved to provide an effective basis for focussing ERAs and ensuring that the associated results support decisions regarding the management of the site. It is recommended that BCMOE endorse the use of the eight-step process and associated guidance documents when relevant B.C.-specific guidance is not available;

- As indicated above, limitations on the availability of financial and/or human resources can lead to evaluation of risks to ecological receptors at sediment contaminated sites with insufficient data. To further address this concern, it is recommended that BCMOE establish minimum data requirements for conducting DERAs. Such guidance should specify the suite of COPCs to be measured by site type, the suite of toxicity tests to be included, and the minimum sampling densities for site investigations using different study designs (i.e., gradient vs. reference). It is understood that the DERA guidance document that is currently under development is the appropriate vehicle for addressing these concerns;

- The reference envelope approach was one of the methods that were recommended by workshop participants for interpreting the results of whole-sediment and/or pore-water toxicity tests. To facilitate broader application of this approach, it is recommended that BCMOE establish a definition of reference sediment, the selection of reference sites, and identify the minimum number of reference sediment, samples that should be collected to support calculation of a reference envelope. Consideration should also be given to evaluating existing data on reference conditions within various regions of the province and determining if regional reference envelopes can be established with existing data;

- Workshop participants indicated that ongoing sources of COPCs (e.g., discharges from sewage treatment plants, combined sewage outfalls, industrial outfalls) exist in the vicinity of many contaminated sites in British Columbia. In addition, many sites are located in areas that are affected by wide-area contamination (e.g., Burrard Inlet, Esquimalt Harbour). Due to the potential for recontamination, little effort is currently being focussed on remediating these sites, even if risks to ecological receptors are found to be unacceptable. To address this issue, it is recommended that BCMOE provide guidance on the assessment and management of contaminated sediments at these types of sites. Importantly, such guidance should address the question as to whether clean-up goals should reflect the SQC, site-specific sediment quality standards, and/or area-specific background levels of COPCs. Such guidance should be developed in conjunction with federal regulatory agencies to assure harmonization with federal and provincial requirements for these types of sites; and,

- While BCMOE has established narrative sediment management objectives (SMOs) to support decisions on the management of contaminated sediments, a number of workshop participants indicated that further clarification of the SMOs would be beneficial. To address this concern, it is recommended that BCMOE further clarify the intent of the SMOs (i.e., these should define the protection goals for benthic communities at typical and sensitive sites in a manner that will support the derivation of both narrative and numerical sediment quality standards; i.e., what levels of response in which toxicity tests reflect the threshold for unacceptable risks to sediment-dwelling organisms).

7.0 References Cited

ASTM (American Society for Testing and Materials). 2006a. Test method for measuring the toxicity of sediment-associated contaminants with estuarine and marine invertebrates. E1367-03E01. ASTM 2006 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2006b. Test method for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. E1706-05. ASTM 2006 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania

Bay, S., D. Greenstein, and D. Young. 2007. Evaluation of methods for measuring sediment toxicity in California bays and estuaries. Technical Report 503. Southern California Coastal Water Research Project. Costa Mesa, California.

BCWLAP (British Columbia Ministry of Water, Land and Air Protection). 2004. Criteria for contaminated sites: Criteria for managing contaminated sediment in British Columbia. Technical Appendix. Environmental Management Branch. Environmental Protection Division. British Columbia Ministry of Water, Land and Air Protection. Victoria, British Columbia.

Golder Associates. 2007. Guidance for detailed ecological risk assessment (DERA) in British Columbia. Submitted to the Science Advisory Board for Contaminated Sediments in British Columbia. Submitted by Golder Associates. North Vancouver, British Columbia.

Ingersoll C.G. and D.D. MacDonald. 2003. A guidance manual to support the assessment of contaminated sediments in freshwater, estuarine, and marine ecosystems in British Columbia. Volume III: Interpretation of the results of sediment quality investigations. Prepared for Pollution Prevention and Remediation Branch. British Columbia Ministry of Water, Land, and Air Protection. Victoria, British Columbia.

MacDonald, D.D. and C.G. Ingersoll. 2003a. A guidance manual to support the assessment of contaminated sediments in freshwater, estuarine, and marine ecosystems, Volume 1: An ecosystem-based framework for assessing and managing contaminated sediments. Prepared for the Environmental Management Branch. British Columbia Ministry of Water, Land and Air Protection. Victoria, British Columbia.

MacDonald, D.D. and C.G. Ingersoll. 2003b. A guidance manual to support the assessment of contaminated sediments in freshwater, estuarine, and marine ecosystems, Volume 2: Design and implementation of sediment quality investigations. Prepared for the Environmental Management Branch. British Columbia Ministry of Water, Land and Air Protection. Victoria, British Columbia.

MacDonald, D.D. and C.G. Ingersoll. 2003c. A guidance manual to support the assessment of contaminated sediments in freshwater, estuarine, and marine ecosystems, Volume 4: Supplemental guidance on the design and implementation of detailed site investigations in marine and estuarine ecosystems. Prepared for the Environmental Management Branch. British Columbia Ministry of Water, Land and Air Protection. Victoria, British Columbia.

MacDonald, D.D., C.G. Ingersoll, D.R.J. Moore, M. Bonnell, R.L. Breton, R.A. Lindskoog, D.B. MacDonald, Y.K. Muirhead, A.V. Pawlitz, D.E. Sims, D.E. Smorong, R.S. Teed, R.P. Thompson, and N. Wang. 2002. Calcasieu Estuary remedial investigation/feasability study (RI/FS): Baseline ecological risk assessment (BERA). Technical report plus appendices. Contract No. 68-W5-0022. Prepared for CDM Federal Programs Corporation and United States Environmental Protection Agency. Dallas, Texas.

Macfarlane, M.W., D.D. MacDonald, and C.G. Ingersoll. 2004. Criteria for contaminated sites: Criteria for managing contaminated sediment in British Columbia. Technical Appendix. Environmental Management Branch. Environmental Protection Division. British Columbia Ministry of Water, Land and Air Protection. Victoria, British Columbia.

Menzie, C., M.H. Henning, 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

USEPA (United States Environmental Protection Agency). 1997. Ecological risk assessment guidance for Superfund: Process for designing and conducting ecological risk assessments. Environmental Response Team. Edison, New Jersey.

USEPA (United States Environmental Protection Agency). 1998. Guidelines for ecological risk assessment. Risk Assessment Forum. EPA/630/R-95/002F. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 2006. Guidance on systematic planning using the data quality objectives process. EPA QA/G-4. EPA/240/B-06/001. Office of Environmental Information. United States Environmental Protection Agency. Washington, District of Columbia.

Table 1. List of work group members at the workshop to support the development of guidance on the assessment of contaminated sediments in British Columbia, Victoria, B.C. - September 24-26, 2007.

Work Group	Facilitators	Work Group Members
1	Don MacDonald	Jack Word Remi Odense Deanna Lee Marc Cameron Scott Becker Jay Field Burt Shephard Chris Ingersoll Doug Bright
2	Beth Power/Chris Ingersoll	Glenn Harris Linda Porebski Bruce Williamson Gary Lawrence Geoff Wickstrom Bill Duncan David Charters Scott Carr
3	Dawn Smorong	Gary Mann Tara George Adrian deBruyn James Elphick Michael McLeay Sam Reimer Steve Bay Dave Mount Marc Greenberg

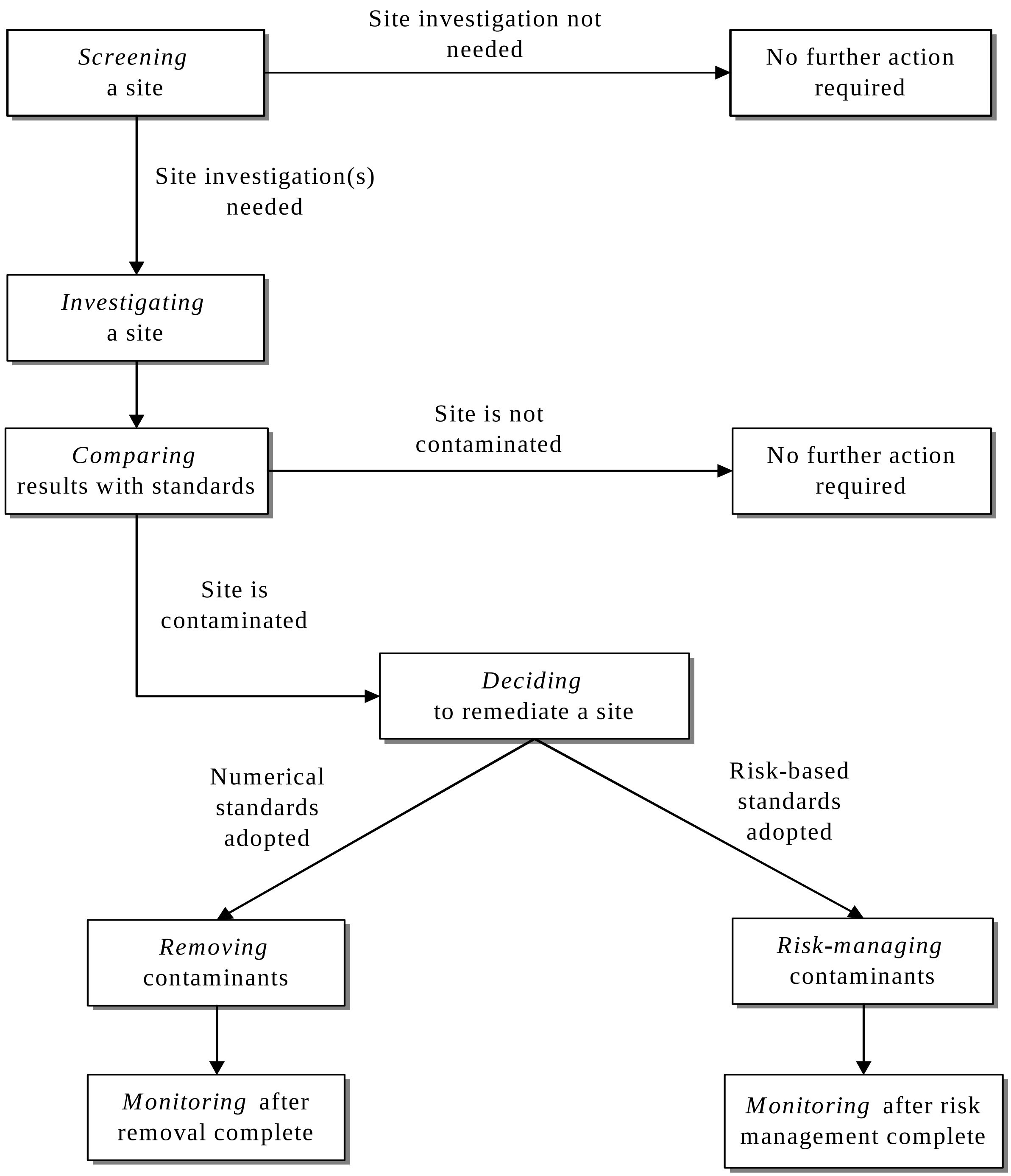
Table 2. Candidate whole-sediment and pore-water toxicity tests for assessing freshwater sediments in British Columbia.

Test Organism	Species	Test Duration	Exposure Medium	Endpoints Measured
Amphipod	*Hyalella azteca*	10- to 14-day	Whole sediment	Survival, growth, biomass
	Hyalella azteca	28- to 42-day	Whole sediment	Survival, growth, reproduction, biomass
Midge	*Chironomus dilutus* or *C. riparius*	10-day	Whole sediment	Survival, growth, biomass
	Chironomus dilutus	20-day	Whole sediment	Survival, growth, biomass
	Chironomus dilutus	60-day	Whole sediment	Survival, reproduction
Oligochaete	*Lumbriculus variegatus*	28-day	Whole sediment	Survival, biomass
	Tubifex tubifex	28-day	Whole sediment	Survival, reproduction
Cladoceran	*Ceriodaphnia dubia*	7-day	Whole sediment/pore water	Survival, reproduction
	Daphnia magna	7-day	Whole sediment/pore water	Survival, reproduction
	Chydorus spp.	10-day	Whole sediment	Survival
Mussel	*Lampsilis siliquoidea*	28-day	Whole sediment	Survival, growth
Microtox	*Vibrio fisheri*	5-30 minutes	Solid phase/pore water	Bioluminescence
Mayfly	*Hexagenia limbata*	21-day	Whole sediment	Survival, growth, molting frequency

Table 3. Candidate whole-sediment and porewater toxicity tests for assessing marine and estuarine sediments in British Columbia.

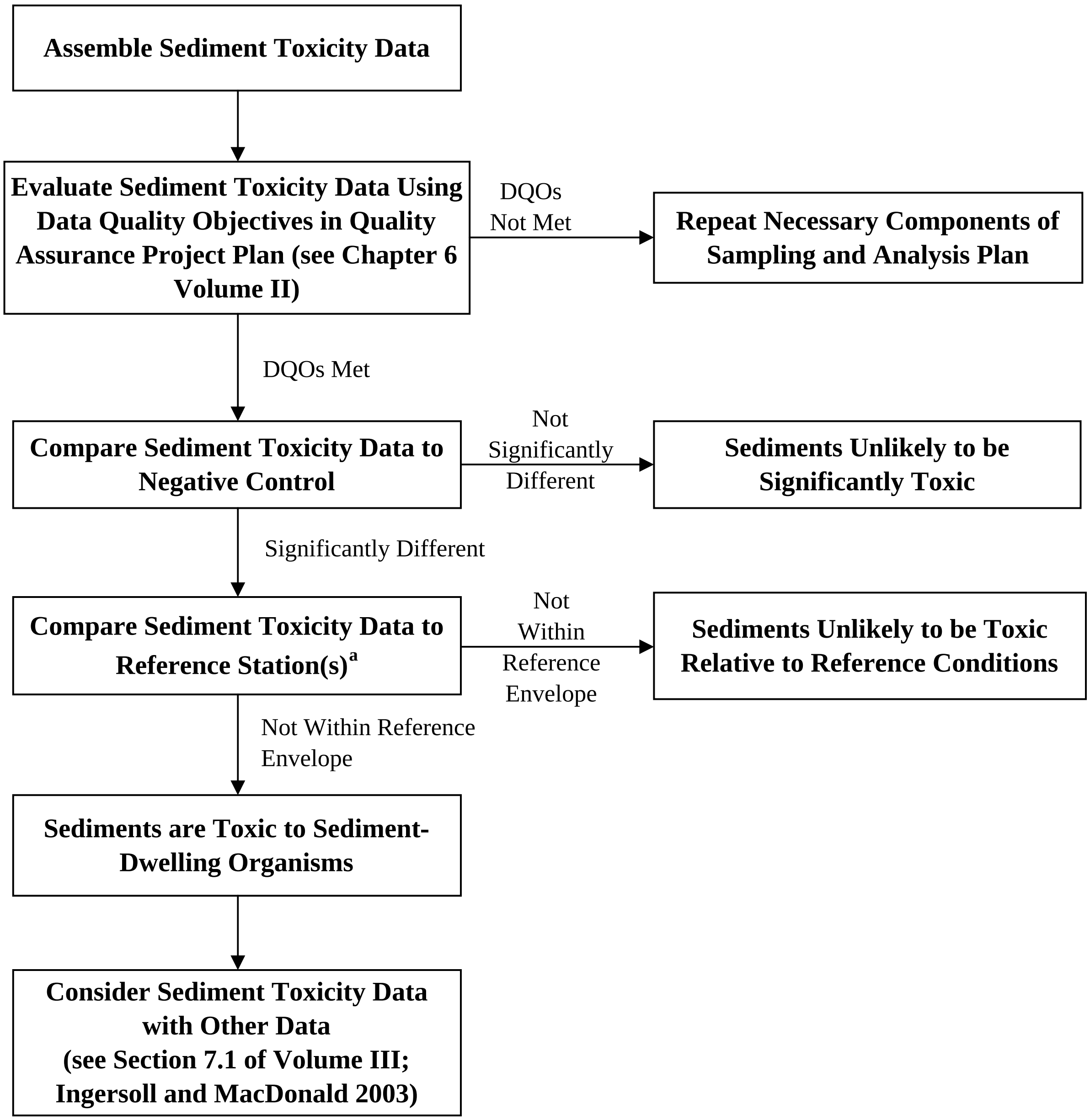
Test Organism	Species	Test Duration	Exposure Medium	Endpoints Measured
Amphipod	*Eohaustorius estuarius, E. washingtonius*	10-day	Whole sediment	Survival, reburial
	Ampelisca abdita	10-day	Whole sediment	Survival
	Rhepoxynius abronius	10-day	Whole sediment	Survival, reburial
	Grandidierella japonica	10-day	Whole sediment	Survival
	Leptocheirus plumulosus	10-day	Whole sediment	Survival
	Leptocheirus plumulosus	28-day	Whole sediment	Survival, growth, reproduction
Polychaete	*Neanthes arenaceodentata*	10-day	Whole sediment	Survival
	Nephtys incisa	10-day	Whole sediment	Survival
	Polydora cornutata	14-day	Whole sediment	Survival, growth
	Neanthes arenaceodentata	20- to 28-day	Whole sediment	Survival, growth, reproduction, biomass
	Capitella capitata	20- to 28-day	Whole sediment	Survival, growth, reproduction, biomass
Mysid	*Americamysis bahia*	10-day	Whole sediment	Survival
	Holmesomysis spp.	10-day	Whole sediment	Survival
Echinoderm	*Arbacia punctulata*	<1-hour	Pore water/sediment-water interface	Fertilization
	Arbacia punctulata	48- to 96-hour	Pore water/sediment-water interface	Development
	Dendraster excentricus	<1-hour	Pore water/sediment-water interface	Fertilization
	Dendraster excentricus	48- to 96-hour	Pore water/sediment-water interface	Development
	Strongylocentrotus purpuratus	<1-hour	Pore water/sediment-water interface	Fertilization
	Strongylocentrotus purpuratus	48- to 96-hour	Pore water/sediment-water interface	Development
	Strongylocentrotus droebachiensis	<1-hour	Pore water/sediment-water interface	Fertilization
	Strongylocentrotus droebachiensis	48- to 96-hour	Pore water/sediment-water interface	Development
Bivalve Mollusc	*Crassostrea gigas*	48-hour	Pore water/sediment-water interface	Fertilization,development
	Crassostrea virginica	48-hour	Pore water/sediment-water interface	Fertilization,development
	Mercenaria mercenaria	48-hour	Pore water/sediment-water interface	Fertilization,development
	Mytilus edulis	48-hour	Pore water/sediment-water interface	Fertilization,development
Microtox	*Vibrio fisheri*	<1-hour to 48-hour	Pore water/solid phase	Bioluminescence

Figure 1. General process for managing contaminated sites in British Columbia.



Note: Both numerical standards and risk-based standards can be applied at any site.
Source: MacFarlane *et al.* (2004).

Figure 2. Recommended procedure for assessing sediment toxicity data.



[a]Comparison to reference sites is only appropriate if reference sites have been well charactized and satisfy criteria for negative controls (i.e., response in reference sediments should not be significantly different from that in negative controls).

Source: Modified from Ingersoll and MacDonald (2003).

Appendices

Appendix 1 Workshop Agenda

The workshop agenda is designed to encourage participants to share technical information and engage in debates on issues related to the assessment of contaminated sediments in British Columbia. The agenda for the workshop is presented below.

Monday, September 24, 2007

9:00 - 9:20 **Welcome and Introductions**

9:20 - 9:30 **Workshop Goals and Objectives**

In addition to presenting the workshop goals and objectives, some background information on use protection goals in B.C. will be provided. Additional context for the workshop will also be provided, including a listing of underlying assumptions.

9:30 - 12:00 **Work Group Session 1:** *Selection of Toxicity Tests for Assessing Contaminated Sediments in British Columbia*

9:30 - 9:45 ***Presentation:*** Review and Evaluation of Candidate Whole-Sediment Toxicity Tests for Assessing Freshwater Sediments (Chris Ingersoll, *U.S. Geological Survey*)

9:45 - 10:00 ***Presentation:*** Evaluation and Selection of Whole-Sediment Toxicity Tests for Assessing Marine Sediments in California (Steve Bay, *Southern California Coastal Water Research Project*)

10:00 - 10:15 ***Presentation***: Considering Statistical Power in the Selection of Toxicity Tests for Assessing Contaminated Sediments (Adrian de Bruyn, *Golder Associates*)

10:15 - 10:30 **Refreshment Break**

10:30 - 10:45 ***Presentation:*** Review and Evaluation of Candidate Pore-water Toxicity Tests for Assessing Freshwater, Estuarine and Marine Sediments (Scott Carr, *U.S. Geological Survey*)

10:45 - 11:00 ***Presentation***: Evaluation of the Role of *in-situ* Toxicity Testing in Contaminated Site Assessment (Marc Greenberg, *U.S. Environmental Protection Agency*)

11:00 - 11:15 ***Presentation***: Approaches to Sediment Assessment under Canada's Disposal at Sea Program (Linda Porebski, *Environment Canada*)

Work Group Discussions: Participants will be split into three work groups to discuss the selection of toxicity tests.

The Focus Questions will include:

1. Which whole-sediment, pore-water, and/or elutriate toxicity tests and test endpoints should be considered for assessing risks posed to benthic invertebrates utilizing benthic habitats, freshwater, estuarine, and marine ecosystems)?
2. What are the strengths and limitations/uncertainties of each of the candidate toxicity tests for this purpose?
3. Which candidate toxicity tests have been demonstrated to support the development of concentration-response relationships, the establishment of risk-based sediment quality standards (or functionally equivalent values), and/or evaluations of risks to sediment-dwelling organisms?
4. Which toxicity tests are most relevant for evaluating risks to aquatic receptors (i.e., microbiota, aquatic plants, benthic invertebrates, and fish) associated with exposure to contaminated sediments at freshwater, estuarine, and marine sites?
5. Develop a matrix which identifies the primary and secondary toxicity tests that could be applied at sites with sediments contaminated by metals, PAHs, PCBs, pesticides, and/or other substances.

12:00 - 1:00 **Lunch (provided)**

1:00 - 3:00 **Work Group Session 1:** *Selection of Toxicity Tests for Assessing Contaminated Sediments in British Columbia (continued)*

3:00 - 3:20 **Refreshment Break**

3:20 - 4:00 **Work Group Session 1:** *Summary Reports and Discussion*

4:00 - 5:00 **Work Group Session 2:** *Use of Sediment Toxicity Test Results to Evaluate Risks to Sediment-Dwelling Organisms Exposed to Contaminated Sediments*

4:00 - 4:15 ***Presentation:*** Designation of Sediment Samples as Toxic or Not Toxic Based on the Reference Envelope Approach (Chris Ingersoll, *U.S. Geological Survey* and Don MacDonald, and *MacDonald Environmental Sciences Ltd.*)

4:15 - 4:30 ***Presentation:*** Designation of Sediment Samples as Toxic or Not Toxic Based on the Results of Power Analyses (MSD approach; Scott Carr, *U.S. Geological Survey*)

4:30 - 4:45 ***Presentation:*** Toxicity Identification Evaluations in Risk Assessments and Their Role in Determining Relevance of Adverse Effects Observed in Laboratory Toxicity Tests (James Elphick, *Nautilus Environmental*)

4:45 - 5:00 ***Presentation:*** Beyond Thresholds: Thoughts About Classification of Toxicity (Jay Field, *National Oceanic and Atmospheric Administration*)

Tuesday, September 25, 2007

8:30 - 12:00 **Work Group Session 2:** *Use of Sediment Toxicity Test Results to Evaluate Risks to Sediment-Dwelling Organisms Exposed to Contaminated Sediments*

8:30 - 8:45 ***Presentation:*** Getting Real: Scientific/ecological misconceptions and dangers in cramming continuous response relationships into two little boxes (toxic or not), and the role of causality in sediment risk assessments (Doug Bright, *UMA Engineering Ltd*)

8:45 - 9:00 ***Informal Presentations:*** Comparison of Various Methods of Designating Sediment Samples as Toxic or Not Toxic

9:00 - 12:00 ***Work Group Discussions:*** Participants will be split into three work groups to discuss the designation of sediment samples representing unacceptable risks to aquatic receptors.

The Focus Questions will include:

1. What are the candidate approaches for designating freshwater, estuarine, and marine sediment samples as toxic and not toxic?
2. What are the strengths and limitations of each approach to designating sediment samples as toxic and not toxic?
3. Which approach (or approaches) is the most relevant for designating individual sediment samples as toxic and not toxic in British Columbia?
4. What should be done when the toxicity test do not concur in their finding of "toxic" for a sediment sample?
5. What should be done when controls fail and/or there are suspected test problems (e.g., grain size problems, unusual test results, such as no relationship between chemistry and toxicity)?

12:00 - 1:00 **Lunch (provided)**

1:00 - 2:00 **Work Group Session 2:** *Summary Reports and Discussion*

2:00 - 5:00 **Work Group Session 3:** *Integration of Multiple Lines-of-Evidence to Assess Risks to Ecological Receptors Associated with Exposure to Contaminated Sediments*

2:00 - 2:15 ***Presentation:*** *Case Study 1:*
Integrating multiple lines-of-evidence to assess contaminated sediments in B.C. (Gary Mann, *Azimuth Consulting Group*)

2:15 - 2:30 ***Presentation***: *Case Study 2*:
Challenges associated with developing WOE frameworks (Gary Lawrence, *Golder Associates*)

2:30 - 2:45 ***Presentation:*** *Case Study 3:*

Integrating multiple lines-of-evidence to assess contaminated sediments in California (Steve Bay, *Southern California Coastal Water Research Project*)

2:45 - 3:00 ***Presentation:*** *Case Study 4:*
Using EqP theory as a means to rationalize seemingly disparate data (Dave Mount, *U.S. Environmental Protection Agency*)

3:00 - 3:15 ***Presentation:*** *Case Study 5:*
Integrating multiple lines-of-evidence to assess contaminated sediments in Louisiana. (Chris Ingersoll, *U.S. Geological Survey* and Don MacDonald, *MacDonald Environmental Sciences Ltd.*)

3:15 **Working Refreshment Break**

3:30 - 5:00 ***Work Group Discussions:*** Participants will be split into three work groups to discuss the integration of multiple lines-of-evidence in sediment risk assessment. Participants will be encouraged to review the results of the SETAC Pellston Workshop on this topic prior to the meeting.

The Focus Questions will include:

1. What are the candidate approaches for integrating multiple lines-of-evidence related to sediment quality conditions?
2. What are the strengths and limitations of the various approaches to integrating multiple lines-of-evidence?
3. What are the guiding principles that should be considered during the integration of multiple lines-of-evidence?
4. How should the results for multiple measurement endpoints and multiple lines-of evidence be integrated to obtain a weight-of-evidence for evaluating effects on aquatic receptors at sites with contaminated sediments?

Wednesday, September 26, 2007

8:30 - 10:10 **Work Group Session 3:** *Integration of Multiple Lines-of-Evidence to Assess Risks to Ecological Receptors Associated with Exposure to Contaminated Sediments*

10:10 - 10:30 **Refreshment Break**

10:30 - 11:30 **Work Group Session 3:** *Summary Reports and Discussion*

11:30 - 12:00 **Workshop Wrap-Up and Path Forward**

Appendix 2 List of Participants at the Workshop to Support the Development of Guidance on the Assessment of Contaminated Sediments in British Columbia, Victoria, B.C. - September 24-26, 2007.

Participant	Affiliation	Email Address
Bay, Steve	Southern California Coastal Water Research Project	SteveB@sccwrp.org
Becker, Scott	Integral Corp	sbecker@integral-corp.com
Bright, Doug	UMA Engineering Ltd.	Doug.Bright@uma.aecom.com
Cameron, Marc	Hemmera	mcameron@hemmera.com
Carr, Scott	U.S.Geological Survey	scott.carr@tamucc.edu
Charters, David	U.S. Environmental Protection Agency	Charters.DavidW@epa.gov
deBruyn, Adrian	Golder Associates	adebruyn@golder.com
Duncan, Bill	Teck Cominco Metals	bill.duncan@teckcominco.com
Elphick, James	Nautilus Environmental	James@nautilusenvironmental.com
Field, Jay	National Oceanic and Atmospheric Administration	jay.field@noaa.gov
George, Tara	B.C. Ministry of Environment	Tara.George@gov.bc.ca
Greenberg, Marc	U.S. Environmental Protection Agency	Greenberg.Marc@epamail.epa.gov
Harris, Glenn	B.C. Ministry of Environment	Glenn.Harris@gov.bc.ca
Ingersoll, Chris	U.S.Geological Survey	cingersoll@usgs.gov
Lawrence, Gary	Golder Associates	glawrence@golder.com
Lee, Deanna	Environment Canada	Deanna.Lee@ec.gc.ca
MacDonald, Don	Sustainable Fisheries Foundation	mesl@shaw.ca
Mann, Gary	Azimuth Consulting	gmann@azimuthgroup.ca
McLeay, Michael	Seacor Environmental	mmcleay@seacorcanada.com
Mount, Dave	U.S. Environmental Protection Agency	mount.dave@epamail.epa.gov
Odense, Remi	B.C. Ministry of Environment	remi.odense@gov.bc.ca
Porebski, Linda	Environment Canada	linda.porebski@ec.gc.ca
Power, Beth	Azimuth Consulting	bpower@azimuthgroup.ca
Reimer, Sam	Seacor Environmental	sreimer@seacorcanada.com
Shephard, Burt	U.S. Environmental Protection Agency	Shephard.Burt@epa.gov
Smorong, Dawn	MacDonald Environmental Sciences Ltd.	mesl@shaw.ca
Wickstrom, Geoff	Hemmera	gwickstrom@hemmera.com
Williamson, Bruce	Geosyntec	bwilliamson@geosyntec.com
Word, Jack Q	NewFields Northwest	jqword@newfields.com

Appendix 3A Selection of Toxicity Tests for Assessing Contaminated Sediments in British Columbia - *Summary of the Results for Work Group 1*

1.0 Candidate Whole-Sediment and Pore-Water Toxicity Tests

Work group members identified a variety of whole-sediment and pore-water toxicity tests that could be used to evaluate sediment quality conditions at freshwater, estuarine, and marine site. For freshwater sites, these toxicity tests included (note: each participant was able to place three votes to identify their preferences for the various candidate toxicity tests):

- 10 to 14-day *Hyalella azteca* - survival, growth, biomass (4 votes)
- 28 to 42-day *Hyalella azteca* - survival, growth, reproduction (several metrics) biomass (8 votes)
- 10-day *Chironomus dilutus* - survival and growth (can use for bioaccumulation), biomass (8 votes)
- 60-day *Chironomus dilutus* - not enough data to support use yet
- 28-day *Lumbriculus variegatus* - bioaccumulation and survival (mass sometimes shrink) (4 votes)
- 20 day *Chironomus dilutus* - survival, growth, biomass (1 vote)
- 10-day *Chironomus riparius* - survival, growth, biomass
- 21-day mayfly *(Hexagenia limbata*) - no lab culture, hard to get "good" eggs;
- 28-day *Tubifex tubifex* - survival, reproduction
- 7- day *Ceriodaphnia dubia* or *Daphnia magna* - survival, reproduction;
- 28-day *Diporea* sp. - survival
- Microtox - solid phase
- 10-day *Chydorus* sp. - survival; freshwater benthic cladoceran used in European studies (Burt Shepherd for more information)
- 28-day freshwater mussel (exposed to deep sediments - really exposed to pore water sediments) - survival and growth (2 votes).

Discussion points related to toxicity test identification included:

- Sediment-feeding fish should also be considered (e.g., catostomid)
- Bioaccumulation tests need to be considered as well
- There are still data gaps on relative sensitivities of species to groups of contaminants (USEPA is conducted some evaluations using five chemical classes to evaluate sensitivity in water-only exposures
- Prefer to choose species that are relevant to site (mayfly for stream)
- Selection of a consistent suite of tests at many sites provides a basis for having fair treatment across sites (i.e., by regulators)
- Pesticides - Midge see some of these much more than amphipods
- Metals - Sometimes midge are more sensitive, sometimes amphipods are more sensitive
- Copper/NH_3 - Mussels may be more sensitive
- Where questions about availability exist use tiering (i.e., chemistry/toxicity; not toxicity with elevated chemistry; bioaccumulation to confirm that COPCs are not bioavailable
- Select locations with medium and high chemistry to conduct bioaccumulation testing and/or use TIE
- Elutriates - Low priority of assessing in-place sediments - may be useful for questions regarding effects at disposal site
- Pore-water tests - Can be used in Tiered Framework; like TIE/Bioaccumulation;
- *In-situ* tests and benthic invertebrate community structure - Also used in Tiered Framework (2nd or 3rd tier)
- 2nd/3rd Tiers - Contribute to weight-of-evidence where more information is needed to make decisions on sediment management
- Is nitrogen blanket appropriate for storing whole sediment? Probably not because mixture goes anaerobic and changes conditions.

The most highly recommended toxicity tests (i.e., those that should be included in investigations at all freshwater sites) included:

- 10 to 14-day *Hyalella azteca* - survival, growth, biomass (4 votes)
- 28 to 42-day *H. azteca* - survival, growth, reproduction (several metrics) biomass (8 votes)
- 10-day *Chironomus dilutus* - survival and growth (can use for bioaccumulation), biomass (8 votes).

The candidate toxicity tests that were identified by work group members for assessing marine and estuarine sites included:

- 48 to 96-hour echinoderm, bivalve mollusc or abalone larvae - fertilization, survival, and abnormality (3 votes); pore-water exposure or whole-sediment with suspended particulate
- 10-day toxicity tests with amphipods - survival and/or reburial (11 votes) - *Eohaustorius estuarius*, *Ampelisca abdita*, *Rhepoxynius abronius, Grandidierella japonica*, or *Leptocheirus plumulosus*
- 28-day *L. plumulosus* - survival, growth, reproduction (1 vote), 28-day test with *A. abdita* is still under development
- 10-day polychaete - survival; *Neanthes arenoceodentata*; *Nephtys* sp.; *Polydora* sp.
- 20-day to 28-d polychaete - survival, growth, biomass (7 votes); *N. arenoceodentata*, *Capitella capitella* sp. — also reproduction
- 28-day bioaccumulation (2 votes); *Macoma nasuta*, *Macoma balthica*, *Nereis virens*, *Nephtys* sp., Sandabs
- 10-day Mysid - survival; 2 groups - *Americamysis* sp., *Holmesomysis* sp.;
- Microtox - solid-phase test
- Elutriate Test - Algae (red) – *Parvula champea*; Macrocystis; Fish – *Menidia* sp.; Sandab; English sole
- 48 to 96-hour sediment-water interface testing with echinoid or bivalve larvae (Note: Virtually all species of benthos/fish can be tested like this).

The most highly recommended toxicity tests (i.e., those that should be included in investigations at all marine and estuarine sites) included:

- 10-day amphipod - survival (11 votes)

Discussion relative to the selection of toxicity tests for marine and estuarine sites in British Columbia included:

- 28-day bioaccumulation provides, perhaps different information than polychaete toxicity and probably more useful
- 28-day amphipod tends to be less powerful for discriminating among site (less sensitive for survival than 10-day; more variable for growth and reproduction)

- Selection of amphipods for 10-day tests should include salinity tolerance, grain size tolerance, availability, exposure/sensitivity (burrowers vs. tube dwellers), chemicals - chlorinated pesticides (insecticides) - amphipods increased(?) sensitivities
- Variable Conditions - Selection of Toxicity Tests

Salinity - Low – *Eohaustorius estuarius/Leptocheirus plumulosus/ Ampelisca abdita*

High (33%) – *E. estuarius/L. plumulosus/A. abdita/ Rhepoxynius abronius/Grandidierella japonica* sp.

Variable – *E. estuarius/L. plumulosus/A. abdita*

Note: Acclimation period is essential to follow (ASTM protocol). Acclimation period for sediment (microbial community - 7-21 days)

Grain Size - Fine – (This is where chemistry is). Tube-dwellers do great (*A. abdita/L. Plumulosus/Grandidierella japonica*) Still consider *E. estuarius/R. abronius*, as preferred species because they do not isolate themselves

Coarse -

Variable – *R. abronius*

- Research needs - Other indigenous amphipods that live (freely) in fine-grained sediments
- All of the amphipod species are useful for sites with metals, PAHs, PCBs, OC pesticides, and COPC mixtures
- Consider blood pigment to sequester (e.g., metals). Molluscs accumulation PAHs better
- Second solid phase test would be desirable; but we don’t really have one that would provide data that is complimentary to amphipod data
- Pore water larval test? Sometimes work/sometimes not
- Pore water chemistry is correlated with amphipod toxicity. Relevant exposure medium for free-living amphipods
- Toxicity Identification Evaluations are better for higher tiers in the assessment process to evaluate bioavailable fractions.

2.0 Summary of Recommendations

Tier 1 - Freshwater

- 10-day *Chironomus dilutus* - survival and growth
- 28-day *Hyalella azteca* - survival and growth.

Tier 2 - Freshwater

- 28-day *Lumbriculus variegatus* -survival
- TIE of pore water/whole sediment
- Benthic invertebrate community structure/*in-situ* test.

Tier 1 - Saltwater

- 10-day amphipod - survival
- 96-hour mussel larval test (sediment-water interface when standard methods become available).

Freshwater

- 10-day *Chironomus dilutus*
- 28-day *Hyalella azteca*.

Our experience suggests that the recommended Tier 1 toxicity tests are reflective of responses in the field.

Appendix 3B Use of Sediment Toxicity Test Results to Evaluate Risks to Sediment-Dwelling Organisms Exposed to Contaminated Sediments - *Summary of the Results for Work Group 2*

Scenario:

- Sediments are in-place
- Historical contamination
- Contaminated sites scenario
- We exceed sediment standards (schedule 9) and we are applying "risk-specific standards" under the CSR.

What is our question?

How would we choose tests for prioritization?

1. For prioritization? - bad, (grey), good - acceptable vs unacceptable risks?
2. Is active sediment management needed? - yes, no
3. For developing clean-up goals? -

Our Group decided that to evaluate the toxicity tests, we had to review the selection criteria.

As one moves from chemistry to toxicity to field measures… shifting from false positive errors… towards the "right" answer.

When one exceeds standards, how to do we choose toxicity tests that minimize the chances of false positives?

Are the tests going to describe bioavailability?

Tox tests for DUMMIES!

<u>To pose for wider discussion:</u>

1. Cause vs correlation: lots of debate about ecological relevance and link between test and real-life
2. Pore-water vs. direct exposure
3. Trading off various aspects of ecological relevance is difficult because we are dealing with non-comparables. Exposure route, life stage, chronic, etc.

Pick a test battery:

- Use the PF and CSM to set the primary drivers
- Pick tests to match those drivers

What are those primary drivers?

- Exposure route?
- Taxonomic representation?
- What should NOT be drivers?
    - Chronic vs acute?
    - Test length

Invertebrates – worms							
Rank?	**Criteria**	***Neanthes***	***Polydora***	***Dinophilus* pw**			
High	Ecological relevance	High (growth)	High	High			
High	Confounding factors (grain size, oc, salinity, ammonia, H2S)	Oc, avoid high grain size	TBD	Porewater			
High	Sensitivity to contaminants (???)	TBT	Relatively low sensitivity relative to other assays	Most sensitive relative to other worms			
Med	Reproducibility						
Low	Availability	High	Low	High			
Med	Field validation	?	?available	?			
Med	Precision	Mod	Mod	Mod			
Med	Practicality	High	Mod?	High			
Low	Route of exposure/contact	Ingestion & dermal	Sed-h2o interface	porewater			
Low	Test duration	20d or 28d	14d	7d			
Low	Cost						
Mod	Test success rate/reliability	High	High	High			
Low-High	Standardized method/peer review/round robins?	High	Med	Med			
High	Range of endpoints?	Survival & growth	Survival & growth	Full life cycle-reprod			
COMMENTS:							

Invertebrates – amphipods							
Rank?	**Criteria**	***Eohaustorius estuarius***	***Eohaustorius washingtonius***	***Ampelisca***	***Leptocheirus***	***Amphiporeia***	***Rhepoxynius***
High	Ecological relevance (good surrogate?)	Med - endpt is mortality, high exposure	Med - endpt is mortality, high exposure	Med - endpt is mortality, high exposure	High – but b/c of hi variability in repro endpt makes use diff.		
High	Confounding factors (grain size, oc, salinity, ammonia, H2S)	Robust to salinity & grain size	Sens to grain size (fines)	Sens to oc and grain size (fines)	Expect less precision for repro endpt	See method	Grain size sensitive (fines)
High	Sensitivity to contaminants (???)	Varying sensitivity to different classes					
High	Range of endpoints?	Surv, emergence, reburial	Surv, emergence, reburial	Surv, emergence, reburial	Growth & repro (28d), surv, emergence, reburial	Surv, emergence, reburial	Surv, emergence, reburial
COMMENTS	In choosing amphipod tests, consider selecting multiple tests due to their varying sensitivity to difference chemical classes. Either know its sens to the chemical, or run more than one (future R&D). These tests are also more responsive to confounding factors and habitat. Pacific vs Atlantic; north vs south.						

Appendix 3C Selection of Toxicity Tests for Assessing Contaminated Sediments in British Columbia - *Summary of the Results for Work Group 3*

Focus Question #1

Which whole-sediment, pore-water, and/or elutriate toxicity tests and test endpoints should be considered for assessing risks posed to benthic invertebrates utilizing benthic habitats, freshwater, estuarine, and marine ecosystems)?

Marine Toxicity Tests

- Whole-sediment toxicity tests recommended (i.e., preferential compared to pore-water, elutriate, or lifecycle toxicity tests)
- Candidate recommended toxicity tests
  - 28-d *Leptocheirus plumulosus* - survival, growth, reproduction
    - Advantage: the only standardized chronic amphipod test
  - 10-d *Rhepoxynius abronius* - survival
  - 10-d *Eohaustorius estuarius* - survival
    - Advantage: tolerant to grainsize and salinity ranges
  - 10-d *Ampelisca abdita* - survival
    - Less responsive
    - Exposure to interstitial water limited
    - Less robust to lab testing
    - Does not live in immediate area (geographic distribution)
    - Challenging to culture in the lab
  - Molluscs
    - Disadvantage is that the test has not been established yet (because it is relatively new)
  - Polychaetes
    - *Neanthes arenoceodentata* sp.
    - 20-d or 28-d durations (20-d test more established; 28-d test is newer)
    - 20-d growth endpoint less sensitive than 28-d growth endpoint
    - survival endpoint is not sensitive

        - Good control of organism in laboratory
    - Polydora
        - Issues exist with availability/culturing
    - Echinoderms
        - Sediment-water interface test
        - Endpoint is embryo-development
        - Advantages - short test duration; sub-lethal endpoint; uses a sensitive lifestage; deformity endpoint useful
        - Disadvantages - less environmentally realistic; mortality endpoint less useful because the sediment settling time is too short

General Notes:

- Among these, replicate variability is similar
- Generally, no preference between the candidate toxicity tests (although the *Ampelisca abdita* test has several limitations and should perhaps not be recommended)
- *Eohaustorius estuarius* has a greater sensitivity to metals
- 10-day reburial endpoint - not recommended because the animals have to be almost dead to show a response for this endpoint

<u>Freshwater Toxicity Tests</u>

- Whole-sediment toxicity tests recommended (i.e., preferential compared to pore-water, elutriate, or lifecycle toxicity tests)
- Candidate recommended toxicity tests
    - Amphipod
        - *Hyalella azteca* or *Diporea* sp.
        - 10-d, 28-d, 42-d durations
        - Survival, growth, reproduction endpoints (survival only for *Diporea*)
        - Disadvantage is that there are availability issues with *Diporea* sp.
    - Midge
        - *Chironomus riparius* or *dilutus*
        - 10-d, 14-d, 20-d durations
        - Survival, growth, reproduction, emergence endpoints
        - Emergence endpoint necessary? Early life stage, therefore a more sensitive test

- Oligochaetes
    - *Lumbriculus variegatus* or *Tubifex tubifex*
    - Disadvantage is grain size intolerances
- Microtox
    - Not recommended

General notes:

- Variety of tests should be conducted
- Phylogenetic diversity is limited for *Hyalella azteca* and *Chironomus* sp.
- Modify test methods to address environmental realism (e.g., develop tests for marine plants, diatoms)

Focus Question #2

What are the strengths and limitations/uncertainties of each of the candidate toxicity tests for this purpose?

General Discussion (regarding preferred features/strengths of recommended toxicity tests):

- Environmental realism (e.g., realistic exposure scenario relative to grainsize, species, etc)
- Acceptable test variability
- Responsiveness to COPCs (i.e., sensitivity, as all species are responsive at some level)
- Phylogenetic/taxonomic diversity
- Test feasability (e.g., organism availability; several laboratories capable of conducting test effectively)
- Existing standardized methodology, which provides the following advantages:
    - Includes round robin testing
    - Reproducible between labs
    - Includes performance criteria
    - Consideration is the test method jurisdiction (e.g., US vs. Canada)
- Ecological significance (much discussion about this, but agreement that we don't know with certainty the ecological significance of laboratory toxicity tests)
- Tolerance to a broad range of physical/chemical properties (e.g., grainsize, salinity)

- Direct association with sediment (exposure realism; e.g., epi-benthic vs. infaunal) organisms)
- Samples should be tested with multiple toxicity tests; tests that use both infaunal and epibenthic species should be included
- Cost

Focus Question #3

Which candidate toxicity tests have been demonstrated to support the development of concentration-response relationships, the establishment of risk-based sediment quality standards (or functionally equivalent values), and/or evaluations of risks to sediment-dwelling organisms?

This question was not addressed directly. It was assumed that this exists for the short list of recommended toxicity tests (i.e., a good toxicity test will show a dose-response relationship; i.e., the test animal is responsive and sensitive).

Focus Question #4

Which toxicity tests are most relevant for evaluating risks to aquatic receptors (i.e., microbiota, aquatic plants, benthic invertebrates, and fish) associated with exposure to contaminated sediments at freshwater, estuarine, and marine sites?

This question was not addressed directly.

- Certain receptor groups are not relevant for assessment relative to sediment
- Core benthic toxicity tests should be standard, and then use other tests that may address site-specific questions/risks to other receptor groups
- However, if effects to benthos then there is often an assumption that there are effects to other receptor groups
- Need to look at site-specific COPCs to assess if other receptor groups need to be individually assessed (i.e., if bioaccumulative substances present or suspected)

Focus Question #5

Develop a matrix which identifies the primary and secondary toxicity tests that could be applied at sites with sediments contaminated by metals, PAHs, PCBs, pesticides, and/or other substances.

This question was not addressed directly. If bioaccumulative COPCs are present, sediment toxicity tests will not address the risks presented at the site.

The presentation that Work Group 3 presented to workshop participants follows.

"Good Science" Clause

- These are generalized recommendations based on a general scenario
- Nothing herein should be viewed as:
  - A reason to blindly include a test if it doesn't make sense for a particular site or assessment
  - A reason to not include other tests if it does make sense
- There are very good reasons to gather data using test types that we have not emphasized here

Desirable Characteristics (1 of 5)

- Environmental realism
  - Coherence of nature of exposure in test with that occuring in the field
  - For question at hand, focuses us on tests with bedded sediments
  - This is not to suggest that other types of tests can't yield important information, or might not be useful for different risk questions

Desirable Characteristics (2 of 5)

- "Acceptable" test variability
  - To date, a matter of "what we can live with"
  - Needs careful evaluation to maximize information from tests
- Phylogenetic diversity
  - Important to insuring good ecological coverage
  - Characteristic of a test battery, not of a single test
  - Emphasize need for further development of some currently second tier species/tests

Desirable Characteristics (3 of 5)

- Sensitivity to contaminants of concern
  - Related to phylogenetic diversity
  - The "most sensitive species" myth
  - The "pollution-tolerant" oligochaete is most sensitivt to some contaminants
- Laboratory availability/experience
- Standardized procedures
  - Important of course, but shouldn't be permanently constrained by what has been done thus far

Desirable Characteristics (4 of 8)

- Tolerance of variety of sediment characteristics
  - Important, of course, but shouldn't let the tail wag the dog (e.g., oligochaetes desirable in many respects, but do need to be concerned about grain size

Desirable Characteristics (5 of 5)

- Cost
  - Has to be implementable, but better information may cost more
  - Reasonable tie breaker
- Ecological significance
  - Endpoints have relevance to health of field populations (e.g., survival, growth, reproduction)

Freshwater Tests

- Focus on those dealing with bedded sediments and direct sediment contact
  - Leaves out daphnid and perhaps fish tests
- For amphipod, focus on *Hyalella*
  - *Diporeia* not clearly superior and running behind on experience, availability, etc.
  - Inclusion of growth as an endpoint important
  - Choosing between 10-d and 28-d tests to be based on larger evaluation of cost/benefit (others have done?)

Freshwater Tests

- Midge tests an important component of base battery
  - No real preference for *dilutus* v *riparius*
  - Uncertainty of importance of emergence endpoint
  - Emergence protocols may be preferable more because they start with younger life stages
  - Consider survival/growth-based methods that start with younger organisms?

Freshwater Tests

- No other freshwater tests have the information base that midge and amphipod studies do
- However this leaves a pretty slim taxonomic coverage
- May point to need for more evaluation/development of other species
  - Oligochaetes
  - Mayfly (*Hexagenia*)

Freshwater Tests

- Oligochaetes
  - *Tubifex* and *Lumbriculus* methods exist
  - May need some more experience to address issues of grain size, etc.
  - More obligate benthic organism than any other
- Mayfly
  - Nice link to EPT focus of many benthic surveys
  - Are organism availability issues resolved?

Marine Tests

- Focus on bedded sediment and direct sediment contact tests
- Core tests to be considered for each study:
  - 10-d amphipod survival
  - 28-d amphipod growth
  - 20/28 d polychaete growth
  - 2-4 d embryo development

Marine Tests

- 20/28 d polychaete growth
  - Neanthes only viable test method
  - Polydora not available
- 2-4 d embryo development at sediment water interface
  - Bivalve embryos
  - Echinoderm embryos

Marine Tests

- Pore water/elutriate tests
  - Not recommended for core

Marine Tests

- Oligochaetes

Appendix 4A Interpretation of Toxicity Tests for Assessing Contaminated Sediments in British Columbia - *Summary of the Results for Work Group 1*

1.0 Glossary of Terms

ERM - effects range median
LC_{50} - lethal concentration affecting 50% of population
LOEC - lowest observed effect concentration
MSD - minimum significant difference
NT - not toxic
RA - risk assessment
SQC - sediment quality criteria
SQG - sediment quality guidelines
T - toxic
WSC - whole-sediment chemistry
WST - whole-sediment toxicity
PWT - pore-water toxicity

2.0 Background for Contaminated Sediment Assessment

- Tier 1 - WSC vs. SQC, SQGs, etc.
- Tier 2 - Suite of Toxicity Tests
- Tier 3 - Other tools;
- BC Framework - Site is contaminated if WSC > SQC
  - Option 1 - Cleanup to SQC
  - Option 2 - Risk-Based Approach (Additional Sampling; Toxicity/chemistry)

3.0 Toxicity Test Interpretation

Work group members identified several consideration for the interpretation of whole-sediment and pore-water toxicity tests that included:

- Comparison to a predefined benchmark (e.g., 30% reduction in survival) (e.g., MSD, Policy decision - professional judgement
- Verify that test is acceptable

  - Negative control
  - Positive control - status of test in population (LOEC vs. EC_{50} target LOEC - this hasn't worked well in the past
- Use multiple benchmarks - e.g., SCCWRP 4 category model and carry results into weight-of-evidence assessment; Yes/No - Doesn't differentiate between samples with different magnitudes of toxicity
- H, M, L Bin and No Effect.

4.0 Regulatory Perspective

- Simple is better
- T/NT is helpful for decision making
- Currently CA (?) Survival of <80%
- Compare to Reference
- Start with concentration-response plots (collective professional experience; identify the EC_x that corresponds to threshold for benthic response (EC_x for benthos) or MSD - statistical analysis of large data sets
- Need rules for minimum number of samples for concentration-response relationships
- Need to understand biology of tests (e.g., polychaete male (much larger)/female - sex-ratio affects growth endpoint)
- Note: Need more formalized Industry Policy on the following two points
  - Reference condition - Pristine vs. Urbanized Harbour
  - Need for watershed approach to address sources as a whole within the watershed OR set reference conditions - contemporary background.

5.0 Role of Reference Conditions in Interpreting Results of Toxicity Tests

- Ministry Policy: Don't clean-up to below background levels
  - Regional (?) Background for metals in soil
  - Define local background due to natural process
- Need to incorporate reference data into interpretation framework

6.0 Recommended Framework for Establishing Toxicity Designations

- Step 1: Conduct whole-sediment toxicity test
- Step 2: Evaluate test validity
  - Negative(?) control
  - Reference toxicant
  - Data quality objectives
- Step 3: Compare Results to Negative Control Data
  - MSD (Be prepared to flag samples)
  - EC_x-based on comparison to benthic response
  - Evaluate power of test (variance)
  - MSD + Sig = Tox
- Step 4: Compare Results to Reference Results
  - More guidance needed on selection of reference sites;
- Step 5: Designate as Toxic if negative control + reference

7.0 Recommended Framework for Interpreting Toxicity Test Results

- Option 1:
  - Designate preliminary toxicity (using above framework)
  - Generate concentration-response relationships
  - Reclassify samples in risk categories based on concentration-response relationships (e.g., low, moderate, and high);

- Option 2:
  - Designate samples into multiple categories at outset (like SCCWRP);

- Don't lose information on toxicity as we work through the process

Appendix 4B Use of Sediment Toxicity Test Results to Evaluate Risks to Sediment-Dwelling Organisms Exposed to Contaminated Sediments - *Summary of the Results for Work Group 2*

Our Question

At contaminated sites, what tools and approaches should be used to interpret the results of whole sediment and porewater toxicity tests on decisions for management of contaminated sediments?

Debated

- How to interpret toxicity test data (surrogate effects data) on the continuum of responses in a species sensitive distribution (SSD)?
- Group underlined that toxicity test data are surrogate data are for the real world
- Toxicity data measure effects on individuals, which we implicitly link to populations/communities
- Toxicity test data provide insight - but are not a direct measure of community responses.

In a Perfect World

- We would know where toxicity test results fall in terms of their relevance - to field measures, etc. - the answer matters because it will influence the way we use and evaluate toxicity test data
- Recognize that there is not a bright line of effect/no effect - but rather a continuum… but regulators need yes/no.

Contaminated Sites - Protection Goal

- Goal is narrative - Question: How to apply to sediment toxicity tests?
- At this point - discussion diverged (rich discussion about defining populations, dilution of effects, spatial scale, etc.)
  - 20% of species?

- 20% effect/impairment (growth/reproduction) on population of a site?
- Need to define a population
- Interpretation of guidance has been left to regulators
- Can't have lethal endpoint, must be chronic
- Can't use dilution to resolve the issue
- Onus is on the responsible party
- Technical input to MoE
- EC_{20} is a starting point - onus is on the site owner to make an argument that something else should apply, or that other lines of evidence should be brought to bear
- Bright line - that can then be interpreted…

Fundamental difference of opinion in the group.

1. Whether to interpret toxicity test data as toxic/non-toxic vs a continuum of response.
2. Decided to disagree (and what we presented is more consistent with what regulators need, with the understanding that the interpretation would be part of risk management or could change with additional data).

Do we evaluate the tests one-by-one, or together (and then the link to other LOE). Conceptually, does one consider these to be independent of each other (lumpers or splitters).

Assume that three or more toxicity tests have been conducted…

1. Choose the appropriate reference site(s) (this should be done a priori and as part of PF. Default should be to avoid laboratory negative control for statistical purposes. Confirm that reference site is suitable before using it. Confirm that the test results are valid with respect to quality assurance/quality control.
2. Compare the concentrations of measured confounding factors to benchmarks (e.g., ammonia).
3. Assume initially that the laboratory tests will be equally weighted. If this turns out not to be the case (e.g., confounding factor, realism of sample, BPJ etc.), then the risk assessor should provide a rationale and alter the weighting.
4. Look at concentration response relationship - look at contaminants that are driving the potential risk (use HQs) - could use PCA or like analyses. Don't have high expectations for a clear outcome, but try this anyways.
5. The effects data should be corrected for reference site results (or another correction that is technically defensible for the site) and then evaluated against BC MoE's "EC20 protection goal". This concept is in essence a policy-driven decision point. If you are

below it, then this line of evidence "passes", if you are above it - either you need more information or risk management is applied.

6. If sufficient data exist to calculate MSDs, then this should be explored.
7. Once the bright line has been evaluated, then it is further evaluated as part of the overall weight-of-evidence.

General Guidelines for Test Interpretation

Do's

- Use a battery of tests that cover exposure routes and endpoints/taxa
- Present statistical results completely - power, ECx, statistical findings
- Up front, review the data in view of confounding factors - to determine the likelihood these factors are affecting results
- Consider the population on the site
- Uncertainty assessment.

Don't

- Eliminate a test species, just because it doesn't live at your site, or in your region (it is a surrogate)
- Rely solely on statistical significance
- Don't present the results as toxic / non-toxic (overstates).

Spirited discussion on Yes/no vs. some kind of fractured continuum (negligible, low, med, high)

Appendix 4C Use of Sediment Toxicity Test Results to Evaluate Risks to Sediment-Dwelling Organisms Exposed to Contaminated Sediments - *Summary of the Results for Work Group 3*

General comments (regarding preferred features for designating samples as toxic/not toxic):

- Assessment endpoint related to measurement endpoint related to toxicity determinations
- Sample size - affects how tests interpreted
- Information required to interpret growth tests
- Yes/no determination - lose too much scientific information (therefore a multiple category system is preferred)
- To determine ‘accuracy’ of toxicity test designations - revisit sites to see how effective decisions were
- Amphipod toxicity tests typically used to establish relationship between toxicity test results and benthic community health
- Number of replicates in toxicity test is a consideration
- Should choose a method that provides the lowest level of uncertainty
- Method should consider ecological relevance of toxicity test
- Framework should have two components - look at magnitude of response, and response compared to control. Increases confidence in designations
- Beneficial to relate studies from different areas.

General comments (regarding Tier 1 Guidance in BC)

- Preference given to non-lethal effects
- Based on $EC_{20/50}$ type values (is this ecologically relevant?); can actually measure, therefore operationally effective
- Somewhat arbitrary (i.e., based on distribution; develop dose-response curve and then determine the 20% effect level)
- Usually not enough samples to incorporate background/reference response.

Focus Question #1
What are the candidate approaches for designating freshwater, estuarine, and marine sediment samples as toxic and not toxic?

Focus Question #2
What are the strengths and limitations of each approach to designating sediment samples as toxic and not toxic?

Reference Envelope

Strengths
- Good for hazardous waste sites
- Statistical
- Ecologically relevant
- Addresses the magnitude of results
- Could establish a permanent reference site or sites for a region (to address costs associated).

Limitations
- Hard to identify acceptable reference sites
- Not specific to protection goals (e.g., EC_{20})
- Expensive - requires a large sample size.

Minimum Significant Difference (MSD)/Minimum Detectable Difference (MDD)

Strengths
- Removes statistical constructs/artifacts (e.g., false +'ves, skinny hits)
- Statistical.

Limitations
- Not ecologically relevant?
- Necessary to establish so that the MSD is relevant to a region (multi-lab analysis)
- Test specific
- Generated based on a large data set then applied to other studies
- Make recommendation to the ministry that there should be a study funded to develop MSD for various toxicity tests?

Comments
- After generating the MSD number(s), these can be refined as additional data are collected that can inform the analysis. This allows both for decision-making in the near-term, and allows for flexibility/refinement
- MSD concept is not dissimilar to $EC_{20/50}$ approach.

Statistical Significance

Strengths
- Speaks to the certainty/uncertainty of result.

Limitations
- Not ecologically relevant?

Gradient Approach (example discussed was the California method)

General Discussion/Info:
- All toxicity designations are control-adjusted
- Not toxic - within acceptable control response range (whether statistically significant or not)
- Low - Statistically significant but within MSD value
- Moderate and High - based on magnitude of response.

Strengths
- Multiple category approach - retains more information (i.e., toxicity data not broken down to a yes/no answer).

National Sediment Inventory Approach

Strengths
- Based on round robin testing results
- Based on a large data set
- Tiered system
- Test specific.

Limitations
- Somewhat arbitrary.

Focus Question #3

Which approach (or approaches) is the most relevant for designating individual sediment samples as toxic and not toxic in British Columbia?

General Comments on Recommended Approach:

- Incorporate magnitude of response
  - Cutoffs derived from a separate analysis and have ecological significance
- Incorporate statistical significance (provides a measure of certainty/uncertainty)
- Ideally, would be ecologically relevant (but no way to establish this with confidence)
- Incorporate a multi-classification system for grading test results
  - What do these categories represent?
    - Arbitrary (i.e., 20% / 50%)
    - Or establish a process (the process should be the same for multiple toxicity tests; have the same philosophical basis)
- Handling of Type I/II errors should be consistent with management goals.

Recommended Approach:

- System used in California

	Response within Control Performance Range	Outside Control but Less than MSD, 20%, etc.	Greater than MSD, 20%, etc. but less than severe effect	Excceds severe effect
Statistically Significant	"OK"	"Low"	"Moderate"	"High"
Not Statistically Significant	"OK"	"OK"	"Low"	"High"

Focus Question #4

What should be done when the toxicity test do not concur in their finding of "toxic" for a sediment sample?

Not addressed.

Focus Question #5

What should be done when controls fail and/or there are suspected test problems (e.g., grain size problems, unusual test results, such as no relationship between chemistry and toxicity)?

Not addressed.

The presentation that Work Group 3 presented to the workshop participants follows.

Focus Question

- At hazardous waste sites, how should the results of toxicity tests be interpreted to support decisions on the management of contaminated sites?

Overarching Principles

- Acceptable v. unacceptable risk is not intrinsically the same as toxic v. non-toxic
- Need to avoid losing information in process of reducing outcomes to toxic v. non-toxic
- The fundamental concept of a toxicity test is the demonstration of an interpretable exposure (dose) response relationship over a range of exposure (dose)

Overarching Principles

- The endpoints in toxicity tests are generally measurement endpoints, not assessment endpoints
- Properly defining "effect" in a toxicity test requires:
  - A definition of what is an unacceptable ecological effect
  - An understanding of how unacceptable ecological effect relates to toxicity as measured in a toxicity test

Interpreting Toxicity Tests

- Information comes from many approaches
  - Statistical significance within test
  - Statistical significance among tests (MSD)
  - Reference envelope
  - Point of comparison (reference, control, etc.)
- Preserve magnitude of response and exposure response range information as far into assessment as possible
  - Don't simplify to "acceptable/unacceptable" until the very end of the assessment

What is the Question?

- Classification criteria need to be tightly linked to the question being asked (purpose of the assessment)
- The question being asked is not always the same
  - Chemicals concs are elevated – are there likely to be toxicological impacts?
  - Where within this site are toxicological stresses most severe?
  - What level of biological effect is "acceptable" (i.e., remedial goals)

Need to Think About Consequences of Errors

- Definition of toxic depends on how error is to be handled
  - Screening in or screening out?
  - Would we rather make type 1 or type 2 errors?
- Explicit recognition that errors exist, and we need to decide how the consequences play out

Single Test is Part of a Larger Picture

- Determining what an EC20 is in an exposure gradient is different that deciding whether an individual sample caused an effect greater or less than 20%

Are the Tests Up to the Challenge?

- Need to be cognizant of the ability of a test procedure to detect the level of effect considered to be unacceptably adverse

"Toxic" is a Loaded Term

- "Toxic/not toxic" determinations on single sediment samples are not definitive risk statements – they are generally intermediate calculations in a broader analysis
  - e.g., T v NT on individual samples may be needed to generate a PEL, but the PEL is really the outcome, not the T/NT on individual samples.

Interpretation Categories

- Statistical significance within a single toxicity test is a statement about experimental variability relative to the magnitude of response – it is not directly related to expected ecological impact
- MSD type approaches help to put a context on the statistical significance (e.g., is the finding an artifact of unusually tight variance), but is still primarily a statistical finding, not a biological finding

Interpretation Categories

- Reference envelope approaches begin to make statements about whether the degree of biological effect lies outside what is considered "normal" biological response (not that they are immune to measurement variability)
- Exposure gradient approaches begin to relate a degree of response to a degree of exposure – but, they assume you know the appropriate exposure expression

Example Response Categorization

	Response within Control Performance Range	Outside Control but Less than MSD, 20%, etc.	Greater than MSD, 20%, etc. but less than severe effect	Excceds severe effect
Statistically Significant	"OK"	"Low"	"Moderate"	"High"
Not Statistically Significant	"OK"	"OK"	"Low"	"High"

Take Home Points

- A good approach should
  - Consider statistical significance, but not as a sole definition
  - Consider magnitude of response
  - Relate categorization of response magnitude relative to protection goals (which are, hopefully, related to ecological effects)
  - Have flexibility to revisit protection goals as understanding improves

Appendix 5A Weight-of Evidence Approaches for Assessing Contaminated Sediments in British Columbia - *Summary of the Results for Work Group 1*

1.0 Defining the Weight-of-Evidence Approach

Definition:

- A tool or mechanism to improve understanding, interpretation of, and inferences to be drawn from multiple lines-of-evidence to inform recommendations to be made to risk and site managers
- Facilitates prioritization for decision making
- Improves confidence regarding decisions about management of contaminated sediment
- Used within a Risk Assessment Framework
- Integrates multiple lines-of-evidence to assess risks to ecological receptors
- Can be done on a per sample or per site basis
- Used to determine if something needs to be done at site and, if so, where
- For bioaccumulation - If CBR in benthic assessment - aquatic: ERED DB (1400 records); terrestrial UDMEED; If food chain consider separately.

2.0 Applications of Weight-of-Evidence Assessment

- How to weight various lines-of-evidence and determine which ones should most influence decision-making
- Better understand mechanisms and relationships between chemistry and effects
- There a several candidate approaches, including quantitative, semi-quantitative or qualitative
- Should be linked to the COPC(s) at a site
- Should link degree of effects to ecosystem responses
- Candidate lines-of-evidence for benthic invertebrate assessments include:
    - Whole-sediment chemistry (metals, etc) - Tissue residues/critical body residues
    - Whole-sediment toxicity (multiple toxicity tests) - Others
    - Pore-water chemistry

    - Pore-water toxicity
    - Benthic invertebrate community structure
- Weight-of-evidence depends on the level of understanding/data needed to make good decisions
- Need to establish chemical causation?
- Desirable to identify chemicals or chemical classes causing effects, but not absolutely necessary (very important to resolve allocation issues)
- Are there off ramps? (e.g., certain concentrations of COPCs may get you out of process or generate a management decision).

3.0 Risk Assessment Applications at Small Sites

- Maybe better to focus of criteria-based approach
- Monitor small-craft harbours rather than active management
- Improve operating practices in harbours.

4.0 Prerequisites for Conducting Weight-of-Evidence Assessments

There are a number of steps that need to be undertaken prior to developing a WOE approach for assessing risks to benthic invertebrates, including:

- Data Quality Assessment must be done (compare to DQOs and power analysis)
- Data Sufficiency Assessment must be done (power analysis)
- Evaluate data applicability
- Relevance of data for addressing the questions being asked (DQOs)
- Do we have the right data to make decisions with required degree of certainty;
- Ensure that sampling sites are located appropriately, relative to sources, transport mechanisms, depositional areas, etc.
- Utilize systematic DQOs to process in study design (as per USEPA guidance).

5.0 Guiding Principles for Weight-of-Evidence Assessments

- Demonstrate that the WOE approach provides a predictive/reliable - relative to effects
- Demonstrate that weight-of-evidence approach provides a basis for understanding the results in a manner supported by body of knowledge about effects of COPCs in sediments (e.g., Koch Postulate to show causation)
- Consistent with narrative intent for sediment quality standards
- Reproducibility
- Transparency - clearly define the methods/decisions/language
- Reflects results of uncertainty assessment
- Documents confidence that can be placed in individual lines-of-evidence (consider variability vs. true uncertainty)
- Apply greater weight to lines-of-evidence in which we have more confidence;
- Design to lead to conclusions that are stronger as multiple lines-of-evidence (even weaker ones) point in the same direction
- Complexity of assessment should match the complexity of the site, but must meet minimum data quality and quantity
- Should requirements support sediment management decision (who gets to participate in decisions).

6.0 Other Considerations for Weight-of-Evidence Assessments

- Should lines-of-evidence be added? Multiplied?
- Should be divided evenly?
- Equal weighting of lines-of-evidence initially - Move off this if weighting improves interpretation of data (depending! E.g., DQOs)
- Don’t be prescriptive about weight-of-evidence methods
- Consistency with guiding principles is most important
- Three main approaches are available, including
    - Narrative Approach - Easy communication/explanation; More flexible
    - Numerical Quantitative Approach - Requires quantitative component to lines-of-evidence; May not be much different than narrative; Rule of 5
    - Bin Approach (as per California example).

Appendix 5B Weight-of Evidence Approaches for Assessing Contaminated Sediments in British Columbia - *Summary of the Results for Work Group 2*

The presentation that Work Group 3 presented to the workshop participants follows.

Workgroup 2
Seven Guiding Principles for Integrating WoE Decisions

Guiding Principles

1. Reproducible
2. Transparent
3. Addresses uncertainty
4. Reliability
5. Consistent with the narrative intent
6. Facilitate management decisions
7. Scientifically defensible

Prioritized Guiding Principles

1. Facilitate management decisions
2. Scientifically defensible
3. Reproducible
4. Transparent
5. Reliability
6. Addresses uncertainty
7. Consistent with the narrative intent

Candidate WoE Approaches (Q1)

- Best Professional Judgment
- Tiered Approaches
- Decision Matrix (e.g., Bay approach)
- Semi-quantitative (e.g., Calcasieu)
- Fully quantitative (e.g., Menzies)

Matrix Ranking of WoE Approaches

	BPJ	Tiered	Matrix	S-Q	FQ
1. Facil. Mngmt. Decis.	0				
2. Scientific. Defensible	0				
3. Reproducible	-				
4. Transparent	-				
5. Reliability	0				
6. Uncertainty	+				
7. Consistent w/ NI	+				

Matrix Ranking of WoE Approaches

	BPJ	Tiered	Matrix	S-Q	FQ
1. Facil. Mngmt. Decis.	0				0
2. Scientific. Defensible	0				0
3. Reproducible	-				+
4. Transparent	-				+
5. Reliability	0				+
6. Uncertainty	+				-
7. Consistent w/ NI	+				0

Matrix Ranking of WoE Approaches

	BPJ	Tiered	Matrix	S-Q	FQ
1. Facil. Mngmt. Decis.	0		+		0
2. Scientific. Defensible	0		+		0
3. Reproducible	-		0		+
4. Transparent	-		+		+
5. Reliability	0		+		+
6. Uncertainty	+		0		-
7. Consistent w/ NI	+		+		0

Matrix Ranking of WoE Approaches

	BPJ	Tiered	Matrix	S-Q	FQ
1. Facil. Mngmt. Decis.	0		+		0
2. Scientific. Defensible	0		+		0
3. Reproducible	-		0		+
4. Transparent	-		+		+
5. Reliability	0		+		+
6. Uncertainty	+		0		-
7. Consistent w/ NI	+		+		0
Overall Score	0		5		2

Matrix Ranking of WoE Approaches

	BPJ	Tiered	Matrix	S-Q	FQ
1. Facil. Mngmt. Decis.	0		+		0
2. Scientific. Defensible	0		+		0
3. Reproducible	-		0		+
4. Transparent	-		+		+
5. Reliability	0		+		+
6. Uncertainty	+		0		-
7. Consistent w/ NI	+		+		0
Overall Score	0		5		2

Appendix 5B Integration of Multiple Lines-of-Evidence to Assess Risks to Ecological Receptors Associated with Exposure to Contaminated Sediments - *Summary of the Results for Work Group 3*

Focus Question #1
What are the candidate approaches for integrating multiple lines-of-evidence related to sediment quality conditions?

Calcasieu (risk-based; dynamic)
California system (monitoring-based; prescriptive)
Best Professional Judgement (subjective)

Focus Question #2
What are the strengths and limitations of the various approaches to integrating multiple lines-of-evidence?

Calcasieu Example of Weighing Multiple Lines of Interest

- Objective of system: Risk-based; designed to integrate risk over different AoIs, reaches
  - Narrative to support scoring numbers
  - Convenient for large data sets (i.e., to summarize risks for each sample)
  - Dynamic approach
  - Iterative approach
  - Incorporates uncertainty
  - Deals with disconcordance.

California

- Objective of system: monitoring; assessment (limited)
  - Back-check on outcome - does the results of the integrated assessment make sense, relative to best professional judgement? (Test preconceptions about how the system performs)
  - Decision criteria based on good quantitative science

- Focusses on effect (not causes)
- System is prescriptive; benefit of this is consistency in classification over a large region, and consistency in calibration between lines of evidence
- The process for integrating multiple lines of evidence influences study designs (i.e., helps identify the need for data and types of data)
- Results help to identify and prioritize areas that require further study and/or require management decisions
- Results also help in assessment and making management decisions.

Best Professional Judgement

- Difficult to reproduce
  - Potentially introduces a lack of consistency between sites/projects.

Focus Question #3

What are the guiding principles that should be considered during the integration of multiple lines-of-evidence?

- DERA guidance has good/relevant guiding principles
- Reproducibility
  - Does this require a cook book?
  - Strongly affected by consultant experience
- Consistency
  - Establish clear decision criteria that
    - Allow for flexibility
    - Encourages reproducibility
  - Decision criteria should address which LOE are necessary
    - To identify data that are informative to decision-making process (e.g., 28-d *Hyalella azteca* survival vs. function studies)
    - To identify data that are oriented to assisting with remedial options
  - Decision criteria should address how to interpret the results of data collected; i.e., how to interpret toxicity testing data, benthic community assessment data, chemistry data (this exists - chemistry and administrative rules)
  - Decision criteria should address how to integrate LOE

    - Decision criteria should be established a priori (in DERA guidance and administrative rules?)

- Transparency

Focus Question #4

How should the results for multiple measurement endpoints and multiple lines-of evidence be integrated to obtain a weight-of-evidence for evaluating effects on aquatic receptors at sites with contaminated sediments?

- Criteria for weighing/integrating should be established *a priori*
- BC puts more emphasis on benthic invertebrate community assessments (compared to Superfund)
- Strengths of DERA guidance:
    - Provides relevant approaches/guiding principles
    - Is not overly prescriptive
    - Allows flexibility to risk assessor
- What is the target audience of the DERA guidance? (i.e., experienced vs. non-experienced risk assessor)
- Weighing/integration system should:
    - Relate to risk drivers and remedial decisions
    - Assess samples in a consistent way
    - Allow for the classification system to be refined to address site-specific conditions
- Don't need to use all LOE at all sites
- Weight-of-evidence - what is the information needed to reach decisions
- Don't want to lose sight of other receptor groups
- How many LOE? Risk assessment requires rigor.

General Discussion:

- Weighing and Integrating are different - weighing incorporates the uncertainty
- Unique situations (e.g., bioavailability issues)
    - Requires professional judgement
    - California "unknown" boxes

- Conflicting data - more data/information required (although often not an option) for Calcasieu, incorporated issues associated with conflicting data into weighing system
- Causality is the purpose of a risk assessment (if this is not addressed, then it's just an impact assessment). If causality is not established, then the results of the assessment will be to clean everything, walk away, or the results are simply a guess
- Superfund - there are some regulators currently (unofficially) recommending to not conduct benthic community assessments
- Weighing - if there is a high degree of confidence (low uncertainty), then weighing process is simplified; if there is high uncertainty, the weighing process provides a process for eliminating LOE, identifying if more data need to be collected, etc.
- USEPA Superfund 8-step process is tiered
  - Screening-level risk assessment does not incorporate weight-of-evidence (process to eliminate pathways that require assessment)
  - Baseline risk assessment does incorporate weight-of-evidence
  - Identify risk drivers
  - Identify causality
  - Criteria for weight-of-evidence decided during problem formulation stage (in a baseline RA) - guidance for this process is NOT prescriptive
  - This process is relevant to B.C. also
- Comment for the BC process - guidance (policy) at this stage would be beneficial (provide consistency) (current guidance is very open, not-prescriptive; this has advantages/disadvantages)
- Weighing/integration system addresses how to get a "number" from a suite of data, and also identifies where that "number" is relevant.

<u>Parking Lot Topics</u>

- Sediment resuspension - risks to water column receptors
- How to use/interpret indigenous species data (i.e., surrogacy)
- Control acceptability QA criteria
- Sediment-fish toxicity tests - have methods been developed?
- Identifying reference areas
  - How to do it?
  - Establish permanent/multiple sites?
  - Develop guidance on acceptable reference sites.

The presentation that Work Group 3 presented to the workshop participants follows.

Compare/Contrast California and Calcaa;lsdkjfa;ldkf

- While one numeric and the other narrative, they have more in common that difference
- California focused more on individual locations; Calcaalksdfj; a site/regional assessment

Different WOE Analyses

- Goal of determining the existence and/or extent of impact/risk
- Goal of defining key stressors and potential remedial goals for those stressors
- Location oriented (SCWRB) v. regional/aggregated (Calcaalsdkja;las)
- e.g., California approach neutral on cause
- May be sequential

Guiding Principles

- Transparency, consistency, reproducibility important
  - Don't want a system that is assessor-dependent
  - Want a system that has comparability among sites (e.g., severe impairment means the same across sites)
  - Clear decision criteria (both for individual lines of evidence and for integration of lines)
- However, need flexibility to incorporate unique information, circumstances

Elements of Weighting LOE

- Relationship to assessment endpoint
  - e.g., benthic community may have high weight because tightly linked to assessment endpoint
- Confidence in "accuracy" of line of evidence
  - How certain are you that the "message" from the line of evidence is what you presume it to be (e.g., if benthic community is confounded by other factors)

Pros and Cons of Prescriptiveness

- Prescriptiveness pushes the system toward consistency across sites and assessments
- Concern is that precriptiveness would remove flexibility to recognize unique issues, additional information, etc.
  - Don't want to throw away data
  - Need a feedback loop

WOE an Ongoing Process

- WOE evaluation often brings gaps in data and/or understanding to the fore
  - Don't wait too long to start thinking about this
  - Wherever reasonable, pursue opportunities to resolve apparent discordance, rather than "averaging" discordant information
- Feedback to risk management; may be more cost effective to just pursue remedial action than to invest in more assessment
- Communicate potential impact of resolving discordance on management decisions

Need for a Ground-Truthing

- To avoid having the WOE process take on a life of its own, need a backcheck on the outcome
  - e.g., California independent expert assessment
  - others likely to be less involved, but inclusion of a "reasonableness assessment" is important