THE Coastal Resource Coordinator's Bioassessment Manual

Report No. HAZMAT 93-1 (revised)

Prepared by

Hazardous Materials Response and Assessment Division
Office of Ocean Resources Conservation and Assessment
National Oceanic and Atmospheric Administration

Purpose and Use of This Guidance

This manual and any internal procedures adopted for its implementation are intended solely as guidance. They do not constitute rulemaking by any agency and may not be relied upon to create a right or benefit, substantive or procedural, enforceable by law or in equity, by any person. Any agency or person may take action at variance with this manual or its internal implementing procedures.

This document should be cited as: MacDonald, D.A., M.B. Matta, L.J. Field, C. Cairncross, and M.D. Munn. 2003. The coastal resource coordinator's bioassessment manual. Report No. HAZMAT 93-1 (revised). Seattle, WA. National Oceanic and Atmospheric Administration. 160 pp + appendices.

THE Coastal Resource Coordinator's BIOASSESSMENT MANUAL

Donald A. MacDonald
Mary B. Matta
L. Jay Field
Charles Cairncross
and
Mark D. Munn*
(E.V.S. Consultants, Inc.
Seattle, Washington)

December 1992 – Revised July 2003

Report No. HAZMAT 93-1

NOAA / Hazardous Materials Response and Assessment Division
7600 Sand Point Way NE
Seattle, Washington 98115

*Current address: U.S. Geological Survey
Water Resources Division
Tacoma, Washington

TABLE OF CONTENTS

CHAPTER 8 SUMMARY OF RECOMMENDED BIOASSAY PROTOCOLS

LIST OF TABLES AND FIGURES

TABLES

FIGURES

PREFACE

This manual has been designed for the use of NOAA's Coastal Resource Coordinators as an introduction to and a general guide for using bioassessment techniques for evaluating conditions at hazardous waste sites. It is meant to be a ready reference to help them make suggestions for or evaluate proposed ecological work plans for these sites. It is not meant to be a cookbook for work plan development. More detailed information on the use and applicability of these techniques is available from the CRC technical support staff in Seattle.

The manual has been designed in loose-leaf form so that it can be readily modified as changes occur in the rapidly developing field of bioassessment or gaps in the information are identified. An early planned modification will be an expansion of the discussion on laboratory bioaccumulation methodologies. The users should be aware that this guide is not all inclusive of the currently available bioassessment methodologies but does include those methodologies that have been most commonly used at sites of concern to NOAA.

At this time I would like to thank the people, who in addition to the authors, made this manual possible. Reviewers Edward R. Long, Rebecca Hoff, Nancy Beckvar, Peter Knight, Sean Morrison, Diane Wehner, Karen Wurst, Waynon Johnson, and Chris Mebane for their critical comments on earlier drafts of the manual; NOAA Hazardous Materials Response and Assessment Division librarian, John Kaperick for helping to obtain pertinent documents; technical editor Charlene Swartzell for making the final version readable; graphic artist Virginia Curl for cover graphics, and CRC Branch Manager, Alyce Fritz, for her support and guidance.

Donald A. MacDonald
November 20, 1992

CHAPTER 1
BIOASSESSMENT: AN OVERVIEW

DEFINITION

Bioassessment is the characterization of environmental conditions through the use of biological organisms. The major objectives of this document are to provide some general guidelines on: the application of bioassessment procedures to the different stages of the hazardous waste site remedial process, the design of bioassessment studies, the use of specific bioassessment methodologies, and the concurrent physico-chemical measurements needed. In addition, a summary of recommended toxicity testing protocols, most of which represent the present state-of-the-art, is provided for use in specific situations. The bioassessment methodologies discussed in this document are generally restricted to those applicable to aquatic environments since this is the environment of concern to the National Oceanic and Atmospheric Administration (NOAA).

There are three levels of questions concerning contamination of the environment near a hazardous waste site:

1. Are contaminants present?
2. Are contaminants bioavailable?
3. Are contaminants causing or have the potential to cause bioeffects?

While chemical analyses are an important first step in the characterization of hazardous waste sites, by themselves they yield little information on impacts or potential impacts to biological resources. The presence of a contaminant does not mean it is bioavailable, and the fact that it is bioavailable does not necessarily mean it is capable of causing bioeffects. In the context of this document, **bioeffect** means an effect on a biological organism, population, or community that is detrimental to the health of the organism, population, or community. Determining the potential for causing bioeffects is a necessary component in the determination of the environmental impact of a hazardous waste site.

Bioassessment methodologies use well-defined tests with biological organisms to determine biological sensitivities to contaminated soil, sediment, or water samples from hazardous waste sites. These methodologies can be grouped into four general categories: toxicity tests, bioaccumulation, biomarkers (biochemical effects, physiological effects, incidence of disease), and community studies. Individual bioassessment methodologies are incapable of proving cause and effect at hazardous waste sites. However, an integrated approach using

a number of assessment methodologies, both biological and chemical, can provide a preponderance of evidence linking observed bioeffects to a hazardous waste site. Chemical analysis of environmental media indicates if contaminants are present in the environment. Bioaccumulation studies indicate if the contaminants are bioavailable; this is especially important for chemicals that are not acutely toxic. Toxicity tests indicate if contaminated media are capable of causing bioeffects, in particular, acute toxicity. Community studies and biomarkers indicate if indigenous organisms, in the vicinity of the waste site, are experiencing bioeffects. Taken singly the data generated from each of these methods gives some information about conditions in the vicinity of the waste site, but, taken as part of a properly integrated bioassessment study, the same data can strongly suggest that the hazardous waste site is the cause of the observed bioeffects (Figure 1-1).

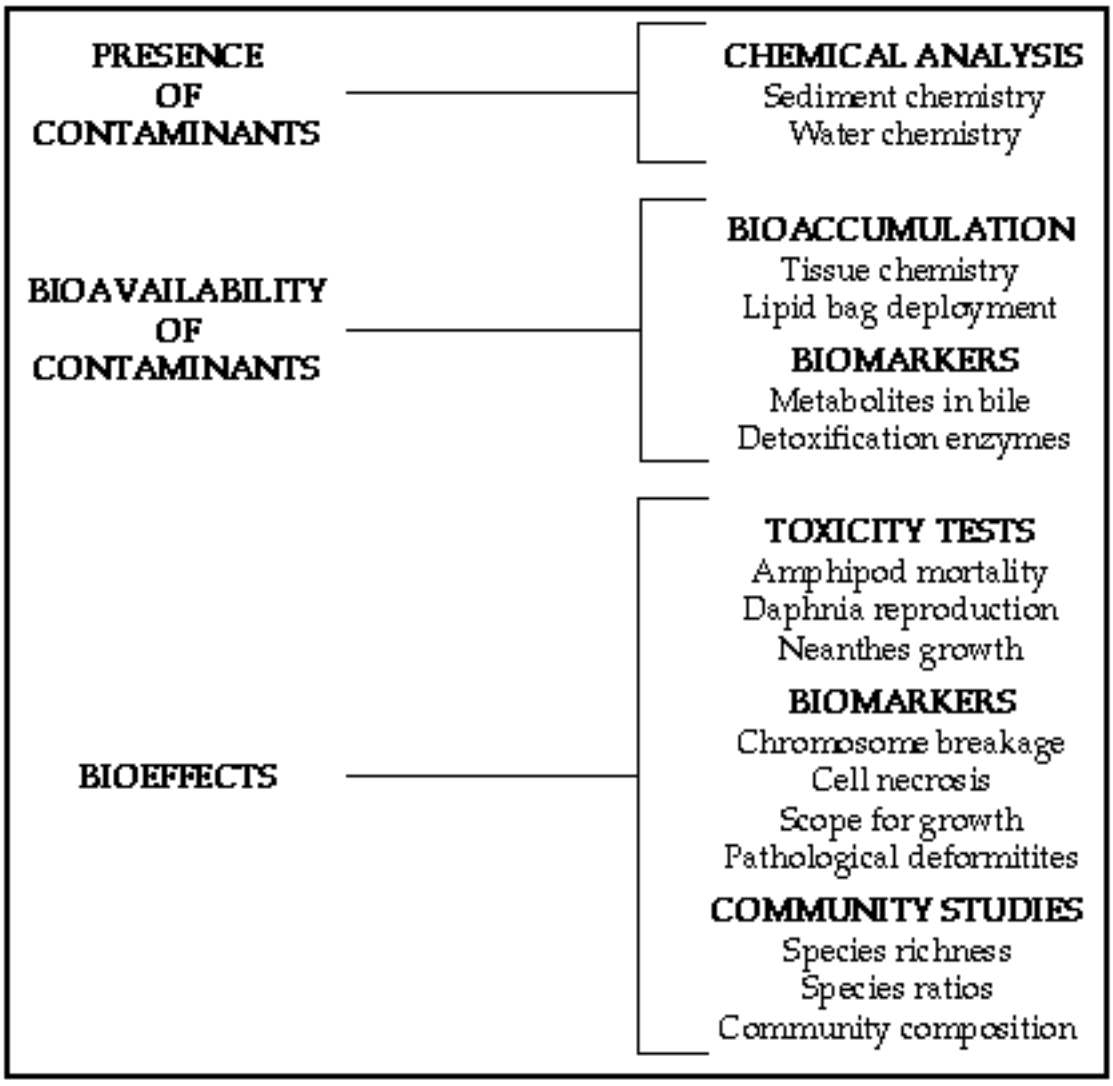

Figure 1-1. Assessment categories and example methods for addressing various levels of concern at a hazardous waste site.

CHAPTER 2
THE ROLE OF BIOASSESSMENT IN THE REMEDIAL PROCESS

INTRODUCTION

The selection of the appropriate bioassessment procedure(s) is dependent on the objectives of the particular stage in the remedial process. The amount of information required to make necessary decisions regarding further activities at the site tends to increase as the overall assessment of a site proceeds from the preliminary screening to the endangerment assessment and selection of remedial action. Regardless of the stage in the remedial process, it is of primary importance to clearly establish the intended application of bioassessment data prior to such data collection, including the specific questions that are being addressed, the proposed decision criteria, and appropriate sampling strategy and statistical design. If properly planned, data collected at each stage can be used to guide the development of work plans for subsequent stages. The early incorporation of an overall sampling design that adequately addresses the potential risks to natural resources reduces the possibility of additional sampling later that might delay or hinder the remedial process. The role of bioassessment approaches during the different stages of a site remediation is discussed in more detail in this chapter.

PRELIMINARY ASSESSMENT

The data needs for the preliminary assessment are usually directed toward broad-scale screening for evidence of release of contaminants to the environment and the presence or absence of toxicity. Bioassessment procedures, particularly bioaccumulation studies and toxicity tests, can play an important role in site screening.

The bioaccumulation of chemical contaminants in tissues of resident organisms effectively demonstrates the bioavailability of contaminants associated with the site. For contaminants known to bioaccumulate to a greater extent in higher trophic level organisms (e.g., PCBs, mercury), determining the degree of contamination in a representative higher trophic level organism may provide a worst-case evaluation appropriate for the preliminary assessment.

Combined with information on the site history, visual observations of site characteristics and chemical analyses, toxicity tests can provide qualitative information on the distribution

of toxic substances and the relative toxicity of the site. Toxicity tests can also provide an indication of the relationship between toxicity and chemical concentrations that may be particularly useful in defining areas where contamination is of concern. For this purpose, the use of two toxicity tests is recommended. As resources permit, additional acute toxicity tests could be employed to broaden the range of toxic substances that can be detected. This is especially important at sites with complex mixtures of contaminants or where the site history indicates the possible presence of substances for which analyses are not normally done (e.g., pharmaceutical or dye manufacturing). At sites where the presence of substances known to be especially toxic to a specific organism or type of organism is suspected, additional toxicity tests for those substances should be included. For example, since fish are particularly sensitive to the pesticides endosulfan, rotenone, and toxaphene, inclusion of an appropriate laboratory fish toxicity test would be recommended if the presence of one or more of these substances was suspected.

REMEDIAL INVESTIGATION

During the remedial investigation (RI) stage, bioassessment tools provide important information for the ecological risk assessment. They can be useful in characterizations of both the exposure and ecological effects. Chemical information obtained during the preliminary assessment can be used to help guide the site characterization during the RI phase of the study. Chemical analyses should provide information on the types of substances present and the possible ranges of their concentrations. Using basic information available for many compounds on the hazardous substance list regarding their toxicity to different organisms, chemical information can indicate which bioassessment approaches may be most fruitful. At sites where only one, or a few closely related substances of concern are predominant, it may be possible to select a bioassessment technique known to be sensitive to those substances. This is important when it is necessary to limit the number of different tests run on each sample in order to increase the total number of samples that can be tested. For example, acute Crustacea toxicity tests may alone be an adequate bioassessment option in situations where trace metals are the primary concern, since Crustacea used in common toxicity tests are reasonably sensitive to toxicity from most metals. This would allow an increase in the spatial or temporal coverage of sampling, while maintaining confidence that the toxic substances present were being detected. Where chemical information during the preliminary assessment indicates the presence of a complex mixture of contaminants, two or more different toxicity tests should be run on each sample to increase the likelihood that toxicity from the different toxic materials would be

identified. If the compounds such as PCBs or dioxin are present that are known to bioaccumulate and are not acutely toxic to most organisms, bioaccumulation studies may be the best approach.

Bioassessment procedures such as toxicity tests or benthic community studies provide a direct evaluation of the spatial distribution of toxic areas and an indication of the degree to which toxicity is associated with the distribution of one or more toxic substances. In the simplest case, the bioassessment acts as an adjunct to the chemical analyses. The chemical measurements provide the link between the spatial areas and the source(s) of contamination, while the bioassessment measurements determine the zones where the chemical contamination is sufficient to be toxic. In other situations, toxicity detected in the bioassessment may not vary spatially in the same manner as the majority of the chemicals measured; possibly because the biological test is responding to substances that are not detected in the standard chemical tests. In both cases, synoptic surveys of bioassessment and chemical measurements should be made. (Note that the spatial heterogeneity at hazardous waste sites is often very high, both horizontally and vertically. As a result, the bioassessment and chemical samples must, if possible, be taken as aliquots from the same homogenized sample.)

Toxicity tests have been used effectively to determine the extent of contamination at the Rocky Mountain Arsenal waste site in Colorado (Thomas et al., 1986). Using the results from a lettuce seed soil toxicity test and a statistical mapping technique called kriging, the extent of contamination at the site was determined solely by mapping the levels of toxicity.

In summary, the selection of the numbers of samples and the types of tests to perform must be considered on a case-by-case basis, weighing the trade off between collecting greater numbers of samples to improve spatial coverage and conducting more tests at each station to broaden the types of toxicity that will be detected. In general, a minimum of two toxicity tests should be performed. Additional toxicity tests should be included if the available data indicate the possible presence of organism-specific toxins or complex suites of substances. As resources allow, *in situ* bioassessments (e.g., benthic community assessments, incidence of disease, bioaccumulation) may be included in the site characterization. Although these types of tests tend to be less precise in their ability to define the spatial extent of problem areas, they are important in demonstrating impact to natural resources and supporting the results of the toxicity tests.

The data needed for the characterization of ecological effects from a site differ from those needed for the earlier phases of the remedial process. Impact evaluation should be based on

in situ measures of response, i.e., community effects and/or disease incidence. Toxicity tests, even *in situ*, are measures of the presence of toxicity and hence indicate possible effects, but only direct assessment of resident organisms can establish and quantify those effects. (Note, however, that the toxicity test and chemical data that is obtained for other purposes may be invaluable in demonstrating that any measured *in situ* effects are related to the contamination, and are not a result of natural fluctuations.) The characteristics of the habitats and populations that may be affected are major determinants in selecting bioassessment procedures for impact evaluation. The procedure selected must consider whether sufficient individuals are present at the sampling locations to obtain the statistical precision capable of allowing toxic effects to be resolved from natural variability. For example, small streams may have too limited a fish population to make creditable population or disease measurements possible.

The scope of work for ecological assessments is dependent on the balance that must be achieved between the resources available and the quantitative precision of the assessment. Both the selection of multiple assessment procedures and the number of samples collected directly affect the costs of the overall assessment. The selection task at most sites is generally simpler than implied above, however, because in most situations:

a) The populations suitable for assessment are in fact limited.

b) The area over which the exposure is expected is also usually limited, particularly with respect to the ranges of many organisms that would be tested.

c) At least the minimum numbers of samples required to obtain statistically meaningful data can usually be readily determined.

FEASIBILITY STUDY: BIOASSESSMENT AS A TOOL TO ESTABLISH SITE CLEANUP TARGET LEVELS.

The ultimate goal of site remedial activities is to eliminate exposing biological organisms to any toxic materials. Bioassessment procedures have obvious uses in meeting that goal by providing site-specific information on the distribution of toxic areas and the levels of site-related contaminants that are toxic. Bioassessments can be used in two ways to help determine target levels for cleanup.

The most frequently approach for using bioassessment procedures to establish cleanup levels is to determine the toxicity of a range of concentrations of the substances of concern at a site. The range of concentrations to be tested can be prepared by diluting a single sample

of contaminated media from the site in the laboratory with clean media, or by collecting a number of samples of contaminated media from different spatial areas that are known to have different chemical concentrations. The samples representing the range of chemical concentrations obtained by either approach are tested using one or more bioassessment techniques to determine which concentrations are toxic and which are not. The lowest contaminant concentration found to be toxic (induce the designated biological endpoint) is determined to be the target level for cleanup.

One example of this approach is the apparent effects threshold (AET), which uses field chemistry data (concentrations of toxic substances in sediments) and at least one biological indicator of injury (sediment toxicity tests, altered benthic infaunal abundance, bioaccumulation, histopathology, etc.) to determine the concentration of a given contaminant above which statistically significant biological effects would always be expected (Tetra Tech, 1986). The AET approach was developed in Puget Sound, Washington to establish chemical criteria for disposal of dredged material, and is being considered by U.S. Environmental Protection Agency (EPA) Region 10 for use in establishing cleanup target levels in that area. As currently developed, the AET approach should be used on a site-specific basis and can require extensive data collection. However, by carefully selecting one or a few appropriate bioassessment procedures, the approach can be cost effective at many sites. Further, if similar data are collected at different sites within a region that have similar habitat characteristics, the data may be suitable to combine into a regional database that might be used to develop standards with broad application.

A second approach that is particularly useful where the suite of chemicals is complex (e.g., landfills and hazardous waste recycling sites) is to use the biological responses measured by one or more of the bioassessment techniques directly to determine which areas of a waste site need to be cleaned up. For example, it might be agreed that all sediments in a stream that were significantly toxic in the applicable toxicity test (irrespective of any chemical measurement) would be treated to eliminate the toxicity. No remedy would be applied to sediments that were not significantly toxic. This approach was suggested by investigators at the Rocky Mountain Arsenal waste site where toxicity tests results showing toxicity above a pre-selected level were used with a statistical mapping technique (kriging) to define contaminated areas for potential cleanup. (Thomas et al., 1986).

In both approaches discussed above, it is important to note that there are currently no guidelines for selecting appropriate tests to determine target levels or to select the endpoint of those tests that have more than one possible endpoint. To use bioassessment procedures

most effectively for this purpose, at least general agreement among the parties involved must first be obtained regarding the acceptability of the approach and the bioassessment measurements that will be the basis for decisions. Ideally, this agreement should be reached early in the investigations when the scope of work is planned.

In many cases, the data obtained from the surveys to determine the extent of contamination can also be used to determine target levels, if the data originally collected included sufficient spatial coverage and/or sampling across gradients of chemical concentrations present. If these data needs are not met, additional sampling will be required. It is also possible that a biological endpoint or bioassessment procedure completely different from that used in site characterization may be selected to define the cleanup procedure.

REMEDIAL DESIGN/REMEDIAL ACTION: BIOASSESSMENT AS A TOOL TO EVALUATE THE SUCCESS OF REMEDIAL ACTIONS

In the same way that bioassessment procedures can be used to establish and map the toxicity of media associated with a site, they can also be used after remediation to confirm that toxicity has been eliminated. Bioassessment may also be used as a tool for monitoring any ongoing activities that are part of the remedy, such as discharges of treated groundwater. Toxicity tests are commonly used by EPA and other regulatory agencies to monitor effluents for National Pollutant Discharge Elimination System permits and for testing dredged sediments for open-water disposal.

Where the remedy has involved treating or isolating contaminated soils/sediments, toxicity tests may be the only useful immediate measure of the success of such remedial activities in those areas where the remedy has resulted in substantial disruption (i.e., from construction or dredging) of the natural system. In any case, the most logical bioassessment procedures to use to determine the "cleanliness" of the site are the ones employed to determine the toxicity of the site prior to cleanup. This should be especially clear in the situations where a particular test was used in the selection of the target cleanup level.

If post-remedy assessments are to be performed on sites where no previous bioassessment has been done, the selection of appropriate procedures can be developed following the guidelines for determining the extent of contamination (Preliminary Assessment and Remedial Investigation sections of this chapter). The level of effort will depend on the resources available and the level of confidence in the result that is needed.

REFERENCES

Thomas, J.M., J.R. Skalski, J.F. Cline, M.C. McShane, and J.C. Simpson. 1986. Characterization of chemical waste site contamination and determination of its extent using bioassays. Environmental Toxicology and Chemistry 5: 487-501.

Tetra Tech. 1986. Recommended protocols for measuring selected environmental variables in Puget Sound. U.S. Environmental Protection Agency, Puget Sound Estuary Program, Seattle, WA.

CHAPTER 3
TOXICITY TESTS

INTRODUCTION

A toxicity test is a procedure that involves the exposure of organisms to complex environmental samples (water, sediment, or sediment extract) under controlled conditions to determine if adverse effects occur. Test samples usually contain unknown amounts of mixtures of contaminants. This procedure is sometimes referred to as a bioassay, but 'toxicity test' is the more appropriate term because a bioassay is a test to determine the toxicity threshold of a specific substance, while this test is used to determine the toxicity of a whole sample, not its chemical components. Toxicity tests may be performed in the laboratory or with caged organisms in the field; however, protocols for *in-situ* toxicity tests are still in the developmental stage. Many organisms, from bacteria to mammals, have been used in such testing, and recently, toxicity tests using cell cultures and biomarker type indicators have been developed (see Chapter 2). These tests measure a variety of organism responses (endpoints) to determine the toxic effects of substances on biological organisms. While the most commonly measured endpoint is death of the organisms, other endpoints frequently measured include developmental abnormalities, behavioral changes, changes in reproductive success, and alteration of growth. Although some programs also measure the bioaccumulation of contaminants at the end of toxicity testing, the use of bioaccumulation as an endpoint is not discussed in this chapter. Toxicity tests are becoming increasingly important in bioassessment, in part, because they are relatively inexpensive and numerous tests can be conducted quickly. Although the following discussion applies to toxicity testing in general, most of the emphasis is on sediment toxicity tests since these are often recommended at waste sites.

Objectives of Toxicity Testing

Toxicity tests put environmental chemical data into perspective by acting as a measure of bioavailability. They can be used to document the extent of bioavailable contamination and to illustrate the potential for adverse effects at a waste site. Toxicity test results can be used early in the investigation of a site as a screening tool to indicate whether further bioassessment (for example, benthic community analyses) should be conducted. Results of toxicity tests can provide some information to indicate whether (and where) remediation should occur. Toxicity testing may help determine target levels for remediation. Results of toxicity tests alone will not be adequate for making remediation decisions; however, when

combined with chemical analyses and other bioassessment results, they are a useful tool. Toxicity testing can also be useful in monitoring the success of remediation.

Advantages of Toxicity Testing

The principal advantage and reason for conducting toxicity tests at a site is to provide a direct quantifiable measure of the potential for the occurrence of bioeffects at the site (Table 3-1). Because toxicity tests measure the relative toxicity of a mixture of chemicals, any synergistic or antagonistic effect between chemicals is automatically taken into account. Used in conjunction with chemical analysis, toxicity tests can be used to correlate toxic effects with concentrations of specific contaminants and thus serve as an indirect measure of bioavailability. Unlike chemical analyses, toxicity tests are not limited by a predetermined list of chemicals to be tested for. Therefore, if toxic effects are found and there is no correlation between the effects and the contamination levels, this could indicate that a chemical substance not analyzed for was causing the effect (i.e., you can't find something you don't look for). In general, toxicity test organisms are intended as generic representatives of sensitive organisms that could be present at a site, and thus can provide indications of whether conditions are toxic enough to kill or otherwise impact sensitive species. In other cases, a species of particular concern at a site can be used as the test organism to provide a more direct indication of potential effects to this species.

Since toxicity tests are based on controlled procedures that minimize natural variability of conditions such as temperature, it is easier to detect differences between sites than with other more variable bioassessment methods, such as benthic community studies. Unlike measures of benthic community structure, toxicity tests are not dependent on the presence of any particular *in-situ* population. Because the test samples can be collected from small, well-defined areas, the spatial resolution of toxicity test results is better than for most other assessment approaches. Many toxicity tests have well-developed procedures for which widely accepted protocols have been developed. Finally, they are quick, relatively inexpensive and are available from a number of agency and contract laboratories. As a result of these factors, toxicity tests are often the first choice for bioassessment at waste sites.

Table 3-1. Advantages and disadvantages of toxicity tests.

Advantages	Disadvantages
Provides quantifiable information about the potential for bioeffects at a site.	Not designed to mimic natural exposure, so may be difficult to relate directly to actual responses at a site.
Indirect indicator of bioavailability of contaminants.	Response not necessarily directly related to specific contaminant(s).
Response not restricted by predetermined list of contaminants.	If test organisms do not naturally occur at the site it may be difficult to relate effects on test organisms to organisms occurring naturally at the site being tested.
Indicates potential effects to sensitive species or to species of particular concern.	Tests are difficult to perform correctly by inexperienced laboratories.
Performed under controlled test conditions (i.e., minimizes natural variability).	Not surrogates for determining population changes.
Not dependent on the presence of any particular *in-situ* population.	Not appropriate for contaminants that cause subtle effects over long periods, or for those where the major concern lies in their potential to bioaccumulate or biomagnify.
Spatial resolution of toxicity test results is better than for most other assessment approaches.	May observe toxicity in unexpected places (i.e., 'clean' sites) due to unknown or unquantified factors.
Many have well-developed and widely accepted protocols.	
They are quick and relatively inexpensive.	

Disadvantages of Toxicity Testing

The principal disadvantage of toxicity tests is that while they may be a good measure of the potential for adverse environmental effects, they are rarely designed to precisely mimic natural exposure. As a result it may be difficult to relate the results directly to actual

responses at a site (Table 3-1). Toxicity tests primarily determine whether the samples tested can interfere with the biota, not whether they have actually caused effects at a site. They are not true *in-situ* measures of toxicity. Toxicity tests by themselves give no information on what contaminants are present in a sample. Care must be taken when running a toxicity test to assure that any measured toxicity is the result of the presence of contaminants and is not due to naturally occurring substances such as ammonia and sulfides. Although toxicity tests may be a quick and inexpensive bioassessment technique, many of the tests are difficult to perform correctly by inexperienced laboratories. Toxicity tests are not surrogates for determining changes in benthic or other populations. They are not appropriate for contaminants that cause subtle effects over long periods, or for those where the major concern lies in their potential to bioaccumulate or biomagnify.

TEST SELECTION

Before selecting a toxicity test for application at a waste site, the questions to be answered by the test must be clearly defined. If a specific hypothesis can be formulated, test selection and interpretation become much easier. Identifying specific concerns at the site will help define the desirable test parameters. These parameters will include: sample matrix (soil, sediment, water); test species and, if appropriate, life stage; duration of the test; and measurement endpoints (e.g., death, growth, reproduction). The character of the environment (e.g., wetland, freshwater, estuarine), the receptors of concern, contaminants thought to be present at the site, and known exposure pathways will all contribute to the selection of an appropriate test or tests. Table 3-2 provides a general outline for selecting appropriate toxicity tests. While it is usually preferable to select a test that uses generally accepted protocols (e.g., ASTM), new or modified tests that are more suitable to the specific site conditions should not be dismissed Selecting a battery of tests will greatly improve the chances of correctly interpreting the potential for toxic effects at a waste site.

Testing Water, Soil, or Sediment

Toxicity tests have been developed that expose test organisms to water, soil, or sediment. In water toxicity tests, the test organisms are placed directly in a sample of the water of concern. However, there are four different exposure scenarios for soil or sediment toxicity tests. First, organisms can be exposed to whole, intact soil or sediments ("bulk sediment" or the "solid phase") with, in the case of sediments, overlying clean water. Second, they can be exposed to soil or "sediment elutriate" where clean water is mixed with the test material. The mixture is then either allowed to settle or is centrifuged and the water phase is poured off to become the test sample. This elutriate sample contains the "suspended phase" if it still contains particulates, or is the "liquid phase" if the particulates have been removed by centrifugation or filtration.

Table 3-2. Steps in toxicity test selection.

DEFINE TEST OBJECTIVES

DETERMINE ECOSYSTEM TYPE
- **Terrestrial**
- **Freshwater**
- **Brackish (Estuarine)**
- **Marine**

DETERMINE MATRIX TO BE TESTED
- **Soil**
- **Water**
- **Sediment**
  - **Bulk**
  - **Elutriate**
  - **Extract**
  - **Pore Water**

DETERMINE TYPE OF TEST DESIRED
- **Acute**
- **Chronic**

SELECT TEST ORGANISM

DEFINE ENDPOINTS TO BE OBSERVED
- **Death**
- **Growth**
- **Reproduction**
- **Etc.**

The third scenario is exposing organisms to soil or sediment extracts, using a chemical extraction procedure similar to that used for sediment chemistry analyses. This extraction process isolates specific classes of contaminants (neutral, non-ionic organic compounds) while failing to extract others (metals and highly acidic and basic organic compounds). The fourth, relatively new technique involves exposing organisms to the interstitial or pore water present in a sediment sample. Pore water can be collected from wet bulk sediment through the use of centrifuges, squeezers and filters, or dialysis chambers. After the pore water is collected, water toxicity tests can be used. Although

techniques have been developed to use caged animals to measure changes in growth rates in the field, methods to expose organisms to sediment *in-situ* are still under development.

The choice of whether to perform either soil, sediment, or water toxicity tests is dependent on conditions at the site under investigation. Surface water in wetland areas, pools, streams, and rivers is often present near waste sites. Water toxicity tests are well developed, have standard protocols, use a variety of species and endpoints, and are relatively easy to perform compared to sediment toxicity tests. However, contaminants in surface water may not occur at acutely toxic levels because many contaminants are hydrophobic and are rapidly removed from the water column by adsorption to suspended particulate matter and sediments. Also, levels of contaminants in surface water are often highly variable over time; this is especially true if inputs are periodic rather than continuous. Bulk soil toxicity tests, like those for water, are relatively well developed, fast, and inexpensive and have standard protocols.

Sediments often act as reservoirs of contamination near waste sites. They generally have higher concentrations of contaminants than the overlying water and contaminant levels are less variable over time. Bulk sediment tests expose the test organisms to contaminated sediments covered with clean water. The test organisms can be either free swimming, recieving their exposure from contaminants that diffuse from the sediment into the overlying water, or they can be burrowing benthic organisms exposed through direct contact and sometimes ingestion of the contaminated sediments. However, there are a limited number of standard protocols currently available for sediment toxicity tests.

Sediment elutriate and extraction tests involve the transfer, by either simple mixing or chemical extraction, of contaminants from the sediments to clean water or solvent that is then used as the test medium. Equilibrium partitioning theory[1] suggests that the bioavailable toxic phase of sediments is the pore water phase, therefore it follows that a direct measure of sediment toxicity is one that is performed on the pore water phase (Long, personal communication). Since the resulting test matrix for elutriate and pore water tests is water, water toxicty test protocols have been adapted for use with these tests.

Toxicities determined for the different routes of exposure to sediment samples (bulk sediment, elutriate, extracted or pore water) may not necessarily agree with each other.

[1] Equilibrium partitioning theory states that the chemicals in sediments are distributed between the particulate (solid) phase and the aqueous (pore water) phase; this can be expressed mathematically as a partitioning coefficient.

Elutriate tests were originally developed for testing dredged material to simulate conditions occurring during open-water disposal and are not considered appropriate for testing the toxicity of *in-situ* sediments. The chemical extraction techniques, used for extraction testing, remove only certain contaminants so the test organisms are not exposed to the full suite of contaminants that are actually present in the contaminated sediments. Pore water techniques require specialized laboratory equipment, and contaminant concentrations may vary depending on the extraction technique. There are currently no generally accepted protocols for pore water extraction.

Studies reviewed by Ankley et al. (1991) indicated that pore water exposures provide more information on sediment toxicity than elutriate exposures. They found that pore water samples were consistently more toxic than sediment elutriate samples. However, they also found that pore water samples were sometimes more toxic than bulk sediment samples, possibly due to pH differences or to the dilution of toxicants by the addition of clean water to bulk sediment samples (Ankley et al., 1991). Chapman and Fink (1984) also noted differences between toxicity of bulk sediment and sediment elutriate. Toxicity of bulk sediment to larval polychaetes was generally greater than that of sediment elutriate. However, elutriate samples from some stations were toxic, while bulk sediment from the same stations were not (Chapman and Fink, 1984). Contaminants that have low solubility in water generally have lower toxicities determined by elutriate tests. Bulk sediment testing is currently the preferred method for testing sediments at hazardous waste sites to determine the potential for biological effects at the site. The sediment toxicity tests in the list of recommended toxicity test protocols (Table 8-1) are based primarily on exposures to bulk sediments.

Selecting a Test Organism

A wide variety of organisms have been used in toxicity tests. The most commonly used soil toxicity tests are the seed germination test (typically using common crop species), the root elongation test (most often performed with lettuce), and the earthworm test. Freshwater organisms used for water and sediment toxicity tests include algae (*Selenastrum*), daphnids, chironomids, amphipods, and fish (especially fathead minnow and rainbow trout). The most commonly tested marine and estuarine organisms are amphipods, mysids, and bivalve or echinoderm larvae. Luminescent bacteria have also been used in tests of water, sediment, and soil.

The choice of a test species need not consider whether the species is native to the area where the sample was taken. However, if an organism is available that is also native to the test site, its use can increase the ecological relevance of the results. If a test has not been developed using a species native to the test area, the use of a surrogate species still can provide valuable information. For example, a test using an amphipod native to the study area may be available, but the test may not evaluate reproductive effects. Combining the results of the amphipod test with results of a reproductive test using a surrogate species can provide a more useful suite of information.

The relative sensitivity to specific contaminants varies greatly among different organisms (see Table 8-2). The physiology and behavior of the species probably influences its response to contaminants. This selective sensitivity to toxic materials should be considered when selecting an appropriate test organism. If the presence of a particular group of contaminants is known, a species thought to be sensitive to those contaminants can be used in a toxicity test. When a complex mixture or unknown mixture of contaminants is suspected, it is generally advantageous to test two or more different test organisms to improve the chances of correctly identifying the presence of toxic materials. Also, in the advent of an unexpected failure of one test, there will still be useful information available to assess the potential for toxicity.

Some examples of test organisms are: *Photobacterium phosphoreum*., *Selenastrum capricornutum*, *Daphnia magna,* chironomid larvae, *Pimephales promelas*, *Neanthes* sp., and *Arbacia punctulata*. Each of these organisms have their advantages and disadvantages. The use of the Microtox® bacterial assay (*P. phosphoreum*) utilizing an organic solvent to extract contaminants appears to be consistently sensitive to some organic compounds (but not insecticides and herbicides). However, it is not very sensitive to metals (Munkittrick et al., 1991) and was found to be relatively insensitive to Prudhoe Bay crude oil (Buchman, personal communication). Also, it may be hard to show the environmental relevance of Microtox® test results. Algae such as *Selenastrum capricornutum* appear to be sensitive to metals and some organic contaminants, especially herbicides (Giesy and Hoke, 1990). Daphnids, especially *Daphnia magna*, are very sensitive to metals (Munkittrick et al., 1991; Giesy and Hoke, 1990). Chironomid larvae are also thought to be very sensitive to metals, especially when growth is measured as the response (Giesy and Hoke, 1990). Fathead minnows (*Pimephales promelas*) may be sensitive to PAHs and creosote, cyanide, and some metals. Amphipods and bivalve larvae are thought to be similar to each other, but less sensitive than the Microtox® test, in their overall sensitivity to contaminants (Williams et al.,

1986; Long et al., 1990; Pastorak and Becker, 1989). However, this relative sensitivity now appears to be site or species specific (Long, personal communication). Amphipods may be more sensitive to creosote and PAHs than the Microtox® test (Pastorak and Becker, 1989). The polychaete *Neanthes* appears to be sensitive to metals and PAHs (Johns, 1988). The embryos and sperm cells of the sea urchin *Arbacia punctulata* appear to be similar to oyster larvae in their sensitivity to metals (Nacci et al., 1986).

In general, only healthy organisms of similar size and life history stage should be used in toxicity tests. The organism should be appropriate for the material to be tested. For example, a planktonic or pelagic organism would be appropriate for a water toxicity test and a benthic organism would be appropriate for a sediment toxicity test. Taxonomic identifications of organisms must be confirmed by a qualified taxonomist.

Acute or Chronic Testing

Toxicity tests can be classified as "acute" or "chronic" tests. These terms refer to the duration of the test with regard to the life cycle of the organism being tested, not to the endpoint of the test. However, there is no universal agreement on the precise definition of these terms. For short-lived organisms, for example, daphnids, choronimids, and amphipods, the definitions are fairly straight forward and universally accepted: an **acute** test is one that is completed within a small portion of the organism's life cycle, while a **chronic** test exceeds at least one life cycle (Chapman, 1989). The problem with the definitions arises when long-lived organisms, such as fish, are tested. This is because it is impractical to run life-cycle-length toxicity tests when the life cycle is a year or more. Even if the life cycle is only a few months, life-cycle-length tests would be impractical on a routine basis. Therefore the question is: how long must the duration of a test be for it to be considered a chronic test when it is impractical to run it for one or more life cycles? There is currently no generally accepted answer to the question.

Endpoints

The **endpoint** of a toxicty test is the response of the organism that is used as a measure of toxicity. Endpoints can be classified as "lethal" or "sublethal." The only **lethal** endpoint is the death of the test organism and is reported as either percent survival or percent mortality. **Sublethal** endpoints do not involve the death of the test organism but are responses that could affect the survival of *in situ* populations. Sublethal endpoints include: developmental abnormalities, behavioral changes, changes in reproductive success, and physiological responses reflecting changes in enzyme activity and growth rates. While

acute toxicity tests often test for lethality, they may also use sublethal endpoints. Conversely, chronic tests often are designed to measure sublethal endpoints, but lethality could be one of the observed results. Therefore it needs to be kept in mind that the terms acute and chronic are not interchangeable with the terms lethal and sublethal.

Organisms can respond in many different ways to toxic materials. In selecting the response to be used as the endpoint of a toxicity test, consideration must be given to the ease with which it is quantifiable (counted, measured) and whether the response actually reflects an adverse impact on the organism. The use of a continuous response endpoint, such as growth, can increase the discriminatory power of a test (Giesy and Hoke, 1990). A range of responses then allows samples to be ranked by their relative toxicity. When death is the measured endpoint, samples can only be classified as toxic or nontoxic compared to control samples. In addition, growth and reproductive effects are usually more sensitive endpoints than lethality and may reflect the potential for community alterations.

Sediment Test Conditions: Static or Flow-Through

Sediment toxicity tests can be performed under "static" conditions where the entire test takes place in a closed container with test sediment and overlying water added only at the beginning of the test. These tests are simple measures of effects of acute (or sometimes chronic) exposures. However, waste product buildup could produce a toxic effect and cause erroneous results. Chronic tests can be done under static conditions, "flow-through" conditions, or with "static-renewal" conditions.

Flow-through: **Flow-through** sediment toxicity tests have a constant flow of clean water through the container of test material. While this may better simulate actual test site conditions than static tests, the actual toxicity of bedded sediments may be underestimated by these tests due to the dilution of contaminant concentrations.

Static renewal: Tests performed under **static-renewal** conditions involve the periodic replacement of water and/or sediment with clean water and/or fresh sediment from the test site during the course of the tests. Static-renewal test protocols have not yet been developed.

Cost Effectiveness

In a review of toxicity tests using seven different species Pastorak and Becker (1989) evaluated cost effectiveness as a combination of low cost and high overall sensitivity to

contaminants. The Microtox® test was determined to be the most cost effective. The use of two amphipod species (*Rhepoxynius* and *Eohaustorius*) and the measurement of developmental abnormalities in echinoderms (*Dendraster)* were determined to be moderately cost effective. Tests using polychaetes (*Neanthes*) and geoduck clams (*Panope*) were determined to be the least cost effective.

Although cost can be an important factor in the selection of toxicity tests, it should not be the primary criterion. Ideal tests should have high discriminatory power, low within-sample variability, and strong positive correlation with measured concentrations of contaminants (Long et al., 1990). Although the Microtox® test is relatively inexpensive and may be very sensitive to some contaminants, this does not necessarily make it the test of choice for all situations.

SAMPLING DESIGN CONSIDERATIONS

Designing a sampling plan for toxicity testing should be done in conjunction with a chemical analysis sampling plan. Ideally the samples to be used for chemical analysis should be split (subsampled) with one portion being chemically analyzed and the other portion being used for toxicity testing. At the very least samples for chemical analysis and toxicity testing must be taken at the same time and location. Without this conjunctive sampling it would be impossible to correlate toxicity to contaminant concentrations.

Another consideration when designing a sampling plan is the holding time for samples; for example, the time between sample collection and the beginning of the toxicity test. Many chemical and toxicity test protocols specify maximum holding times for environmental test samples in order to insure that the properties of the sample do not significantly change between sampling and analysis. For example, the prolonged storage or exposure to air of sediment samples will volatilize acid volatile sulfides (AVS) thus increasing the availability and toxicity of metals which are normally bound to AVS. However, the exact implications of exceedances of holding times are unknown for most chemicals. Resident infauna in sediments will eventually die, decay, and may produce lethal levels of ammonia. Bacteria present in the sample may continue to alter contaminants such as PAHs. Mercury for example, is thought to change more quickly than other metals. Since the different forms of many metals vary in toxicity, the response to test samples might change if holding times are exceeded. A basic rule of thumb is not to exceed holding times of two weeks for sediment toxicity tests when the samples are maintained at 4 degrees Celsius. While sediments for

chemical analysis can be kept frozen for several years before being analyzed, toxicity test sediments should not be frozen since this is thought to alter toxicity.

INTERPRETATION OF RESULTS

The interpretation of toxicity test results can be difficult. An observed toxic response may not correlate with measured chemical concentrations. If toxicity does not correlate with measured contamination it does not necessarily mean that the toxicity results are incorrect. Contaminants present in the sample may not be bioavailable. Also, a response may be caused by contaminants that were not measured in the chemical analyses. Natural factors such as grain size, ammonia, or sulfides also can produce a toxic response in some organisms. When two or more different toxicity tests are conducted, the results may not correlate with each other. This may be due to the differences in sensitivity between species to the mixture of chemicals in the sample. However, the most toxic samples will be those in which all the tests and endpoints showed significant effects.

Comparison with Control Samples

The interpretation of the results of toxicity tests is centered around detecting statistical differences between responses to test materials and to "negative controls."

Negative control: A **negative control** is a sample known to be nontoxic to the test organisms and in which they can function normally. Negative controls are a critically important factor in toxicity test studies. Negative controls should not be confused with reference samples. Reference samples are samples generally taken from the same system (i. e. stream, lake, estuary) as the test samples, but from an area not impacted by the hazardous waste site. Negative controls are used to evaluate the health and viability of the test organisms, the effects of handling the organisms in the laboratory environment, and the proper running of the toxicity tests. Therefore, they must not cause a significant response in test organisms (for example, death should occur in less than 10 percent of control organisms). Water toxicity tests can use distilled or clean seawater as negative control samples. The choice of an uncontaminated site to provide acceptable control sediment or soil is critical and often difficult. If the test organisms are collected from the wild, as opposed to cultured in a laboratory, then the sediments from their collection site can be used as a negative control.

Positive controls: Also, test organisms are usually exposed to "positive controls." **Positive controls** consist of a dilution series of water spiked with a toxic compound that produces a

response (the endpoint) in the species used for the test. Positive control samples help establish a dose-response relationship for the test species. These tests demonstrate the responsiveness of the test organisms for the endpoint of the test.

EC_{50} and LC_{50}

Toxicologists often report the toxicity of a substance as either an EC_{50} or an LC_{50} value. An **EC_{50}** is the concentration of a particular chemical associated with a sublethal response in 50 percent of the test organisms; EC stands for effective concentration. An **LC_{50}** is the concentration associated with death of 50 percent of test organisms; LC stands for lethal concentration. The results of toxicity tests on environmental samples are rarely reported as either EC_{50} or LC_{50} values because the samples consist of a mixture of chemicals, often unknown, and the toxicity cannot be associated with one specific chemical. However, if tests are run on a series of dilutions of the test sample, then the results can be reported as either an EC_{50} or an LC_{50} with regard to the sample dilution, not the chemical concentrations in the sample. For example, if toxicity tests were performed on a dilution series of a contaminated water sample and 50 percent of the test organisms were killed by the solution containing 40 percent test sample, the LC_{50} would be 40 percent.

Other Factors Influencing Toxicity Tests

In all experimental measurements, and especially those involving living organisms, it is important to identify outside factors that may interfere with a correct interpretation of the test results. The effects of unknown co-contaminants, impurities, and degradation products (such as ammonia or sulfides) in the test material can further complicate interpretation of the toxicity of the test material. Another factor that can confuse interpretation of toxicity test results is seasonal variation in the test organism's sensitivity to the substances being tested. The physical characteristics of the sample matrix can influence toxicity test results both by controlling the bioavailability of the contaminants and by directly affecting the test organisms. The latter is of special concern when the test organisms do not occur naturally in the area from which the test samples were taken. In the case of water these physical characteristics include pH, salinity, and temperature, while in sediments grain size, total organic carbon (TOC), and water content must also be considered. Very high concentrations of fines in apparently uncontaminated sediments have been found to be toxic to the the amphipod, *Rhepoxynius abronius*, who prefers fine sandy sediments (DeWitt et al., 1988). The toxicity was believed to be due to either the fine grain size, the high sediment water

content or high TOC, but because of the interrelatedness of these characteristics, the exact cause could not be identified.

There has been some research conducted to determine factors that control bioavailability (and presumably toxicity) of contaminants in marine sediment samples. AVS concentrations appear to reduce and preclude the bioavailability and, therefore, the toxicity of cadmium, and possibly other divalent metals, to two amphipod species (DiToro et al., 1990). Sediments containing a high percentage of fine grain material (silts and clay) and/or a high percent of TOC have the potential for containing higher contaminant concentrations than do coarser or lower TOC sediments. However, the contaminants may be less bioavailable due to binding by the fine grains and TOC material. These factors and the oxidation-reduction potential (Eh) in sediment can influence the distribution of chemicals between solid and aqueous phases, and can therefore influence availability to organisms. Disturbing a sample can also change the distribution of contaminants between solid and liquid phases. Other factors such as lighting, temperature, and pH influence the behavior of the test organism and can increase or decrease apparent toxicity. A well-developed protocol will include the control and measurement of as many of these factors as appropriate (especially grain size, TOC, AVS, ammonia, pH, dissolved oxygen, and temperature) to correct for influences that might lead to misinterpretation of the results. Accepted protocols should specify the use of both positive and negative control samples to show that response is caused by some toxic agent present in the sample, and not due to the laboratory environment or defects in the test organisms.

Determining the Cause of Toxicity

When attempting to determine which contaminants are causing toxicity at a site, it is imperative that chemical analyses are conducted on portions of the samples tested for toxicity. Correlation between toxicity and chemical concentrations may provide some first-order clues as to which chemicals are most highly associated with the observed toxicity. However, cause and effect relationships are not determined by correlation analysis. It may be impossible to identify the toxic component of a sample that contains high levels of a variety of contaminants. In such cases, toxicity identification evaluation (TIE) techniques can be used to identify the class of contaminant most responsible for observed toxicity (Figure 3-1). TIE procedures use chemical and physical fractionation techniques and toxicity testing to isolate the chemical fraction most responsible for observed toxicity. The fraction with the greatest observed toxicity can be chemically analyzed to determine compounds that are present at high levels. Although these techniques have been most widely used with

complex wastewater effluents, similar approaches have been taken with sediment pore water or sediment elutriate. At the present time, there are no TIE procedures available to directly test bulk sediment (Ankley et al., 1992). When the cause of toxicity in bulk sediment is to be determined, TIE procedures can be conducted on sediment elutriates or pore water samples. However, toxicity of elutriate or pore water must first be confirmed. TIEs are currently only in the research and development stage.

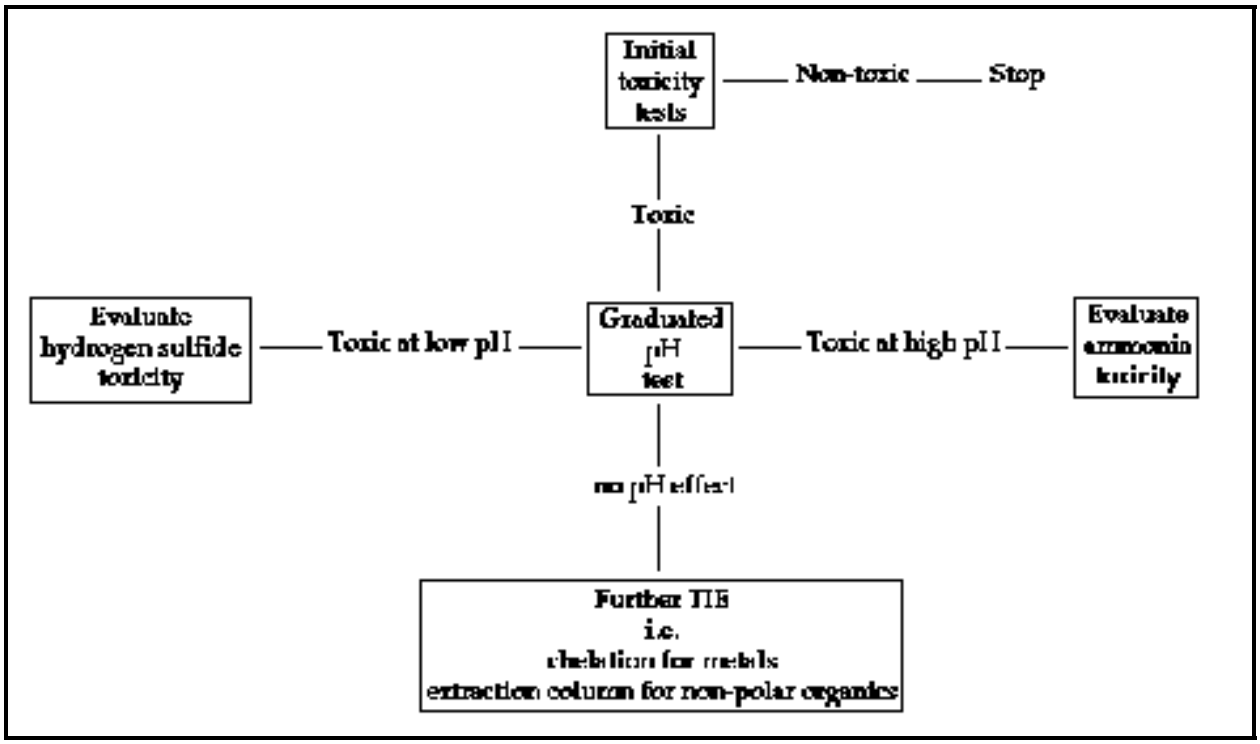

Figure 3-1. TIE strategy to evaluate contributions of contaminant groups (Ankley et al., 1991).

For sediment elutriates and pore water, a phased approach to eliminate possible groups of chemicals causing toxicity can be useful (Ankley et al., 1992; Giesy and Hoke, 1990). Ammonia can be eliminated as the cause of toxicity if toxicity does not occur in samples after pH has been increased, or if measured concentrations of ammonia are known to be below toxic levels. Similarly, hydrogen sulfide is more toxic at low pH values and toxic levels of hydrogen sulfide have been identified for many species. Cationic metals can be implicated as a cause of toxicity by testing the toxicity of a chelated sample. If toxicity decreases after chelation of the sample, cationic metals are implicated and further chemical analyses of the sample may indicate which metals are responsible. Nonpolar organic compounds such as PAHs, pesticides, and PCBs are implicated as a cause if toxicity is

reduced after the sample is passed through an extraction column to remove organic compounds (Ankley et al., 1992).

SUMMARY

Toxicity tests are a relatively quick and inexpensive means of determining if the environmental media (soil, water or sediment) in the vicinity of a hazardous waste site has the potential for adversely affecting biological organisms. The sensitivity of the various available tests is dependent on site conditions, species being tested, and duration of the test. The selection of an appropriate test and test organism at a particular waste site is dependent on the contaminants of concern, habitat types (terrestrial, freshwater, estuarine, and marine) present, and the media being tested. Once appropriate tests have been selected, both negative and positive controls as well as reference site samples need to be incorporated as part of the testing program. Finally, because test sensitivity is highly variable, depending on the test and test organism, a minimum of two different tests should be performed at each site, and when there is a wide range of contaminant types (metals, organics, etc.) more tests should be used.

REFERENCES

Ankley, G.T., M.K. Schubauer-Berigan, and J.R. Dierkes. 1991. Predicting the toxicity of bulk sediments to aquatic organisms with aqueous test fractions: pore water vs. elutriate. Environmental Toxicology and Chemistry 10(10) 1359-1366.

Ankley, G.T., M.K. Schubauer-Berigan, and R.A. Hoke. 1992. Use of toxicity identification evaluation techniques to identify dredged material disposal options: A proposed approach. Environmental Management 16(1). 1-6.

Buchman, M.F. 1992. personal communication. Hazardous Materials Response and Assessment Division, Office of Ocean Resources Conservation and Assessment, National Oceanic and Atmospheric Administration. Seattle, WA.

Chapman, P.M. and R. Fink. 1984. Effects of Puget Sound sediments and their elutriates on the life cycle of *Capitella capitata*. Bulletin of Environmental Contamination and Toxicology. 33(4). 451-459.

DeWitt, T.H., G.R. Ditsworth, and R.C. Swartz. 1988. Effects of natural sediment features on survival of the phoxocephalid amphipod, *Rhepoxynius abronius*. Marine Enviromental Research 25. 99-124.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environmental Toxicity and Chemistry 9(12). 1487-1502

Giesy J.P. and R.A. Hoke. 1990. Freshwater sediment quality criteria: toxicity bioassessment. In: R. Baudo, J.P. Giesy and H. Muntau, eds. Sediments: Chemistry and Toxicity of In-Place Pollutants. Chelsea, Michigan, Lewis Publishers, Inc. pp 265-348.

Johns, D.M. 1988. Puget Sound dredged disposal analysis sublethal test demonstration. Report for U.S. Army Corps of Engineers, Seattle, by PTI Environmental Services. Contract Number DACW67-86-D-0047. 94 pp.

Long, E.R. 1992. personal communication. Coastal Monitoring and Bioeffects Division, Office of Ocean Resources Conservation and Assessment, National Oceanic and Atmospheric Administration. Seattle, WA.

Long, E.R., M.F. Buchman, S.M.Bay, R.J. Breteler, R.S. Carr, P.M. Chapman, J.E. Hose, A.L. Lissner, J. Scott, and D.A. Wolfe. 1990. Comparative evaluation of five toxicity tests with sediment from San Francisco Bay and Tomales Bay, CA. Environmental Toxicology and Chemistry 9(9). 1193-1214.

Munkittrick, K.R., E.A. Power, and G.A. Sergy. 1991. The relative sensitivity of Microtox®, Daphnid, rainbow trout, and flathead minnow acute lethality tests. Environmental Toxicology and Water Quality, 6. 35-62.

Nacci, D., E. Jackim, and R. Walsh. 1986. Comparative evaluation of three rapid marine toxicity tests: Sea urchin early embryo growth test, sea urchin sperm cell toxicity test, and Microtox. Environmental Toxicology and Chemistry. 5(6). 521-525.

Pastorak, R.A. and D.S. Becker. 1989. Comparisons of bioassays for assessing sediment toxicity in Puget Sound. Final Report to EPA Region 10. EPA 910/9-89-004. 85 pages.

Williams, L.G., P.M. Chapman, and T.C. Ginn. 1986. A comparative evaluation of marine sediment toxicity using bacterial luminescence, oyster embryo, and amphipod sediment bioassays. Marine Environmental Research 19(3). 225-249.

CHAPTER 4
BIOACCUMULATION

INTRODUCTION

Bioaccumulation is the net result when the uptake of a chemical by a biological organism exceeds the depuration of the chemical from the organism. Uptake may occur directly from the air, water, soil, or sediment via absorption or indirectly through the ingestion of food containing the chemical. **Bioconcentration** is the process by which a chemical is directly taken up (by absorption only) from water and is accumulated to levels greater than those found in the surrounding water. **Biomagnification** is the increase in tissue concentrations of a bioaccumulated chemical as the chemical passes up through two or more trophic levels. A chemical is usually considered capable of being biomagnified if concentrations of the chemical increase by more than an order of magnitude at each step up the food chain. DDTs, polychlorinated biphenyls (PCBs) and mercury are among the few chemicals for which there is evidence of biomagnification.

Bioaccumulation studies take many forms. They may simply involve the measurement of tissue residues in indigenous organisms, or they may involve the measurement of tissue residues in test organisms exposed to contaminated environmental media (water or sediment) for a specific length of time. This exposure may be accomplished either by transplanting the test organisms to a contaminated area or by exposing them in a laboratory. The results of these types of studies are generally reported as the concentration of the chemical per unit weight of the organism (**body burden**) or some component of the organism.

Another type of bioaccumulation study uses artificial organisms (e.g., lipid bags). These artificial organisms are exposed to contaminated water or sediments for a specific length of time and then the concentration of the chemical(s) of concern in the lipid are measured. Since artificial organisms only take up chemicals through absorption, they only measure the potential for bioconcentration of the chemical(s). In studies with artificial or real organisms when the only route of uptake is through absorption (not through ingestion) the results are often reported as a unitless **bioconcentration factor (BCF)**. The BCF relates the concentration of a chemical in an organism (real or artificial) or component of an organism to the average concentration found in the surrounding water.

Advantages of Bioaccumulation Studies

The principal advantage and reason for conducting a bioaccumulation study at a site is to provide a direct measure of the bioavailability of contaminants there (Table 4-1). The chemical analysis of sediments may indicate that a site is highly contaminated, but it does not indicate that the contaminants are available to the biota. In the case of metals, they may be bound up in the crystal lattice of minerals making them totally unavailable for biological uptake. Also, like sediments, bioaccumulation analyses are an integrated measure of contamination levels over time, while water analyses give concentrations for instants in time. In addition to being extremely variable over time, contaminant concentrations in water are also relatively low. The measurement of these low concentrations requires laborious multistep techniques that are not only expensive, but increase the risk of laboratory contamination and dilution. Biota that concentrate contaminants with respect to the surrounding water, permit less complex and therefore less expensive analyses. Bioaccumulation studies can also indicate the potential for human health risks by analyzing organisms consumed by humans.

Disadvantages of Bioaccumulation Studies

Among the disadvantages of bioaccumulation studies (Table 4-1) is the lack of direct correlation between body burdens and bioeffects. **Bioeffect**, as used here, is a change in the condition or functioning of an organism resulting from exposure to a toxic chemical(s) that reduces its potential viability. While bioaccumulation does not necessarily indicate bioeffects, at least in the case of metals, bioeffects cannot occur without bioaccumulation (Phillips, 1977). This simply means that unless metals are retained in organisms at concentrations higher than normally found in a healthy organism (many metals are micronutrients), the organism will not be negatively affected by the metal. The same is not true for organic contaminants, some of which are metabolized with the metabolites being more toxic than the original chemical, e.g., polycyclic aromatic hydrocarbons (PAHs) in fish.

The high natural variability between individuals and species can sometimes make the interpretation of bioaccumulation data difficult. Some contaminants are readily metabolized by some organisms (PAHs in fish), while others are actively regulated so the uptake by an organism is not related to the environmental concentration of the contaminant (copper [Cu] in mussels, Phillips, 1977). Also, different contaminants compete with each other for uptake by organisms. Therefore, the presence of a particular contaminant or group of contaminants may inhibit the bioaccumulation of another contaminant. For example, the

uptake of Cu in mussels is influenced by the concentrations of zinc (Zn), cadmium (Cd), and lead (Pb) present (Phillips, 1977). Contaminants can inhibit the rate of biological processes that can then reduce the uptake of contaminants. For example, generally the faster an organism grows, the more rapidly it accumulates contaminants; so, if the growth rate is slowed by the presence of contamination, the rate of bioaccumulation will also be decreased. While this seems a formidable list of disadvantages, many can be minimized by the proper selection of test organisms and methodologies.

Table 4-1. Advantages and disadvantages of bioaccumulation studies.

Advantages	Disadvantages
Direct measure of bioavailability.	Relationship between body burdens and bioeffects uncertain.
Integrates contamination levels over time.	High natural variability between individuals and between species.
Concentrates chemicals from water allowing easier and less expensive analyses.	No direct relationship between body burdens and environmental levels for some contaminants due to bioregulation or metabolism.
Potential for determining human health risks.	Difficult to associate contamination in mobile species to area of environmental contamination.
	Uptake of one contaminant may be inhibited by the presence of other contaminants.
	Rates of biological processes maybe reduced by contamination thus reducing rates of bioaccumulation.

TEST ORGANISMS

Bioaccumulation field studies can be conducted by analyzing indigenous organisms, transplanted caged organisms, or artificial organisms. Each of these individual methodologies has advantages and disadvantages (Table 4-2), and the choice of which one to use at a particular site will be dependent on site conditions, contaminants of concern, and the specific objectives or purpose of the study. In addition, standard protocols currently exist for bioaccumulation studies with resident and transplanted organisms (e.g., NOAA's National Status and Trends (NS&T) Mussel Watch and California's Mussel Watch, respectively), while the use of artificial organisms is still in the experimental stages.

Type of Organism

The simplest and most straightforward bioaccumulation study is one involving the use of organisms that are indigenous to both the site of concern and the proposed reference site. The use of indigenous organisms permits the correlation of site contamination and the bioavailability of contaminants to the resident biota. While this method may appear to be ideal, the use of indigenous organisms does have certain drawbacks. The first and probably the most serious drawback is that the use of indigenous organisms requires that an appropriate organism (see below) is present and sufficiently plentiful at the contaminated site as well as at an appropriate reference site. There is also a high degree of variability between individuals of indigenous populations, and the locations of test stations are restricted by the presence or absence of the chosen test organism.

The use of transplanted organisms in bioaccumulation studies can overcome some of the drawbacks of indigenous organisms. A transplanted organism can be selected based on its ability to bioaccumulate the chemical or chemicals of concern at the study site and is not dependent on its actual presence at the site. Sufficient biomass for analysis is assured from the start by transplanting excess organisms to allow for any mortality. With transplanted organisms, investigators have more freedom in the selection of sampling locations. For example, organisms can be randomly or evenly spaced along a transect moving away from the site to determine the existence of a gradient.

Table 4-2. Advantages and disadvantages of the various types of organisms used in bioaccumulation studies.

Type of Test Organism	Advantages	Disadvantages
Indigenous	Indicator of what *in situ* organisms are actually accumulating Integrates accumulation over an extended time period	Locating an appropriate organism on-site and at reference site Collecting sufficient biomass for analysis Possibility of higher individual variability Test stations restricted by presence of organism
Transplanted	Direct indication of bioavailability Reduced variability between individuals Organism can be selected based on environment and contaminant(s) of concern Broader coverage of area of concern, gradients more readily determined	Relatively short exposure time More involved (expensive) methodology More difficult to relate results to *in situ* organisms
Artificial	Minimal variability between individuals Broader coverage of area of concern, gradients more readily determined Does not actively select for contaminants	Only measures bioconcentration component Only a simplified model of an organism May be difficult to relate data to real world

The principal drawbacks of transplanted organisms can be the relatively short exposure time and more involved and expensive methodology (increase in labor costs). While indigenous organisms may be accumulating site-related contaminants during their entire life cycle, transplanted organisms are generally only exposed during a small portion of their life cycle. The results of transplant bioaccumulation studies can indicate whether the transplanted organism is accumulating more contaminants at the site than at a reference site and thus infer that there are more bioavailable contaminants at the test site than at the reference site. However, the data may not be representative of the degree of bioaccumulation by indigenous organisms. When transplanted organisms are used for bioaccumulation studies, it is mandatory that a sample of the test organisms taken directly from the source of the organisms be analyzed for levels of the contaminants of concern. While not mandatory, it is highly desirable to have chemical data for the water and/or sediments from the source of the transplanted organisms.

Whether using indigenous or transplanted organisms, other parameters, in addition to chemical concentrations, need to be recorded. These parameters include: species, organism size, age, sex, spawning condition, lipid content, moisture content, and specific tissue type analyzed. All of these parameters will affect contaminant body burdens. For example, if the contaminant of concern is lipophilic, an organism with a higher lipid content will have higher contaminant concentrations than one with a lower lipid content (if they were both exposed to the same level of contamination). A large proportion of the lipids in a ripe female are located in the egg mass; when the female spawns the lipids are lost resulting in a reduced contaminant body burden for the female. Therefore, if recently spawned females were sampled without recording their reproductive condition the results would suggest that contaminant levels were lower than they actually were.

Artificial organisms are usually semipermeable membrane bags containing a lipid. These artificial organisms are passive accumulators while biota are active accumulators, and what the artificial organisms are actually measuring is the partitioning of contaminants between a water and lipid phase. Therefore, the results of artificial organism studies indicate the potential for the bioconcentration of lipophilic contaminants, i.e., the quantity of lipophilic contaminants that are available for direct uptake by absorption. The principal advantage of artificial organisms is the lack of variability between individual organisms in a study. As a result, all differences in accumulation of contaminants between test sites are due to site differences and not test organism differences. The major drawback to artificial organisms is that they only measure the level of contaminants available through passive absorption and

do not measure the level of contaminants that are available through active uptake mechanisms and through the ingestion of particulates, sediments, and food. Particulates and sediments may be significant sources of contaminants for filter and deposit feeders, respectively. Also, food is a particularly important source for those contaminants that biomagnify. Therefore, it must be kept in mind that artificial organisms are just simplified models of real organisms and, as such, are not able to indicate exactly what real organisms accumulate.

Test Organism Selection

The selection of an appropriate organism, whether indigenous or transplanted, for a bioaccumulation study is extremely important. Table 4-3 gives a list of criteria that should be considered when selecting an organism for a bioaccumulation study. The ideal organism should meet all of these criteria, but since bioaccumulation studies are conducted in the real world, that organism doesn't exist. Possibly the two most important criteria are numbers 1 and 8. The test organism should be able to tolerate the expected levels of contamination without any significant changes in its viability, including changes in its metabolic rate. Any such changes may reduce the bioaccumulation rate. The organism should also concentrate the contaminant or contaminants of concern. For example, while leeches meet most of the selection criteria, they would be a poor choice of organism for PCB studies because they do not accumulate organochlorides (Environment Ontario, 1988).

Sedentary organisms are preferred (criterion 2) because they can be associated with a specific location and the bioaccumulation data can be related to sediment chemistry. If biomagnification or routes of human exposure are of concern, then non-sedentary organisms from higher trophic levels might be preferable. However, it is difficult to associate non-sedentary organisms (e.g., fish) with specific locations and thus specific levels or sources of contamination. If they are used, then this lack of precision should be noted. The organism needs to be hardy enough (criterion 7) to transport to and survive in a laboratory in case depuration before analysis is necessary; this is particularly true of deposit feeders whose gut content could be a significant proportion of total body concentration of a contaminant (Chapman, 1985). In the case of transplant organisms, they need to be hardy enough to survive collecting, handling, and caging. Some organisms actively regulate levels of certain contaminants, especially those that are also essential micronutrients. This regulation can confound the correlation between organismal and water concentrations (criterion 9). For example, Cu is regulated by mussels, and, therefore mussels are not a good indicator of its bioavailability (Phillips, 1977).

Table 4-3. Organism selection criteria for bioaccumulation studies, adapted from Phillips, 1977.

1	The organism should accumulate the pollutant without being killed by the levels encountered.
2	The organism should be sedentary in order to be representative of the area of collection.
3	The organism should be abundant in the study area.
4	The organism should be long lived in order to allow sampling of more than one year class.
5	The organism should be of reasonable size, giving adequate tissue for analysis.
6	The organism must still be growing, i.e., producing new tissue.
7	The organism should be easy to sample and hardy enough to survive in the laboratory allowing depuration before analysis (if desired).
8	The organism should exhibit a high concentration factor for contaminants of concern, allowing direct analysis without preconcentration.
9	A simple correlation should exist between contaminant content of the organism and the average contaminant concentration in the surrounding waters.
10	All organisms in a survey should exhibit the same correlation between their contaminant contents and those in the surrounding water at all locations studied.

If a specific tissue of an organism is going to be analyzed for bioaccumulation then criteria 5, 8, and 9 should be applied in the selection of the tissue. Many organic contaminants are lipophilic and tend to accumulate more in liver than in muscle tissue because of the higher lipid content in liver. Because of this lipophilicity of some contaminants, lipid content of test organisms should be determined, especially if organic contaminants are of concern, and the chemical concentrations normalized for lipids. Since different metals tend to accumulate in different tissues depending on the organism (Hawker, 1990), it may be desirable to analyze more than one tissue type if a suite of metals is of concern.

While bivalves are possibly the closest thing to a perfect class of bioaccumulation organism and are frequently used in biomonitoring studies throughout the world, they are not appropriate for all hazardous waste sites. There is no one organism that is appropriate for all waste sites; the selection of an appropriate organism needs to be done on a site-specific basis.

SUMMARY

Bioaccumulation studies are a relatively straight forward method of measuring the bioavailability of specific contaminants. While there are three basic approaches to these studies, using indigenous organisms, transplanted organisms, or artificial organisms, the preferred approach at hazardous waste sites is the use of indigenous organisms. However, decisions on the approach as well as the specific organism or organisms to use must be based on site specific conditions.

Finally, bioaccumulation studies should not be performed in a vacuum; contaminant concentrations in sediments and/or water should be determined at the same sites where organisms are collected. A full-scale bioassessment study should include:

- Chemical analysis of pertinent media to determine levels of contamination present.
- Bioaccumulation studies to determine the availability of contaminants; toxicity testing to determine the toxic effects of contaminants.
- Benthic community studies to determine if the indigenous community has been affected.

REFERENCES

Chapman, P.M. 1985. Effects of gut sediment contents on measurements of metal levels in benthic invertebrates–a cautionary note. Bulletin of Environmental Contamination and Toxicology 35. 345-347.

Environment Ontario. 1988. The investigation, evaluation, and recommendations of biomonitoring organisms for procedures development of environmental monitoring. Ontario, Canada, Environment Ontario. 95 pp plus appendices.

Hawker, D. W. 1990. Bioaccumulation of metallic substances and organometallic compounds. D.W. Connell, (ed) Bioaccumulation of Xenobiotic Compounds. Boca Raton, LA CRC Press, Inc. pp 187-205.

Phillips, D. J. H. 1977. The use of biological indicator organisms to monitor trace metal pollution in the marine and estuarine environments–A review. Environmental Pollution 13. 281-317.

CHAPTER 5
BIOMARKERS

INTRODUCTION

A **biomarker** is a direct biological measure of the response of an organism to exposure to a contaminant. More specifically, biomarkers are "...biochemical, physiological, or histological indicators of either exposure to, or effects of, xenobiotic chemicals at the suborganismal or organismal level" (Huggett et al., 1992).

Exposure to a contaminant means that an organism has contact with and has taken up a contaminant. Under this definition four conditions must be fulfilled for an organism to be exposed to a contaminant:

1. The contaminant must be present in the same environment as the organism
2. The contaminant must be in a bioavailable form.
3. The organism must be able to come into physical contact with the contaminant.
4. Uptake of the contaminant by the organism must have occurred.

Uptake of a contaminant is defined as the passing of the contaminant from the external environment surrounding the organism, across a cell boundary layer, and into the internal environment of the organism. There are three potential routes of uptake: through the skin, through the lining of the respiratory organs, and through the lining of the gastrointestinal tract of the organism. Ingestion, the intake of substances into the gastrointestinal tract is not considered uptake.

Once contaminant uptake has occurred the organism can respond in various ways at various levels. The first level of response involves the accumulation[1] or excretion of the relatively unchanged contaminant; or, it involves the metabolism of the contaminant, with the metabolites being either accumulated or excreted. Accumulation of the contaminant or its metabolites can result in higher level responses by the organism. These responses can range from the molecular (e.g., genetic abnormalities) to the histopathological (e.g., lesions). It is these responses to contaminant uptake that are measured by the use of biomarkers. Figure 5-1 illustrates the general exposure-effects pathway and examples of measures of exposure, response, and effects that can be performed on indigenous aquatic organisms after contaminant exposure.

[1] Some environmental scientists consider bioacummulation of contaminants, as well as population, commmunity, and ecosystem measures of response or effects as biomarkers, but the most common usage restricts the term to suborganismal and organismal measures (Huggett *et al*., 1992; McCarthy and Shugart, 1990), and that is how the term will be used in this manual.

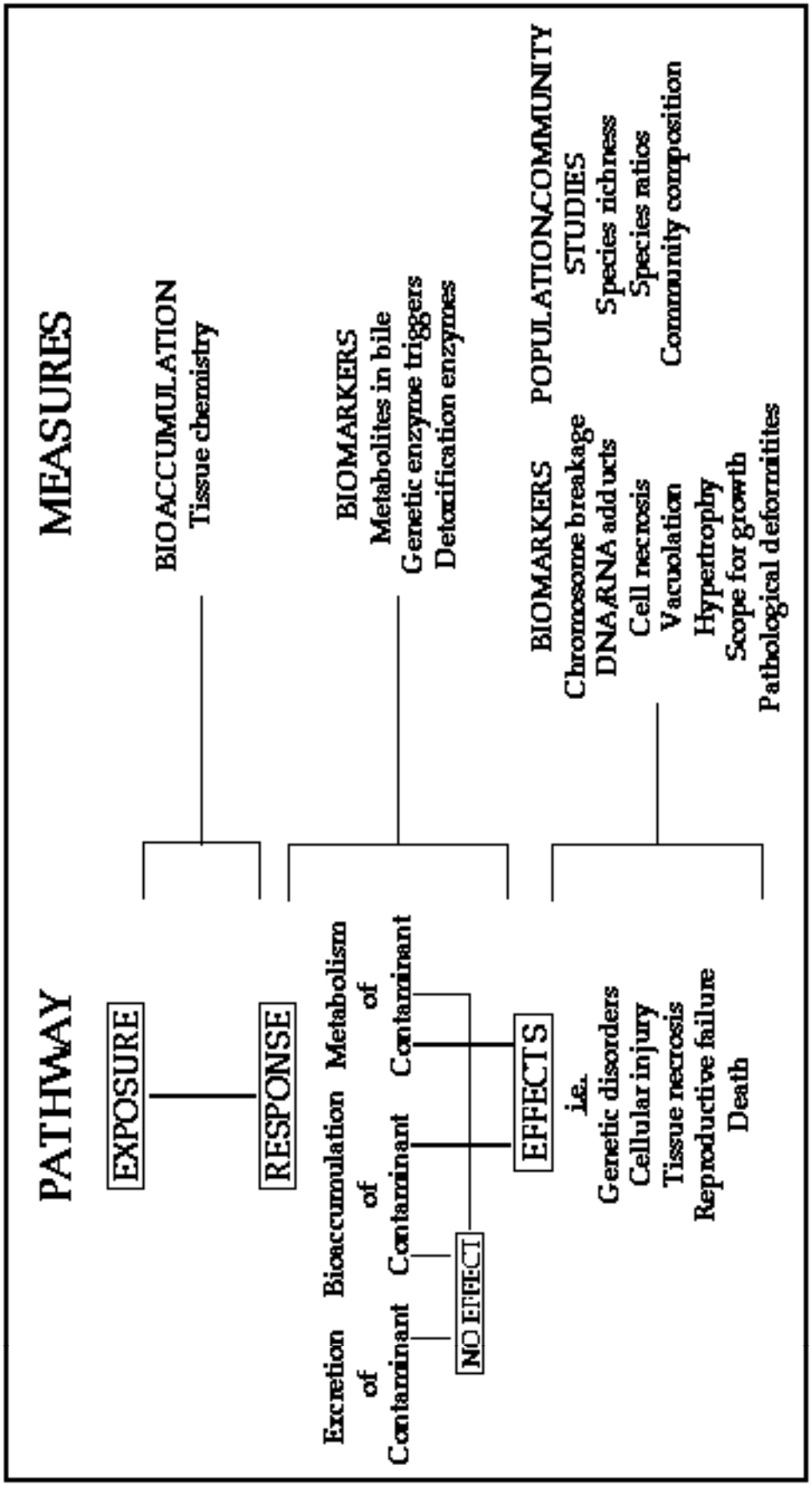

Figure 5-1. Exposure - effects pathway and potential indigenous organism measures of response and adverse effects.

It is important to remember that a biomarker is a measure of a response of an organism to exposure to a contaminant or class of contaminants; it may or may not be a measure of an adverse effect caused by the contaminant. However, the response measured may be a potential precursor of adverse effects, such as reduced fecundity or death, should exposure continue or increase. Exposure to contaminants may trigger the release of enzymes and other proteins that can catalyze further reactions. Measurement of these reaction products may indicate that an animal's natural detoxification mechanisms are functioning normally or, conversely, that they are overloaded. Furthermore, these reactions can indicate that the animal is stressed, that genetic material has been altered, or that reproduction might be affected. It should be noted that detoxification and metabolism are not synonymous since some metabolites are more reactive and more toxic than the parent compound. This is the case with metabolites of PAH, many of which are potential mutagens (Melancon et al., 1992). Histopathological changes within organs and other, higher-level physiological disorders such as skin tumors, fin erosion, skeletal defects, reductions in growth rate, and measures of reproductive health are effects biomarkers that have also been correlated with exposure to contaminants. These biomarkers have been measured in feral or caged organisms from contaminated areas to document exposure to contamination or actual adverse effects.

There are wide ranges of pathways by which contaminants can interact with biological organisms. Some contaminants (metal ions, for example) are free to bind with sensitive cellular components directly after uptake. Many organisms actively regulate uptake of essential metals. Other contaminants are transformed into more reactive components (the first phase of biotransformation), some of which are more toxic than their parent compounds. These toxic metabolites may then be free to bind with DNA or other proteins. Contaminants may be excreted through bile or urine, or may accumulate in fat stores or as granules (Luoma et al., 1991). A second phase in the biotransformation of xenobiotic compounds links metabolites to water-soluble conjugating compounds naturally present in cells. The binding of reactive contaminants with DNA and other proteins probably provides the subcellular mechanism for many adverse effects. Although the progression of the development of physiological disorders from subcellular reactions through gross pathologies has not been conclusively demonstrated, some evidence exists to link such conditions to exposure to toxicants.

Because of their recent development, most work with biomarkers has been done at the research level and not at the practical application level, although NOAA has a history of

involvement with their use in Puget Sound and elsewhere (Long, personal communication). This lack of a history of practical application coupled with a lack of specific protocols for their use and the controversy over the exact relationships between contaminants, biomarkers, and significant ecological effects, has resulted in the infrequent use of biomarkers at waste sites. However several biomarkers show great promise as bioassessment tools for evaluating effects of specific groups of contaminants found at many waste sites. This chapter attempts to identify those biomarkers with the greatest promise for issues that apply to waste sites.

Objectives and Purpose

There are two purposes for the measurement of biomarkers at waste sites. The use of biomarkers can indicate that organisms have been exposed to bioavailable contaminants and may indicate that adverse effects are occurring. Biomarkers can also illustrate the extent of contamination through the measurement of biological responses. Organisms must be exposed to bioavailable contaminants in order to produce a response. Some biomarkers (for example, liver tumors in fish) can be considered a direct measurement of an adverse effect. A demonstrated reduction in growth rates of organisms near a site when compared with a reference site also would be considered an adverse effect. When used in combination with other assessment tools, biomarkers can help evaluate the need for remediation. If indicators of exposure and adverse effects are seen prior to remediation, these measures can be repeated during and/or after remediation to monitor its success.

Advantages of Biomarkers

Biomarkers have the advantage of measuring actual biological responses to environmental conditions (Table 5-1). They can provide indications of biological effects occurring near a site. These measures can integrate the patchy temporal nature of exposure, and can provide information that is ecologically relevant. Some biomarkers (e.g., aryl hydrocarbon hydroxylase (AHH) induction) are very dose dependent and diminish rapidly upon removal of the toxicant(s) (Long, personal communication), thus providing temporal information on contaminant distribution. They can account for actual environmental conditions that cannot be reproduced in laboratory toxicity tests. Biomarkers can indicate sensitive subcellular effects that may not be measured after short-term laboratory exposures (McCarthy and Shugart, 1990). For contaminants that are rapidly metabolized or not accumulated, biomarkers may offer the only direct measures of their uptake. The greatest advantage of biomarker measurements is that they can demonstrate that organisms actually

present near a waste site have been adversely affected by contamination, i.e., they can answer the 'So what?' question.

A further advantage is that a variety of levels of severity of effects can be assessed. Some measures simply indicate that exposure has occurred. Other measures can indicate a broad spectrum of responses. Liver lesions, for example, can indicate mild or severe biological effects depending on the type of lesion that is most prevalent. Different measurements can also indicate that animals are stressed or diseased, or that reproduction or growth has been impaired. In addition, biomarkers can provide a more sensitive measure of response than community changes and can provide some indication of the cause of the effects (McCarthy and Shugart, 1990). Biomarkers can be chosen for their known sensitivity to particular groups of contaminants and for their applicability to some particular species of concern. For example, mixed function oxidase (MFO) enzymes are induced by hydrocarbons and metallothioneins are induced by metals. Biomarkers that are known to be sensitive only to certain classes of contaminants offer advantages over other measures (such as community or population measures) that are known to be affected by a broad spectrum of toxicants and natural factors.

Disadvantages of Biomarkers

Although biomarkers and disorders discussed here have been well researched and applied in the natural environment, most of these measures have not been widely used at waste sites, and there are no accepted protocols for their use (Table 5-1). The perception that these measures are experimental will probably continue to limit their use at most waste sites. A major limitation of these approaches is the difficulty in determining the actual extent of exposure to specific contaminants, particularly for fish and other mobile organisms. It should be noted that for most biomarker measurements, there is no absolute measure of unacceptable response. Because some responses are also produced by natural conditions such as reproductive state and season, interpreting the significance of results depends very heavily upon the expertise of the investigator and the availability of an uncontaminated reference area that is otherwise very similar to the test site. Responses may vary by species, even between closely related species. Some disorders can take years to develop within the organism, and possibly years to disappear. Biomarker tests can be expensive and require specialists to conduct, although some tests are less expensive to conduct than evaluations of community effects (McCarthy and Shugart, 1990). Also specific tests are not available for all groups of contaminants and for all groups of organisms.

Table 5.1. Advantages and disadvantages of using biomarkers at waste sites.

Advantages	Disadvantages
Measure actual biological responses to contaminants.	Little history of use at waste sites.
May integrate pathcy temporal exposure.	No existing EPA or other accepted protocols.
Demonstrate effects on indigenous organisms.	No absolute mesure of unacceptable response.
Assess a variety of severity levels.	Responses may be caused by natural factors.
Measure more sensitive responses than other bioassessment methods.	May require experienced expert investigators.
Selective for particular contaminant or class of contaminant.	Not always a known relationship between response and significant ecological effects.
Selective for a particular species of concern.	Responses may take years to develop or disappear (after remediation).
May be cheaper than higher level ecological studies.	Not yet feasible for all groups of organisms or contaminants.
	Few commercial laboratories can perform the tests.

SELECTING A TEST

The response of aquatic organisms to contaminants provides opportunities to measure actual environmental exposure before bioeffects have occurred as well as measuring actual biological effects. Selecting an appropriate test will depend entirely on the questions to be answered at the site. Certain biomarkers can demonstrate that contaminants are bioavailable, and can help determine the extent of contamination. Other measures are more useful for indicating specific biological effects (e.g., growth, reproductive, or disease). Knowledge of the contaminants present at the site will help define which tests might be useful. Site-specific concerns for particular types of effects will also aid in the choice of a test. Finally, the species to be tested must be present at the site in adequate numbers. The species chosen can also provide a link to human health or to other organisms of concern

through food-web pathways. Selection of a test will thus be based on the contaminants of concern (Table 5-2) and the species to be tested (Table 5-3).

TYPES OF BIOMARKERS

Enzyme and Protein Systems

Cells are known to respond to environmental stress with the production of enzymes and other proteins. Most of these proteins are normally found in cells that are not stressed, where they play a role in normal cellular function (Sanders, 1990). These proteins include enzymes responsible for catalyzing transformation reactions, for metabolizing and excreting steroids produced during gametogenesis, and proteins that bind with contaminants, making them unavailable for further reactions with sensitive cellular components. These biochemical changes are usually the first detectable responses to changes in the environment (Stegeman et al., 1992). These changes are usually highly sensitive and may lead to more severe effects within an organism.

Table 5.2. Sensitivity of biomarkers to specific contaminant groups.

Contaminant Group	Tests
PAHs	Kidney, intestinal, and liver lesions. DNA adducts. MFO enzymes. Bile Metabolites. Reproductive hormone measures.
PCBs and DDTs	MFO enzymes. Binding proteins. DNA adducts. Growth reduction. Reproductive hormones.
Metals	Binding proteins. Growth reduction. Lesions, fin erosion. MFO enzymes (weak).

Detoxification Enzymes

Under normal conditions, animals have some capacity to process contaminants so that they are kept from sensitive cellular sites of toxic action. The first phase in the metabolism of absorbed lipid-soluble contaminants is the production of enzymes to catalyze transformation reactions. The most important enzymes that catalyze detoxification

processes are those using the cytochrome P-450 electron transfer system. These enzymes are known as MFO enzymes. MFO enzymes synthesize and degrade natural compounds such as steroids and fatty acids and also transform the structure of foreign compounds (Stegeman et al., 1992). This family of enzymes includes AHH and ethoxyresorufin-O-deethylase (EROD). MFO enzymes are present in the endoplasmic reticulum of cells of many organs (gills, liver, intestines, and kidneys for example). They are usually most active in the liver on a whole organ basis and per gram of tissue.

Table 5-3. Biomarker tests currently applied to specific groups of organisms.

Mollusks	Fish	Other Species
MFO enzymes (weak)	MFO enzymes	MFO enzymes
Binding proteins	Binding proteins	Binding proteins
Neoplasms		
Lesions and other histopathological disorders	Lesions and other histopathological disorders	
Growth reduction		
	DNA adducts	
	Skeletal defects	
	PAH metabolites in bile	
	Reproductive measures	

The activity of some specific MFO enzymes (such as AHH and EROD) has been shown to increase in response to exposure to many lipid soluble organic contaminants such as PAHs, PCBs and dioxin (Stegeman et al., 1992). Fish liver and kidney tissue have shown increases in MFO enzyme levels in response to 3-methylcholanthrene and similar compounds, including PAHs, DDTs, dioxins, and some PCBs (Leech et al., 1982; Stegeman, 1981; Gruger et al., 1977; Kurelec et al., 1977; Stegeman et al., 1992). However, Cd apparently decreases

the activity of MFO enzymes by interfering with protein groups on the enzymes (Sorenson, 1991). DDT does not appear to be an active inducer of MFO enzymes (Stegeman et al., 1992), but was associated with elevated MFO induction in San Francisco Bay starry flounder (Spies et al., 1988).

Other enzymes are known to be affected by exposure to metals. Cd, Zn, and silver (Ag) increase the activity of delta-aminolevulinic acid dehydratase (ALAD), a liver enzyme controlling formation of hemoglobin; while lead (Pb) decreases activity of ALAD (Johansson-Sjobeck and Larsson, 1979). Mayer et al. (1992) consider ALAD measurements in fish and birds to be an efficient and inexpensive measure of Pb exposure (but not necessarily a measure of Pb toxicity). However, measures of enzyme levels are not considered to be reliable indicators of exposure to other metals because of inconsistent responses (Sorenson, 1991).

MFO enzymes are produced to metabolize foreign compounds following exposure. Their activity can show that organisms have been both exposed to and have responded to contaminants. They are not a good indicator of actual adverse effects, however, because the organism may be effectively metabolizing, binding, and/or excreting contaminants, thus preventing adverse effects. Another difficulty in the interpretation of enzyme activity levels is that some of these enzymes may also be produced in response to changes in temperature, reproductive state, or capture stress. In spite of these difficulties, under conditions where exposure to contaminants cannot be demonstrated with other measures, MFO enzyme activity is a useful measure of contaminant exposure that may be applied at waste sites. When compared to reference areas, elevated levels of these proteins can indicate that contaminants are bioavailable near the site. Some examples of these conditions include sites where contaminants do not bioaccumulate (PAHs in fish, for example), sites where concentrations in water are not measurable and fine grained sediments are not available for sampling, or as an indicator of contaminant bioavailability. Currently, MFO enzyme measurements are not recommended for invertebrates because rapid induction of appropriate enzymes has not yet been conclusively demonstrated (Stegeman et al., 1992), and more importantly, the physical mechanism of MFO induction in invertebrates is not understood (Stegeman personal communication to Long).

Binding Proteins

The production of MFO enzymes generally increases the organism's detoxification capacity for the type of chemicals to which it is exposed. Detectable increases in an organism's

enzymatic activity may therefore be an indicator of its recent exposure to contaminants. The second phase in the transformation of contaminants (conjugation reactions) depends on the availability of conjugating substances produced within the organism, for example, the production of MFO enzymes, the concentration of binding or scavenging proteins. For example, in rats, death of liver cells occurred after GSH levels were depleted (Reid et al., 1973). This was attributed to binding of oxygenated metabolites with genetic material and other proteins. However, levels of binding proteins do not indicate that actual adverse effects are occurring because the organism may be effectively binding, and/or excreting contaminants, thus preventing adverse effects.

For organic contaminants, one of the more important conjugating compounds is the peptide, glutathione (GSH), which sequesters oxygenated metabolites away from sensitive cellular sites. The activity of this and other conjugating substances may be increased or induced by exposure to various classes of organic compounds (Meister, 1983; Thomas and Wofford, 1984; Varanasi et al., in press). Enzymes that catalyze reactions with GSH have also been studied as biomarkers. These glutathione transferases (GST) also bind to contaminant metabolites and appear to be elevated in fish, crabs, and mussels from sites contaminated with PAHs (Stegeman et al., 1992). However, further research is needed to identify compounds that can induce GSH and GST levels before these measures should be applied at waste sites.

For trace metals, the most important group of scavenging proteins is metallathioneins (MT), of which a variety of forms exist. MT is involved in a variety of processes concerned with metal metabolism including the regulation of the uptake of essential metals (especially Zn and Cu) and metal detoxification (Stegeman et al., 1992). MT appears to be a promising indicator of exposure to Cd, Cu, Zn, and mercury (Hg) (Stegeman et al., 1992). However, the concentration of MT in cells may be increased not only by exposure to contaminants, but also by sexual maturation, temperature, and nutritional status (Benson et all, 1990; Stegeman et al., 1992). MT synthesis and induction has been well studied in fish, bivalves, and other organisms. However, further research is probably required before MT concentrations will be useful as biomarkers (Melancon et al., 1992).

The activity of conjugating substances such as MT and GSH can indicate that organisms have been exposed to contaminants. Some contaminants are known to be more effective at inducing the production of conjugating substances. The consequences of inadequate production of binding proteins on an organism's health can be severe. When binding protein levels are depleted, contaminants are free to bind with more sensitive proteins. For

example, in rats, death of liver cells occurred after GSH levels were depleted (Reid et all, 1973). This was attributed to binding of oxygenated metabolites with genetic material and other proteins. However, levels of binding proteins do not indicate that actual adverse effects are occurring because the organism may be effectively binding, and/r excreting contaminants, thus preventing adverse effects.

Contaminant Metabolites in Bile

Some contaminants are rapidly metabolized in higher organisms and do not accumulate in tissues. PAHs for example, do not generally accumulate in fish tissue. However, PAH contamination of sediment has been correlated with some adverse effects in benthic fish. Although the development of these adverse effects may take years, the metabolites of PAH compounds can be detected in fish bile very quickly after exposure to PAHs (a matter of hours to days) and the metabolites remain detectable for weeks after an exposure (Melancon et al., 1992). Therefore, the measurement of these metabolites can serve as a good indicator of recent exposure to PAHs. It should be kept in mind however, that these methods are only semi-quantitative and appear to be most responsive under conditions with relatively high levels of contamination (Varanasi et al., in press). These measurements have been applied near waste sites contaminated with creosote (e.g., Eagle Harbor, in Puget Sound).

Genetic Disorders

Mutagenic, promutagenic, and carcinogenic compounds may not be detected in acute toxicity tests, but may be equally lethal as any acute toxin in the long run. There is a family of subcellular biomarkers that can be used to detect the effects of these compounds. The metabolism of many contaminants results in the creation of oxygenated metabolites that are known to react with genetic material. The outcome and implications of interactions with DNA are difficult to predict, especially since organisms have some limited ability to repair DNA. However, interactions with DNA are usually evidenced by abnormalities at the chromosome level and the effects of these types of disorders can be far reaching. Studies of liver pathologies in rats show that cellular death was associated with excess metabolites binding to DNA, RNA, or other proteins (Reid et al., 1973).

It is possible to measure outcomes of reactions between DNA and contaminants. One outcome, known to be the result of chromosome breakage, is the formation of micronuclei (MN) in red blood cells of fish (Schmid, 1976). MN formation is also a very widely used biomarker in mammalian, industrial hygiene research (Long, personal communication). MN are small cytoplasmic chromatin masses that resemble small nuclei. These irregularities

are the result of lagging chromosomes during cell division or from unbalanced chromosomal fragments. MN formation has been induced in the laboratory by exposure to chemicals such as benzo(a)pyrene and ethylmethanesulfonate (Hooftman and de Raat, 1982; Hose et al., 1984). However, because of questions about other causes of MN formation, it has not been widely used and is not recommended for waste site investigations at this time.

A more direct approach to assess contaminant effects on genetic material has been recently developed. The binding of contaminant metabolites with DNA molecules (the formation of DNA adducts) is measured through the labeling of DNA with ^{32}P (Randerath et al., 1981) and autoradiography. DNA adducts have been measured in English sole from Puget Sound (Stein et al., 1990), and in winter flounder from Boston Harbor and Long Island Sound (Stein et al., 1989). DNA adducts have been demonstrated for metabolites of PCBs and PAH compounds. These direct measures of adverse effects have great potential for application at waste sites contaminated with PAHs or PCBs.

Other measures of genetic effects include chromosome analysis (recording visible chromosomal aberrations) and sister chromatid exchange (SCE). SCE have been detected and quantified in polychaetes, mollusks, and fish in field surveys and laboratory exposures. It was also one of the methods used by the United States Environmental Protection Agency (EPA) at the Davisville, Rhode Island Superfund site (Munns et al., 1991) These are both labor intensive procedures that may have some applications to further document effects where other screening methods indicate the potential for effects (Shugart et al., 1992).

Immune System Responses

A variety of measures of immune system response are under investigation for use as indicators of contamination. These include blood cell counts, kidney macrophage function (known to be sensitive to PAHs in fish), and specific antibody counts (Benson and DiGiulio, 1992). NOAA's NS&T Program is currently testing immune responses in oysters from Tampa Bay and mussels from San Diego Bay (Long, personal communication). Since these measures are still in the research stage of development, they are not recommended for analysis at waste sites.

Histopathology

Histopathology, the study of tissue responses to injury or disease, can indicate early warning conditions in organisms such as fish and shellfish long before more advanced abnormalities appear. The nature of disorders can also help identify contaminants

responsible for them. For field bioassessment, histopathology is the most rapid method of detecting adverse acute and chronic effects of exposure (Hinton et al., 1992). Histopathological changes reflect prior alterations in biochemical function (Hinton et al., 1992). Most histopathological studies have focused on the liver of higher organisms such as fish. Although there have been many attempts to discover reliable histopathological biomarkers in bivalve mollusks, further research is needed to support associations between lesions and contaminant exposure.

After an exposure that causes cell death, the cell is acted upon by enzymes that cause nuclear and cytoplasmic changes that are easily detected. These alterations are referred to as degeneration and necrosis. Other types of injuries result in disturbances to the normal growth, repair, or replacement of cells. Growth anomalies may be non-cancerous or neoplastic (resulting in tumors). Neoplasms, or cancerous tumors, are abnormal masses of tissue that are obviously different from surrounding tissues and grow excessively even after removal of the causative factor. Tumors absorb nourishment at the expense of normal tissues, yet provide no beneficial service. Non-neoplastic growths are characterized by disturbances in tissue differentiation, diminished growth, or excessive, but not limitless, growth. This latter category is distinct from neoplasms and is referred to as "proliferative disorders."

There are some histopathological conditions seen in fish that are highly suspected to be caused by chemical contamination. Those conditions are:

1. Neoplasms (cancerous tumors), including both adenomas (benign tumors) and carcinomas (malignant tumors).
2. Foci of cellular alteration (FCA), locations of altered, pre-neoplastic cells.
3. Proliferative disorders (areas of excessive growth).
4. Specific degenerative or necrotic (SDN) lesions (Myers et al., 1990).

This latter category includes nuclear pleomorphism (where cell nuclei are misshapen), and megalocytic hepatosis (where cells and nuclei are abnormally enlarged and other degenerative effects are noted). Only liver and kidney conditions are highlighted here because these organs play the major role in eliminating waste products and subsequently have a greater susceptibility to lesions. Other conditions, such as, vacuolation (where cells contain empty space) and hypertrophy (where cells and surrounding tissues are enlarged) may also be related to contamination, but no direct correlation with sediment contamination

has been demonstrated. A small percentage of fish from uncontaminated areas also exhibit some signs of these lesions.

Results from a growing number of studies indicate that there is a causal link between certain fish liver lesions, especially tumors and pre-tumorous conditions, and exposure to some contaminants, especially PAHs, and possibly PCBs. A sequential scheme of lesions that progress toward neoplasms, similar to that developed from laboratory rat or mice studies, is now being demonstrated for some species of benthic fish (Myers et al., 1990). Laboratory exposures of healthy bottom fish to chemical extracts of contaminated sediments have successfully induced liver lesions similar to those observed in feral fish (Varanasi et al., 1987).

Evidence exists that the frequency and type of liver lesions changes as the fish ages (Varanasi et al., in press). In fact, the prevalence of most lesions increases with the age of the fish population, so histopathological data analyses must take into account the age of the fish when comparing sites (Long, personal communication). Liver lesions have been demonstrated to adversely affect liver function (Casillas et al., 1985). Liver disorders have been measured at waste sites contaminated with PAHs (in Eagle Harbor, for example) to demonstrate that site related contamination has caused adverse biological effects. These measurements show great promise for other marine or estuarine sites contaminated with PAH compounds. However, the detection of histopathological disorders, such as liver lesions, requires dissection of the organ, thin tissue section preparation, and examination by an expert While several Federal laboratories have this capability, few commercial or academic laboratories are able to perform these analyses (Long, personal communication).

Pathology

When cellular injury proceeds unchecked and major portions of a tissue or an entire organ become affected, it is sometimes possible to observe the damage without dissecting an organism. This is certainly the case for conditions affecting the skin, lips, and eyes of fish. Since the skin is one route of exposure to contaminants for fish, it is not surprising that some pathological conditions may be observed in this tissue.

Fin erosion is one of the most easily detected external abnormalities in fish, and therefore, is one of the most commonly documented. It is characterized by loss of epidermal, dermal, and fin ray tissues. This condition must be accurately distinguished from net damage or wounds from predators. Because its absolute identification must be confirmed with

histological examination, the significance and accuracy of observations are difficult to determine.

Though commonly reported, the cause of fin erosion is poorly understood. Laboratory tests have verified that exposure to contaminated sediments can induce fin erosion in healthy fish (Sinderman, 1979). Causative agents identified to date include PCBs, crude oil, Pb, Zn, and Cd (Sinderman, 1979; Tetra Tech, 1986). Fin erosion is most likely the result of a combination of factors including chemical contamination, secondary bacterial infection, mechanical injury, and poor water quality (low-dissolved oxygen) (Sinderman, 1979). Skin and lip lesions in bullheads have been induced by exposures to PAH-contaminated sediments in laboratory tests (Black, 1983). Because of the controversy over the causes of fin erosion, this measure of effects would be difficult to attribute to contaminants from a specific waste site. Until further research illuminates the causes of this disorder, it is not recommended for application at waste sites.

The cause of epidermal tumors is also not completely understood. Some lesions are thought to be caused by contaminants, while others are more clearly related to viruses or parasites. Three particular skin lesions (papillomas, squamous carcinomas, and chromatophoromas) have been recommended for further study in contaminated areas (Hinton et al., 1992). As with fin erosion, this disorder would currently be difficult to attribute to contamination from a waste site.

Skeletal anomalies are found in fish from areas of highly contaminated sediments. Most observed skeletal anomalies involve the spinal column and include fusion, flexures, and vertebral compressions. Skeletal anomalies also include abnormalities of the head, fins, and gills. Skeletal anomalies have been induced in fishes after lab exposure to the chlorinated pesticide kepone and heavy metals (Sinderman et al., 1980). However, some of these deformities can also be caused by nutritional deficiencies (Hinton et al., 1992). Mayer et al. (1992) recommend that further research be conducted to assess the degree of abnormality compared to normal occurrences that could be attributed to the effects of contaminants.

Growth

Metabolic energy is required for movement, active transport of substances across cellular membranes, biosynthesis of compounds, and reproduction. An organism expends a great deal of energy just maintaining its normal functioning. This metabolic energy is produced and stored in the form of chemical bonds within a high-energy molecule (adenosine triphosphate). Energy from nutrition in excess of the normal maintenance requirement is what is available for the organism to use in growth processes. However, if an organism is responding to some environmental stress (including exposure to contaminants) it must expend energy to compensate for the stress, thereby reducing the amount available for growth.

Linear growth rates in caged animals have been used as a simple indication of physiological stress. Growth measurements of organisms caged near waste sites can be a useful indicator of effects due to contaminants. Caged juvenile mussels have been used in San Diego Bay to evaluate effects of marinas and naval facilities on biota (Salazar and Salazar, 1991). This technique has been applied at two sites in the Puget Sound area, Commencement Bay and Elliott Bay.

A scope for growth (SFG) index, defined as the theoretical amount of energy available to an organism for growth and reproduction (Warren and Davis, 1967), may also be used as an indicator of physiological stress. The SFG index is expressed as the difference between the energy value of all the food consumed and the energy value of all processes other than growth (respiration and excretion). It is an extremely labor-intensive method. This technique has not previously been used at waste sites, but has been used a great deal in Europe and the United Kingdom, especially on or near oil platforms and, in this country, in Narragansett and San Francisco bays and in southern California (Long, personal communication). SFG may have some potential for application using transplanted bivalves during environmental assessment investigations.

Reproduction

Reproduction in female fish is governed by an intricate system of hormones, proteins, and external signals (Nagahama, 1987). During maturation, eggs pass through two phases: primary growth and vitellogenesis. Primary growth results in large increases in oocyte volume and development of numerous sub-cellular structures or organelles. Vitellogenesis is the sequestering of a protein, vitellogenin, into the yolk. Vitellogenin is produced in the

liver and released into the blood in response to the reproductive hormone, estradiol. Once vitellogenesis ends and the oocytes have reached their fully developed size, a surge of hormones induce the follicles to release steroids. This in turn stimulates resumption of oocyte cell division, hydration of the oocytes, and eventually, ovulation (discharge of the oocytes). This entire process is initiated in response to external signals of light and temperature.

Because of the complexity of the reproductive system, there are a number of disturbances that could be a result of contamination. Not all vitellogenic oocytes mature and ovulate. The ovarian follicle may lose its integrity and its oocyte may not be released (known as atresia). This condition is frequently observed in species that spawn more than once per year, during the end of the spawning season when unspawned eggs are resorbed (Braekevelt and McMillan, 1967). Stress, altered photoperiod, temperature regimes, and poor diet have all been correlated with atresia (Cross et al., 1984). However, any factor that lowers gonadotrophin[2] levels (including exposure to contaminants) could induce atresia (Braekevelt and McMillan, 1967). Oocyte atresia has great potential to indicate adverse effects that are highly significant, but it has not been verified in contaminated areas.

Contaminants that induce production of MFO enzymes can also affect enzymes that regulate the production of steroids controlling spawning (Spies et al., 1984). Exposure of English sole to extracts of contaminated sediment (containing PAHs and PCBs) reduced plasma levels of estradiol (the steroid controlling vitellogenesis) in sexually maturing females (Stein et al., 1991). Similar effects have been seen in Atlantic croaker, *Micropogonias undulatus*, (Thomas, 1988). Depressed plasma estradiol levels have been linked with impaired ovarian development and reduced ability to spawn (Johnson et al., 1988; Varanasi et al., in press). Other laboratory studies have also linked decreased fertility with exposures to contaminants, including PAHs, PCBs, and pesticides (Nagler et al., 1986; Cross et al., 1984; Hose et al., 1981). Some investigators theorize that MFO enzymes produced after exposure to PAHs, PCBs, and dioxins result in the metabolism of steroids (such as estradiol and testosterone) that control reproduction; however, the exact cause of depressions in estradiol levels are not known (Johnson et al., 1988; Johnson et al., in press).

Measures of reproductive success that may be affected by contamination include hormone levels in plasma (estradiol, for example) and oocyte condition (atresia, for example) in feral fish. Fish can also be taken from the environment and held in the laboratory to induce

[2] Gonadotrophins are a class of gonad-stimulating hormones required for the development and maintenance of the gonads in seasonal breeders.

spawning. Fertilization success, hatching success, egg condition, and estradiol and vitellogenin levels can then be measured and correlated to contamination at the site where the fish were taken, or they can be correlated with contaminants measured in fish tissue or bile. However, the measurement of fertilization success is very labor intensive, has a high chance of failure, and takes a long time, therefore it is not recommended for application to problems at waste sites. Measurements of reproductive hormone levels in fish near waste sites may be a useful technique for predicting reproductive effects (Johnson et al, 1988). They may be particularly useful at estuarine sites contaminated with PAHs or PCBs where white croaker, Atlantic croaker, English sole, flathead sole, or rock sole are present. Although these tests have been applied to winter flounder from Boston Harbor and Raritan Bay, there was little evidence that steroid metabolism was altered by contaminants in this species (Johnson et al., in press).

DATA INTERPRETATION

The interpretation of results of biomarker studies rests heavily on comparisons between test sites and appropriate reference sites. These indicators of exposure and response to contamination can also reflect stress generated by seasonal factors, reproductive state, and other natural factors. For these reasons, it is essential to collect organisms from reference sites at the same time and with exactly the same methods as from test sites. Natural variability can be more easily evaluated if age, sex, and reproductive state are noted when organisms are collected. The statistical significance of responses will vary by the parameters measured by individual tests. In general, however, results of tests will indicate that organisms at the site are (or are not) different from animals upstream, downstream, or at the reference site.

SUMMARY

Biomarkers are direct biological measures of the response of an organism to exposure to a contaminant. They consist of a wide range of measurable responses from the biochemical to the physiological. Biomarkers not only indicate that contaminants are present and bioavailable, but, they also indicate that the organisms are responding to the contaminants. However, while they can indicate that the contaminants are causing adverse biological effects, they do not necessarily do so. They may just indicate that the organisms are detoxifying and eliminating the contaminants. There are numerous biomarker tests being worked with and some even used at waste sites (e.g., SCE, growth), but the majority are still in the research and development stage. Therefore, while biomarkers hold great promise for

use at future hazardous waste sites, only a few are sufficiently developed to be currently considered for use at hazardous waste sites.

REFERENCES

Benson, W.H., K.N. Baer, and C.F. Watson. 1990. Metallothionein as a biomarker of environmental metal contamination. In: McCarthy, J.F. and L.R. Shugart. 1990. Biomarkers of Environmental Contamination. Boca Raton, FL, Lewis Publishers. pp. 255-266.

Benson, W.H. and R.T. DiGiulio. 1992. Biomarkers in hazard assessments of contaminated sediments. In: Burton, G. A. (ed). Sediment Toxicity Assessment. Boca Raton, FL, Lewis Publishers. pp. 241-266.

Black, J. 1983. Field and laboratory studies of environmental carcinogenesis in Niagara River fish. Great Lakes Research 9: 326=334.

Braekevelt, C. R. and D. B. McMillan. 1967. Cyclic changes in the ovary of the brook stickleback, *Eucalia inconstans*. Journal of Morphology 123: 373-396.

Casillas, E., M.S. Myers, L.D. Rhodes, and B.B. McCain. 1985. Serum chemistry of diseased English sole, *Parphrys vetulus* Girard, from polluted areas of Puget Sound, Washington. Journal of Fish Diseases 8. pp 437-449.

Cross, J.N., V.E. Raco, and D.W. Diehl. 1984. Fish reproduction around outfalls. In: Bascom, W (ed). Southern California Coastal Water Research Project Biennial Report, 1983-84. Long Beach, CA, Southern California Water Research Project. pp. 211-228.

Gruger, E. H., M. M. Wekell, P. T. Numoto, and D. R. Craddock. 1977. Induction of aryl hydrocarbon hydroxylase in salmon exposed to petroleum dissolved in seawater and to petroleum and polychlorinated biphenyls, separate, and together, in food. Bulletin of Environmental Contamination and Toxicology 17: 512-520.

Hinton, D.E., P.C. Baumann, G.R. Gardner, W.E. Hawkins, J.D. Hendricks, R.A. Murchelano, and M.S. Okihiro. 1992. Histopathologic Biomarkers. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman (eds). Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, CO, July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp. 155-210.

Hooftman, R. N. and W. K. de Raat. 1982. Induction of nuclear anomalies (micronuclei) in the peripheral blood erythrocytes of the eastern mudminnow (*Umbra pygmaea*) by ethyl methanesulphonate. Mutation Research 104: 147-152.

Hose, J. E., J. B. Hannah, M. L. Landolt, B. S. Miller, W. T. Iwaoka, and S. P. Felton. 1981. Uptake of benzo(a)pyrene by gonadal tissue of flatfish (family *Pleuronectidae*) and its effects on subsequent egg development. Journal of Toxicology and Environmental Health 7: 991-1000.

Hose, J. E., J. B. Hannah, H. W. Puffer, and M. L. Landolt. 1984. Histologic and skeletal abnormalities in benzo(a)pyrene treated rainbow trout alevins. Archives of Environmental Contamination and Toxicology 13(1): 675-684.

Hugget, R.J., R.A. Kimerle, P.M. Mehrle, H.L. Bergman, K.L.Dickson, J.A. Fava, J.F. McCarthy, R. Parrish, P.B. Dorn, V. MacFarland, and G. Lahvis. 1992. Introduction. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman (eds). Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, CO, July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp 1-4.

Johansson-Sjobeck, M. L., and A. Larsson. 1979. Effects of inorganic lead on delta-aminolevulinic acid dehydratase activity in the rainbow trout, *Salmo gairdneri*. Archives of Environmental Contamination and Toxicology 8:419-431.

Johnson, L., E. Casillas, S. Sol, T. Collier, J. Stein, and U. Varanasi. In press. Contaminant effects on reproductive success in selected benthic fish species. Marine Environmental Research.

Johnson, L., E. Casillas, D. Misitano, T. Collier, J.E. Stein, B. McCain, and U. Varanasi. 1988. Bioindicators of reproductive impairment in female English sole (*Parophrys vetulus*) exposed to environmental contaminants. In: Oceans '89 Volume 2: Ocean Pollution. An international conference addressing methods for understanding the global ocean. September 18-21, 1989, Seattle, WA. Marine Technology Society; Oceanic Engineering Society of the Institute of Electrical and Electronics Engineers. pp. 391-396.

Kurelec, B., S. Britvic, M. Rijavec, W. E. G. Muller, and R. K. Zahn. 1977. Benzo[a]pyrene monooxygenase induction in marine fish -- molecular response to oil pollution. Marine Biology 44: 211-216.

Leech, J. J., M. J. Vodicnik, and C. R. Elcombe. 1982. Induction of monooxygenasae activity in fish. In: L. J. Weber, ed. Aquatic Toxicology. New York: Raven Press. pp 107-148.

Long, E.R. 1992. personal communication. Coastal Monitoring and Bioeffects Division, Office of Ocean Resources Conservation and Assessment, National Oceanic and Atmospheric Administration. Seattle, WA.

Luoma, S. N. and J.L. Carter. 1991. Effects of Trace metals on aquatic benthos. In: Newman, M.C. and A.W. McIntosh (eds). Metal Ecotoxicology: Concepts and Applications. Chelsea, MI: Lewis Publishers. 399 pp.

Mayer, F.L., D.J. Versteeg, M.J. McKee, L.C. Folmar, R.L. Graney, D. C. McCume, and B.A. Rattner. 1992. Physiological and nonspecific biomarkers. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman (eds). Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, CO, July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp. 5-86.

McCarthy, J.F. and L.R. Shugart. 1990. Biomarkers of Environmental Contamination. Boca Raton, FL, Lewis Publishers. 457 pp.

Meister, A. 1983. Selective modification of glutathione metabolism. Science 220: 472-477.

Melancon, M.J., R. Alsher, W. Benson, G. Kruzynski, R.F. Lee, H.C. Sikka, and R.B. Spies. 1992. Metabolic Products as Biomarkers. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman (eds). Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, CO, July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp. 87-124.

Munns, W.R. Jr., C. Mueller, D.J. Cobb, T.R. Gleason, G.G. Pesch and R.K. Johnston. 1991. Marine ecological risk assessment at Naval Construction Battalion Center, Davisville, Rhode Island. Naval Ocean Systems Center Technical Report 1437. 129 pp+.

Myers, M.S., J.T. Landahl, M.M. Krahn, L.L. Johnson, and B.B. McCain. 1990. Overview of studies on liver carcinogenesis in English sole from Puget Sound; Evidence for a xenobiotic chemical etiology I: Pathology and Epizootiology. The Science of the Total Environment 94:33-50.

Nagahama, J. 1987. Gonadotropin action on gametogenesis and steroidogenesis in teleost gonads. Zoological Science 4: 209-222.

Nagler, J. J., P. Aysola, and S. M. Ruby. 1986. Effect of sublethal pentachlorophenol on early oogenesis in maturing female rainbow trout (*Salmo gairdneri*). Archives of Environmental Contamination and Toxicology 15: 549-555.

Randerath, K., M. Reddy, and R.C. Gupta. 1981. 32P-postlabeling analysis for DNA damage. Proceedings. National Academy of Science United States of America 78. 6126-6129.

Reid, W. D., G. Krishna, J. R. B. Gillette, and B.B. Brodie. 1973. Biochemical mechanisms of hepatic necrosis induced by aromatic hydrocarbons. Phamacology 10: 193-214.

Sanders, B. 1990. Stress Proteins: Potential as multitiered biomarkers. In: McCarthy, J.F. and L.R. Shugart (eds). Biomarkers of Environmental Contamination. Boca Raton, FL, Lewis Publishers. pp. 165-192.

Salazar, M. H., and S. M. Salazar. 1991. Assessing site-specific effects of TBT contamination with mussel growth rates. Marine Environmental Research 32:131-150.

Schmid, W. 1976. The micronucleus test for cytogenetic analysis. In: A. Hollaender (ed) Chemical Mutagens: Principles and Methods for Their Detection. New York: Plenum Press. pp 31-53.

Shugart, L.R., J. Bickham, G. Jackim, G. McMahon, W. Ridley, J. Stein, and S. Steinert. 1992. DNA alterations. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman (eds). Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, Co July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp. 125-154.

Sinderman, C. 1979. Pollution-associated diseases and abnormalities of fish and shellfish: A review. U.S. Fisheries Bulletin 76: 717-749.

Sinderman, C.J., F.B. Bang, N.O. Christensen, V. Dethlefsen, J.C. Harshbarger, J.R. Mitchell, and M.F. Mulcahy. 1980. The role and value of pathobiology in pollution effects monitoring programs. Rapport. Proefstation voor de Groenteteelt in de Vollegrond in Nederland 179: 135-151.

Sorenson, E. M.B. 1991. Metal Poisoning in Fish. Boca Raton, FL, CRC Press. 374 pp.

Spies, R. B., D. W. J. Rice, and R. R. Ireland. 1984. Preliminary studies of growth, reproduction, and activity of hepatic mixed-function oxydase in *Platichthys stellatus*. Marine Environmental Research 14: 426-428.

Spies, R.B., D.W.J. Rice, and J.W. Felton. 1988. The effects of organic contaminants on reproduction of starry flounder, *Platichthys stellatus* (Pallas), in San Francisco Bay. I. Hepatic contamination and mixed-function oxidase (MFO) activity during the reproductive season. Marine Biology 98: 191-200.

Stegeman, J. 1981. Polynuclear aromatic hydrocarbons and their metabolism in the marine environment. In: H. Gelboin, and P. O. P. Ts'o, (eds) Polycyclic Hydrocarbons and Cancer. New York: Academic Press. pp 1-60.

Stegeman, J.J., M. Brouwer, R.T. DiGiulio, L. Forlin, B.A. Fowler, B.M. Sanders, and P.A. Van Veld. 1992. Enzyme and protein synthesis as indicators of contaminant exposure and effect. In: Huggett, R.J., R.A. Kimerle, P.M. Mehrle, and H.L. Bergman. Biomarkers: Biochemical, Physiological, and Histological Markers of Anthropogenic Stress. Proceedings of the Eighth Pellston Workshop. Keystone, Co July 23-28, 1989. Boca Raton, FL, Lewis Publishers. pp. 235-336.

Stein, J.E., W.L. Reichert, M. Nishimoto, and U. Varanasi. 1989. 32P-postlabeling of DNA: A sensitive method for assessing environmentally induced genotoxicity. In: Oceans '89 Volume 2: Ocean Pollution. An international conference addressing methods for understanding the global ocean. September 18-21, 1989, Seattle, WA. Marine Technology Society; Oceanic Engineering Society of the Institute of Electrical and Electronics Engineers. pp. 385-390.

Stein, J.E., W.L. Reichert, M. Nishimoto, and U. Varanasi. 1990. Overview of studies on liver carcinogenesis in English sole from Puget Sound; Evidence for a xenobiotic chemical etiology II: Biochemical Studies. The Science of the Total Environment 94:51-69.

Stein, J.E., T. Hom, H.R. Sanborn, and U. Varanasi. 1991. Effects of exposure to a contaminated sediment extract on the metabolism and disposition of 17B-Estradiol in English sole (*Parophrys vetulus*). Comparative Biochemistry and Physiology Vol. 99C, No 1/2, pp. 231-240.

Tetra Tech. 1986. Recommended protocols for measuring selected environmental variables in Puget Sound. U.S. Environmental Protection Agency, Puget Sound Estuary Program, Seattle, WA.

Thomas, P. 1988. Reproductive endocrine function in female Atlantic croaker exposed to pollutants. Marine Environmental Research 24:179-183.

Thomas, P. and P. T. Wofford. 1984. Effects of metals and organic compounds on hepatic glutathione, cysteine, and acid-soluble thiol levels in mullet (*Mugil cephalus*). Toxicology and Applied Pharmacology 76: 172-182.

Varanasi, U., J.E. Stein, L.L. Johnson, T. K. Collier, E. Casillas, and M.S. Myers. in press. Evaluation of bioindicators of contaminant exposure and effects in coastal ecosystems. In: Marine Environmental Research.

Varanasi, U., D. W. Brown, S.-L. Chan, J. T. Landahl, B. B. McCain, M. S. Meyers, M. H. Schiewe, J. E. Stein, and D. D. Weber. 1987. Etiology of tumors in bottom-dwelling marine fish. Final report submitted to National Cancer Institute. Seattle, WA: National Oceanic and Atmospheric Administration. 250 pp.

Warren, C. E. and G. E. Davis. 1967. Laboratory studies on the feeding, bioenergetics, and growth of fishes. In: S. D. Gerking, ed. The Biological Basis of Freshwater Fish Production. Oxford, England: Blackwell Scientific Publications. pp 175-214.

CHAPTER 6
BENTHIC COMMUNITY STUDIES

INTRODUCTION

Most coastal CERCLA sites are located next to or near aquatic systems (e.g., streams, rivers, estuaries, bays). These aquatic systems play an important role in site investigations because they are ecologically important to NOAA trust resources or because they serve as contaminant pathways to habitats used by NOAA trust resources. Assessing aquatic biological communities permits the determination of the ecological condition or 'status' of an aquatic environment (Hunsaker and Carpenter, 1990). Aquatic biological community studies can expand and improve the assessment of the impacts of CERCLA sites, help determine sites of major concern, identify impaired beneficial uses, determine program priorities, and help monitor the overall success of a remediation program.

While any combination of taxonomic group (algae, invertebrates, or fish) and level of biological organization (individual, population, community, or ecosystem) can be used to assess the biological health of an aquatic system, this chapter focuses on the use of benthic invertebrate communities. Benthic invertebrates are commonly used in environmental monitoring because they show cumulative effects of present and past conditions, they have low mobility and relatively long life cycles, and their ecological relationships are relatively well understood (Herricks and Cairns, 1982; Wilhm, 1975). In addition, sampling procedures are relatively well developed, the group is heterogeneous in that a single sampling technique collects a considerable number of species from a wide range of phyla, and macroinvertebrates are generally abundant (Mason, 1981). **Macroinvertebrates** are those invertebrates retained by a 0.5-millimeter (mm) mesh screen (McIntyre et al., 1984) and are preferred for environmental monitoring studies over microinvertebrates because their taxonomy is better known. However, there are many sites where invertebrates <0.5 mm may need to be targeted, e.g., lakes and large rivers.

Community level studies provide the greatest amount of information on the biological integrity of an aquatic system, while at the same time permitting the examination of individual taxa that occur within the benthic invertebrate community. Communities are assessed from either a structural or functional perspective. Community structure is the measurement of biotic characteristics (e.g., abundance, diversity, and species composition) at a point in time; whereas, community function is the measurement of rate processes (e.g.,

species colonization rates) of the ecosystem. The use of biological communities in environmental monitoring is normally done from a structural perspective because structural studies normally take less time, are more conventional, and permit comparisons to be easily made with data from other studies (Mattews et al., 1982). However, it must be kept in mind that contamination is not the only factor capable of changing community structure. Changes in salinity, temperature, sediment texture, and shading, to name a few, can all effect community structure.

This chapter reviews three of the four main elements of a benthic invertebrate study: study objective, sampling methods, and data analysis. The fourth element, study design, will be covered in Chapter 7. The last section of this chapter reviews the use of EPA's Rapid Bioassessment Protocols for benthic invertebrates in streams.

STUDY OBJECTIVES

The most important component of a benthic invertebrate study is a well-defined objective. Numerous studies conducted at CERCLA sites have included a benthic invertebrate component, but have commonly lacked an overall study design discussing the need or appropriateness of benthic invertebrate community assessments and how the information will be used. The objective(s) of a study must be determined before an appropriate sampling program is designed.

Benthic invertebrate communities can be used at CERCLA sites to determine:

- Benthic resources using the habitats of interest and potential risks to these communities associated with the release of site-related contaminants,
- Appropriate organisms for bioassays or bioaccumulation studies,
- Whether the stream of interest has a benthic community that can be efficiently used to monitor a site,
- If benthic invertebrate communities in the vicinity of contaminated sediments and/or water are statistically different from non-impacted (reference) areas.
- Effectiveness of remediation activities.

SAMPLING METHODS

This section presents information on qualitative, semiquantitative, and quantitative sampling, along with a review of sampling devices used for collecting benthic invertebrate samples.

Quantitative vs Non-Quantitative Sampling

There are numerous sampling methods that can be used to assess benthic invertebrate communities. The major considerations prior to sampling are whether one needs quantitative or qualitative data and what sampling device will best collect the information required. Quantitative sampling is time consuming and therefore more expensive, but is required when one needs to determine statistical differences among stations;, whereas qualitative sampling is more rapid, less expensive, and typically used for surveys or to supplement quantitative sampling.

The concepts of qualitative, semi-quantitative, and quantitative sampling are well illustrated in an example presented by Holme and McIntyre (1984):

qualitative:	Collecting samples from different benthic habitats to produce a list of taxa or species in the area.
quantitative:	Estimating the number of individuals or biomass per unit area, using grabs, corers or other quantitative samplers.
semi-quantitative:	(Qualitative): determining the relative abundance of species using a dredge.

In this example, qualitative sampling is used for preliminary surveys, which may be followed by more complex qualitative sampling (semi-quantitative), and by quantitative studies.

Qualitative Sampling: The objective of simple qualitative sampling is to compare sites using the presence or absence of benthic invertebrates. Because statistical analysis is not conducted on qualitative data, there are no set guidelines for how or where the samples should be collected. Samples can be collected using most any type of method and gear. Commonly used equipment includes dip, kick or push nets, grab samplers, dredges, trawls, and artificial substrate samplers. For comparability, the same approach and sampling equipment should be used at each station. While this type of method is acceptable for

general surveys, it is highly subjective and therefore will limit the degree of confidence associated with statements made about the community of interest.

Quantitative Sampling: Quantitative sampling provides an estimate of the abundance and/or biomass (standing crop) of various components of the benthic invertebrate community per unit area, volume, or sampling unit. It needs to be conducted using statistically appropriate methods, paying particular attention to sample randomization. Quantitative sampling also provides information on the composition of the community and the distribution of the various taxonomic groups. There are a wide variety of sampling devices (e.g., stream net, grab, and core samplers) that can be used for sampling benthic invertebrate communities quantitatively. It is critical that the sampling device be appropriate for the habitat of interest.

Semi-quantitative Sampling: The rather nebulous term 'semi-quantitative' sampling is often used to refer to sampling that involves collecting benthic invertebrates by level of effort, as in time expended per habitat rather than by area or volume sampled. It can also refer to sampling involving the use of quantitative sampling devices in a non-random manner. The principal difference between quantitative and semi-quantitative sampling is in the degree of confidence that the replicate samples are comparable and free from investigator-induced bias. By comparable is meant that each replicate is sampling identical portions of the population, e.g., the portion present in 1 m^2 of bottom sediment. The nonrandom use of quantitative sampling devices (e.g., grabs, cores) introduces bias into the sampling. The lack of comparable replicates or sampling bias both preclude the statistical analysis of the data. Semi-quantitative sampling is basically just qualitative sampling with more rules.

Examples of commonly used semi-quantitative freshwater sampling methods include the traveling kick method (Hornig and Pollard, 1978; Pollard, 1981) and the U.S. EPA Rapid Bioassessment Protocols II and III (Plafkin et al., 1989). Estuarine and marine semi-quantitative methods involve the use of push nets (intertidal) or trawls (subtidal) operated for specified units of time at specified speeds (Eleftheriou and Holme, 1984). The results from these sampling methods are based on unit of effort as opposed to area or volume sampled. These sampling methods provide only semi-quantitative results because the investigator cannot be sure that equal units of effort sample equal proportions of the desired population. For example, during a 20-minute bottom trawl, the trawl may not actually be on the bottom for the entire time (e.g., it may bounce on and off the bottom) and each replicate trawl may be on the bottom for different proportions of the 20 minutes.

Whether sampling is quantitative or qualitative, habitats sampled should be as similar in relation to physical and chemical parameters as possible. In streams this means more than just comparing riffles with riffles and pools with pools; other factors need to be taken into account such as substrate, water velocity, grain size, depth, pH, hardness, and degree of shading from adjacent terrestrial plants. An important additional parameter effecting benthic communities in estuarine and marine environments is salinity. Intertidal community sampling needs to take into account tidal elevation because duration of submergence will effect the composition of the community. Submergence time will also effect the duration of exposure to waterborne contaminants. Intertidal benthic communities are also effected by their degree of exposure to wave action; for example, a sheltered rocky intertidal community will be different from an exposed intertidal community. The physical and chemical parameters of the sampled sites need to be taken into account when comparing sites for degree of exposure to contaminants so you can be reasonably confident that differences between sites are due to differences in contamination and not just differences in normal environmental variables.

Sampling Devices

Benthic macroinvertebrates can be collected from either natural or artificial substrate with each type offering advantages, depending on site-specific conditions. Natural substrates should be sampled wherever possible. Artificial substrates should be used when natural substrates cannot physically be sampled or when the substrate is so highly variable that heterogeneity needs to be removed as a variable. For example, when upstream stations are dominated by sand and downstream stations are dominated by gravel, the use of artificial substrates would permit the sampling of uniform substrates in both areas. More detailed information on sampling design, sampling devices and their use are reviewed extensively in Klemm et al. (1990), Mudroch and MacKnight (1991), and Holme and McIntyre (1984).

Stream-Net Samplers: Stream-net samplers are fitted with a fine mesh net (210 to 500 μ) and collect benthic invertebrates from flowing water as it passes through the sampler. A horizontal, rectangular frame upstream of the vertical net mouth marks off the area to be sample. Generally, the procedure involves stirring up the substrate within the rectangle and scraping stones to remove clinging organisms, the current then sweeps the organisms into the net. These samplers are typically used in shallow waters (< 0.5 meters) with coarse substrate. In the majority of cases this means shallow riffle habitats. The dominant stream-net samplers include the Surber, Invertebrate Box Sampler, and Hess Sampler (Table 6-1).

Table 6-1. Stream-net samplers used to assess benthic communities in freshwater streams.

SAMPLER	SIZE	HABITAT	EFFECTIVENESS	ADVANTAGES	LIMITATIONS
Surber	1 ft^2	Shallow, flowing streams, less than 32 cm deep with good current; rubble substrate, mud, sand and gravel.	Relatively quantitative when used by experienced biologist; performance depends on current and substrate.	Encloses area sampled; easily transported or constructed; samples a unit area.	Difficult to set in some substrate types, that is, large rubble; cannot be used efficiently in still, slow-moving waters.
Hess	0.09 m^2	Shallow, flowing streams, less than 32 cm deep with good current; rubble substrate, mud, sand and gravel.	Relatively quantitative when used by experienced biologist; performance depends on current and substrate.	Completely encloses area sampled; easily transported or constructed; samples a unit area; can be used in weed beds.	Difficult to set in some substrate types, that is, large rubble; cannot be used efficiently in still, slow-moving waters.

Intertidal Samplers: The intertidal zone is unique among those aquatic environments that are routinely sampled for environmental assessments. It can be sampled while exposed to air. On sandy or muddy shores a square sheet-metal frame (e.g., 0.1 or 0.25 m^2) is driven into the substrate and the sediments within the frame are excavated to the desired depth and sieved (Eleftheriou and Holme, 1984). On rocky shores, a square frame of heavy gauge wire is laid on the substrate and the animals within the frame are counted, weighed, or estimated in terms of percent cover (Eleftheriou and Holme, 1984). This latter operation can either be done i*n situ* or by taking a photograph.

Grab Samplers: A grab sampler is any device that is lowered vertically to collect a sample by penetrating the substrate and obtaining a discrete quantity of bottom sediment (Table 6-2). All grab samplers have some type of jaw mechanism that closes upon impact with the sediments, and usually cover a surface area of 0.1 or 0.2 m^2. Grab samplers are most commonly used in medium to fine sediments and in deeper waters, i.e. nonwadable. There are numerous types of grabs, with the more common ones including the Ponar, Ekman,

Peterson, van Veen, and Smith-McIntyre (Table 6-2). The Ponar is the most commonly used grab sampler for pools and lakes, while the van Veen is generally the grab of choice in estuaries and sheltered marine waters. The van Veen has been adopted as the standard sampler in Puget Sound (Simenstad et al., 1991) and the Baltic Sea (Eleftheriou and Holme, 1984). For more exposed marine waters the more stable Smith-McIntyre or Day samplers are preferred.

Core samplers: Core samplers are tubes that are vertically inserted into the substrate and when withdrawn contain the enclosed material and associated fauna. They are best suited for sampling soft homogeneous substrates and are commonly used for assessing benthic invertebrates in lake, estuarine, and marine environments; however, they are occasionally used in streams. Their major advantage over grabs is they sample to a deeper, more uniform depth. There are numerous types of cores ranging from small homemade, hand-held cores for sampling low-order streams to large complex cores for deep-water, open-ocean sampling.

Air-lift samplers: Air-lift samplers use compressed air to scour the substrate and raise water, lighter substrate material, and fauna as the air ascends in a delivery pipe. The material is passed through a net to collect the benthic invertebrates. The air-lift sampler is most efficient in static or slowly moving water. While this type of sampler is not used extensively due to cost and complexity, they are beginning to receive more use in larger rivers and estuaries with fine substrate.

Trawl and Dredge Samplers: Trawl and dredge samplers are towed horizontally across the substrate and their use is generally restricted to lake, estuarine, and marine environments. Trawls consist of nets with their mouths held open by a frame, towed across the surface of the benthic substrate with little if any penetration. Dredges, like trawls, are towed horizontally across the benthic substrate, but they are of heavy construction, often with metal mesh nets or metal collection boxes, and are designed to break off pieces of rock, scrape organisms off hard surfaces, or for limited penetration into the substrate (Eleftheriou and Holme, 1984). Both samplers are often towed for a specified period of time at a uniform speed to permit a better qualitative comparison of sites. Trawls and dredges are good supplements to cores and grabs because they sample a larger area of the benthos.

Table 6-2. Dominant grab samplers used in assessing benthic communities.

SAMPLER	SIZE	HABITAT	EFFECTIVENESS	ADVANTAGES	LIMITATIONS
van Veen	0.1 m^2 & 0.2 m^2	Estuaries, bays, oceans, and other open water habitats with sand, gravel, mud, or clay.	More effective than Peterson grab or Ponar grab in deeper water.	Long arms give increased stability and better digging efficiency; adopted as standard sampler by Puget Sound Protocols.	A very heavy grab that requires use of a boat with winch and cable; stones, pebbles and other debris can hold jaws open causing loss of sample.
Smith-McIntyre	0.1 m^2	Estuaries, bays, oceans, and other open-water habitats with sand, gravel, mud, or clay.	More effective than van Veen under rough conditions.	Grab is contained in a frame giving greater stability; trigger plates insure grab is not activated before resting on the bottom; activated by heavy springs giving more uniform penetration.	A very heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample; considered by some as complicated and dangerous because of the spring operation.
Standard Ponar	523 cm^2	Freshwater lakes, rivers, estuaries, and reservoirs with hard and soft sediments such as clay, hard pan, sand, gravel, and muck; somewhat less efficient in softer sediments.	Not entirely adequate for deep burrowing organisms in soft sediments; very efficient for hard sediments; collects both qualitative and quantitative samples.	Better penetration than other grabs; side plates and screens reduce washout, shock waves, and substrate disturbance; best quantitative grab sampler for freshwater use.	A very heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample.
Petite Ponar	232 cm^2	Freshwater lakes, rivers, estuaries, and reservoirs with moderately hard sediments such as sand, silt, and mud; will not penetrate clay; somewhat less efficient in soft sediments and coarse gravel.	Not entirely adequate for deep burrowing organisms in soft sediments; not useful in clay.	Good penetration for such a small grab; side plates and screens reduce washout, shock waves, and substrate disturbance; can be operated by hand without boat or winch.	Jaws can be blocked by stones, sticks, and other debris causing loss of part of the sample; not efficient in swiftly flowing water of over one meter per second velocity.
Ekman		Freshwater rivers, lakes ,and reservoirs where there is little current; soft sediments such as muck and silt.	Efficient only in soft sediments but weights can be added for deeper penetration in fine sand; collects both qualitative and quantitative samples.	Easy to operate by hand without winch, can be pushed into substrate in shallow water; hinged doors at top reduce washout, shock waves, and disturbance of the substrate; comes in a range of sizes.	Light weight so that jaw will not penetrate hard substrates; jaws often do not close completely due to blocking of jaws or failure of closing mechanism; inefficient in deep water or where there is even moderate current.
Peterson	0.6 to 0.099 m^2	Freshwater lakes, reservoirs, rivers, and estuaries with sand, gravel, clay, and hard pan substrates.	Less effective in most substrates than the Ponar.	Can give quantitative samples if used properly; ranges of sizes available.	Standard grab is heavy and requires boat with winch; can cause washout if dropped rapidly to the bottom; shallow bite by jaws; jaws are easily blocked by debris causing loss of sample; of questionable value as a quantitative sampler.

Cameras: Cameras are coming into increasingly wider use for analyzing benthic communities. They can either be hand or remotely operated. One type of camera analysis, REMOTS® (Remote Ecological Monitoring of the Seafloor), couples photographs of the benthos with a video digitizer and computer image analysis system (Science Applications International Corporation (SAIC), 1985). The camera is located in a wedged shape prism mounted on a frame that is lowered to the bottom. When the frame rests on the bottom, the prism is lowered several centimeters into the sediment and a photograph is taken through the plexiglass window of the prism. The resulting negative is directly subjected to computer imaging analysis producing information on grain size, sediment surface boundary roughness, erosional and deposition criteria, depth of apparent redox potential discontinuity, and infaunal successional stage. REMOTS® data has been used by EPA as part of a risk assessment study for a site in Narragansett Bay (Munns et al., 1991).

Miscellaneous Qualitative Devices: Any sampling equipment presented in this report can be used for qualitative sampling; however, there are a few devices that are used explicitly for qualitative sampling. These include hand-held screens, dip nets, sweep nets, push nets, kick nets, etc. These sampling devices can be used in most any type of shallow water habitat and, if used by an experienced biologist, can provide an adequate assessment of the relative health of a stream. To allow for a better qualitative comparison of the sites with these devices, they can be used to sample a predetermined area or for a predetermined time interval.

Artificial Substrates: An artificial substrate is defined as any device used to mimic specific features of the aquatic environment into which it is placed. Artificial substrates are typically used to sample aquatic habitats that can not be sampled effectively using conventional devices due to physical limitations or substrate heterogeneity (Table 6-3). Artificial substrates remove substrate heterogeneity as a variable by providing identical substrates for organisms to settle on at each station sampled. The two most commonly used artificial substrates include rock-filled wire baskets and Hester-Dendy multiplate samplers. Both effectively sample 'drift' communities but are poor for sampling infauna. NOAA's Marine Ecosystem Analysis Project used horizontal artificial substrates successfully in Puget Sound to assess pollution effects in Commencement Bay, Elliott Bay, and Dyes Inlet in the mid-1980s (Long, personal communication). It should be noted that the purpose of artificial substrates is to monitor changes in invertebrate communities over time and space and do not necessarily reflect the benthic invertebrate community that resides on the natural

substrate. However, this does not negate their usefulness under specific conditions. There are numerous advantages and disadvantages to artificial substrates.

The overall sampling design for artificial substrates is similar to the design for natural substrates described above. Stations selected should be as similar as possible to reduce variability. Artificial substrates can either be anchored to a float near the water surface, suspended within the water column, or set on the stream bottom. Substrates should be positioned at all stations in the same manner. The artificial substrates should remain in the stream for a six- to eight-week colonization period.

Table 6-3. Advantages and disadvantages of artificial substrates.

Advantages	Disadvantages
Allows collection of data from locations that cannot be sampled effectively by other means. Permits standardized sampling. Reduces variability compared with other types of sampling. Requires less operator skill than other methods. Convenient to use. Permits nondestructive sampling of an environment. Permits greater flexibility in sampling programs. Provides effects data for a specific time period (period of deployment).	Colonization dynamics not fully known. Nonrepresentative sampling under either natural or polluted conditions. Artificial substrates require long exposure time (6-8 weeks) to obtain a sample. Loss of fauna on retrieval of samplers. Unforeseen losses of artificial substrates. Inconvenient to use and logistically awkward.

DATA ANALYSIS

This section presents a review of commonly used ecological and statistical approaches for assessing benthic invertebrate communities. Analysis of biological data should focus on approaches that enhance biological interpretation. Statistical techniques should not be the primary assessment approach, but should be used in conjunction with other methods to

obtain a clearer understanding of the biological communities and factors controlling community structure. Statistical analyses may indicate a significant difference in a particular community parameter, but that difference may have little, if any, ecological meaning or relevance.

Ecological Analysis

There are numerous approaches to assess the ecological health of benthic invertebrate communities, including community structure and function. Measurements of community structure include total abundance, diversity, and species composition. Measurements of community function include species colonization rates and feeding guilds. The use of biological communities in environmental monitoring is normally assessed from a structural perspective (Matthews et al., 1982).

The most common ecological analyses used to assess benthic invertebrate communities are listed below. There are several community measures that should always be part of any benthic invertebrate study, including total abundance, taxa richness, percent contribution of the dominant taxa, community similarity, and a listing of all taxa collected by station. Species diversity indices, based on total number of taxa present and relative abundance of each taxa, have been used in the past to characterize sites, but they can be very deceiving depending on community structure and should no longer be used. Furthermore, reports should include all raw data by replicate and a description of the physical habitat at each station. Other analyses can be used on a site by site basis. The ecological metrics selected to assess benthic invertebrate communities in the vicinity of a CERLCA site should be realistic for the type of habitat and the natural resident community. This is particularly important in low-gradient coastal streams since they can naturally have a substrate dominated by fine materials that may support taxonomic groups typically considered stress-tolerant (e.g., Oligochaeta and Chironomidae). Therefore, when interpreting data from aquatic systems it is important to consider the natural status of the stream and not assume that a community dominated by stress-tolerant groups has been negatively impacted by contaminants.

Abundance: Abundance is the number of individuals collected in a specific area, typically 1 m^2. Abundance can be assessed in relation to total number of invertebrates or number of individuals of specific taxonomic groups, usually ranging from order to species. In general, variability between samples increases from total abundance down the taxonomic scale to species. Therefore, four replicate samples may be sufficient for assessing total abundance, whereas, up to fifty samples could be required to have any accuracy on a particular species.

Taxa Richness: Taxa richness is the total number of taxa or taxonomic units recognizable as individual species, whether or not identified to the species level, collected from a station. This is the most basic measure of community diversity, and is the strongest measurement available for assessing community health. Generally, richness increases with increasing water quality and/or habitat diversity, but not always.

Percent Contribution of Dominant Taxa: This provides an indication of community balance. A community dominated by a few taxa is considered to indicate environmental stress, whereas a community with a more even distribution of taxa is generally indicative of better water quality. Percent contribution of dominant taxa should always be part of any benthic invertebrate study.

Community Similarity Indices: These indices are used to determine the similarity in benthic invertebrate communities among stations. Community similarity indices provide a measure of how similar two benthic communities (stations) are in relation to taxa present at both stations, or by the taxa present and their relative abundance. Community similarity indices are widely used and are extremely useful in assessing benthic invertebrate communities in polluted rivers. The similarity matrix is then used to perform a cluster analysis that groups stations based upon their similarity.

EPT Index; This is the total number of distinct taxa within the orders Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies) which are orders of freshwater insects, classified as pollution sensitive.

Ratio of EPT to Chironomidae: This index uses the abundances of these indicator groups as a measure of community balance. Ephemeroptera, Plecoptera, and Trichoptera are pollution-sensitive indicators; whereas, Chironomidae (midges) are pollution-tolerant indicators. A healthy community has an even distribution among all four groups; whereas, an unhealthy system can have a community dominated by the Chironomidae.

Hilsenhoff Biotic Index (HBI): This index summarizes overall pollution tolerance of a benthic invertebrate community with a single value. The index was originally developed to detect organic pollution in riffle/run habitats; therefore, it is uncertain at this point how sensitive the index is to other forms of contamination.

Statistical Approaches

Statistical techniques should be used to help examine and evaluate data, but, there are situations where a nonstatistical approach will provide a better answer to a particular question for far less money and time. Statistical analyses used to assess invertebrate communities can include graphical presentation of data, descriptive statistics, hypothesis testing, correlation and regression statistics, and classification and grouping methods. Graphing raw data is useful for a cursory review of the data and can provide insight to the types of analyses that may be most useful.

The most common statistical approach to assessing benthic invertebrate communities are descriptive statistics, which can include mean, variance (S^2), confidence intervals (CI) (95 percent CI), and range of values. For a survey study, this level of data analysis may be sufficient for assessing status and trends of benthic communities. However, more specific questions, such as, "Is there a statistically significant difference among stations in relation to parameter x ?," require more sophisticated analyses such as analysis of variance (ANOVA). Power analysis may be required to determine the proper sample size for these types of statistical tests. Selection of the statistical analyses to be applied to the data set must be done at the design stage of the study, not after the data is collected (see Chapter 7). A correlation analysis can be used to determine if there is a relationship between two variables (e.g., taxa richness and Cu) near a site. Regression analysis can be used to demonstrate that there is a dependency of one variable on another.

There are a number of multivariate approaches that can be used for pattern analysis. Some of the more common approaches include cluster analysis, principal component analysis, and discriminate analysis. For more detail on these and other multivariate methods see Green (1979), Johnson and Wichern (1982), Gauch (1982), and Pielou (1984).

Chapter 7 of this document has a more detailed discussion of statistical concerns with regards to bioassessment studies.

EPA'S RAPID BIOASSESSMENT PROTOCOLS

In 1987 the EPA released a report entitled "Surface Water Monitoring: A Framework for Change" that concluded that the present monitoring programs be restructured to address more current environmental problems. In response to this emphasis, the Assessment and Watershed Protection Division of the EPA developed and published a report entitled "Rapid

Bioassessment Protocols for Use in Streams and River: Benthic Macroinvertebrates and Fish" (EPA/444/4-89-001). These protocols (Protocols I through V) are to be used for planning and management purposes as screening, site ranking, and trend monitoring.

Protocols I through III use benthic invertebrate communities with each higher level involving a more complex sampling method and analysis. Protocol I is a qualitative survey method that can be used to establish the general condition of a stream using the presence and/or absence of particular invertebrates; taxonomic identification is done in the field with organisms identified to the order or family level. Protocol II is a more detailed program that includes more habitats and the assessment of the community using a variety of community metrics; taxonomic identification is done in the field with organisms identified to the order or family level. Protocol III is the most complex of the benthic invertebrate protocols and includes taking all invertebrates collected to the laboratory for identification of genus/species. This method also involves more complex community analysis. Protocols IV and V involve fish communities. Protocol IV is based on a questionnaire survey and published literature on the area of interest. Protocol V involves field sampling fish communities with a more complex analysis of the information. All five protocols involve assessing physical and chemical parameters for a more integrated analysis.

The EPA Rapid Bioassessment Protocol Manual is an extremely useful document. However, there are a number of things one must remember when using it. All five protocols require a biologist to collect and analyze the data, preferably one that has extensive training in stream ecology. Furthermore, it is important to remember that these protocols and the metrics used to analyze the data have not been tested in all the ecoregions of the United States, and therefore may not be the correct approach in every situation. One needs to select the protocol and/or metrics best suited for the type of habitat of interest.

The EPA Rapid Bioassessment Protocol is useful at CERCLA sites for preliminary assessment, low priority sites, screening and prioritization of numerous habitats at a CERCLA site, and monitoring success of site remediation. In all cases the method should only be used in wadable freshwater streams that have adequate reference stations. The Rapid Bioassessment Protocol should not be used for other types of habitats or for sites where one needs a high degree of confidence in whether there has been a statistically significant change in the benthic community.

SUMMARY

Study Objective

The first step in designing a benthic invertebrate study is to determine the objective(s) of the study. Field studies at CERCLA sites are observational and not experimental; therefore, while statistical techniques can be used, there are limitations in the interpretation of results from some tests. Observational studies cannot determine cause and effect, but only differences among stations, spatial patterns, and correlations among variables. For example, if there is a statistically significant difference among stations in a particular community parameter (e.g., taxa richness), it is acceptable to state that there was a difference among stations. It would be inappropriate to state that the difference was due to contaminants released from a CERCLA site. However, benthic studies used in conjunction with sediment chemistry and toxicity tests, as part of a sediment quality triad approach (Long and Chapman, 1985), can provide a preponderance of evidence that strongly infers that the CERCLA site is responsible for the differences.

Sampling Methods

- Quantitative sampling should be used for statistically based programs.
- Qualitative sampling should be used for a general survey of a site.

Sampling Devices

- Natural substrates should be sampled whenever possible, with artificial substrates used when natural substrates cannot physically be sampled or when high substrate variability needs to be controlled as a factor.
- Stream-net samplers (e.g., Hess or Surber) should be used for assessing benthic invertebrate communities in riffle habitats; whereas, a grab sampler (e.g., Ponar) should be used for assessing benthic invertebrate communities in pool habitats.
- Qualitative sampling can be done with most any device available.
- Artificial substrates can either be anchored to a float near the water surface, suspended within the water column, or set on the bottom.

- Artificial substrates should remain in the water for a six- to eight-week colonization period.

Data Analysis

- All benthic invertebrate studies should at a minimum include total abundance, taxa richness, percent contribution of dominant taxa, and community similarity indices. Additional ecological analysis would depend on the nature of the community and aquatic system.

- Statistical analysis used to assess benthic invertebrate communities should be based upon the objective of the study.

REFERENCES

Eleftheriou A. and N.A. Holme. 1984. Macrofauna techniques. In: N.A. Holme and A.D. McIntyre (eds). Methods for the Study of Marine Benthos (2nd Ed.) London. pp 1-26.

Gauch, H.G. 1982. Multivariate Analysis in Community Ecology. Cambridge Studies in Ecology. New York, Cambridge University Press. 298 pp.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. New York, John Wiley and Sons. 257 pp.

Herricks, E.E. and J.Cairns, Jr. 1982. Biological Monitoring. Part III - Receiving system methodology based on community structure. Water Resources 16:141-153.

Holme, N.A. and A.D. McIntyre (eds.). 1984. Methods for the Study of Marine Benthos (2nd Ed.). London, Blackwell Scientific Publications. 387 pp.

Hornig, C.E. and J.E. Pollard. 1978. Macroinvertebrate sampling techniques for streams in semi-arid regions. Comparisons of the Surber method and a unit-effort traveling kick method. EPA-600/4-78-040. Las Vegas, NV. U.S. Environmental Protection Agency.

Hunsaker, C.T. and D.E. Carpenter, eds. 1990. Ecological indicators for the Environmental Monitoring and Assessment Program. EPA 600/3-90/060. Triangle Park, NC, U.S. Environmental Protection Agency, Office of Research and Development, Research.

Johnson, R.A. and D.W. Wichern. 1982. Applied Multivariate Statistical Analysis. New Jersey, Prentice-Hall. 594 pp.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA 600/4-90/030. Cincinnati, OH, U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory.

Long, E. R. and P. M. Chapman. 1985. A sediment quality triad: Measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Marine Pollution Bulletin (16):10. pp 405-416

Long, E.R. 1992. personal communication. Coastal Monitoring and Bioeffects Division, Office of Ocean Resources Conservation and Assessment, National Oceanic and Atmospheric Administration. Seattle, WA.

Mason, C.F. 1981. Biology of Freshwater Pollution. New York, Longman Inc. 250 pp.

Matthews, R.A., A.L. Buikema Jr., J.Cairns Jr., and J.H. Rogers Jr. 1982. Biological monitoring: Part IIA. receiving system functional methods, relationships and indices. Water Resources. 16:129-139.

McIntyre, A.D., J.M. Elliott, and D.V. Ellis. 1984. Introduction: Design of sampling programmes. In: N.A. Holme and A.D. McIntyre.(eds). Methods for the Study of Marine Benthos (2nd Ed.) London. pp 1-26.

Mudroch, A. and S.D. MacKnight. 1991. Handbook of Techniques for Aquatic Sediment Sampling. Boca Raton, FL. CRC Press. 210 pp.

Munns, W.R., Jr., C. Mueller, D.J. Cobb, T.R. Gleason, G.G. Pesch, and R.K. Johnston. 1991. Marine Ecological Risk Assessment at Naval Construction Battalion Center, Technical Report 1437. Davisville, RI. Naval Ocean Systems Center. 129 pp+.
Pielou, E.C. 1984. The interpretation of ecological data: A primer on classification and ordination. New York, John Wiley and Sons.. . 263 pp.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. EPA/440/4-89-001. Washington, DC, U.S. Environmental Protection Agency, Assessment and Watershed Protection Division. 135 pp+ appendices.

Pollard, J.E. 1981. Investigator differences associated with a kicking method for sampling macroinvertebrates. Journal. Freshwater Ecology 1 (2) 215-224.

Scientific Applications International Corporation (SAIC). 1985. REMOTS® Survey of Zones of Siting Feasibility Puget Sound, Washington. Newport, RI, SAIC pp 154.

Simenstad, C.A., C.D. Tanner, R.M. Thom, and L.L. Conquest. 1991. Estuarine habitat assessment protocol. U.S. EPA 910/9-91-037.

Wilhm, J.L. 1975. Biological indicators of pollution. In: B.A. Whitton (ed) River Ecology Berkeley CA, University of California Press. .pp. 375-402.

CHAPTER 7

STUDY DESIGN AND STATISTICAL ANALYSIS

INTRODUCTION

This chapter discusses the major elements of designing a bioassessment study to assess conditions at a hazardous waste site, including the statistical analysis of the resulting data. Before beginning to design an environmental study at a hazardous waste site it is necessary to understand the general limitations of this type of study. There are basically two types of studies, experimental and observational[1]. Experimental studies permit the researcher to control all the variables and to replicate the treatments of concern (e.g., contaminant concentrations). With observational studies the only variables the researcher can control are time and space (location). While statistical analysis can be applied to both types of studies, the experimental study permits conclusions regarding cause and effect, but the observational study can only determine spatial and temporal patterns and correlations among variables. Environmental studies at hazardous waste sites are observational studies, therefore, they do not permit the drawing of absolute conclusions with regards to the causes of any bioeffects that might be observed. For example, in a benthic community study, if there is a statistically significant difference among stations in a particular community parameter (e.g., taxa richness), it is acceptable to state that there was a difference among stations. It would be inappropriate to state that the difference was due to contaminants released from the hazardous waste site. However, a properly designed bioassessment study can provide evidence that strongly suggests a cause for the observed bioeffects. Waste site studies are further limited by the fact that contamination has already occurred so there is no opportunity for before- and after-contamination sampling.

If properly misused, statistical analysis can prove anything you wish to prove, regardless of the facts. If properly used, statistics become a very helpful decision-making tool. However, even when properly used, statistics can give unwarranted credence to biological data sets, especially in the eyes of the statistically naive. Therefore, whether applying or reviewing statistical analyses you need to know that the data analyzed is in an appropriate form and

---

[1] Hurlbert, 1984 uses the terms manipulative and mensurative experiments, while Eberhardt and Thomas, 1991 talk about controlled and uncontrolled events.

that you understand exactly what the statistical results mean. The purpose of this chapter is not to make the reader a statistician, but to familiarize the reader with some of the requirements and problems of statistical analysis. This will be accomplished by defining the more commonly encountered statistical terms, briefly discussing the differences between descriptive and inductive statistics, as well as parametric and nonparametric methodologies for both comparative and correlative analyses. Other topics of discussion include hypothesis formulation, probability, and replication and pseudoreplication.

Before continuing, some terms as they apply to bioassessment studies and statistical analysis, in particular, need to be defined. A **statistical population** consists of the totality of all possible observations of the variable with which we are concerned. While a **biological population** consists of a group of individuals of the same species between which genetic material freely flows. Since the terms 'statistical' and 'biological' rarely precede the term 'population,' the context in which the term 'population' is found must be used to determine the term's exact meaning. If we are concerned with the mercury concentration in the mussels in the Bay of Belfalas, then the 'statistical' population would consist of the measurements of mercury concentration in each of the five million mussels making up the 'biological' population in the bay. Since we would not want to analyze all five million mussels, thus wiping out the mussel population in the bay, we would take a **sample,** which is a subset of the population. In this case, the sample would consist of something less than five million observations. In environmental studies the term sample often refers to a single observation and in discussions of the statistical analysis of these studies, the term often refers to single observations as well as to population subsets consisting of several observations. In the context of this chapter, individual observations making up a sample will be referred to as **sample replicates** or simply **replicates**. The characteristics of a statistical population, such as the population mean or variance, are referred to as **parameters**; the characteristics of the sample (e.g., sample mean, sample variance) are referred to as **statistics**.

STUDY DESIGN

Question Formulation

The first step in designing any environmental study is to determine the question or questions you wish the study to answer. The more concisely you formulate your

question(s), the more precisely you can design a study that will provide appropriate answers. At hazardous waste sites the general questions of concern to NOAA are:

1. Are contaminants in or entering the aquatic environment?
2. Are contaminants bioavailable?
3. Are contaminants causing or have the potential to cause bioeffects?

The first question also includes concern over what contaminants are present and what is the extent of the contamination (i.e., how far away from the site does contamination extend). No single study methodology is capable of answering all of these questions. Therefore, an environmental study at a hazardous waste site should consist of a multiplicity of methods. The use of a multiplicity of study methods helps relate environmental contamination and bioeffects to the waste site through a preponderance of evidence.

The presence of contaminants in the environment can be determined by chemical analysis of various media (e.g., water, sediment). Samples taken near the hazardous waste site can be statistically compared with samples from an area that represents a similar habitat but is not influenced by the site (i.e., a reference site). Such comparisons can show that higher concentrations of contaminants exist near the waste site than at the reference site, however, they cannot show that the waste site is the source of the contaminants. If the study design correlates decreasing contaminant concentrations with increasing distance from the waste site, then it would be reasonable to infer that the waste site is the source of the contaminants.

The presence of contaminants in the environment does not mean that they are bioavailable (i.e., able to be taken up by biological organisms). Therefore, the next step is to determine the bioavailability of the contaminants. Currently no acceptable method of chemical analysis can determine contaminant bioavailability. Therefore, chemical analysis can answer question number one, but supplies little if any information with regard to the other two questions.

The most direct way to determine if contaminants are bioavailable is by performing tissue analysis on indigenous organisms to determine if contaminants or metabolites of contaminants are bioaccumulating in these organisms. Bioaccumulation studies can also be conducted with transplanted or artificial organisms as well as by collecting contaminated media and exposing laboratory organisms to it. To determine whether or not the bioaccumulation is related to the waste site, it needs to be correlated to media contamination

levels. This is accomplished by collecting indigenous organisms from or placing transplanted or artificial organisms in the same locations from which the chemically analyzed media samples were taken. For conducting correlation analyses, transplanted and artificial organisms have an advantage in that the contaminant levels measured in these organisms are solely due to current levels of media contamination. Contaminant levels in indigenous organisms may reflect past media contamination and therefore may not show a correlation with current media contamination levels. Laboratory bioaccumulation studies could also be conducted, but the media samples to be used should be split, with one portion being chemically analyzed and the other used in the bioaccumulation study.

While bioaccumulation of contaminants implies their presence in the environment, it does not necessarily mean that there are any bioeffects related to the contamination. Determining whether contaminants from a hazardous waste site have caused or have the potential to cause bioeffects can be the most difficult question to answer. Hazardous waste sites can fall into three basic categories with regard to contamination of aquatic environments:

- Sites that do not release contaminants into aquatic environments.
- Sites that release such high concentrations of contaminants into aquatic environments that no biological organisms exist near the site.
- Sites that release some intermediate concentrations of contaminants.

Sites in the first category are of no concern to NOAA. At sites in the second category the effects are obvious. Sites in the third category, which make up the majority of sites with which NOAA is concerned, present the biggest challenge with regard to determining the existence of or potential for bioeffects and relating the effects to the waste site.

Toxicity testing and benthic community studies are the two most commonly used bioassessment methods for determining bioeffects or the potential for bioeffects at hazardous waste sites. A third approach for determining bioeffects is the use of biomarkers. These methods by themselves cannot tell you that the waste site is causing bioeffects; they can only tell you that conditions at the waste and reference sites are different. However, when these methods are used in conjunction with media chemical analysis and bioaccumulation studies, a body of evidence can be built up from which a cause and effect can be inferred. For example, suppose a bioassessment study, using bioaccumulation, toxicity testing, benthic community studies, and metallothionein analysis (a biomarker test indicating a response to metal exposure), was conducted at a riverine hazardous waste site in conjunction with sediment chemistry, and the following results were obtained:

- Sediment chemistry showed high concentrations of metals near the site with a gradient of decreasing concentrations moving downstream from the site.
- Tissue analysis of clams showed elevated levels of metals in areas of high sediment metals concentrations.
- Toxicity tests correlated increased mortality of waterfleas with high sediment metals concentrations.
- Benthic communities near the site had reduced biomass when compared to more distant and upstream sites.
- MT levels were elevated in crayfish.

By themselves each of these bioassessment techniques indicate only a difference between sampling stations near the waste site and reference sites. Taken together and in conjunction with the sediment data, they present strong evidence that metals are being released from the waste site and are causing bioeffects in the adjacent river. This approach is often referred to as a 'preponderance of evidence.'

So far this discussion has only dealt with general questions that could relate to any waste site, but before a study design can be completely developed, questions specifically related to the waste site of concern and the techniques to be employed must be formulated. It is at this early stage of question formulation that decisions should be made with regard to the type of statistical analysis to be performed on the collected data. Statistical analysis selection should be driven by the biological concerns at the site, not the other way around (Skellam, 1969; Green, 1979). However, the precise design of a study must be appropriate for the desired statistical analysis. More information on question specificity and statistical analysis will follow in the section on statistical analysis.

Reconnaissance Survey

Before finalizing a study design, a reconnaissance survey, which should include some preliminary sampling, should be conducted. A reconnaissance survey provides information on the types of habitats present, sources of pollution (point and/or nonpoint), access points, and any other factors that may influence either the natural community or the ability to sample those communities. A reconnaissance survey can also be used to determine the best sampling device for the habitats of concern. In addition, if the objective of the study is to determine if there is a significant difference among stations, then preliminary samples can be collected for determining the sample size required for a particular level of statistical

power (Gilbert, 1987; Green, 1979). While a reconnaissance survey may appear to add time and expense to the overall study, it more often than not will save both (Green, 1979). It would be exceedingly costly, in both time and money, to conduct a full-scale bioassessment sampling program only to find out the organism you were trying to sample did not exist at the site or you were in the tidally influenced section of a river and your upstream reference station was actually impacted by the site and could not be used as a reference station. Either of these problems, as well as many others, could cause you to completely redo your sampling and could have been avoided by conducting a reconnaissance survey.

Station Selection

The selection of inappropriate stations can reduce the efficiency of the sampling devices and the data analysis and interpretation, thereby leading to misinformation and poor decisions about the health of an aquatic ecosystem and future actions to be taken at a site. The selection of appropriate stations will minimize the influences of the natural variability that occurs in aquatic ecosystems and increase the confidence for making specific statements about biological integrity.

A wide variety of non-contaminant related abiotic factors, which can influence the results of bioassessment studies, should be considered when designing a study. The combination of physical and chemical factors can determine whether a particular taxa will successfully inhabit a given area or whether an organism will survive a toxicity test irrespective of human influences, i.e., anthropogenic contamination. All sampling stations to be compared should be as similar as possible with respect to all physical and chemical parameters, (e.g., sediment grain size, TOC, salinity, current velocity, and depth) excluding contamination. Physical and chemical factors of concern should be measured at each station to determine any abiotic differences among stations. This is necessary to determine whether biological differences between stations are caused by differences in contaminant levels or by some other abiotic factor.

The number and locations of stations will depend on the objective of the study, the nature of the aquatic system, and the funds available for the assessment. Stations should be positioned in similar habitats with a spatial distribution that will permit an assessment of background conditions as well as cover a gradient from the most-to the least-contaminated areas. A minimum of one station should be located where contamination is expected to be the highest, for example, in a stream one station location should be just downstream of the hazardous waste site. Several stations may be required to identify the area of highest

contamination, because of high spatial variability of contamination resulting from incomplete mixing of contaminants and water or sediments. Therefore a haphazardly placed station may fall outside the contaminant plume. In attempting to determine a gradient, stations should be spaced farther apart as one moves away from the site. This spacing will permit a clearer delineation of the extent of biological effects and identification of "recovery" zones. Particularly in the case of benthic community studies, stations should not be located in areas influenced by atypical conditions or structures, such as bridges, channelized areas, dredging activities, or culverts. To facilitate interpretation of relationships among the data, stations used for bioassessment methodologies should be the same as those used for water or sediment chemistry.

When selecting a habitat for establishing stations, it is important to consider whether the benthic invertebrate community being sampled is being exposed to the contaminant of concern. One of the major areas of disagreement in station location, related to benthic community studies in streams, is whether they should be located in a riffle (erosional habitat) or pool (depositional habitat). Traditionally, the majority of benthic invertebrate monitoring has focused on riffle communities due to ease of sampling, increased sampling precision, higher species diversity, presence of pollution-sensitive taxa, and knowledge of riffle communities and how they respond to stress. This bias towards sampling riffle communities may have inherent problems in the field of contaminant assessment. Because contaminants can be transported over a riffle habitat to settle out in a depositional area, invertebrates in riffle habitats may not be exposed to contaminants in the same way that invertebrates in pool habitats are exposed. For bioassessments at hazardous waste sites, pool habitats will have a higher concentration of contaminants and benthic community data would be more easily correlated with sediment chemistry than in riffles. Therefore, pool habitats would be the preferred habitats to sample, however, an optimal design would be to sample benthic invertebrate communities in both riffle and pool habitats.

Reference stations are critical to assess the biological condition of aquatic systems in the vicinity of a hazardous waste site. An optimal sampling design would be to have the number of reference stations equal the number of treatment stations (i.e., potentially impacted stations). As this is not practical for most studies, a minimum design should include two reference stations. The reference stations should be located in an area not influenced by the site or any additional sources of contamination and should be as similar to treatment stations in all other respects as possible. A single reference station should not be used since it can be misleading if it does not truly represent background conditions. If

reference stations are not available in the same aquatic system (i.e., same stream or estuary), select stations from a reference system that is physically and biologically similar to the system of interest. This could be a stream in an adjacent watershed or one in the same ecoregion.

Random versus Non-random Designs

Sample locations can be determined by either random or nonrandom means. Random sampling is used when quantitative data is required for statistical analysis; whereas nonrandom sampling is sufficient for qualitative or semiquantitative sampling.

Two types of random sampling designs are simple random and stratified random sampling. Simple random sampling is when every unit of the population has an equal chance of being sampled. This is done by using random numbers to select sample locations. Simple random designs are not often used in waste site studies because contamination at the site is invariably heterogeneously distributed throughout the environment, therefore the relatively small number of samples or replicates taken at most sites might miss the areas of highest contamination. This heterogeneity of spatial distribution also applies to benthic organisms whose presence at any one location is dependent on physical factors such as substrate, current, etc. Therefore, the results of simple random sampling at a hazardous waste site may not be representative of conditions at the site. To characterize a Superfund site based on simple random sampling, the sample size must be greatly increased above what is currently the norm at these sites.

Stratified random sampling is preferred in bioassessment studies at hazardous waste sites because it increases sampling efficiency by dividing the habitat into similar strata thus reducing natural variability between stations. Physical features commonly used to stratify an aquatic system are substrate, depth, flow, and duration of submergence (in tidally influenced areas). Several substrata can also be used to ensure a complete assessment of the entire aquatic system of interest, but only similar strata should be compared to determine differences that might be caused by contamination. After the system is stratified, the number of samples can be proportioned among the various strata based upon areal coverage or taken from a single strata.

If quantitative data are not required for a study, then a nonrandom sampling program can be used. A nonrandom design should not be equated with an inadequate sampling design. The design used should depend on the question of interest, which may or may not require a statistical design. Nonrandom designs can either consist of systematic or nonstructured

sampling. Systematic sampling involves using transects or grids with samples collected at predetermined or random intervals across the transect/grid. This type of sampling can be used to determine if there is a gradient of change away from a source of contamination; however, some consider it nonrandom and therefore there are inherent limitations on the type of data analysis that can be performed.

Samples can also be collected using a nonstructured sampling program where there is no predetermined pattern. This approach provides strictly qualitative data.. The Rapid Bioassessment Protocol is technically a nonrandom design, the method states that samples should be collected from riffle and run habitats only, therefore there is a type of stratification to the design (see Chapter 6, EPA's Rapid Bioassessment Protocol).

Sample Replication

For quantitative bioassessments, replicate samples need to be taken for each time and location sampled. Because, to determine that the differences between sampling stations or areas of concern are real, the within station or area variability needs to be determined. Within station variability is determined by the variability of the replicate samples. For the simplest and most straight forward statistical analysis, an equal number of randomly allocated replicate samples should be taken for each time and location sampled.

One commonly occurring problem in environmental studies is pseudoreplication. **Pseudoreplicates** are not independent replicates of the population they are claimed to represent. The easiest way to understand the concept is through an example. A simple example is the case of subsampling. Suppose you took a single grab sample from each of two sites (X and Z) and then took three random subsamples from each grab for analysis. The sampling methodology in and of itself is not in error, but now suppose you try to statistically compare the two sites based on the analysis of these subsamples. Statistical comparison tests assume that the variability between the replicates of the samples being compared are representative of the within population variability (i.e., between replicate variability in sample X is representative of within population variability at site X) of the populations being compared. For this to be true, the replicate samples must be independent of each other, i.e., the location of any one replicate sample must not influence the location of any other replicate sample However, the subsamples that were analyzed actually only represent the variability within the individual grabs, not within the entire population at the site. The subsamples of the grabs are true sample replicates for the individual grabs, i.e., they are random independent observations comprising samples of the statistical

populations consisting of all possible observations of the individual grabs. However, they become pseudoreplicates if used to compare the two sites because they lack independence with respect to the site population; they came from the same grab. Therefore you can statistically compare grab X to grab Z but you cannot statistically compare site X to site Z based on the subsamples.

As a second example of pseudoreplication, assume you wish to determine whether or not submergence time affects Pb levels in mussels from the Bay of Belfalas where the average tidal range is three meters. You decide to sample mussels from one meter below mean high tide (BMHT) and from two meters BMHT. The hypothesis you wish to test would be: there is no significant difference in lead concentrations in mussels from the two depths. You then go out to Mussel Rock and with the aid of a random number table you take eight replicates of mussels from one meter BMHT; then you proceed to Bird Rock and in the same manner take eight replicates of mussels from two meters BMHT. After performing an appropriate statistical analysis you declare that lead concentrations in mussels from the Bay of Belfalas are higher at two meters BMHT than at one meter BMHT with α=0.05.

By stating that the mussels from the one-meter BMHT on Mussel Rock were representative of the entire population of one-meter mussels in the bay and that those from Bird Rock were representative of the entire population of two-meter mussels in the bay you have committed the error of basing your conclusions on pseudoreplicates. The replicates taken from one meter BMHT were independent representatives of the population of one-meter mussels on Mussel Rock, but were not independent representatives of the baywide one-meter mussel population. Likewise the two-meter Bird Rock replicates represented the two-meter population on Bird Rock, not the two-meter baywide population. You can validly conclude from your data that mussels from two meters BMHT on Bird Rock have higher concentrations of lead than do mussels from one meter BMHT on Mussel Rock. You can't conclude that depth is a factor in the difference because location may be a factor; Bird Rock may be located near a point source and all the mussels there may have elevated Pb levels. If both samples were taken from Mussel Rock you could relate the difference to depth, but you still couldn't relate the results to baywide conditions. In order to draw baywide conclusions, the replicates for the two depth samples must be randomly selected from throughout the bay not from just one or two locations. As Hurlbert (1984) states: "Pseudoreplication thus refers not to a problem in experimental design . . . per se but rather to a particular combination of experimental design . . . and statistical analysis which is inappropriate for testing the hypothesis of interest."

The previous example on pseudoreplication raises another problem, which is really more a sampling design problem than an actual statistical analysis problem, although the solution does effect the type of statistical analysis used. The example suggested that Bird Rock might be near a point source and that this was the cause of the difference between samples. When using relatively small sample sizes, completely independent random selection of the two sets of sample replicates could result in the sample being unduly influenced by a single replicate being located near a point source. If Pb concentrations really do increase with depth and the replicate influenced by the point source was from the shallower sample, then no difference might be indicated by the test results. Probably the easiest solution in this case would be to pair the replicates; randomly select sites throughout the bay then take a one-meter and two-meter sample at each site. The statistical analysis used would be designed for paired replicates.

STATISTICAL ANALYSIS

In the past many environmental biologists have had a fear and distrust of statistics; this was in large part probably due to the backwards approach to statistical analysis that was often taken in environmental studies. These biologists would too often conduct a study, compiling quantities of data only to find that it could not be properly analyzed statistically. Because a data set cannot be statistically analyzed does not mean that it contains no useful information, but if data is to be statistically analyzed, determine the type of analysis to be performed during the early planning stages of the study not upon the completion of data collection.

The first step in designing a scientific study is to determine what question or questions you want the study to answer. The questions you want answered will determine whether statistical analysis is necessary, and, if so, the appropriate type of statistical analysis needed. If statistical analysis is desired then the sampling plan must insure that the data collected is quantitative. To answer questions regarding the existing conditions at a site, **descriptive statistics** would be sufficient. Descriptive statistics include things like: means, medians, standard deviations, and ranges. However, to make predictions or inferences about conditions at the site, **inductive statistics** are required. Inductive statistics may use descriptive statistics to permit the determination of whether conditions at a waste site are different from those at a control site with a known probability (as in the case of t-tests). Inductive statistics may utilize the raw data to determine the probability of relationships or dependencies between different parameters, as in correlation and regression analyses. The

key word with regard to inductive statistics is probability. Inductive statistics tell the probability of something being the case; they do not tell whether or not it is the case.

There are some instances when descriptive statistics are sufficient for the task at hand. For example, if the mean concentration of Hg in lobster muscle at a site is two parts per million (ppm) and the U. S. Food and Drug Administration limit for seafood is one ppm, you know that lobsters at the site represent a potential human health risk. Most of the time bioassessment studies will be concerned with whether or not a hazardous waste site has produced a significant adverse impact on the environment and to what extent this impact exists. Inductive statistics will permit you to answer these questions with a specified probability of being right.

Hypothesis Formulation and Testing

Before determining the type of inductive statistics best suited to answer your question, you need to make your question as specific as possible. Your general question may be, 'does the hazardous waste have any potential impact on the local aquatic environment?' You may decide to test for impact by conducting a bioaccumulation study with mussels. Your question would then become, 'do mussels at the site have higher levels of contaminants than do mussels not influenced by the site (e.g., at a reference site)?' Once you have posed the question specifically, including the type of test or study to be conducted, it needs to be put in the form of a **null hypothesis** (H_0). A H_0 is one we hope to reject in favor of an **alternative hypothesis** (H_a) with a known probability of being correct. However, failure to reject H_0 only implies you have no statistically valid evidence to believe otherwise. A H_0 can be disproven with a known probability but cannot be proven.

In the above Bay of Belfalas mussel example the H_0 would be: there is no difference between contaminant levels in mussels at the site and the control site. A statistical way of phasing the H_0 would be that the samples are from the same parent population or are from different parent populations with the same population mean. A possible H_a would be there is a difference, or statistically speaking, the samples are not from the same parent population or the parent populations have different means. The most common H_0 in bioassessment studies is a no difference or no effect hypothesis.

Testing a H_0 is generally accomplished by using one of two types of statistical methods: **parametric** or **nonparametric**. **Parametric** methods are concerned with the characteristics of population parameters (e.g., population mean and variance), while **nonparametric** methods are not dependent on the characteristics of population parameters

for their validity. Parametric tests require that certain assumptions pertaining to the population parameters be met for the tests to be valid. These assumptions include that the samples are drawn from a normally distributed population(s), the parent populations have the same variance, and the mean and variances are independent (i.e., the size of the variance is independent of the size of the mean). These assumptions can (if sample size is sufficiently large) and should be tested prior to applying any statistical tests to your data. Green (1979) describes in detail an appropriate approach for such testing, and today, many computer statistical software packages will test for these assumptions.

Populations in environmental studies rarely meet these conditions. One alternative is to transform the data in such a manner that the conditions are met and then perform a parametric test on the transformed data. The most common transformations are logarithmic or natural logarithmic. However, transformations are not a cure all and it must be kept in mind that if the sample statistics calculated from the transformed data are transformed back, the resulting sample statistics will be different than those calculated from the untransformed data. For example: Hg concentrations in three mussels from the Bay of Belfalas were 2.5, 3.7, and 1.2 with a mean concentration of 2.47; if the concentrations are log transformed they become 0.398, 0.568 and 0.079 with a mean of 0.348. When the antilog of 0.348 is calculated it is 2.23, not the 2.47 arithmetically derived from the original data. Transformations are one more step removed from reality and must be used with caution.

The parametric tests most commonly used in environmental studies are the Student's t-test for the comparison of two samples and ANOVA for the comparison of multiple samples (Table 7-1).

If the parametric test conditions are not met and transformation is undesirable, then nonparametric tests can be considered. Nonparametric tests are essentially independent of the population parameters. While nonparametric tests are not quite as precise as parametric tests when parametric test conditions are met, they lose little precision when these conditions are not met, and may therefore be preferable to parametric tests of transformed data. Nonparametric tests involve less complex calculations than do parametric tests and are also available in commercial software packages. It is perfectly legitimate to apply nonparametric tests to data that meet parametric test conditions as long as it is kept in mind that the results are slightly less precise. Commonly used nonparametric tests for environmental studies are the Mann-Whitney U-test and Kolmogorov-Smirnov two-sample test for comparing two samples and the Kruskal-Wallis test for multiple samples (Simenstad et al., 1991) (Table 7-1).

Because you are dealing with probabilities and not absolutes, there is still the possibility that you might reject the H_0 when in fact it was true, a **type I error**; or you might fail to reject H_0 when it is false, a **type II error**. At hazardous waste sites the most common form of H_0 is: there is no difference between conditions (e.g., contaminant concentrations, benthic

Table 7-1. Examples of parametric and nonparametric tests for the comparison of samples adapted from Elliott (1977) (see Elliott for examples of applications of the tests).

Conditions	Parametric Test	Transformation necessary	Nonparametric Test
Comparison of two samples			
A. Means of large samples (n>50)	Normal deviate (d)	no	U-test Kolmogorov-Smirnov
Small samples (n<50) from a:			
1. random distribution	d for Poisson	no	U-test Kolmogorov-Smirnov
2. contagious distribution	t-test	yes	U-test Kolmogorov-Smirnov
B. Variances of large samples (n>50)	F-test	no	
Small samples (n<50)	F-test	yes	
Comparison of three+ samples from random or contagious distributions	ANOVA	Yes	Quenouille Kruskal-Wallis Friedman
Correlation between two variables (dependent variable y and independent variable x)	Correlation coefficient Regression analysis	Yes (x & y) Yes (y only)	Rank correlation coefficient
Correlation between three+ variables:			
Variables analyzed in pairs	Regression analysis	Yes (y only)	
Variables analyzed simultaneously	Multiple regression	Yes	

community structure, etc.) at the waste site compared with those at a clean reference site. If this H_0 is true then the waste site is clean and no remediation needs to be performed; to the contrary, if it is proven to be false then possibly very expensive remediation would have to be perform. As trustee for certain natural resources, NOAA would like the H_0 to be true (i.e., the site is clean and no threat to NOAA's resources). However, NOAA would prefer to err on the conservative side and commit a type I error, saying that a clean site is contaminated and needs to be cleaned, rather than commit a type II error, saying that a contaminated site is clean thus leaving the contamination in place. Polluters would prefer to commit a type II error because it would mean that they would not have to spend any money on remediation even though the site was actually contaminated.

The results of both parametric and nonparametric tests are generally reported with a probability value (α or p), usually 0.1, 0.05, or 0.01. These values represent the probability of committing a type I error, rejecting the H_0 when it is true. If the statistical test indicates that the samples are significantly different (the H_0 is rejected) with α=0.05 then based on the available evidence there is still a 1 in 20 or 5 percent chance that the samples are not significantly different (the H_0 is true). If a 5 percent chance of being wrong in rejecting the H_0 is too great, then α can be lowered to 0.01. However, by reducing the probability of committing a type I error, without changing the sampling design, you increase the probability of committing a type II error, accepting the H_0 when it is false. The probability of committing a type II error is expressed as 'β'. The **power** of a test is the probability of rejecting the H_0 and accepting the H_a when the H_0 is in fact false; it is represented as '1-β'. Power is an important concept, however, it can only be calculated for a specific H_a (e.g., Hg concentrations at Site 1 are twice as high as at the control site) not a general one (i.e., there is a difference between the sites). NOAA would prefer a powerful test with a concurrently high α and low β.

Whether you use parametric or nonparametric tests, any comparison of samples is dependent on a knowledge of the variability within the individual samples. To accomplish this, each sample must consist of a minimum of three and preferably more sample replicates. The more replicates per sample the smaller the detectable difference for a given value of α. Increasing the number of replicates is also one way of reducing the probability of both type I and type II errors. Since unlimited sampling is extremely impractical with the limited resources available for bioassessment studies, you need to determine the most efficient allocation of your sampling resources (e.g., ten sampling sites with three replicates or three sites with ten replicates per site). Formulas for determining the ideal sample size

based on the desired type of analysis, degree of precision and costs can be found in texts (e.g., in Gilbert, 1987) or by consulting a statistician.

Randomization

Precision is a measure of the closeness of agreement of individual measures of the same quantity (closeness of replicate values), while **accuracy** is the closeness of a measured or computed value to its true value (closeness of sample mean to the true population mean) (Sokal and Rohlf, 1983). It is possible to have precision without accuracy with a biased sampling methodology, but it would only be pure chance to have accuracy without precision. If your sampling methodology called for a four-centimeter mesh size for determining the size of a fish population based on the number of fish caught per unit of effort, then your sample would be biased for fish too large to pass through four-centimeter mesh. While your sampling may be precise, i.e., you catch roughly the same number of fish per unit effort in each replicate, your conclusions about the size of the fish population would not be accurate because your sampling methodology systematically excluded the smaller fish from your calculations.

Statistical analyses take into account random error but not systematic error. Systematic error, or **bias**, is the systematic distortion of a statistic due to sampling methodology. Therefore, for a given value of α to be an accurate indication of potential error, the samples must be unbiased representations of their respective populations. To fulfill this requirement, the selection of any particular replicate must be independent of the selection of any other replicate; in other words, the population must be randomly sampled. Random selection does not mean haphazard selection; each sample replicate must be selected "...using a consistent, standardized technique" (Simenstad et al., 1991). While the selection of each replicate must be random, the treatment (e.g., how it is taken and how it is measured) of each replicate must be identical. The importance of replication and randomization are summed up by Hurlbert (1984): "Replication reduces the effects of 'noise' or random variation or error, thereby increasing the *precision* of estimates. . . Randomization eliminates possible bias on the part of the experimenter, thereby increasing the *accuracy* of such estimates.

Correlation and Regression Analysis

So far we have been discussing how to compare populations to determine whether or not they are different. Suppose the question you wish to ask is, "does the concentration of lead in mussels from the Bay of Belfalas decrease with increasing distance from the Mordor, Inc. lead smelter?" There are two approaches you can take to this problem: you can take samples of *n* replicates each at various distances from the smelter and use some comparison test to test for differences, or you could perform either a correlation or regression analysis on the individual replicates or the sample means with respect to distance from the smelter. A **correlation analysis** determines whether changes in two or more variables are related. A **regression analysis** determines whether changes in a dependent variable are the result of changes in an independent variable. Correlation or regression analyses do not require multiple replicates at each distance nor do they require random selection of sampling locations. At the same time they do not prove that a relationship exists; they only suggest that it does. The easiest approach to determine an environmental gradient is to establish a transect in the direction of concern (e.g., downstream in rivers and streams) and select sampling locations increasingly farther apart as they move away from the site (e.g., 10, 25, 50, 100, 200, 500, and 1000 meters). Then take samples consisting of one random replicate at each distance and perform a regression analysis of Pb concentrations with distance from the smelter. As with the comparison tests, the larger the sample size the smaller the detectable gradient.

SUMMARY

Here are a few final thoughts on the place of statistical analysis in bioassessment studies. The design of bioassessment studies should be driven by the question or questions you want to answer, not by the statistical test you want to perform. Statistical analysis is a tool to help you properly interpret collected data. It is not a magic wand to turn bad data into good answers. Hypothesis testing only tells you the probability of the difference between samples being real or being due to the natural variability of the populations sampled. It cannot tell you whether or not the statistical difference is ecologically significant. Finally, if you have to design a bioassessment study and you are not comfortable with statistical analysis, consult a biostatistician **before** you conduct your study, **not after**.

Table 7-2 presents ten principles in the design of an environmental study as set forth by Green (1979). These are the same principals, expressed in a slightly different format as have been covered in the preceding discussion of bioassessment study design.

Table 7-2. Principals of environmental study design from Green (1979).

1. Be able to state concisely to someone else what question you are asking. Your results will be as coherent and as comprehensible as your initial conception of the problem.
2. Take replicate samples within each combination of time, location, and any other controlled variable. Differences among can only be demonstrated by comparison to differences within.
3. Take an equal (preferably) number of randomly allocated replicate samples for each combination of controlled variables. Putting samples in "representative" or "typical" places is not random sampling.
4. To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent but all else is the same. An effect can only be demonstrated by comparison with a control.
5. Carry out some preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. Those who skip this step because they do not have enough time, usually end up losing time.
6. Verify that your sampling device or method is sampling the population you think you are sampling, and with equal and adequate efficiency over the entire range of sampling conditions to be encountered. Variation in efficiency of sampling from area to area biases among-area comparisons.
7. If the area to be sampled has a large-scale environmental pattern, break the area up into relatively homogeneous subareas and allocate samples to each in proportion to the size of the subarea. If it is an estimate of total abundance over the entire area that is desired, make the allocation proportional to the number of organisms in the subarea.
8. Verify that your sample unit size is appropriate to the size, densities, and spatial distributions of the organisms you are sampling. Then estimate the number of replicate samples required to obtain the precision you want.
9. Test your data to determine whether the error variation is homogeneous, normally distributed, and independent of the mean. If it is not, as will be the case for most field data, then (a) appropriately transform the data, (b) use a distribution-free (nonparametric) procedure, (c) use an appropriate sequential sampling design, or (d) test against simulated H_0 data.
10. Having chosen the best statistical method to test your hypothesis, stick with the result. An unexpected or undesired result is *not* a valid reason for rejecting the method and hunting for a "better" one.

REFERENCES

Eberhardt, L.L. and J.M. Thomas. 1991. Designing environmental field studies. Ecological Monographs 61(1). pp 53-73.

Elliott, J.M. 1977. Statistical analysis of samples of benthic invertebrates. Freshwater Biological Association Scientific Publication No.25. 2nd ed. Kendal, Titus Wilson & Son Ltd. 160 pp.

Gilbert. R.O. 1987. Statistical Methods for Environmental Pollution Monitoring. New York, Van Nostrand Reinhold Co. 330 pp.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. New York, J. Wiley & Sons. 257 pp.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs, 54(2), pp 187-211.

Simenstad, C.A., C.D. Tanner, R.M. Thom, and L.L. Conquest. 1991. Estuarine habitat assessment protocol. Seattle, U.S. Environmental Protection Agency Region 10, Office of Puget Sound. 212 pp.

Skellam, J.G. 1969. Models, inferences, and strategy. Biometrics 25. 457-475.

Sokal, R.R. and F.J. Rohlf. 1983. Biometry. 2nd ed. New York. W.H. Freeman and Co. 859 pp.

CHAPTER 8

SUMMARY OF RECOMMENDED BIOASSAY PROTOCOLS

INTRODUCTION

Summaries of recommended bioassay protocols and protocol documents are presented in this chapter to provide a brief overview of each test procedure. Table 8-1 presents the various protocols by media tested (soil, water, or sediment) and habitat of origin (freshwater, brackish water, or marine). As can be seen from Table 8-1, several of the protocols are applicable to more than one combination of media and habitat. Because of this multiple use of the various protocols, the arrangement of the summaries below into the categories of soil, water, and sediment is only approximate.

SOIL

Lettuce Seed (*Latuca sativa*) Germination Bioassay (Thomas and Cline, 1985).

The lettuce seed germination test is both rapid and economical and has been the preferred test to determine and evaluate phytotoxicity of contaminants in upland soils. However, seed germination is relatively insensitive to chemical insult because it is essentially a self-contained unit that doesn't need external nutrients for germination (Kaputska, personal communication). Growth after germination is far more sensitive to contamination. The lettuce seed germination and initial plant growth test protocol consist of taking 100 grams of air-dried test soil mixed with 50 grams of washed 16 mesh screened silica sand. Place the mixture in the bottom half of a 150-mm plastic petri dish and seeded with 40 lettuce (*Latuca sativa*) seeds. Place 35 milliliters of distilled water onto the soil-sand mixture and spread 50 grams of silica sand evenly over the dampened soil and seed. Place the petri dish ion the bottom of a 0.1-mm polyethylene bag expanded to contain air, closed at the top with a rubber band, and incubated in a controlled-environment chamber. Seed germination is determined visually by counting the shoot spikes when they extend 1 cm above the soil's surface. Stem lengths are determined by measuring each plant with a ruler after 10 days. In most cases, overall germination appears to be the most sensitive, but there are contaminants that will permit germination but do not allow shoot growth.

Table 8-1. Bioassay tests by habitat and media.

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Eisenia foetida [C. Oligochaeta-worm]	Common in U.S. Soils w/considerable organic matter; sewage beds, trickling filters	Porcella. 1983 (19)						7,14D=LT AT=	Survival	Cocoons commercially purchased; can be cultured in lab	Worms naturally occur in soils rich in organic materials. Worms may be protected from toxic chemicals due to adsorption of toxins to soil particles and organic matter (19).
Latuca sativa [Lettuce]		Thomas and Cline. 1985 (25)						10D=LT AT=	Germination		In most cases, overall germination is most sensitive, but there are contaminants that will still permit germination but do not allow shoot growth (25). Seed germination least sensitive stage of plant development (9).
WATER Fresh											
Ceriodaphnia dubia [O. Cladocera-water flea]	Abundant in lakes, ponds, and quiescent sections of streams and rivers throughout N. America	Horning, W. B. and C. I. Weber. 1985. (7)	Y	Y		Y	C	7D=LT AT=2-3wks	Survival/ Reproduction	Purchase brood stocks from Newton, OH. (22)	Elevated turbidity and dissolved substances problem (22); avoid sudden pH and temperature changes.
Daphnia magna [O. Cladocera-water flea]	Occur in nearly all types of freshwater habitats; lake & pond dweller, restricted to waters in N. & W N. America	Porcella. 1983. (19)	Y	Y		Y		48H=LT AT=1-1.5wk	Immobilization	Cultures available from commercial supply houses & from government labs	
Pimephales promelas [O. Cypriniformes-Fathead minnow]	Warm-water ponds, lakes, & sluggish streams	Horning, W. B. and C. I. Weber. 1985. (7)	Y	Y		Y	C	7D=-LT AT=2-3wks	Survival/ Growth	Acquire from Labs that is known to have disease-free fish; US EPA Environ. Research Lab @ Duluth, Minn.	
Selenastrum capricornutum	Ubiquitous in most freshwaters in N. America	Porcella. 1983. (19)	Y	Y				96H=LT AT=2-3wks	Growth (direct counting or mass)	Strains recommended are ATCC 22662, UTEX 1648, UTCC 37; commercial suppliers-American Type Culture Collection & Culture Collection of Algae	

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Brackish											
Daphnia magna [O. Cladocera-water flea]	Occur in nearly all types of freshwater habitats; lake & pond dweller, restricted to waters in N. & W N. America	Porcella. 1983. (19)	Y	Y		Y		48H=LT AT=1-1.5wk	Immobilization	Cultures available from commercial supply houses & from government labs	
Photobacterium phosphoreum [Luminescent bacterium]	Marine bacterium	Puget Sound Estuary Program. 1991. (20)	N	Y		Y	A	15,30min AT=2-3wks	Decreased luminescence	Purchase from commercial sources; Strain # NRRL B-11177 deposited w/ Northern Regional Research Laboratory in Peoria, Illinois, USA	Organic extraction method–specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Saline extraction method specific for water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from sediments (20).
Selenastrum capricornutum	Ubiquitous in most freshwaters in N. America	Porcella. 1983. (19)	Y	Y				96H=LT AT=2-3wks	Growth (direct counting or mass) EC_{90} , EC_{50} , SC_{20}	Strains recommended are ATCC 22662, UTEX 1648, UTCC 37; commercial suppliers-American Type Culture Collection & Culture Collection of Algae	Microscope counting recommended.(19)
Marine											
Crassostrea gigas *Mytilus edulis* [C. Pelecypoda-clam; mussel]	Puget Sound region; bays, inlets, etc. of the Pacific coast; oyster range-Morro Bay, CA to N. British Columbia; mussel range-Arctic, N. Pacific, N. Atlantic	Puget Sound Estuary Program. 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development	Field collection	Not recommended for sediments with interstitial salinity less than 10ppt *C. gigas* spawns during summer; M. *edulis* during the spring and summer; recovery of larvae from sediments difficult (20).
Photobacterium phosphoreum [Luminescent bacterium]	Marine bacterium	Puget Sound Estuary Program. 1991. (20)	N	Y		Y	A	15,30min AT=2-3wks	Decreased luminescence	Purchase from commercial sources; Strain # NRRL B-11177 deposited w/ Northern Regional Research Laboratory in Peoria, Illinois, USA	Organic extraction method specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Saline extraction method specific for water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from sediments (20).

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
SEDIMENT Fresh											
Chironomus riparius [O. Diptera--midge]	Distribution world-wide; thermophilous; standing water most spp.; eutrophic lakes, ponds, & streams; aquatic during larval & pupal stages	ASTM Designation E 1706-95b (3)	Y		Y		A	10D=LT AT=2-3wks	Survival and growth	Lab stock culture or commercial source	Ammonia accumulation (8). Avoid H_2S_2 if possible; over feeding leads to fungus contamination (3).
Chironomus tentans [O. Diptera--midge]	Holarctic distribution; mid-continental N. America; eutrophic lakes & ponds	ASTM Designation E 1706-95b (3) U.S. EPA 1994 Test Method 100.2 (26)	Y				A	10D=LT AT=2-3wks	Survival, growth and emergence	Lab stock culture or commercial source	Affected by H_2S_2 > 0.3mg/l (ASTM 1706-95b); over feeding leads to fungus contamination (3).
Daphnia magna [O. Cladocera-water flea]	Occur in nearly all types of freshwater habitats	Nebeker et al., 1984 (15)	Y					48H=LT AT=1-1.5wk 10D=LT AT=1-1.5wk	Survival Survival/ Reproduction	Purchase brood stocks-Newton, OH (5) and EPA Western Fish Tox. Station, Corvallis OR. (13)	
Hyalella azteca [O. Amphipoda]	Entire American continent; permanent meso or eutrophic lakes supporting aquatic plants, ponds, streams	ASTM Designation E 1706-95b (3) U.S. EPA 1994 Test Method 100.1 (26)	Y	Y	Y	Y	A C	10D=LT AT=2-3wks 30D=LT AT=2-3wks	Survival/# of young/Growth Reproduction/ Behavior/Sexual development.	Field: collected in freshwater lakes, ponds, and streams	Recovery of young difficult (3). Caution in feeding to avoid bacterial and fungal growth on sediments (3).
Photobacterium phosphoreum [Luminescent bacterium]	Marine bacterium	Puget Sound Estuary Program, 1991. (20)	N	Y		Y	A	15,30min AT=2-3wks	Decreased luminescence	Purchase from commercial sources; Strain # NRRL B-11177 deposited w/ Northern Regional Research Laboratory in Peoria, Illinois, USA	Organic extraction method specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Saline extraction method specific for water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from sediments. (20)

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Pimephales promelas [O. Cypriniformes-Fathead minnow]	Central part of the US, touches northern Mexico; not native west of the Rocky Mountains, nor the three most easterly Atlantic provinces	Horning, W. B , and C. I. Weber, 1985. (7)	Y	Y		Y	C	7D=-LT AT=2-3wks	Survival/ Growth	EPA Fish Tox. Lab Newton, OH (24); ponds, lakes, ditches, & slow muddy streams	Food added during tests may affect results; highly volatile/degradable toxins may not be detected (7); spawning in May or June in northern part of their range when average water temperatures are as low as 13 to 17°C
Brackish											
Ampelisca abdita [O. Amphipoda]	Range from Maine to south-central Florida & eastern Gulf of Mexico; low intertidal zone to 60 m; San Francisco Bay; euryhaline	ASTM Designation E 1367-92 (1)	Y	Y	Y	Y		10D=LT AT=2-3wks	Survival/ Growth	Field: East Coast; Gulf of Mexico; SF Bay. Commercial: John Brezina, San Rafael, CA; SAIC Narragansett, RI	Chronic tests best indicators of environmental impacts (6); Sensitive to coarse grain sediments; exposed to pore water also; flow-through tests recommended and very sensitive to O_2 stress (4); aeration recommended to break up surface contamination (29)
Crassostrea gigas *Mytilus edulis* [C. Pelecypoda-clam; mussel]	Puget Sound region; bays, inlets, etc. of the Pacific coast; oyster range-Morro Bay, CA to N. British Columbia; mussel range-Arctic, N. Pacific, N. Atlantic	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development	Purchase from commercial sources	Not recommended for sediments with interstitial salinity less than 10ppt *C. gigas* spawns during summer; *M. edulis* during the spring and summer; recovery of larvae from sediments difficult. Unknown high mortalities sometimes in controls (20).
Hyalella azteca [O. Amphipoda]	Entire American continent; permanent lakes, ponds, streams	ASTM Designation E 1706-95b (3)	Y	Y	Y	Y	A C	10D=LT AT=2-3wks 30D=LT AT=2-3wks	Survival/# of young/Growth Reproduction/ Behavior/ Sexual development	Field: collected in freshwater lakes, ponds, and streams	Feed cautiously to avoid bacterial and fungal growth on sediments (3).

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Leptocheirus plumulosus [O. Amphipoda]	Chesapeake Bay & other East Coast estuaries; found in both oligohaline & mesohaline portions; Cape Cod to N. Florida; fine sand to very fine mud	ASTM Designation E 1367-92 (1) Schlekat, C. E. et al. 1992. (23) McGee, B. L. et al. 1993. (14)	Y				A	10-d, 20-d, & 28 or 30-d LT AT =	Survival	Field collected-Chesapeake Bay, Chester River	Tests--static (10-d) & static renewal (20-28-d). It appears that this amphipod is a generalist; able to thrive in sediments differing in particle size and organic content. Absence of salinity-related effects on the survivorship of L. plumulosus allows for testing of sediments at ambient salinities. Two broods--early to midspring & in the fall; timing may vary. Limitations: •pH needs to be monitored/adjusted by adding spring water (average pH, hardness, & alkalinity are 7.7, 110 ppm as CaCO3, and 100 ppm as CaCO3 respectively) to distilled water at a ratio of 1:4 (v/v). •Effects on numbers of young per female indicate the potential for the development of endpoints more sensitive than adult mortality. •Temperature should be @ 20°C; potential effects on toxicity. •Juveniles should be feed during long-term exposures; potential effects on toxicity. •Sediment organic carbon needs investigation-may have supplemented the amphipod nourishment.
Photobacterium phosphoreum [Luminescent bacterium]	Marine bacterium	Puget Sound Estuary Program, 1991. (20)	N	Y		Y	A	15,30min AT=2-3wks	Decreased luminescence	Purchase from commercial sources; Strain # NRRL B-11177 deposited w/ Northern Regional Research Laboratory in Peoria, Illinois, USA	Organic extraction method specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Saline extraction method specific for water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from sediments. (20)
Rhepoxynius abronius [O. Amphipoda]	West coast N. America from central California to Puget Sound, WA; clean, fine, sandy sediments; lower intertidal to 274 m	ASTM Designation E 1367-92 (1)	Y	Y		Y		10D=LT AT=2-3wks	Survival/ Emergence/ Reburial	Field: West Coast, central California to Puget Sound	Fine sediments adversely affect; less than 10D tests not recommended (24); do not freeze sediments (18); aeration recommended to break up surface contamination (29). Emergence and reburial not very sensitive (13)

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Rhepoxynius abronius [O. Amphipoda]	Same as ASTM 1367; fine sands low intertidal to 60 m	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y		10D=LT AT=2-3wks	Survival/ Emergence/ Nonreburial	Field: West Beach, Whidbey Is., WA. Comm.: Yaquina Bay, OR (10)	Interstitial salinity @ • 25ppt; grain size affects negative @ extremes of fine and coarse material. Sensitivity varies with seasons (20). Emergence and reburial not very sensitive (13)
Crassostrea gigas *Mytilus edulis* [C. Pelecypoda-clam; mussel]	Puget Sound region; bays, inlets, etc. of the Pacific coast; oyster range-Morro Bay, CA to N. British Columbia; mussel range-Arctic, N. Pacific, N. Atlantic	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development	Field collection	Not recommended for sediments with interstitial salinity less than 10ppt *C. gigas* spawns during summer; *M. edulis* during the spring and summer; recovery of larvae from sediments difficult (20).
Dendraster excentricus [C. Echinoidea-sand dollar]	Puget Sound region; sandy beaches @ low tide	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development/ % fertilization	Field collected on Puget Sound beaches during low tide	Spawn from April to October; recovery of larvae from sediments difficult. Not recommended for sediments with < 10ppt; unknown high mortalities might occur in controls (20).
Strongylocentrotus drobaciensis [C. Echinoidea-green sea urchin]	Puget Sound region; shallow subtidal zone	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development/ % fertilization	Field collected on coast of Washington	Spawn from December to April; recovery of larvae from sediments difficult. Not recommended for sediments with interstitial salinity less than 10ppt ; unknown high mortalities might occur in controls (20).
Strongylocentrotus purpuratus [C. Echinoidea-purple sea urchin]	Puget Sound region; intertidal zone	Puget Sound Estuary Program, 1991. (20)	Y	Y		Y	A	48H=LT AT=3wks	Survival/ Abnormal larval development/ % fertilization	Field collected on coast of Washington	Spawn from December to April; recovery of larvae from sediments difficult. Not recommended for sediments with interstitial salinity less than 10ppt ; unknown high mortalities might occur in controls (20).
Eohaustorius estuarius [O. Amphipoda]	N. American west coast from British Columbia to central California; low tide zone to 10 m; fine sand	Puget Sound Estuary Program, 19 91. (20)	Y	Y	Y	Y		10D=LT AT=2-3wks	Survival/ Emergence/ Nonreburial	Field collected	Interstitial salinities from 2-28ppt. Do not freeze sediments (6); seasonal factors (20).

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Eohaustorius estuarius [O. Amphipoda]	N. American west coast from British Columbia to central California; low tide zone to 10 m; fine sand	ASTM Designation E 1367-92 (1)	Y	Y	Y	Y		10D=LT AT=2-3wks	Survival	Field collected BC (Canada) to Central California	Less sensitive than *R. abronius* (1).
Grandidierella japonica [O. Amphipoda]	San Francisco Bay & other N. Calif. bays; sand, silt, or clay sediments	ASTM Designation E 1367-92 (1)	Y	Y		Y		10D=LT AT=2-3wks	Survival/ Emergence/ Reburial	Field collected northern California bays, SF Bay, and southern California bays. (16)	Females carrying embryos should not be used in tests. Aeration recommended to break up surface contamination and abnormal swimming behavior affect results (18); grain size affected growth and molting affects survival results (16). Most reliable source, upper Newport Bay CA, is highly polluted reducing sensitivity of test (10)
Rhepoxynius abronius [O. Amphipoda]	West coast N. America from central California to Puget Sound, WA; clean, fine, sandy sediments; lower intertidal to 274 m	ASTM Designation E 1367-92 (1)	Y	Y		Y		10D=LT AT=2-3wks	Survival/ Emergence/ Reburial	Field: west coast, central California to Puget Sound	Fine sediments adversely affect; less than 10D tests not recommended (24); do not freeze sediments (18); aeration recommended to break up surface contamination (29). Emergence and reburial not very sensitive (12)
Rhepoxynius abronius [O. Amphipoda]	Same as ASTM 1367; fine sands low intertidal to 60 m	Puget Sound Estuary Program. 1991. (20)	Y	Y		Y		10D=LT AT=2-3wks	Survival/ Emergence/ Nonreburial	Field: West Beach, Whidbey Is., WA. Comm.: Yaquina Bay, OR (10)	Interstitial salinity @ • 25ppt; grain size affects negative @ extremes of fine and coarse material. Sensitivity varies with seasons (20) Emergence and reburial not very sensitive (12)
Photobacterium phosphoreum [Luminescent bacterium]	Marine bacterium	Puget Sound Estuary Program. 1991. (20)	N	Y		Y	A	15,30min AT=2-3wks	Decreased luminescence Decreased luminescence	Purchase from commercial sources; Strain # NRRL B-11177 deposited w/ Northern Regional Research Laboratory in Peoria, Illinois, USA	Organic extraction method specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Saline extraction method specific for water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from sediments. (20)
Neanthes arenaceodentata [C. Polychaeta-polychaete worm]	Marine; world-wide distribution; N. England, Florida, Calif. Europe, & central Pacific ocean; Conn. & Calif. populations; estuarine intertidal sand or muddy sand beaches, bays, harbors	Puget Sound Estuary Program. 1991. (20) ASTM Designation E 1611-94 (2)	Y Y	Y	Y	Y	A	20D=LT AT=2wks 10D=LT AT=	Survival/ Biomass & Avg. individual biomass/ Fecundity Survival (10-d)	California species from commercial source; Reish (21). Environ. Division, U.S. Army Corp of Engineers, Vicksburg, MI	Adjust for tests below salinities @ less than 20 ppt. Size of juveniles critical to test; do not freeze sediment (17) Survival and growth about equal sensitivity, reproductive fecundity (# and size of eggs in coelom) more sensitive (10) Chromosome numbers differ in populations of worms from California, Connecticut, and Hawaii--thus far, only the California worms have been used as test organisms; watch for avoidance to the sediments-replace worms if occurs.

Media Habitat / Test Organism Species	Indigenous Region for Species	Protocol References	Contaminant Sensitivity Mtl	Pest.	PCB	PAH	Acute vs Chronic	Lab time & analysis duration	Endpoints	Availability	Comments
SOILS Upland											
Neanthes virens [C. Polychaeta- polychaete worm-sand- or clamworm]	N. Atlantic Ocean from Arctic to Virginia, Iceland, Europe from British Isles, N. Sea, & France; shallow intertidal; fine sands, clay, peat & water soaked wood	ASTM Designation E 1611-94 (2)						10D=LT AT=	Survival	Field collection @ low tides from intertidal sand-mud flats	Worm also known as *Nereis virens* can not be cultured in lab because sexual maturity is only a short period of time (hours) & their life-cycle is long (1-2 years). The protocols state that the mortality endpoint for *N. virens* is rather insensitive and thus makes the organism suitable for bioaccumulation studies.
Oncorhyunchus mykiss [O. Salmoniformes-Rainbow trout]	Gonad cells of rainbow trout; steelhead range-N. Baja, Bering Sea, Japan; steelhead trout range-central Calif. to coastal streams of Alaska	Puget Sound Estuary Program. 1991. (20)						48H=LT AT=3wks	Anaphase aberration	Media for culture from commercial sources; Cells from trout gonads	Specific for neutral, nonionic organic compounds (PAH, PCB); not for metals and highly acidic or basic organic materials. Natural genotoxicity may influence results (20).

Key:

LT = lab time (time for the actual protocols to be completed)

AT = analysis time (time to complete the analysis and interpretation of data) (Times estimated by EVS, C. A. McPherson, Lab Supervisor 2/7/92)

H = hour(s)

D = day(s)

Min = minutes

Y = Yes, implying that test is sensitive to at least some contaminants in this category

N = No

S = Some

(#) indicates number of reference.

Earthworm (*Eisenia foetida*) Lethality Bioassay (Porcella, 1983).

The earthworm lethality bioassay is used to determine toxicity in upland soils. Earthworms were selected as an indicator species because they are representative of the terrestrial environment and are of considerable importance in improving soil aeration, drainage, and fertility. *Eisenia foetida* is a species that generally occurs in soils with very high organic content. It is also a species with a short life cycle and can breed readily in a wide range of organic wastes. This means that a standard strain can be used because laboratories could easily breed their own stock if supplied with cocoons from a central source. In this test, *Eisenia foetida* must be at least two months old, with a clitellum, and weigh 400 to 800 milligrams. Worm survival is evaluated after a 7- or 14-day exposure to a mixture (3:1) of an artificial soil and the test soil. Since soil is such a variable medium, the addition of a carefully defined artificial loam soil provides a suitable growth medium for the earthworm. For each test, 400 grams of this test mixture containing ten worms is placed in a 500-ml crystalizing dish and kept from dehydrating. Mortality is determined at the end of the test by emptying the soil into a tray, sorting out the worms and testing their reaction to mechanical anterior stimulus.

WATER COLUMN

Selenastrum capricornutum Algal Lethality Bioassay (Porcella, 1983).

The algal *Selenastrum capricornutum* bioassay is used to test the toxicity of freshwater samples. *S. capricornutum* is a unicellular, non-motile chlorophyte that is readily available and easily maintained. Unicellular algae are important producers of oxygen and form the basis of the food web in aquatic ecosystems. The alga *S. capricornutum* is added to test solutions containing various concentrations of the test material and growth is measured at 96 hours. Test material solutions are made up with the test material diluted to the proper concentration to which an algal assay growth medium of macro- and micro-nutrients is added. Between 0.1 and 1 ml of six- to eight-day old *S. capricornutum* culture stock is added to the test container to achieve a concentration of 10,000 cells per ml at the beginning of the test. The test mixture is incubated for 96 hours and algal growth is measured by any of the following methods: electronic particle counting, biomass (dry weight), absorbance, or direct microscopic counts. Because the algal test is designed to provide a comparative response to varying dilutions of sample, it is better to use absorbance or an electronic particle counter to measure growth. Results are expressed in terms of the EC_{50}. It is also recommended that

reference controls with a known toxic substance to check the sensitivity of the organisms be used in the test.

Water flea *Daphnia magna* EC_{50} Bioassay (Porcella, 1983).

The *Daphnia magna* EC_{50} bioassay is the preferred test to determine the toxicity of freshwater samples. Daphnids occur in nearly all types of freshwater habitats and have been recommended as a bioassay organism by the Committee on Methods for Toxicity Tests with Aquatic Organisms (ASTM, 1975; ASTM, 1980) because of their wide geographic distribution, important role in the aquatic food web, wide temperature tolerance, wide pH tolerance, ready availability, and ease of culture. The test consists of placing five *Daphnia magna* (early instar stages 2-4) in a 250-ml container with 200 ml of the test solution and incubating 48 hours at 19°C using a photoperiod of 16 hours light to 8 hours of dark. The test material concentration, which effectively influences 50 percent of the population (EC_{50}) within the 96 percent confidence limits, is determined after 24 and 48 hours. The EC_{50} is determined in this test because of the difficulty in ascertaining death (LC_{50}) for the *Daphnia*. The principal criterion of effect on *Daphnia* is immobilization, defined as lack of movement except for minor activity of appendages. Reference controls with a known toxic substance are also recommended to check the sensitivity of the organisms used in the test.

Ceriodaphnia dubia Survival and Reproduction Bioassay (EPA, 1985).

The *Ceriodaphnia* bioassay can be used to determine toxicity in freshwater environments. Traditionally used to measure the chronic toxicity of industrial whole effluents and receiving water, this test takes into account the synergistic, antagonistic, and additive effects of all the chemical, physical, and biological components that adversely affect the physiological and biochemical functions of the test organisms. The test requires that *Ceriodaphnia* be exposed in a static renewal system for seven days to different concentrations of test solution. Test results are based on survival and reproduction.

One *Ceriodaphnia* (2 to 24 hours old, all within 4 hours of the same age) is placed in each of 10 replicate 30 ml beakers containing 15 ml of test solution. The minimum of five test concentrations should be used. Selection of these test concentrations should be based upon the particular objectives of the study. Two commonly used dilution factors are 0.3 and 0.5. If it is suspected that *Ceriodaphnia* predators are present in the test solution, dilution water, or culture media, water should be filtered through a plankton net with 30 µm mesh

openings. During the test, the *Ceriodaphnia* are fed digested trout chow, yeast, or CEROPHYL^{R} at a rate of 0.1 ml food suspension per 15 ml of test solution.

In the absence of toxic substances, young production may exceed 30 per adult, 10 to 15 young are released every 36 to 48 hours. If toxic substances are present, young may develop in the brood pouch of adults, but may not be released during the exposure period. Adult survival and reproduction are recorded each day in the test chambers until the end of the test (seven days). Using several statistical techniques, the reproductive production and survival of *Ceriodaphnia* are evaluated with the various test concentrations and control to assess the chronic toxicity.

Fathead Minnow (*Pimephales promelas*) Larval Survival and Growth Bioassay (EPA, 1985).

The fathead minnow *Pimephales promelas* larval survival and growth bioassay can be used to determine toxicity in water and sediments in freshwater environments. Traditionally used as a test for estimating the chronic toxicity of industrial whole effluents and receiving water, this test takes into account the synergistic, antagonistic, and additive effects of all the chemical, physical, and biological components that adversely affect the physiological and biochemical functions of the test organisms. *P. promelas* larvae are exposed in a static renewal system for seven days to different concentrations of test solution. Test results are based on the survival and growth (increase in weight) of the larvae.

Ten *P. promelas* larvae (preferably less than 24 hours old) are placed into each of two replicate one-liter beakers containing 500 ml of test solution. A minimum of five different concentrations of the test solution plus a control is needed to conduct the bioassay. The selection of the test concentrations should be based upon the particular objectives of the study. Two commonly used dilution factors are 0.3 and 0.5. The fish larvae in each test chamber are fed 0.1 ml of newly hatched (less than 24 hours old) brine shrimp nauplii, three times daily at four-hour intervals. The numbers of live and dead larvae in each test chamber are recorded daily and dead larvae are discarded. At the end of seven days, the larvae are counted, preserved in 4 percent formalin and dried. Dry weight analysis is conducted to determine growth rates of larvae in each of the test solutions, and chronic toxicity is evaluated using several statistical methods.

SEDIMENTS

Daphnia magna Lethality and Partial Life Cycle Bioassay (Nebeker et al., 1984).

The *Daphnia magna* bioassay is the preferred protocol for determining toxicity in freshwater sediments. Daphnids occur in nearly all types of freshwater habitats and have been recommended as bioassay organisms by the Committee on Methods for Toxicity Tests with Aquatic Organisms (ASTM, 1980) because of their wide geographic distribution, important role in the aquatic food web, temperature requirements, wide pH tolerance, ready availability, and ease of culture.

The 48-h *D. magna* lethality test consists of exposing water fleas to 200 ml sediments in a 1000-ml beaker to which 800 ml of dilution water is added. Into each test beaker, 15 organisms are placed and incubated for 48 hours. After 48 hours, the water and fine suspended sediment in each beaker (not the bulk of the sediment) is poured through a 0.5-mm mesh screen to collect and count the live and dead organisms. Because death is not always easily determined with *D. magna*, a sublethal effect can be used as an endpoint. The principal criterion for effect on *D. magna* is immobilization, defined as lack of movement except for minor activity of appendages. Reference controls with known toxic substance are also recommended to check the sensitivity of the organisms used in the test.

The 10-day *D. magna* partial life-cycle test is conducted using 20, five-day old *D. magna* exposed in 2.5 liters of water to 500 ml of test sediment. Survival and reproductive success are assessed after 10 days (three broods) by screening the water and fine sediments to collect and count the live and dead adult and young *D. magna*.

ASTM Designation: E 1367 - 92. 1996. Standard Guide for Conducting 10-Day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods. (*Rhepoxynius abronius, Eohaustorius estuarius, Ampelisca abdita, Grandidierella japonica, Leptocheirus plumulosus*.)

General:

This guide describes procedures for obtaining laboratory data concerning the short-term adverse effects of contaminated sediment, or test material added to contaminated or uncontaminated sediment on marine or estuarine infaunal amphipods during static 10-day exposures. Procedures for static sediment toxicity tests are described for the following species: *Rhepoxynius abronius, Eohaustorius estuarius, Ampelisca abdita, Grandidierella japonica,*

and *Leptocheirus plumulosus*. Methods outlined in this document should be useful for toxicity testing of other aquatic taxa, and are applicable to sediments containing most chemicals, commercial products, and known or unknown mixtures. These methods can also be used to conduct bioconcentration tests and *in situ*, tests. LC_{50} or EC_{50} values may be determined. Sections of this guide include: general considerations (referenced documents, terminology, summary, significance and use, interferences, safety precautions), apparatus, toxicity test water, test and control sediments, test organisms, experimental design, procedure, analytical methodology, acceptability of test, calculation of results, documentation, tables and appendices.

Rhepoxynius abronius amphipod, emergence, mortality, and reburial bioassay:

The free-burrowing amphipod, *Rhepoxynius abronius*, is used to test sediments from marine areas. Found along the West Coast from central California to Puget Sound, Washington (available year round), it is the most sensitive species of amphipod, (the first to disappear from polluted areas and somewhat adversely affected by very fine-grained sediments). Adults or large immature young (3 to 5 mm mixed sexes) should be used in the tests, 20 per test chamber, exposed for 10 days, monitored each day for emergence, and then separated from the test substrates for analysis. Amphipods in sorting trays are counted as live or dead. Survivors are transferred to dishes and allowed one hour to rebury. Numbers of survivors unable to rebury in clean sediment can be used to calculate an EC_{50}.

Some considerations pertaining to this test include sediment grain size, salinity above 25 parts per thousand (ppt), size of test organisms, controls and sediments from non-toxic areas, and at least five laboratory replicates per test. Emergence from highly toxic sediment may occur, which should be observed and recorded (when the amphipod is completely or partially out of the sediment, on the sediment surface, swimming or floating). Mortality is determined at the end of the test. Total numbers of live and dead are recorded. Amphipods that are inactive but not obviously dead should be observed under a low-power microscope for neuromuscular pleopod twitch, with or without gentle prodding. Control survival is generally 95 percent or greater, and must be at least 90 percent for the toxicity test to be considered valid. Data on the ability of amphipods to rebury in clean sediment is used to detect sublethal effects. The numbers of organisms able to rebury within the time period specified should be recorded and an EC_{50} calculated.

Eohaustorius estuarius amphipod acute toxicity mortality bioassay:

The free-burrowing sand dweller, *Eohaustorius estuarius*, is distributed from British Columbia south to central California (available year round). Salinity tolerance is from 2 to 28 ppt. Twenty organisms per test chamber are used. The size of the animals should be 3- to 5-mm total length, with a caution not to use larger specimens, since they are senescent. A fine sediment control should be used if test sediments are predominantly silt or clays. *Eohaustorius* is only slightly less sensitive than *Rhepoxynius* to contaminants.

The major consideration when interpreting the results of acute toxicity tests is to keep in mind that the reproductive ability or long-term survival may be affected by contaminants at concentrations lower than those that produce a lethal or sublethal response. Control survival should be at least 90 percent.

Ampelisca abdita amphipod 10-day mortality growth and emergence bioassay:

This amphipod is an infaunal tube dweller inhabiting the low intertidal zone from central Maine to north Florida, the Gulf of Mexico, and San Francisco Bay. Salinity tolerance is classified as fully marine to 10 ppt, and is found in fine sand and mud to silt habitats, generally with high organic content. Each replicate should test 20 to 30 amphipods. The endpoint for the 10-day test is mortality, and dead animals (if it does not respond to gentle probing) should be counted and removed daily. Any animals not accounted for when the sieved material is examined are presumed to have died during the test. Other observations include animals out of their tubes, those that only exhibit a muscular pleopod twitch, the presence of molts, and the condition of the tubes built. Emergence from the sediment and the inability to construct a proper tube are sublethal behavioral responses that would ultimately result in death. For the growth test, small juveniles should be selected. Additional organisms should be preserved for later comparisons since this group represents the initial size. Growth is measured by length from the base of the first antennae to the base of the telson. Measurements are done after preservation and counting of test survivors.

Ampelisca is a good organism to test for PAHs, heavy metals, and PCBs. Because this amphipod is a particle feeder and will be exposed to contaminated particles in suspension, on the sediment surface, or through interstitial pore water, and since routes of exposure have not been fully examined, it is difficult to determine how toxicity occurs.

Grandidierella japonica amphipod emergence mortality bioassay:

Some northern and southern California bays and San Francisco Bay are the habitats for these tube-dwelling amphipods. They are marine, with salinity tolerances from 30 to 35 ppt. *Grandidierella japonica* lives in a variety of sediments (sands, silts, or clays) in the intertidal zone. Immature animals, 3- to 6-mm long should be used in the tests. Use 20 organisms per test chamber and allow them to bury. Those that do not bury should be replaced, unless they repeatedly burrow into the sediment and immediately emerge in an apparent avoidance response to the test substrate. In that case amphipods are not replaced. No females carrying embryos in their marsupium should be used. At the termination of the test, the reburial data can be used to determine an EC_{50} for a sublethal measurement. Surviving animals should be allowed one hour to rebury in clean collection site sediment.

When interpreting the data from acute toxicity tests, it should be kept in mind that the reproductive ability or long-term survival may be affected by contaminants at lower concentrations than those that produce a lethal or sublethal effect in a short-term test. *Grandidierella japonica's* ability to live in a burrow in a variety of sediment types gives broad application for its use in research and regulatory applications.

Leptocheirus plumulosus amphipod, emergence, mortality, and reburial bioassay:

The infaunal amphipod, *Leptocheirus plumulosus*, is indigenous to oligohaline and mesohaline regions thus it is used to test sediments from estuarine areas. The species builds U-shaped burrows and is found subtidally along the East Coast from central Cape Cod to Florida. *L. plumulosus* is considered to be as sensitive to contamination as *Hyalella azteca*. Adults or large juveniles (3 to 5 mm mixed sexes) should be used in the tests, 20 per test chamber, exposed for 10 days, monitored each day for emergence, and then separated from the test substrates for analysis. Amphipods in sorting trays are counted as live or dead. Survivors are transferred to dishes and allowed one hour to rebury. Numbers of survivors unable to rebury in clean sediment can be used to calculate a sublethal EC_{50}.

While *Leptocheirus plumulosus* is tolerant of a wide range of sediment grain sizes, a grain size reference using coarse sediments should be included among the controls. Test-water salinity should preferably match the interstitial water salinity of the test sediments but must be between 2 to 32 parts per thousand (ppt). Emergence from highly toxic sediment may occur, which should be observed and recorded (when the amphipod is completely or partially out of the sediment, on the sediment surface, swimming or floating). Mortality is

determined at the end of the test. Total numbers of live and dead are recorded. Amphipods that are inactive but not obviously dead should be observed under a low-power microscope for neuromuscular pleopod twitch, with or without gentle prodding. Control survival must be at least 90 percent for the toxicity test to be considered valid. Partial life cycle tests (28 to 30 days) have been conducted with *L. plumulosus* (Schlekat et al. 1992; McGee et al. 1993), and while no formal protocols exist for these tests, protocols are currently under development.

ASTM Designation: E 1706 - 95b (Replaces 1383-90). 1996. Standard test methods for measuring the toxicity of sediment-associated contaminants with fresh water invertebrates. (*Hyalella azteca, Chironomus tentans, Chironomus riparius*).

General:

This document covers procedures for obtaining laboratory data to evaluate adverse effects of contaminants associated with whole sediment on freshwater organisms. Another source for essentially the same procedures is the US EPA's Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates (U.S. EPA 1994). The methods are designed to assess the toxic effects on invertebrate survival, growth, or reproduction, for short-term (e.g., 10 days) or long-term tests, in static or flow-through water systems. Procedures are described for three species: *Hyalella azteca*, *Chironomus tentans*, and *Chironomus riparius*. With modifications, these procedures may be used for sediment tests with other aquatic species. These procedures are applicable to most sediments, chemicals, materials adhering to sediments, or interstitial water. They might also be used to conduct bioaccumulation tests. The results of these tests may be reported in terms of LC_{50}s, EC_{50}s, no observed ECs (NOEC), and low observed ECs (LOEC).

Sections elaborated on in this document include: a referenced document list, terminology, summary guide, significance and use, interferences, hazards, apparatus, overlying water, sediment characterization, test organisms, experimental design, procedures, analytical methodology, acceptability of tests, calculation, report, and appendices. The appendices section describes the organisms used for sediment toxicity testing.

For each species, the significance, life history and cycle, collection, brood stock, handling, age, acclimation, toxicity test specifications (experimental design, type of tests, initiation of tests, feeding, and biological data) sections are included and described.

Hyalella azteca amphipod survival, behavioral, growth, and reproductive bioassays:

Naturally collected (they occur in permanent lakes, ponds, and streams throughout the entire American continent) or laboratory raised *Hyalella azteca* can be used in a wide range of sediment testing (particle size tolerance ranging from >90 percent silt and clay to 100 percent sand-size particles). The life cycle is divided into three stages: immature (instars 1-5), juvenile (instars 6-7), and the adult stages. It is an epibenthic burrowing detritivore, exhibiting sexual dimorphism.

Tests should be started with organisms (second or third instar) about 2- to 3-mm long. Experimental designs (purpose, procedures, and calculations) dictate the number of treatments, test chambers, and amphipods per treatment. Nebeker et al. (1984) recommend two or more replicate aquaria per treatment with 100 organisms in each. Ingersoll and Nelson (1990) recommend four replicates per treatment with 20 organisms per replicate for a total of 80 amphipods per treatment. Duration can range from •10 days (short-term), continuing up to 30 days (long-term). Adult survival, the number of young, growth, and development can be used as endpoints. The 30-day tests can also measure reproductive capacity, behavior, and sexual development. Flow-through and static tests can be conducted. Feeding needs to be carefully monitored to avoid bacterial and fungal growth on sediments. Also, behavior should be monitored to check for floaters (sediment avoidance) and reproductive activities (amplexus). The amphipods are collected at the end of the test by screening methods; live amphipods should be counted or preserved for later examination using a low-power binocular microscope. Length of the body in millimeters (±0.01mm) from the base of the first antenna to the tip of the third uropod, or wet and dry measurements are used to quantify growth. An *H. azteca* sediment toxicity test is unacceptable if the average survival in any negative control chamber is less than 80 percent.

Chironomus tentans midge larval survival, growth, adult emergence, reproduction bioassays:

This is a fairly large midge with a short generation time, easily cultured in the laboratory. It also has holarctic distribution and is common in the mid-continental areas of North America. Brood stock can be obtained from the wild or a commercial source. The larvae of *Chironomus tentans* burrow into the sediment to build a case. They are collected

from lotic (streams or running water), and lentic (standing ponds) habitats. The life cycle can be divided into three distinct stages: larval (4 instars), pupal, and adult. Sexual dimorphism is observed when the adults hatch.

Tests with *Chironomus tentans* can be started with second instar larvae (10 days old). Experimental designs (purpose, procedures, and calculations) dictate the number of treatments, test chambers, and midges per treatment. Larval survival, growth (assessed at 10 to 14 days when the larvae have reached the third or fourth instar), or adult emergence can be monitored as biological endpoints in static and flow-through tests. Growth determinations using dry weight are preferable to length. Growth measurements can also be estimated from head capsule width, and also used to determine instar development. Nebeker et al. suggests conducting adult emergence tests for 25 days when started with second instar larvae. Emergence begins on day 20 and lasts for five days. Percent emergence is generally less than 60 percent in these tests. Endpoints calculated include (1) percent emergence, (2) mean emergence time, or (3) day to first emergence. Egg hatching studies may also be conducted using egg masses to estimate effects of exposure on either the number of eggs produced or hatched. A *C. tentans* sediment toxicity test, independent of test duration, is unacceptable if the average survival in any negative control chamber is less than 70 percent.

Chironomus riparius midge larval survival, growth, adult emergence, reproduction bioassays:

This is a fairly large midge with a short generation time, easily cultured in the laboratory. The larvae have direct contact with the sediment by burrowing and building a case. The distribution is worldwide (most species are thermophilic and have adapted to living in standing water, but some occur in cold habitats and in running water). The tubiculous larvae frequently inhabit eutrophic lakes, ponds, and streams with a variety of substrates. They tolerate a wide range of grain sizes from <90 percent silt-and-clay particles to 100 percent sand-size particles. The life cycle can be divided into three distinct stages: the larval (4 instars), the pupal, and the adult. Sexual dimorphism is observed when the adults hatch. Brood stock can be obtained from the wild or from a commercial source, or they can be raised in the laboratory.

Tests with *C. riparius* can be started with larvae less than 24 hours old or with 3-day-old larvae. Experimental designs (purpose, procedures, and calculations) dictate the number of treatments, test chambers, and amphipods per treatment. Ingersoll and Nelson (1990)

recommend using 50 specimens for flow-through testing. The duration of the tests can range from a •10-day test to >10 days and continuing up to 30 days. Larval survival, growth, or adult emergence can be monitored as biological endpoints in static or flow-through tests. Larval survival and growth can be assessed by ending the tests on day 10 to day 14 (3d or 4th instar). At this time, the larvae are removed from the sediment. Growth determination using dry weight is preferable to length. Growth and instar development can be determined by measuring head capsule width. Endpoints in emergence tests include: (1) percent emergence, (2) mean emergence time, or (3) day of first emergence. Egg hatching studies may also be conducted, and egg masses can be used to estimate effects of exposure on either the number of eggs produced or hatched. A *C. riparius* sediment toxicity test is unacceptable if the average survival in control chamber is less than 70 percent.

ASTM Designation: E 1611 - 94. 1996. Standard guide for conducting sediment toxicity tests with marine and estuarine polychaetous annelids. (*Neanthes arenaceodentata, Neanthes virens)*

General: This guide describes procedures for obtaining laboratory data concerning the short-term adverse effects of potentially contaminated sediment, or test material added to contaminated or uncontaminated sediment on marine or estuarine infaunal polychaetes during 10-day or 20- to 28-day exposures. Procedures for the 10-day static sediment toxicity tests are described for *Neanthes arenaceodentata* and *Neanthes virens*. Procedures also are described for the 20 to 28-day static-renewal sediment toxicity test for *N. arenaceodentata*. Methods outlined in this document could be used for toxicity testing with other aquatic infaunal taxa (e.g., other polychaetes, crustacea, bivalves), although modification of the procedures appropriate to the test species might be necessary. LC_{50} or EC_{50} values may also be determined. Sections of this guide include: general considerations (referenced documents, terminology, summary, significance and use, interferences, safety precautions), apparatus, test-water toxicity, test and control sediments, test organisms, experimental design, procedure, analytical methodology, acceptability of test, calculation of results, documentation, tables and appendices.

Neanthes arenaceodentata survival and growth toxicity tests: *N. arenaceodentata* is distributed widely throughout the world in estuarine intertidal sand or muddy sand beaches. It grows up to 10 cm in length, resides in mucoid burrows and may be found at concentrations as high as 1000 /m^2. Its ubiquitousness allows for easy collection and its size allows easy handling. It is readily cultured in the laboratory so, whether organisms are obtained through field sampling or laboratory culturing, they are always readily available

for testing. The 10-day static test has survival as the sole endpoint, while the 20- and 28-day tests have both survival and growth as endpoints. Mean control survival must be at least 90 percent and 80 percent or better in individual replicates. After the number of surviving worms are determined in the two longer tests, the worms from each replicate are placed in a clean petri dish, washed, then placed in clean pre-weighed aluminum pans, dried at 50° C to a constant weight and weighed to the nearest 0.1 mg.

Neanthes virens survival toxicity test: *N. virens* is distributed widely in the Atlantic and is found from the Arctic to Virginia in North America. It is usually found in intertidal and shallow subtidal waters associated with a wide variety of sediment types including coarse and fine sands, clay, peat, and water soaked wood. Specimens measure up to 90 cm long and 4.3 cm wide. No mention is made of laboratory culturing so specimens for testing need to be field collected. Tests are usually run in small aquaria with test organisms readily collected at the end of the test with a small net or by simply sieving the sediments. Mean control survival must be at least 90 percent and 80 percent or better in individual replicates. The protocols note that the mortality endpoint is relatively insensitive to contaminants. This insensitivity along with its large size makes *N. virens* desirable for bioaccumulation studies.

Protocols from Recommended Guidelines for Conducting Laboratory Bioassays on Puget Sound Sediments (Puget Sound Estuary Program, 1991)

Amphipod (*Rhepoxynius abronius* or *Eohaustorius estuarius*) sediment bioassay:

Sediment bioassays that test with the amphipods *R. abronius* or *E. estuarius* are used to characterize the toxicity of marine or estuarine sediments. This assay may be used alone as a screening tool in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. The endpoints include mortality, emergence, and nonreburial. For *R. abronius*, certain limitations should be noted including: an interstitial salinity requirement of • 25 ppt, poor survival at grain size extremes (clay at 50 percent and gravel at 35 percent), a temperature requirement of 15°C, and seasonal sensitivity. Seasonal sensitivity can be corrected for by the use of a positive control, e.g., $CdCl_2$. Predator removal is also an important requirement for running successful tests. The limitations for *E. estuarius* include the same parameters, but sediments with interstitial water salinity of 2 to 28 ppt may be used. Both species may be collected using benthic grabs or small dredges. Approximately one-third more *E. estuarius* than *R. abronius* are required for the bioassay. Mature amphipods are used in the sediment bioassay and should be acclimated to laboratory conditions for two to ten days before testing. For

each field sample, five replicate tests are conducted. Approximately 175 ml of test sediment are placed in the bottom of 1-liter test beakers. The beakers are filled to 750 ml with seawater (28 ppt for *R. abronius*, and ambient salinity for *E. estuarius*). Constant illumination is provided and the water in the beakers is aerated without disturbing the sediments. Twenty amphipods are placed in each beaker and the seawater level is brought up to 950 ml. The bioassay is terminated after 10 days of exposure. The primary endpoint is mortality (lack of pleopod twitch observed under magnification or response to gentle prodding) after 10 days exposure and the secondary endpoints of daily emergence and failure to rebury (within one hour) can also be measured.

A typical sediment bioassay involves 50 to 60 beakers. All bioassays include five replicates of the collection-site control sediment. These beakers comprise a negative (clean) control. Mean mortality in this control should be •20 percent for the test to be considered valid. A positive (contaminated) control is also required for all testing. This involves determining 96-hour LC_{50} values (four to five logarithmic concentration series and a control) for organisms exposed to a reference toxicant ($CdCl_2$) in clean, filtered seawater without sediment. The design of field surveys typically includes a reference sediment involving five replicate laboratory tests of samples from an area believed to be free from sediment contamination. The grain size and organic carbon content of the reference area sediment should be matched with the test sediment. This provides a site-specific basis for comparison of potentially toxic and nontoxic conditions while controlling the effects of exposing amphipods to non-native sediments.

Bivalve larvae sediment bioassay:

The bivalve larvae bioassay recommends Pacific oysters (*Crassostrea gigas*) and blue mussels (*Mytilus edulis*) for testing. The toxicity of marine sediments can be characterized by using these tests alone as a screening tool in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. During the first 48 hours of embryonic development, fertilized oyster and mussel eggs normally develop into a prodissoconch I, a free-swimming, fully shelled larval stage. Failure of the eggs to survive or the proportions of larvae developing in an abnormal manner have been used as the primary indicators of toxicity. A combined mortality and abnormality endpoint may also be calculated from the number of normal surviving larvae. Certain limitations should be noted when considering the use of this bioassay. Testing with bivalve larvae is not recommended for sediments that have an interstitial salinity of less than 10 ppt. This assay is primarily an indicator of the relative toxicity among different

samples. Spawning of *C. gigas* occurs naturally in the Puget Sound area in the summer. The natural spawning period for *M. edulis* is late spring to early summer. Both of these bivalves can be induced to spawn at other times of the year, but may show decreased viability of gametes. A positive control is recommended (48-hour exposure to a reference toxicant in seawater only). Other problems that may affect the results include the lack of evaluation of the bivalves' sensitivities to natural physical and chemical factors, which may influence occasionally high mortalities in control or reference tests. The recovery of live abnormal larvae from the sediments is also difficult.

The adult bivalves are induced to spawn. For each field sample, five replicate tests are conducted. Twenty grams (wet weight) of the appropriate sediment is added to each bottle and the volume is brought up to l liter with filtered or UV-treated seawater (28 ± 1 ppt salinity). The reference area sediment chambers each contain 20 grams of clean sediment. Negative and positive controls to determine LC_{50} and EC_{50} are also prepared consisting of clean seawater without sediment. The chambers are vigorously shaken for 10 seconds and allowed to settle for 4 hours. Each container is inoculated with 20,000 to 40,000 developing embryos. The containers are covered and incubated for 48 hours at 20 ± 1°C for oysters and 16 ± 1°C for mussels under a 14-hour light: 10-hour dark photoperiod. The test chambers are not aerated unless the dissolved oxygen concentration declines below 60 percent of saturation. The mean embryo concentration at 0 hour should be determined by collecting five replicate 10-ml samples from control cultures. This is not a direct measurement of the embryos in the test chambers and the resulting density estimates have an unquantified error component associated with them. This error reduces the reliability of larval mortality estimates and may thereby influence the results of statistical analyses. However, it does not affect larval abnormality estimates. When the embryos in the duplicate seawater control have reached the prodissoconch I stage (approximately 48 to 60 hours), the bioassay is terminated. The water and sediment are stirred, and 10-ml aliquots of samples are removed by pipette and preserved in vials containing five-percent buffered formalin.

Preserved samples (equal in volume to those containing 300 to 500 larvae in controls) are examined in Sedgewick-Rafter cells. Normal and abnormal larvae are enumerated to determine percentage survival and percentage abnormality. A minimum sample size of 20 living larvae in each of five replicated bioassay chambers for test sediment and reference area sediment, and 100 larvae in each replicated chamber for the seawater control should be scored for abnormalities. Percentage survival is based on the number of larvae surviving relative to the mean number of survivors in the seawater controls. Larvae that fail to

transform to the fully shelled, straight-hinged, D-shaped prodissoconch I stage are considered abnormal. Percent abnormal is based on the number of survivors that are abnormal. Five replicates of the seawater control are included in all bioassays. These comprise negative (clean) controls for comparisons among experiments and among laboratories (at least 70 percent of the larvae must survive the 48-hour exposure with seawater; of these, at least 90 percent must show no abnormalities). A positive control is also required ($CdCl_2$ or sodium dodecyl sulfate). Bioassays to establish an LC_{50} or an EC_{50} involve a concentration series (four to five logarithmic) and a control.

Echinoderm embryo sediment bioassay:

The echinoderm embryo bioassay is a rapid and sensitive technique for assessing the toxicity of marine sediments. Purple sea urchins (*Strongylocentrotus purpuratus*), green sea urchins (*S. droebachiensis*), and sand dollars (*Dendraster excentricus*) are the recommended species for testing. During the first 48 to 96 hours of development, fertilized echinoderm eggs normally develop into the pluteus stage. Failure of the eggs to survive and the proportions of larvae developing in an abnormal manner are used as indicators of toxicity. These tests may be used alone for screening in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. Certain limitations should be noted when considering the use of this bioassay. Quantitative results for corresponding endpoints may not be strictly comparable since the three species show slightly different levels of sensitivity to various contaminants. This bioassay is not recommended for sediments that have an interstitial salinity of less than 10 ppt. Echinoderm larvae normally reside in the water column and are not intimately associated with sediments. Therefore, this bioassay is primarily an indicator of the relative toxicity among different samples. *Strongylocentrotus* (spp.) spawning occurs naturally in the Puget Sound region from December to April. The natural spawning period for *D. excentricus* is from April to October. Echinoderms can be induced to spawn at other times of the year, but may show decreased viability of gametes. Care must be taken when transporting echinoderms due to epidemic spawning. A positive control is recommended (48-hour exposure to a reference toxicant in seawater only). Other problems that may affect the results include the lack of evaluation of the bivalves' sensitivities to natural physical and chemical factors, which may influence occasionally high mortalities in control or reference tests. The recovery of living abnormal larvae from the sediments is also difficult.

All echinoderm species can be collected off the coast of Washington. Sand dollars can be collected by hand on many Puget Sound beaches during low tide. Adults are conditioned to

the laboratory, and induced to spawn with chemical stimulation (1 ml of 0.5-molar KCl). Fertilization should be initiated within one hour of spawning by adding sperm to the beaker containing the eggs, at a sperm:egg ratio of approximately 2,000:1. Selected densities of embryos are exposed to the test or reference sediments for 48 to 96 hours. During this time, the embryos will normally develop into the four-armed pluteus stage. Data from tests with longer exposures (> 48 hours) may not be comparable to those from tests conducted using the standard 48-hour exposure. Toxicity test endpoints are based on abnormal shell development and larval death. Five replicate tests are conducted for each field sample. Twenty grams of reference or test sediment is added to each beaker. Filtered or UV-treated seawater (28 ppt salinity) is added to each beaker up to one liter to make a final concentration in all containers of 20 grams (wet weight) of sediment per liter of seawater. Each reference area sediment chamber also contains 20 grams per liter of clean sediment. In addition, two control series are prepared containing clean seawater without sediment (one series is used as a duplicate, sacrificial control to monitor embryo development). The sediments are vigorously shaken for ten seconds and allowed to settle for four hours. Within two hours of fertilization, a 1-ml aliquot of the solution of embryos (about 25,000) is added to each chamber by using an automatic pipette. The containers are covered and incubated for 48 hours at 15 ± 1°C under a 14-hour light: 10-hour dark photoperiod. The test chambers are not aerated unless the dissolved oxygen concentration declines below 60 percent of saturation. The mean embryo concentration at 0 hour should be determined by collecting five replicate 10-ml samples from control cultures. This is not a direct measurement of the embryos in the test chambers, and the resulting density estimates have an unquantified error component associated with them. This error reduces the reliability of larval mortality estimates and may thereby influence the results of statistical analyses. However, it does not affect larval abnormality estimates. When the embryos in the duplicate seawater control have reached the four-armed pluteus stage (approximately 48 to 96 hours), the bioassay is terminated. The water and sediment are stirred, and 10-ml aliquots of samples are removed by pipette and preserved in vials containing five-percent buffered formalin.

Preserved samples (equal in volume to those containing 300 to 500 larvae in controls) are examined in Sedgewick-Rafter cells. Normal and abnormal larvae are enumerated to determine percentage survival and percentage abnormality. A minimum sample size of 20 living larvae in each of five replicated bioassay chambers for test sediment and reference area sediment, and 100 larvae in each replicated chamber for the seawater control should be scored for abnormalities. Percentage survival is based on the number of larvae surviving

relative to the mean number of survivors in the seawater controls. Embryos that fail to transform to the four-armed pluteus stage are considered abnormal. Percentage abnormal for each replicate is based on the number of survivors that are abnormal. Five replicates of the seawater control are included in all bioassays. These comprise negative (clean) controls for comparisons among experiments and among laboratories (at least 70 percent of the larvae must survive the 48-hour exposure with seawater; and of these, at least 90 percent must show no abnormalities). A positive control is also required ($CdCl_2$ or sodium dodecyl sulfate). Bioassays to establish an LC_{50} or an EC_{50} involve a concentration series (four to five logarithmic) and a control.

Anaphase aberration sediment bioassay:

This sediment bioassay is used to characterize the genotoxicity of marine sediments. These tests may be used alone for screening in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. This assay can be used with any type of sediment regardless of the interstitial salinity or grain size characteristics. Rainbow trout gonad cells (RTG-2) are recommended (although any cell type can be used). Certain limitations should be noted when considering the use of this bioassay. The assay depends on a chemical extraction procedure that is specific for neutral, nonionic organic compounds (aromatic and chlorinated hydrocarbons). Other contaminants such as metals and highly acidic and basic organic materials are not efficiently extracted. Natural genotoxicity may occur in the marine sediments and may cause positive genotoxic responses. Extractions are prepared by evaporating a volume of the reagent equivalent to that which would be used for the actual extraction. This is added to the cultures in varying amounts dissolved in the solvent (e.g., dimethyl sulfoxide [DMSO]). Once the extracts have been made, gravimetric determinations of their absolute organic content must be made so that comparable organic concentrations from each site can be used in cell cultures exposed to extracts from different locations. RTG-2 cells grow in a variety of commercially available culture media. The Leibovitz L-15 medium was found to be the most consistent in terms of ease of preparation. Generally, heat-deactivated fetal calf serum is added to the culture medium at 10-percent concentrations to ensure proper growth factors are present.

The cells are grown and tested at 18°C on standard, clean microscope slides or on 1- by 5-cm coverslips in Leighton tubes. The cells are placed into the culture one day before the exposure to insure attachment to the substrate and to begin growing. After 18 to 24 hours, the culture medium is removed and the test material is added (consisting of the L-15

medium dissolved in DMSO to which the extract has been added). Exposure time should be 48 hours (maximum exposure) from the time of addition of the treated medium until fixation. An initial screening test must be conducted to determine the actual extract dilutions to be used for the bioassay. Ideally, dilutions tested for anaphase aberrations comprise the highest concentration of extract (ml/L) that permits continued cell proliferation (i.e., is nontoxic) and a second concentration one dilution lower. Previous experience in Puget Sound has shown that the following six extract dilutions should be prepared: 50, 25, 15, 5, 2, and 1 µg/ml. Cells are first exposed to these concentrations for each sediment tested, and then the concentrations that inhibit mitosis are determined. All results are normalized to organic content that has been previously determined.

To determine mitotic effects and anaphase aberrations, the slides or coverslips containing the cells are removed from the culture medium and fixed in methanol:acetic acid (3:1). Following 15 to 60 minutes in the fixative, the slides are air dried and placed in three-percent Gurr's R66 Geimsa stain for 15 to 30 minutes. The slides are observed with a microscope to determine optimum staining time. The staining is selective for the condensed chromosomes undergoing mitosis. Three replicate slides are made of each exposure concentration with two concentrations for each sediment extract. Each slide is then examined at 500X to 1,000X until a minimum of 100 anaphase cells is observed and scored. In this way, there will be three replicates per dose with 100 anaphase cells per replicate. The numbers and percents of normal and abnormal anaphase stages are recorded. Cells are scored as abnormal if they contain any of the described chromosomal lesions reported for this test (see protocols for list of references).

Controls consist of 1) untreated cultures (used as negative controls), 2) a solvent blank, and 3) a positive control consisting of several concentrations of known genotoxic agents (one should be 0.25 µg/ml benzo(a)pyrene resulting in an anaphase aberration frequency of 50 to 65 percent).

Microtox™ sediment bioassay–organic extract:

The Microtox organic extract bioassay is a rapid, sensitive method of toxicity testing based on light emission by the luminescent bacterium *Photobacterium phosphoreum* in the presence and absence of aqueous toxicants. The emitted light is a product of the bacterial electron transport system and thus directly reflects the metabolic state of the cells. Decreased luminescence provides a quantitative measure of toxicity. These tests may be

used alone for screening in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. This assay can be used with any type of sediment regardless of the interstitial salinity or grain size characteristics. Certain limitations should be noted when considering the use of this bioassay. The assay depends on a chemical extraction procedure that is specific for neutral, nonionic organic compounds (aromatic and chlorinated hydrocarbons). Other contaminants such as metals and highly acidic and basic organic materials are not efficiently extracted. Extraction by an organic solvent results in exposure of test organisms to concentrations much higher than those expected to occur in sediment interstitial water. There may also be naturally occurring toxic substances present in the marine sediments that might affect the results of the assay. Luminescence may increase rather than decrease. For now, these samples are considered nontoxic.

The bioassay is performed using a Microtox toxicity analyzer system (a temperature-regulated photometer equipped with a photomultiplier). Freeze-dried bacteria, reconstitution solution, dilutent, and other necessary materials can be purchased from commercial suppliers. The approach to testing organic extracts uses the basic Microtox method described in the *Microtox Operating Manual* (see protocols for references). The dried bacteria are reconstituted with water and placed in a Microtox cuvette (4°C). The dilution concentrations are 5.0, 0.5, and 0.05-percent extract (v/v). Each primary dilution is adjusted to two-percent NaCl. These diluted extracts are used in a range-finding assay to determine an appropriate primary dilution (should cause 65 to 90 percent decrease in bioluminescence in 15 minutes) for the definitive assay. For the definitive assay, two-fold serial dilutions (e.g., 6.0, 2.5, 1.25, and 0.625 percent in two-percent NaCl) are prepared along with a blank (to measure spontaneous decay of light). In each of 10 test cuvettes, a 10 µL aliquot of bacterial suspension added to 500 µL of dilutent and incubated for 15 minutes in the incubation wells. After 15 minutes, initial levels of light emission are measured in each of the 10 test cuvettes. At 30-second intervals, 500 µL aliquots of each concentration extract are added to two of the cuvettes (i.e., two replicates each of the four extract dilutions and the saline blank). Exactly five minutes after addition of the sediment extract, light emission is measured at 30-second intervals and in the same sequence used for extract additions. Light emission is measured again at 15 minutes; additional measurements are sometimes made at 30 minutes. An ethanol-only control is assayed using the same primary dilution sequence as the sediment test. This is used to adjust the sediment extract data for the contribution of the solvent vehicle.

Ethanol, sodium lauryl sulfate, or other suitable reference toxicants should be used as positive controls to assess daily bioassay performance and to determine differences in response among many bacteria. Clean sediment can be evaluated as a negative control. Bioassay repeatability is evaluated by duplicate testing (i.e., extraction and analysis) of 10 percent of the sediment extracts. Report the range-finding results, raw light emission data, and 15-minute EC_{50} values (at 95 percent CI).

Microtox™ bioassay—saline extract:

The Microtox saline extract bioassay is a rapid, sensitive method of toxicity testing based on light emission by the luminescent bacterium *Photobacterium phosphoreum*. The use of saline extracts of sediment for the Microtox bioassay has been described by Williams et al. (1986). The approach to testing saline extract uses the basic Microtox method described in the *Microtox Operating Manual* (see protocols for references). The major difference in the saline approach versus the organic approach is in the preparation of test samples. Each procedure is specific to the classes of contaminants tested for toxicity and, in general, the results for each approach can be viewed as complementary. The saline extracts Microtox bioassay procedure removes only the water-soluble fraction of sediment-adsorbed trace metals and organic pollutants from the sediments. Contaminants with extremely low water solubility (for example PCBs) will tend to be partitioned almost exclusively onto sediment particles and are unlikely to occur in high concentration in the saline extract. Certain limitations should be noted when considering the use of this bioassay. A correction factor needs to be established for changes in bacterial luminescence caused by variation among samples in sediment pore-water salinity. The 100-percent dilution specified in the protocol consists of 58-percent sediment and 42-percent Microtox dilutent, thereby limiting the sensitivity of the test. The use of a standardized dilution series limits the calculation of EC_{50} in some cases. A range-finding test could be conducted to correct this problem. The use of a saline extract may not mimic the actual pore water composition. Luminescence may increase rather than decrease. For now, these samples are considered nontoxic.

The bioassay is performed by hydrating a vial of freeze-dried bacteria with 1.0 ml of reconstitution solution. Serial dilutions are prepared at 100, 50, 25, 12, 5, and 0 percent of sediment supernatant Microtox dilutent. The 0 percent is a reagent blank needed to measure spontaneous decay in bacterial luminescence. In each of 10 test cuvettes, 10 µL of the bacterial suspension is added to 350 µL of dilutent and incubated for 15 minutes at 15°C. After 15 minutes, initial luminescence in each of the 10 cuvettes is measured. At 30-second intervals, 500 µL aliquots of each supernatant dilution are added to two of the cuvettes (e.g.,

two replicates each of the four test dilutions and the saline blank). Exactly 15 minutes after addition of the sediment supernatants, luminescence is measured at 30-second intervals and in the same sequence used for supernatant additions. The percentage decrease in luminescence relative to the reagent blank is then calculated.

Clean reference sediments are used as negative controls. A calibration curve to determine salinity-induced changes in bacterial luminescence is constructed. The use of reference toxicants (e.g., phenol, sodium arsenate) is needed to assess day-to-day performance and determine the differences in toxic response among lot number of bacterial. Percentage decrease in luminescence after 15-minute exposure for each concentration of supernatant tested is reported.

Juvenile polychaete sediment bioassay: This protocol is for conducting a bioassay in which the survival and change in biomass of juvenile polychaetes (*Neanthes* spp.) are determined following a 20-day exposure to test sediments. Parameters measured to determine the effects of exposure include mortality, total biomass, and average individual biomass. Sediments can be naturally occurring (field collected), or sediments that have been experimentally modified by adding chemicals to the samples. This bioassay is conducted as a static renewal exposure, and food is provided to the test organisms during the exposure period to promote body tissue increase. Following the 20-day exposure period, all surviving worms are collected, dried to a constant weight, and total and average individual biomass is determined. This bioassay recommends the use of the California species of *Neanthes*. These worms are sensitive to metals, hydrocarbons, and multicontaminated media (e.g., sediments). *Neanthes* has not been collected from Puget Sound, but are distributed on the West Coast from Mexico to southern California. These tests may be used alone for screening in broad-scale sediment surveys, in combination with sediment chemistry and *in situ* biological indices, and in laboratory experiments. Certain limitations should be noted when considering the use of this bioassay. The bioassay should be conducted with laboratory-cultured juvenile *Neanthes*. Modification of the protocol may be required for tests conducted at salinities less than 20 ppt (interstitial and overlying water). In addition to exposure chambers containing test sediments, exposure chambers containing control sediment (sediment from where the worms were found or the substrate in which they were cultured) should also be tested. Reference sediments are also prepared (to observe the effects associated with physical and chemical characteristics) that should be similar to the test sediments in grain size and organic content. Sand is typically chosen as an appropriate control sediment. Bioassay seawater should be maintained at a salinity of 28

± 2 ppt at a temperature of 20 ± 1°C. The chambers are 1-liter glass containers (diameter of 10 cm) covered with lids. The gentle aeration rate should be at 150 and 300 ml per minute.

A typical *Neanthes* bioassay for testing 10 sediment samples involves 50 to 60 exposure chambers. Five replicates are tested for each field sample. Each exposure chamber consists of a 1-liter jar containing 2 cm of sediment and seawater. The photoperiod during the testing should be continuous. Worms are collected from the holding tanks and placed in cups, five worms to a cup. Enough cups are used to equal three more than the number of exposure chambers. Worms from these cups are used to estimate initial total biomass. Five juvenile worms are randomly placed into each exposure chamber. The size of the worms used in the assay is a critical factor. They should be 0.5 to 1.0 mg dry weight, two to three weeks post-emergence to ensure that they are in a rapid growth phase during the exposure period. They are fed on an every-other-day basis. Seawater is exchanged on every third day. Following the exposure period, the contents of each replicate chamber are sieved and the number of living worms is recorded. The worms are then weighed to the nearest 0.1 mg dry weight.

A control sediment and a reference sediment should be included as part of every test. A positive (toxic) control is also required. This involves determining 96-hour LC_{50} values for *Neanthes* juveniles exposed in clean, filtered seawater without sediment to reference toxicants ($CdCl_2$). The positive control should be conducted with 10 juveniles per exposure chamber. Reporting must include acute lethality, survival, total biomass (dry weight), and average individual biomass (i.e., total biomass divided by the number of surviving worms). Each of these response criteria should be monitored in a "blind" fashion. This means the observer must have no knowledge of the treatment of the sediment in the beakers.

TEST SELECTION

One of the principle concerns in deciding which toxicity test to use for bioassessment is the sensitivity of the test organism and test endpoint(s). Table 8-2 is a brief summary of the relative sensitivity of various test organisms and endpoints; it is based on the review of some forty articles on comparative toxicity tests using different organisms and/or endpoints. The media used in the various studies included both water and sediment from fresh, brackish and marine environment (see Appendix C for a more complete table broken down by study and the list of articles reviewed). As Table 8-2 indicates, of those test organisms involved in more than one test series, no single organism or endpoint was consistently the most sensitive. For example, *Chironomus tentans* survival was sometimes

more sensitive, sometimes less sensitive and sometimes had the same sensitivity as *Hyalella azteca* survival. This variable sensitivity may be due to either the contaminants or some other parameter of the test medium. The bottom line is that no single toxicity test should be relied on to evaluate conditions at a particular site.

Table 8-2. Relative sensitivities of selected test organisms and endpoints. Unless otherwise specified, endpoints for the compared species are the same as listed for them as a test species. When results are based on specific contaminants or groups of contaminants, the contaminants are indicated in parentheses (mtls = metals, org. = organics).

More Sensitive Species		Test Species (endpoints) Species of Equal Sensitivity		Less Sensitive Species
Mytilus edulis *Rhepoxynius abronius* *Strongylocentrotus purpuratus*	>	***Ampelisca abdita*** **(survival)**	>	*Eohaustorius estuarius* *Leptocheirus plumulosus* *Strongylocentrotus purpuratus*
	>	***Armandia brevis*** **(growth)** *Rhepoxynius abronius*	>	*Dendraster excentricus*
	>	***Ceriodaphnia dubia*** **(survival)**	>	*Daphnia magna* *Hyalella azteca*
	>	***Chironomus riparius*** **(emergence)** *Hyalella azteca*	>	
Hyalella azteca *Lumbriculus variegatus* (Cu)	>	***Chironomus tentans*** **(survival)**	>	*Lumbriculus variegatus* *Hyalella azteca* (dieldrin & chloropyrifos)
Rhepoxynius abronius *Photobacterium phosphoreum* *Neanthes* spp. (growth)	>	***Corophium volutator*** **(survival)**	>	*Macoma balthica* *Pseudopleuronectes americanus*
Rhepoxynius abronius *Photobacterium phosphoreum* *Mytilus galloprovincialis* *Dendraster excentricus* (abn. devel.)	>	***Crassostrea gigas*** **(survival/abn development)** *Dendraster excentricus* (survival)	>	Neanthes spp. (survival)
Photobacterium phosphoreum (org.) *Ceriodaphnia dubia* *Daphnia pulex* *Microcystis aeruginosa* *Pimephales promelas* (org.)	>	***Daphnia magna*** **(survival)**	>	*Photobacterium phosphoreum* (mtls) *Hyalella azteca* *Oncorhynchus mykiss* *Chironomus tentans* (Cu) *Gammarus lacustris* (Cu) *Pimephales promelas* (mtls)

More Sensitive Species		Test Species (endpoints) Species of Equal Sensitivity		Less Sensitive Species
	>	***Daphnia pulex*** **(survival)**	>	*Photobacterium phosphoreum* (mtls) *Oncorhynchus mykiss* *Pimephales promelas* (Mtls) *Daphnia magna* (Mtls)
Armandia brevis *Rhepoxynius abronius*	>	***Dendraster excentricus*** **(growth)**	>	
Rhepoxynius abronius *Dendraster excentricus* (abn. devel.)	>	***Dendraster excentricus*** **(survival)** *Crassostrea gigas* (abn. devel.)	>	*Neanthes* spp. (survival)
	>	***Dendraster excentricus*** **(abnormal development)** Rhepoxynius abronius	>	*Dendraster excentricus* (survival) *Neanthes* spp. (survival) *Crassostrea gigas* (abn. devel.)
Ampelisca abdita *Leptocheirus plumulosus*	>	***Eohaustorius estuarius*** **(survival)** *Rhepoxynius abronius* *Photobacterium phosphoreum* *Leptocheirus plumulosus*	>	*Ampelisca abdita* *Neanthes* (spp.) (Surv/biomass)
Hyalella azteca *Chironomus tentans* *Daphnia magna*	>	***Gammarus lacustris*** **(Cu) (survival)**	>	
Leptocheirus plumulosus (surv/grwth) *Eohaustorius estuarius* *Daphnia magna* *Ceriodaphnia dubia* *Gammarus lacustris* (Cu)	>	***Hyalella azteca*** **(survival)** *Chironomus riparius* (emergence)	>	*Chironomus tentans* (surv/grwth) *Lumbriculus variegatus* *Ampelisca abdita*
Ampelisca abdita *Hyalella azteca*	>	***Leptocheirus plumulosus*** **(survival)** *Eohaustorius estuarius*	>	*Leptocheirus plumulosus* (growth) *Hyalella azteca*
Hyalella azteca *Chironomus tentans*	>	***Lumbriculus variegatus*** **(surv./growth/reproduction)**	>	*Chironomus tentans* (Cu)
Photobacterium phosphoreum *Rhepoxynius abronius* *Corophium volutator* *Neanthes* spp. (growth)	>	***Macoma balthica*** **(bioaccumulation)**	>	*Pseudopleuronectes americanus*
Rhepoxynius abronius	>	***Mytilus edulis*** **(embryo surv./development)** *Rhepoxynius abronius*	>	*Ampelisca abdita* *Strongylocentrotus purpuratus*

More Sensitive Species		**Test Species (endpoints)** Species of Equal Sensitivity		Less Sensitive Species
Rhepoxynius abronius	>	***Neanthes* spp. (growth)**	>	*Rhepoxynius abronius* *Macoma balthica* *Pseudopleuronectes americanus* *Corophium volutator*
Rhepoxynius abronius *Photobacterium phosphoreum* *Eohaustorius estuarius* *Crassostrea gigas* *Dendraster excentricus*	>	***Neanthes* spp. (survival)**	>	
Daphnia magna *Daphnia pulex* *Photobacterium phosphoreum*	>	***Oncorhynchus mykiss* (survival)**	>	*Photobacterium phosphoreum* (mtls) *Pimephales promelas*
Pimephales promelas (mtls) *Daphnia magna* (mtls) *Oncorhynchus mykiss* (mtls) *Daphnia pulex* (mtls)	>	***Photobacterium phosphoreum*/Microtox (bioluminesence)** ***Rhepoxynius* abronius (mtls/PAHs)** Eohaustorius estuarius (mtls/PAHs)	>	*Daphnia magna* *Rhepoxynius abronius* *Oncorhynchus mykiss* *Crassostrea gigas* *Neanthes* spp. *Pimephales promelas* *Macoma balthica* *Pseudopleuronectes americanus* *Corophium volutator*
Daphnia magna *Daphnia pulex* *Oncorhynchus mykiss*	>	***Pimephales promelas* (survival)**	>	*Photobacterium phosphoreum* *Daphnia magna* (effluent)
Corophium volutator *Neanthes* (spp.) (growth) *Macoma balthica* *Photobacterium phosphoreum* *Rhepoxynius abronius*	>	***Pseudopleuronectes americanus* (histopathology)**	>	
Neanthes spp. (growth) *Photobacterium phosphoreum*	>	***Rhepoxynius abronius* (survival)** *Dendraster excentricus* (abn. devel.) *Armandia brevis* (growth) *Eohaustorius estuarius* *Mytilus edulis* (abn. devel.) *Photobacterium phosphoreum*	>	*Crassostrea gigas* *Dendraster excentricus* (surv/growth) *Neanthes* spp. (surv/growth) *Corophium volutator* *Ampelisca abdita* *Photobacterium phosphoreum* (mtls/PAHs) *Pseudopleuronectes americanus* *Strongylocentrotus purpuratus* *Macoma balthica* *Mytilus edulis* (survival)
Ampelisca abdita *Mytilus edulis* *Rhepoxynius abronius*	>	***Strongylocentrotus purpuratus* (cell changes/fertilization)**	>	*Ampelisca abdita*

Sensitivity is not the only, and may not be the most important, criteria on which test selection should be based. Long et al. (1990) suggest four additional criteria on which to judge the usefulness of toxicity tests. They are: (1) replicate variability/analytical precision, how consistent are the results in replicates of the same treatment; (2) discriminatory power, an indication of the range of responses with respect to the range of contamination levels; (3) correlation with other tests and endpoints, and (4) correlation between toxicity and chemistry. As an example of how these criteria may be applied, Long (1997) has found in his studies over the years that while *Rhepoxynius abronius* survival is more sensitive than *Ampelisca abdita* survival, *A. abdita* has better precision and is more readily correlated with sediment chemistry than is *R. abronius* . Therefore, he prefers using *A. abdita* survival tests.

When it was decided to produce a manual of protocols for freshwater toxicity and bioaccumulation tests for sediment associated contaminants, the US EPA developed a list of criteria for toxicity tests that they used to select organisms and endpoints for test protocol development (US EPA 1994). Table 8-3, taken from the resulting manual, "Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates",(US EPA 1994) lists these criteria and their assigned ratings. Some of the additional criteria utilized in selecting the test organisms and endpoints for their manual included ecological relevance as well as practical concerns regarding the conductance and acceptability of the test. The test organisms selected were *Chironomus tentans* and *Hyalella azteca* for toxicity test protocols and *Lumbriculus variegatus* for bioaccumulation protocols.

Table 8-3 Rating of selection criteria for freshwater sediment toxicity testing organisms. (US EPA, 1994)

Criterion	Hyalella azteca	Diporeia spp.	Chironomus tentans	Chironomus riparius	Lumbriculus variegatus	Tubifex tubifex	Hexagenia spp.	Mollusks	Daphnia spp. & Ceriodaphnia spp.
Relative sensitivity toxicity database	+	-	+	-	+	-	-	-	-
Round-robin studies conducted	+	-	+	-	-	-	-	-	-
Contact with sediment	+	+	+	+	+	+	+	+	-
Laboratory culture	+	-	+	+	+	+	-	-	+
Taxonomic identification	+/-	+/-	+/-	+/-	+	+	+	+	+
Ecological importance	+	+	+	+	+	+	+	+	+
Geographical distribution	+	+/-	+	+	+	+	+	+	+/-
Sediment physico-chemical tolerance	+	+	+/-	+	+	+	-	+	NA
Response confirmed with benthos populations	+	+	+	+	+	+	+	-	+
Peer reviewed	+	+	+	+	+	+	+	-	+/-
Endpoints monitored	S, G, M	S, B, A	S, G, E	S, G, E	B, S, R	S, R	S, G	B	S, G, R

A "+" or a "-" rating indicates a positive or negative attribute.

S = Survival, G = Growth, B = Bioaccumulation, A = Avoidance, R = Reproduction, M = Maturation, E = Emergence, NA = not applicable.

The selection of a toxicity test, or some other bioassessment tool, should never be based on a single criterion, for example, sensitivity or the desire to use an indigenous test organism. The test(s) that will provide the most reliable, interpretable, relevant data, for the particular job at hand, should be selected, and this may not necessarily be the most sensitive test or utilize an indigenous organism. Finally, toxicity test were never meant as stand-alone tools, they should be conducted in concordance with other bioassessment tools and the physical and chemical analysis of the medium being tested.

REFERENCES

1[1] ASTM. 1996a. Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. E 1367-92, 1996 Annual Book of Standards, Section 11 Water and Environmental Technology, Vol. 11.05, Biological Effects and Environmental Fate; Biotechnology; Pesticides. Philadelphia: ASTM, pp 769-794.

2 ASTM. 1996b. Standard guide for conducting sediment toxicity tests with marine and estuarine polychaetous annelids. E 1611-94. 1996 Annual Book of Standards, Section 11 Water and Environmental Technology, Vol. 11.05, Biological Effects and Environmental Fate; Biotechnology; Pesticides. Philadelphia: ASTM, pp 1046-1069.

3 ASTM. 1996c. Standard test methods for measuring the toxicity of sediment-associated contaminants with fresh water invertebrates. E 1706-95b, 1996 Annual Book of Standards, Section 11 Water and Environmental Technology, Vol. 11.05, Biological Effects and Environmental Fate; Biotechnology; Pesticides. Philadelphia: ASTM, pp 1176-1258

4 Breteler, R. J., K. J. Scott, and S. P. Shepherd. 1989. Application of a new sediment toxicity test using the marine amphipod *Ampelisca abdita* to San Francisco Bay sediments. In: U. M Cowgill and L. R. Williams (eds). *Aquatic Toxicology and Hazard Assessment: 12th Volume*. ASTM STP 1027, Philadelphia: American Society for Testing and Materials pp. 304-314.

5 Cairns, M. A., A. V. Nebeker, J. H. Gakstatter, and W. L. Griffis. 1984. Toxicity of copper-spiked sediments to freshwater invertebrates. *Environmental Toxicology and Chemistry 3*: 435-445.

6 Gentile, J. H., K. H. Scott, S. M. Lussier, and M. S. Redmond. 1987. The assessment of Black Rock Harbor dredged material impacts on laboratory population responses. Report D-87-3. Prepared for: Department of the U.S. Army Corps of Engineers, Washington, D.C. and U.S. EPA, Washington D.C. by: Vicksburg, Mississippi: U.S. Army Engineer Waterways Experiment Station. 3 pp + Appendix

[1] Numbers correspond to reference numbers used in Table 8-1.

7 Horning. W. B., and C. I. Weber, (eds.). 1985. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. Technical Support Document for Water Quality-based Toxics Control. EPA 440/4-85-032. Washington, D.C.: U.S. Environmental Protection Agency, pp 28-41.

8 Ingersoll, C. G. and M. K. Nelson. 1990. Testing sediment toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In: Landis, W. G. and W. H. van der Schalie (eds.). *Aquatic Toxicology and Risk Assessment: Thirteenth Volume*. Philadelphia: ASTM STP 1096, pp. 93-109.

9 Kaputska, L. 1992. Personal communication. Ecological Planning and Toxicology, Inc. Corvallis, Oregon.

10 Long, E.R. 1992. Personal communication. Coastal Monitoring and Bioeffects Division, Office of Ocean Resources Conservation and Assessment, Seattle: National Oceanic and Atmospheric Administration.

11 Long, E.R. 1997. Personal communication. Coastal Monitoring and Bioeffects Division, Office of Ocean Resources Conservation and Assessment, Seattle: National Oceanic and Atmospheric Administration.

12 Long, E. R., M. F. Buchman, S. M. Bay, R. J. Breteler, R. S. Carr, P. M. Chapman, J. E. Hose, A. L. Lissner, J. Scott, and D. A. Wolfe. 1990. Comparative evaluation of five toxicity tests with sediment from San Francisco Bay and Tomales Bay, California. *Environmental Toxicology and Chemistry 9*(9). 1193-1214.

13 Malueg, K. W., G. S. Schuytema, D. F. Krawczyk, and J. H. Gakstatter. 1984. Laboratory sediment toxicity tests, sediment chemistry and distribution of benthos macroinvertebrates in sediments from the Keweenaw Waterway, Michigan. *Environmental Toxicology and Chemistry 3:* 233-242.

14 McGee, B.L., C. E. Schlekat and E. Reinharz. 1993. Assessing sublethal levels of sediment contamination using the estuarine amphipod Leptocheirus plumulosus. *Environmental Toxicology and Chemistry 12*:??. pp 577-587.

15 Nebeker, A. V., M. A. Cairns, J. H. Gakstatter, K. W. Malueg, G. S. Schuytema, and D. F. Krawczk. 1984. Biological methods for determining toxicity of contaminated sediments to freshwater invertebrates. *Environmental Toxicology and Chemistry 3:* 617-630.

16 Nipper, M. G., D. J. Greenstein, and S. M. Bay. 1989. Short- and long-term sediment toxicity test methods with the amphipod *Grandidierella japonica*. *Environmental Toxicology and Chemistry 8*: 1191-1989.

17 Oakden, J. M., J. S. Oliver and A. R. Flegal. 1984. EDTA chelation and zinc antagonism with cadmium in sediment: Effects on the behavior and mortality of two infaunal amphipods. *Marine Biology 84*: 125-130.

18 Pastorok, R. A. and D. S. Becker. 1989. Comparison of bioassays for assessing sediment toxicity in Puget Sound. Bellevue, Washington: PTI Environmental Services, For U.S. EPA Region 10 Office of Puget Sound. Executive Summary + pp. 1-85. Tables D-1 to D-6.

19 Porcella, D. B. 1983. Protocol for bioassessment of hazardous waste sites. EPA 600/2-83-054. Corvallis, Oregon: U.S. Environmental Protection Agency, pp A-29–A-38.

20 Puget Sound Estuary Program. 1991. Interim Final. Recommended guidelines for conducting laboratory bioassays on Puget Sound sediment. Prepared for: U.S. EPA Region 10, Seattle, July 1991. 81 pp.

21 Rogerson. P. F., S. C. Schimmel, and G. Hoffman. 1985. Chemical and biological characterization of Black Rock Harbor dredged material. Technical Report D-85-9. Field Verification Program (Aquatic Disposal) Final Report. Washington, D.C.: Prepared for Department of the Army, U.S. COE, and U.S. EPA.

22 Sasson-Brickson, G. and G. A. Burton, Jr. 1991. *In situ* and laboratory sediment toxicity testing with *Ceriodaphnia dubia*. *Environmental Toxicology and Chemistry Vol. 10:* 201-207.

23 Schlekat, C.E., B.L. McGee & E. Reinharz. 1992. Testing sediment toxicity in Chesapeake Bay with the amphipod Leptocheirus plumulosus: an evaluation. *Environmental Toxicology & Chemistry 11*:2. 225-236.

24 Swartz, R. C., W. A. DeBen, J. K. P. Jones, O. J. Lamberson, and F. A. Cole. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. ASTM STP 854: R. D. Cardwell, R. Purdy, and R. Comotto-Bahner (eds.). *Aquatic Toxicology and Hazardous Assessment.* Proceedings of the Seventh Annual Symposium. Milwaukee, Wisconsin, April 17-19, 1983, Baltimore: American Society for Testing and Materials. pp 284-307.

25 Thomas, J. M. and J. F. Cline. 1985. Modification of the Neubaurer technique to assess toxicity of hazardous chemicals in soils. *Environmental Toxicology and Chemistry 4*: 201-207.

26 U.S. Environmental Protection Agency. 1994. Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates. EPA/600/R-94/024. Duluth, Minnesota: Office of Research and Development. 133 pp.

27 Westerman, A. 1989. Yearly chronic toxicity comparisons of stream sediments and waters. In: G. W. Suter II and M. A. Lewis (eds) *Aquatic Toxicology and Environmental Fate: Eleventh Volume*, ASTM STP 1007. Philadelphia: American Society for Testing and Materials. pp. 204-214.

28 Williams, L. G., P. M. Chapman, and T. C. Ginn. 1986. A comparative evaluation of marine sediment toxicity using bacterial luminescence, oyster embryo, amphipod sediment bioassays. *Marine Environmental Research 19*: 225-249.

29 Word, J. Q., J. A. Ward, B. Brown, B. D. Walls, and S. Lemilich. 1989. Relative sensitivity and cost of amphipod bioassays. In: Oceans '89. An International Conference Addressing Methods for Understanding The Global Ocean, IEEE Publication Number 89CH2780-5. Sept. 18-21, 1989, Seattle: IEEE> pp. 467-473.

CHAPTER 9

AQUATIC PHYSICO-CHEMICAL PARAMETERS

DEFINITION

Aquatic physico-chemical parameters are those characteristics of the aquatic environment that dictate the type of biological community that should be present under uncontaminated conditions. In order to properly interpret both biological and chemical contaminant data it is necessary to understand the physico-chemical characteristics of the environment being assessed. In this case, it is the aquatic environment, which includes both sediment and water. This is because the physico-chemical parameters (e.g., temperature, salinity, grain size, total organic carbon [TOC], dissolved organic carbon [DOC], pH, and redox potential) can affect all aspects of the biological community present, including the presence or absence of specific organisms, organismal behavior (e.g., burrowing and food selection), and the bioavailability of chemical contaminants.

TEMPERATURE

While temperature may be the easiest physico-chemical sediment parameters to measure, it is still a very important one. It influences the rates of both abiotic and biotic chemical reactions, and thus not only affects other physico-chemical parameters but contaminant (particularly metal and metalloid) bioavailability and toxicity as well (Burton, 2000). An example of temperature's influence on these various factors is that an increase in temperature will cause an increase in biological metabolism, which then results in an increase in oxygen consumption leading to reduced or anoxic sediment (lower Eh). The increase in reduced or anoxic sediment will cause an increase in acid volatile sulfides, which then bind with divalent metal ions making them less bioavailable.

The temperature of sediments and associated interstitial water varies much more slowly over time than that of the overlying water. Sediment temperature generally varies on a seasonal scale, so the temperature taken at the time of sampling is usually a good representation of the sediment temperature for that season at that location. The exception to this assumption is tidally influenced sediment, which may show a diurnal variation in temperature.

The benthic community at a site will be adapted to the normal temperature variations occurring there. When evaluating the recorded temperature of sediment it is necessary to know if it falls within the normal range for the site. If it does not, it may cause stress to the biota in addition to that caused by any chemical contaminants present. Sediment toxicity tests need to be run at the appropriate temperature for the test organism; if it is different than the site temperature, the difference must be taken into account when the tests are evaluated.

SALINITY

Salinity is a parameter that defines whether an aquatic system is classified as fresh water, brackish, or salt water. Historically, salinity has been defined as "the total amount of solid material in grams contained in one kilogram of sea water when all the carbonate has been converted to oxide, the bromine and iodine replaced by chlorine, and all organic matter completely oxidized" (Sverdrup *et al.*, 1942) and is reported as parts per thousand (ppt or ‰). In fresh water it is more simply defined as "the total concentration of ionic components" (Reid and Wood, 1976). As suggested by the definition's complexity, the actual measuring of salinity is a complex and difficult process. One of the properties of seawater is described by the "rule of constant proportions"—that no matter how salinity may vary, the proportions of seawater's major components remain constant (Anikouchine and Sternberg, 1973). Thus by measuring one component, one is able to calculate the concentrations of all components. Until the 1950s, salinity calculations usually took advantage of this rule: the concentration of chloride ions could easily and accurately be measured by titration, and salinity calculated from the result (salinity [‰] = 0.030 + 1.8050 × chlorinity [‰]). In the 1950s, scientists began to measure salinity using meters that utilized the electrical or magnetic properties of seawater that vary in proportion to its salinity, electrical conductance being the most common.

In 1978, under the auspices of UNESCO, in order to increase interlaboratory uniformity and accuracy, a new internationally accepted definition of salinity was promulgated based on a potassium chloride (KCl) standard. Salinity (S) of a seawater sample is now defined in terms of the ratio (K) of the sample's electrical conductivity (at 15°C and the pressure of one standard atmosphere) to a KCl solution (in which the mass fraction of KCl is 0.0324356, at the same temperature and pressure). The K value exactly equal to one corresponds, by definition, to a practical salinity equal to 35. Because the value is based on a ratio, it has no assigned units; while it is sometimes given the units "psu" (practical salinity units), they are considered inappropriate.

At freshwater and open ocean sites salinity will play only a minor role, if any, in site evaluation because it is either at insignificant concentration (freshwater sites), or it varies little both spatially and temporally across the study area (open ocean) even though it is present in significant amounts. However, at coastal sites such as embayments and estuaries, salinity will occur in significant amounts as well as vary significantly both spatially and temporally. It should be noted that high salinity inputs can occur in fresh water (e.g., from potash mines) and can be as toxic in their own right as any other contaminant.

The salinity variability will affect both the local biotic community and contaminant/toxicant chemistry. Salinity changes may stress organisms, making them more susceptible to the additional stress imposed by toxic chemicals, or they may change the way an organism processes contaminants taken up. Salinity changes can affect bioavailability and toxicity of both organic and inorganic chemicals; the exact relationship depends on both the chemical and the organism involved. In general, as contaminated fresh water enters an estuary and mixes with seawater, the increase in salinity will cause metals to precipitate out of the water column, making them less bioavailable and less toxic. However, the same increase in salinity can cause the release of ammonia, resulting in ammonia toxicity (Rysgaard *et al.*, 1999).

A change in salinity also affects the toxicity of organic contaminants, but not consistently as it does with metal contaminants. Brecken-Folse *et al.* (1994) showed that the toxicity to sheepshead minnows (*Cyprinodon variegatus*) of both 4-nitrophenol and 2,4-dinitrophenol decreased with increasing salinity. However, in the same experiment they showed that while 4-nitrophenol toxicity to grass shrimp (*Palaemonetes* spp.) decreased with increasing salinity, the toxicity of 2,4-dinitrophenol to grass shrimp increased with increasing salinity. Even closely related species may respond differently to salinity toxicant interactions. Martello *et al.* (2000) found the effect of high-salinity stress plus pentachlorophenol (PCP) to be subadditive among red abalone, while high-salinity stress seemed to potentiate the effect of PCP on black abalone.

In summary, when evaluating the risk posed by contaminants in coastal or estuarine environments, one must take into account that salinity will affect all bioassessment tools.

SEDIMENT GRAIN SIZE

The grain-size of sediments is a fundamental characteristic of the sediment environment for three major reasons. First, the habitat for benthic organisms is determined in part by sediment grain size. For example, the feeding strategies of many organisms require suitable substrate: filter feeders need relatively coarser material and detrital consumers need fine-grained sediments. As a result, different sediment textures support different "normal" communities of benthic organisms. Texture is thus vital when interpreting benthic community data (e.g., composition and diversity) to determine whether an area has been impacted.

Second, grain-size is an important factor in the accumulation of toxic substances in sediments. Exposure of many dissolved substances to sediments and to particulate matter in the water column (which settles to become the sediments) results in the sorption of those substances by the particulate matter. This process is a surface phenomenon, so sorption increases with increasing available surface area. Therefore, because finer-grained particles have larger surface area per mass (dry weight), the finer particles have the potential to accumulate more toxic substances per dry weight than coarser particles. In response to similar inputs, coarse-grained sediments (sands and gravels) may exhibit minimal uptake of toxic substances, while fine sediments (silts and clays) in the same area may accumulate high concentrations. Thus chemical data obtained from coarse sediments at a site may provide low estimates of exposure and the threat to natural resources, while fine-grained sediments may contain much higher concentrations but not necessarily represent "hotter" areas (more contaminated). In addition, grain-size measurements are one way to account for the effects of differences in the accumulation associated with textural differences.

Note that the differences in concentration among spatial areas that can be corrected for by normalizing to the sediment grain-size primarily reflect a response to the input and transport of toxic substances. The data are less clear as to how important the accumulation potential is in determining toxicity. It is considered likely that at least part of the bioavailability of many substances, particularly organic compounds, is basically the reverse of contaminant accumulation in the sediment; that is, the same contaminant concentrations are more bioavailable in coarse sediments than in fine. On the other hand, for those organisms that actually ingest the sediments, it is not clear whether sediment texture affects bio-uptake of contaminants.

Specific grain-size data can be used semi-quantitatively to determine, for example, that an area had too coarse a sediment to provide useful chemical data (the sediments would not be expected to significantly accumulate toxic substances), or that a finer-grained sediment area should have generally higher contaminant concentration. These data can also be used quantitatively, for example, by generating correlation plots of sediment texture versus the concentration of a particular contaminant. (In most systems distribution is a function of proximity to a source and to the effects of differential accumulation and settling.) In areas that are not close to the source, such correlations are usually quite linear. Such a correlation thus allows one to interpret the contamination's spatial distribution as the result of either differential settling (coarse sediments with relatively low concentrations in one area, fine sediments with higher concentrations in another) or due to proximity to a source (much higher, non-linear relationship between concentration and sediment texture).

The third major reason is that grain size is a good indicator of physical energy at the sampling site (i.e., sand ≈ high energy, silt ≈ low energy). When assessing contaminated sites it is recommended that sample locations be biased toward depositional areas (silty, low-energy areas) and away from sandy, high-energy areas if possible. Grain-size analysis is necessary to ensure proper site selection and to properly interpret chemical contaminant data. The preferred reporting method for such analysis is in percentages of sand (>0.063 – 2 mm), silt (>2 – 63 µm), and clay (• 2 µm) (Mudroch and Bourbonniere, 1991).

Another reason for knowing the sediment grain size is that it can affect the choice of test organism for sediment toxicity tests. DeWitt *et al.* (1988) found the percent mortality of the amphipod *Rhepoxynius abronius* was positively correlated with percent fine-grained sediment in unpolluted sediments. Specifically, in sediments with 80% or greater silt-clay content, survival can be reduced by 15%. They concluded that either the fine-grain size was directly causing amphipod mortality, (by such means as clogging the gills) or some other unmeasured covarying parameter was responsible.

TOTAL SOLIDS/MOISTURE CONTENT

The principal difference between soils and sediments is that, by weight, soils are composed almost exclusively of solids, with relatively little if any moisture content, while the major component of sediments by weight is water. Both total solids and moisture content are frequently reported as a percent, and as such they are the complement measurement of each other. The standard approach for both percent total solids and percent moisture is to weigh the wet sediment, dry it at between 50° and 105°C, and

reweigh the dried sediment. The dried weight of the sediment is the total solids content and the difference between the wet weight and the dry weight is the moisture content; to arrive at percent total solid and percent moisture, respectively, each amount is divided by the wet weight.

Percentages of total solids and moisture are measured to allow conversion of wet-weight values of the samples taken for other chemical analyses to be converted to dry weight and to determine the appropriate quantity of sediment needed to ensure sufficient solids for chemical analysis. Very high moisture content (>70%) can actually reduce the reliability of the sediment chemical analysis. One approach to this problem is to accept the chemical analysis results unconditionally if moisture content is less than 70%; give a qualified acceptance of the results if the moisture content is between 70% and 90% and reject the results if the moisture content exceeds 90% (Finkelstein, 2002). Another approach, which has been tried, is to freeze-dry the sediments prior to analysis (Finkelstein, 2002).

The degree of consolidation of sediments, which is important to the biota, also influences the assessment methods used to evaluate sediments. Sediments with low total solids and high moisture content will cause standard grab samplers to penetrate too deeply, and any biotic sampling devices placed on the sediment surface (e.g., bivalve sampling racks for bioaccumulation studies or biota traps) may actually sink into the sediment, rendering them useless. Knowing the degree of sediment consolidation is also useful for engineering and feasibility studies.

As with all sediment parameters, total solids and moisture content affect the benthic community of the sampled sediments and may influence toxicity test results. For example, densely packed sediments (i.e., high total solids/low moisture content) may impede burrowing, tube building, or feeding by test organisms, thus modifying sediment toxicity by either reducing exposure (i.e., organism stays on sediment surface) or enhancing organism stress (i.e., fatigue from burrowing attempts) (Lamberson *et al.*, 2000).

TOTAL VOLATILE SOLIDS AND TOTAL ORGANIC CARBON

Measurements of total volatile solids (TVS) and total organic carbon (TOC) provide data on slightly different aspects of sediments, as discussed below, but their primary intent is to provide a measure of the amount of organic matter in the sediments. TVS is specifically a measure of all volatile components of the sediment that are lost during high-temperature combustion (both organic and inorganic): TOC is specifically a measure of carbon that is

bound up in organic matter. Total organic matter (TOM) in water consists of thousands of components, including macroscopic particles, colloids, dissolved macromolecules, and specific compounds. In **most** cases, the total amount of organic matter in sediments is associated with plant debris, not with the toxic substances themselves. There are exceptions; for example, areas contaminated with coal tar will have high TOC, but it will be associated with the coal tar contamination. Organic content is also one of the fundamental characteristics of sediments and TVS/TOC values are used in a manner very analogous to that of grain size.

In most areas, organic matter in sediment forms the food for many benthic organisms. Organic-poor sediments (TOC <0.25%) can be compared to infertile farm fields; they are not capable of supporting abundant benthic organisms. On the other hand, very organic-rich sediments (TOC >15%) may be inhospitable to many larger organisms because microbial activity consumes all of the available oxygen and may form natural toxic substances (e.g., ammonia and sulfides).

The sorption energetics of some toxic substances, particularly some low-solubility organic compounds but also some metals, strongly favor accumulation on organic surfaces over inorganic particles. As a result, sediments with high TOC tend to accumulate higher concentrations of toxic substances than do low-TOC sediments from the same area. Thus TVS/TOC data help interpret the distribution of contamination and the threat to natural resources. The TOC/TVS data can be used in the same way as grain-size data (discussed above). It should also be noted that because some natural organic matter accumulates on sediment particles in the same way that toxic substances do, and because organic particles tend to be of small size and have a low settling velocity, they settle to the bottom only in quiet-water areas. Thus, high concentrations of TOC tend to correlate with areas of fine-grained sediments.

The equilibrium partitioning theory assumes that sediment organic carbon controls the sorption of contaminants onto sediment particles, and that these sorbed contaminants are not bioavailable. This assumption is not universally accepted because sediment organic carbon is also the principal food source for many benthic organisms. Boese *et al.* (1990), in their work with the clam *Macoma nasuta* and the contaminant hexachlorobenzene (HCB), concluded that 63% to 84% of the HCB tissue residues were derived by uptake through the gut from digested solids and only 11% to 12% were derived by uptake via the gills from interstitial water. Gunnarsson *et al.* (1999), working with the brittle star *Amphiura filiformis* and five different sources of organic carbon, concluded that the type of organic carbon,

specifically its quality as a food source, was critical in the uptake of sorbed contaminants. They found bioaccumulation factors ranging from 2.7 for lignin to 14.6 for sea lettuce (*Ulva lactuca*) and directly correlated with the quality of the organic carbon as a food source. Selck *et al.* (1998), in studying the toxicity and toxicokinetics of cadmium in the polychaete *Capitella* sp., concluded that sediment-bound cadmium contributed 95% of the total cadmium taken up by the feeding worms. This conclusion agrees with that of Wang *et al.* (1999), who concluded that most (>98%) of the cadmium (Cd), cobalt (Co), selenium (Se), and zinc (Zn) in polychaetes was derived from ingested sediment because of the high ingestion rates and the low uptake rates from pore water.

Schlekat *et al.* (2000) also looked at the effect on contaminant uptake of different organic carbon sources. They determined the uptake efficiency of the amphipod *Leptocheirus plumulosus* for silver (Ag), cadmium, and zinc from seven sources: bacterial exopolymeric sediment coatings, polymeric coatings made from *Spartina alterniflora* extract, amorphous iron oxide coatings, the diatom *Phaeodactylum tricornutum*, the chlorophyte *Dunaliella tertiolecta*, processed estuarine sediment, and fresh estuarine sediment. The highest assimilation efficiencies for silver and cadmium were from the bacterial exopolymeric coatings, while for zinc the highest efficiencies were for phytoplankton and processed sediments. Ortego and Benson (1992) found that humic acid reduced the toxicity of the insecticide fenvalerate, but fulvic acid had no effect on its toxicity (as tested by the Microtox® bioassay), while both reduced the toxicity of permethrin. Lee *et al.* (1993) found that the reduction of chemical toxicity due to the presence of organic matter (specifically, dissolved humic material [DHM]) was dependent on the specific chemical contaminant. DHM reduced the toxicity of 4-chloroanilin to *Daphnia magna,* but it failed to reduce the toxicity of both tetrabromobisphenol-A and pentachlorophenol.

On the other side of the issue, DeWitt *et al.* (1992), studying the toxic responses of the amphipod *Rhepoxynius abronius* to fluoranthene, found that differences in organic matter did affect toxicity. But since the absolute range was small, equilibrium partitioning based sediment quality guidelines do not need to be corrected for organic matter quality. Ankley *et al.* (1994), working with *Chironomus tentans,* concluded that within the range of organic carbon tested (3% to 8%) the equilibrium-partitioning model based on organic carbon is appropriate for predicting the bioavailability of sediment-associated chlorpyrifos to benthic invertebrates. It should be noted that individual studies on either side of the issue looked only at a single species, whereas, the importance of type of organic matter may be species dependent. Also, those studies that looked at various types of organic matter did not

generally look at mixtures: in situ sediments, whether anthropogenically influenced or not, contain mixtures of organic matter, and the variability in *in situ* sediments may not be significant. Also, regardless of issues of uptake control by biota, as mentioned earlier, in areas contaminated by substances such as coal tar the TOC is actually associated with the contaminants. Any attempt to determine the toxicity of these sediments by normalizing to TOC will greatly underestimate the toxicity. Since there is currently no resolution of this issue, the best course is to evaluate the individual sediments at each site in order to determine the bioavailability of contaminants.

TVS is measured by the loss of weight from high-temperature combustion of the sediments. This procedure results in the total oxidation of organic matter, oxidation of some inorganic substances, and dehydration of some minerals. TOC is measured by a number of techniques, usually less severe than TVS, and also with detection procedures specific for carbon. The U.S. Environmental Protection Agency's preferred method is found in the Puget Sound Protocols and Guidelines (U.S.EPA, 2001). Depending on the technique, TOC may or may not measure all carbon if it is in large particles (wood fragments) or stable forms (e.g., coal fragments). In most sediments, TVS measurements will yield higher values than TOC, but the two are usually fairly well correlated. Because TOC appears more likely to measure organic components that have biological significance, it is generally the preferred measurement.

OIL AND GREASE

Oil and grease (O&G) measurements determine the concentrations of a specific class of organic substances, those soluble in a non-polar organic solvent. The substances recovered in an O&G extraction include a broad range of natural and anthropogenic compounds, but in most cases high O&G values reflect predominately human impacts. With regard to toxic substances, some of the target compound list substances would be included in O&G measurements, but would rarely be more than a small fraction of the total material recovered. There is some evidence that O&G distributions correlate with those of some toxic substances, probably a reflection of similar accumulation mechanisms and similar input sources. O&G has been used more as a general indicator of anthropogenic contamination and predates most modern analytical procedures, but its use at hazardous waste sites has been minimal.

TOTAL SULFIDE

Sulfides are compounds containing one or more sulfur atoms connected directly to a carbon, metal, or other non-oxygen atoms. In sediments, sulfides exist as insoluble precipitates and as dissolved sulfide compounds. Hydrogen sulfide (H_2S), the toxicologically important form of sulfide, is produced when bacteria reduce sulfates and putrefy proteins. Measurements of sediment sulfide content are relevant primarily in marine sediments, because the sulfate present in seawater provides a major reservoir of reducible material that supports the growth of microbial populations when available oxygen has been consumed. Total sulfide measurements provide an indication of how reducing an environment is (negative Eh values). Sulfides in pore water may be analyzed as total sulfides (TS), as dissolved sulfides (DS), and as hydrogen sulfide. TS consist of acid-soluble metallic sulfides in suspended matter plus dissolved hydrogen sulfide. DS remain after the suspended solids have been removed by flocculation and settling. Hydrogen sulfide may be analyzed directly or calculated from DS concentration, sample pH, and the H_2S ionization constant (U.S. Army COE, 1995).

A number of mineral phases that precipitate and sorb toxic trace metals under aerobic conditions, dissolve under anoxic conditions and may thus make sediment-bound metals much more available. On the other hand, most toxic metals form very insoluble sulfide minerals. In addition, because sulfide is toxic to most organisms and is formed when oxygen is absent, sediments with high sulfide levels are often devoid of any life save specialized microorganisms.

When handling sediments to be analyzed for sulfides, care must be taken to prevent oxygenation of the sediments, which will rapidly convert the sulfides into sulfate. On the other hand, insufficient oxygenation of sediments during toxicity tests may result in a buildup of sulfides to toxic levels, thus confounding the test results. To avoid this problem, it is wise to monitor for sulfides while running toxicity tests.

Interestingly, the human nose is more sensitive to hydrogen sulfide than any available chemical tests. Thus, field observations for the presence of sulfide are important. However, variables such as wind and temperature, as well as the fact that the nose rapidly loses its sensitivity to sulfide after repeated exposure, make it difficult to rely on field measurements for quantitative estimates.

ACID VOLATILE SULFIDES

Acid volatile sulfides (AVS) are operationally defined metal sulfides that react with aqueous acids at room temperature to liberate gaseous hydrogen sulfide (Boothman *et al.*, 2001). In most sediments AVS are composed principally of amorphous ferrous sulfide (FeS) and manganese sulfide (MnS), while in contaminated sediments other divalent metal sulfides can make up a significant portion of the AVS present. It is the ability of these other divalent metals to displace the iron and manganese that make AVS a controlling factor in the concentration of dissolved metals in sediment interstitial water and thus their bioavailability and toxicity. The interpretation of AVS data requires the concurrent measurement of another operationally defined class of sediment components, simultaneously extracted metals (SEM). SEM are defined as those metals whose sulfides are less soluble than ferrous sulfide and are extracted from the sediments during the acid-volitization step of the analysis (Boothman *et al.,* 2001). The divalent metals most commonly found as SEM are nickel, zinc, cadmium, lead, and copper. Silver is also sometimes found, and while mercury can also form an AVS, other factors are more important in determining mercury's bioavailability and toxicity (U.S. EPA, 2000).

The basic theory behind the use of AVS in sediment evaluation is as follows: when the molar concentration of AVS is greater than the combined molar concentration of SEM, the SEM will be bound up as sulfides, making them unavailable to biota and thus nontoxic. When there is more SEM than AVS, the metals are potentially bioavailable and toxic. The theory is based on the assumption that the dissolved divalent metal ion is the only bioavailable and toxic form of SEM and that bound metals are not bioavailable and therefore nontoxic. The earliest work on the AVS/SEM relationship involved cadmium only (Di Toro *et al.*, 1990), and virtually as soon as the cadmium concentration exceeded AVS, toxicity was observed. When studies with other metals or metal combinations were performed, no toxicity was observed when AVS exceeded SEM; but, unlike the work with cadmium, toxicity was not always immediately observed when SEM exceeded AVS (Ankley *et al.,* 1991, 1993; Besser *et al.,* 1996). It was realized that, depending on the metal, other factors in addition to AVS were controlling its bioavailability, in particular organic carbon (Mahoney, 1995). It has been suggested that instead of trying to base sediment quality guidelines for metals on AVS alone, the guidelines can be based on both AVS and organic carbon (Ankley *et al.,* 1996)

Other work has raised questions about the absolute reliability of the AVS/SEM paradigm. Hare *et al*. (1994) found in colonization and growth studies of cadmium contaminated test sediments that the abundance of only one taxa was related to the Cd/AVS ratio, while for the 17 other taxa there was no relationship between the Cd/AVS ratio and either abundance or growth. Lee and coworkers (B-G. Lee *et al.,* 2000a,b) looked at bioaccumulation of metals in five marine benthic species and its relationship to AVS/SEM and porewater concentrations. They ran experiments in which metal concentrations (cadmium, nickel, and zinc) were varied while AVS concentrations remained fixed, and in which AVS concentrations varied while metals concentrations were fixed. They found that metal concentrations in animal tissue correlated with metal concentrations extracted from the sediments, not with AVS/SEM or porewater concentrations, in 13 out of 15 metal-animal combinations. They concluded that the influence of AVS on bioavailability of metals to benthic organisms depends on the organism and the metal. For a good review of these issues, see Mason (2000).

Based on current knowledge, SEM/AVS data can best be used in conjunction with bioaccumulation and toxicity studies as a tool to explain the studies' results, with the understanding that depending on the metals and organisms involved SEM/AVS may not be able to explain the results. SEM/AVS data by itself should not be used to predict the lack of adverse biological effects. One final important point when analyzing and evaluating SEM/AVS data is that AVS can vary both temporally and spatially (Besser *et al*., 1996, and J-S. Lee *et al*., 2000). SEM/AVS data must be collected at the same time as the studies are performed and from the portion of sediment in contact with the biota.

pH

The pH value of sediments (a measure of hydrogen ion concentrations) indicates their acidity (low pH values) or basic/alkaline nature (high pH values) (more accurately, it is a measure of the pH of the interstitial water). Neutral sediments have a pH of 7, and "normal" pH values usually range from 6 to 8. This parameter is important to measure because many substances, particularly inorganic substances, may undergo a number of reactions in natural systems that are pH dependent. Some of these values are relatively simple, phase-change reactions. For example, iron/manganese oxyhydroxides form as solid phases, coating other sediment particles at neutral and high pH but dissolving at low pH. Many other metals form similar oxyhydroxides or are scavenged by the oxyhydroxides. As a result of these processes, the metals tend to be much more mobile (because they are dissolved) at low pH

(acid) and less mobile (because they precipitate) at high pH. Other more complex reactions are also possible.

The solubility, hence mobility, of acidic and basic organic compounds is also affected by pH, although the importance of this effect is usually relatively minor compared to the elements.

Finally, the toxicity of many substances is affected by pH, in part because of changes in mobility/bioavailability and in part because of additional stress on the organisms.

Eh (ELECTRODE POTENTIAL)

The Eh of the sediments (actually of the interstitial water) is a measure of oxidation/reduction potential. The latter indicates the direction of change for substances that can undergo oxidation and/or reduction reactions. Fundamentally, natural systems in which molecular oxygen is present (aerobic) will show high-position Eh values and represent oxidizing environments. As oxygen is consumed by chemical and biological reactions (all metabolism is primarily the oxidation of organic matter), the system becomes increasingly reducing. Systems in which all oxygen has been consumed (anoxic) will show negative Eh values.

Similar to the effects of changes in pH, many substances, particularly the elements, undergo phase and/or speciation (form) changes depending on Eh. For example, at low/negative Eh (reducing environments) the iron/manganese oxyhydroxides are reduced, liberating the iron and manganese as dissolved metals. In addition, depending on the Eh value, iron may change from a valance state of +3 to +2. Other elements undergo similar changes. Further, a number of more dominant substances may also be altered. For example, nitrogen is predominantly oxidized to the nitrate ion in aerobic systems but reduced to ammonia in anoxic systems. Similarly sulfate is reduced to sulfide under anoxic conditions. The latter reaction is particularly important in seawater, where the relatively high concentration of sulfate can form high concentrations of sulfide.

All of these changes are important for three reasons.

1. As with pH, the mobility and availability of substances may be markedly different at different Eh values, but these changes may be complex. For example, the sulfides of many toxic metals are even less soluble than their oxidized form.

2. The form of the substances may change, generating more or less toxic forms. For example, the hexavalent form (i.e., oxidized chromium) is much more toxic than the trivalent form (i.e., reduced chromium).

3. As with pH, low oxygen conditions add additional stress to most organisms and thus may make them more susceptible to the effects of toxic substances.

SUMMARY

When conducting a biological assessment of contaminated sediments it is necessary to understand their physico-chemical parameters and chemical contaminant concentrations in order to properly interpret the results. Many of these parameters (e.g., temperature, AVS, Eh) vary significantly by time scales (e.g., diurnally, seasonally), and it is necessary to understand how such variations affect contaminant bioavailability. This variability must be taken into account when making any conclusions based on the bioassessment.

The parameters discussed in this chapter are not the only ones that can be used to evaluate sediment contamination and toxicity, but they are the ones that are either necessary or most helpful under virtually all conditions. Other parameters that can aid in sediment contamination and toxicity interpretation under certain conditions include ammonia, aluminum, and iron, to name a few.

Finally, one of the best recent references for protocols for many of the measurements discussed above (grain size, total solids, TVS, TOC, O&G, and total sulfides) is the. *Recommended Protocols for Measuring Conventional Sediment Variables in Puget Sound.* (Tetra Tech, 1986). This reference, which is part of the Puget Sound Protocols and Guidelines, was based on modifications to standard procedures following formal review by scientists and laboratory personnel (academic, agency, and commercial) in the Puget Sound region and represents an attempt to provide the most useful methods. The protocols were developed specifically for work in Puget Sound, a marine/estuarine environment, but the protocols for these conventional variables should be usable anywhere. A current version of this document can be accessed on-line at:
http://www.wa.gov/puget_sound/Publications/protocols/protocol.html

ADDITIONAL READING

Buchanan, J.B. 1984. Sediment analysis. In: *Methods for the Study of Marine Benthos*, N.A. Holme and A.D. McIntyre, eds. Boston, Mass.: Blackwell Scientific Publications. pp. 41–65.

Burgess, R.M. and R.A. McKinney. 1999. Importance of interstitial, overlying water and whole sediment exposures to bioaccumulation by marine bivalves. *Environmental Pollution* 104: pp 373–382.

Di Toro, D.M., C.S. Zarba, D.J. Hansen, W.J. Berry, R.C. Swartz, C.E. Cowan, S.P. Pavlou, H.E. Allen, N.A. Thomas, and P.R. Paquin. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. *Environmental Toxicology and Chemistry* 10(12): pp. 1541–1583.

Folk, R.L. 1968. *Petrology of Sedimentary Rocks.* Austin, Tex.: Hemphill Publishing Co. 172 pp.

Harkey, G.A., M.J. Lydy, J. Kukkonen, and P.F. Landrum. 1994. Feeding selectivity and assimilation of PAH and PCB in *Diporeia* spp. *Environmental Toxicology and Chemistry* 13(9): pp. 1445–1455.

Lee, H. II, B.L. Boese, and R.C. Randall. 1990. A method for determining cut uptake efficiencies of hydrophobic pollutants in a deposit-feeding clam. *Environmental Toxicology and Chemistry* 9(2): pp. 215–219.

Lauenstein, G.G. and A.Y. Cantillo (eds.). 1993. *Sampling and Analytical Methods of the National Status and Trends Program National Benthic Surveillance and Mussel Watch Projects 1984-1992: Volume II Comprehensive Descriptions of Complementary Measurements.* NOAA Technical Memorandum NOS ORCA 71. National Oceanic and Atmospheric Administration. Silver Spring, Maryland. 102 pp.

Plumb, R.H. 1981. *Procedures for Handling and Chemical Analysis of Sediment and Water Samples.* Technical Report EPA/CE—81—1. Vicksburg, Miss.: U.S. Army Corps of Engineers.

U.S. Environmental Protection Agency (EPA). 1983. *Methods for Chemical Analysis of Water and Wastes.* Technical Report EPA 600/4—79—020 (Revised 1983). Cincinnati, Oh.: Environmental Monitoring and Support Laboratory.

REFERENCES

Anikouchine, W.A. and R.W. Sternberg. 1973. *The World Ocean An Introduction to Oceanography.* Englewood Cliffs, N.J.: Prentice-Hall, Inc. 338 pp.

Ankley, G.T., D.J. Call, J.S. Cox, M.D. Kahl, R.A. Hoke, and P.A. Kosian. 1994. Organic carbon partitioning as a basis for predicting the toxicity of chlorpyrifos in sediments. *Environmental Toxicology and Chemistry* 13(4): pp. 621–626.

Ankley, G.T., D.M. Di Toro, D.J. Hansen, and W.J. Berry. 1996. Technical basis and proposal for deriving sediment quality criteria for metals. *Environmental Toxicology and Chemistry* 15(12): pp. 2056–2066.

Ankley, G.T., V.R. Mattson, E.N. Leonard, C.W. West, and J.L. Bennett. 1993. Predicting the acute toxicity of copper in freshwater sediments: Evaluation of the role of acid-volatile sulfide. *Environmental Toxicology and Chemistry* 12(2): pp. 315–320.

Ankley, G.T., G.L. Phipps, E.N. Leonard, D.A. Benoit, V.R. Mattson, P.A. Kosian, A.M. Cotter, J.R. Dierkes, D.J. Hansen, and J.D. Mahoney. 1991. Acid-volatile sulfide as a factor mediating cadmium and nickel bioavailability in contaminated sediments. *Environmental Toxicology and Chemistry* 10(10): pp. 1299–1307.

Besser, J.M., C.G. Ingersoll, and J.P. Giesy. 1996. Effects of spatial and temporal variation of acid-volatile sulfide on the bioavailability of copper and zinc in freshwater sediments. *Environmental Toxicology and Chemistry* 15(3): pp. 286–293.

Boese, B.L., H. Lee II, D.T. Specht, R.C. Randall, and M.H. Winsor. 1990. Comparison of aqueous and solid-phase uptake for hexachlorobenzene in the tellinid clam *Macoma nasuta* (Conrad): A mass balance approach. *Environmental Science and Technology* 9(2): pp. 221–231.

Boothman, W.S., D.J. Hansen, W.J. Berry, D.L. Robson, A. Helmstetter, J.M. Corbin, and S.D. Pratt. 2001. Biological response to variation of acid-volatile sulfides and metals in field-exposed spiked sediments. Environmental Toxicology and Chemistry 20:2, pp. 264-272.

Brecken-Folse, J.A., F.L. Mayer, L.E. Pedigo, and L.L. Marking. 1994. Acute toxicity of 4-nitrophenol, 2,4-dinitrophenol, terbufos and trichlorfon to grass shrimp (*Palaemonetes* spp.) and sheepshead minnows (*Cyprinodon variegatus*) as affected by salinity and temperature. *Environmental Toxicology and Chemistry* 13(1): pp. 67–77.

Burton, G.A., Jr. 2000. Sediment collection and processing: Factors affecting realism. In: *Sediment Toxicity Assessment,* G.A. Burton, Jr., ed. Chelsea, Mich.: Lewis Publishers, Inc. pp. 183-211.

Dewitt, T.H., G.R. Ditsworth, and R.C. Swartz. 1988. Effects of natural sediment features on survival of the phoxocephalid amphipod, *Rhepoxynius abronius*. *Marine Environmental Research* 25: pp. 99–124.

Dewitt, T.H., R.J. Ozretich, R.C. Swartz, J.O. Lamberson, D.W. Schults, G.R. Ditsworth, J.K.P. Jones, L. Hoselton, and L.M. Smithe. 1992. The influence of organic matter quality on the toxicity and partitioning of sediment-associated fluoranthene. *Environmental Toxicology and Chemistry* 11(2): pp. 197–208.

Di Toro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: The role of acid volatile sulfide. *Environmental Toxicity and Chemistry* 9(12): pp. 1487–1502.

Finkelstein, K., Coastal Resource Coordinator, NOAA Coastal Protection and Restoration Division. 2002. Personal communication.

Gunnarsson, J.S., M.E. Granberg, H.C. Nilsson, R. Rosenberg, and B. Hellman. 1999. Influence of sediment-organic matter quality on growth and polychlorobiphenyl bioavailability in Echinodermata (*Amphiura filiformis*). *Environmental Toxicology and Chemistry* 18(7): pp. 1534–1543.

Hare, L., R. Carignan and M.A. Huerta-Diaz. 1994. A field study of metal toxicity and accumulation by benthic invertebrates; implications for the acid-volatile sulfide (AVS) model. Limnology & Oceanography 37:7. pp. 1653-1668.

Lamberson, J.O., T.H. DeWitt, and R.C. Swartz. 2000. Assessment of sediment toxicity to marine benthos. In: *Sediment Toxicity Assessment*, G.A. Burton, Jr., ed. Chelsea, Mich.: Lewis Publishers, Inc. pp. 183–211.

Lee, B-G., S.B. Griscom, J-S. Lee, H.J. Choi, C-H. Koh, S.N. Luoma, and N.S. Fisher. 2000a. Influences of dietary uptake and reactive sulfides on metal bioavailability from aquatic sediments. Science 287:5451, pp. 282-284.

Lee, B-G., J-S. Lee, H.J. S.N. Luoma, Choi, and C-H. Koh. 2000b. Influence of acid volatile sulfides and metal concentrations on metal bioavailability to marine invertebrates in contaminated sediments. Environmental Science & Technology 34:21, pp. 4517-4523.

Lee, J-S., B-G. Lee, S.N. Luoma, H.J. Choi, C-H. Koh, and C.L. Brown. 2000. Influence of acid volatile sulfides and metal concentrations on metal partitioning in contaminated sediments. Environmental Science & Technology 34:21, pp. 4511-4516.

Lee, S.K., D. Freitag, C. Steinberg, A. Kettrup, and Y.H. Kim. 1993. Effects of dissolved humic materials on acute toxicity of some organic chemicals to aquatic organisms. *Water Research* 27(2): pp. 199–204.

Mahony, D.J. 1995. Certain chemical aspects of heavy metal-sediment interactions: The role of organic carbon and the kinetics of metal sulfide oxidation. In: *The Utility of AVS/EqP in Hazardous Waste Site Evaluations,* D.A. MacDonald and S. M. Salazar, eds. NOAA Technical Memorandum NOS ORCA 87. Seattle, Washington. pp 23–26.

Martello, L.B., C.S. Friedman, and R.S. Tjeerdema. 2000. Combined effects of pentachlorophenol and salinity stress on phagocytic and chemotactic function in two species of abalone. *Aquatic Toxicology* 49(3): pp. 213–225.

Mason, R.P. 2000. The bioaccumulation of mercury, methylmercury, and other toxic elements into pelagic and benthic organisms. In: *Sediment Toxicity Assessment,* G.A. Burton, Jr., ed. Chelsea, Mich.: Lewis Publishers, Inc. pp. 127–149.

Mudroch, A. and R.A. Bourbonniere. 1991. Sediment preservation, processing and storage. In: *Handbook of Techniques for Aquatic Sediment Sampling,* A. Mudroch and S. MacKnight, eds. Boca Raton, Fla.: CRC Press. pp. 131–170.

Ortego, L.S. and W.H. Benson. 1992. Effects of dissolved humic material on acute toxicity of selected pyrethroid insecticides. *Environmental Toxicology and Chemistry* 11(2): pp. 261–265.

Reid, G.K. and R.D. Wood. 1976. *Ecology of Inland Waters and Estuaries.* New York: D. Van Nostrand Co. 485 pp.

Rysgaard, S., P. Thastum, T. Dalsgaard, P.B. Christensen, and N.P. Sloth. 1999. Effects of salinity on NH4+ adsorption capacity, nitrification, and denitrification in Danish estuarine sediments. *Estuaries* 22(1): pp. 21–30.

Schlekat, C.E., A.W. Decho, and G.T. Chandler. 2000. Bioavailability of particle-associated silver, cadmium, and zinc to the estuarine amphipod *Leptocheirus plumulosus* through dietary ingestion. *Limnology and Oceanography* 45(1): pp. 11–21.

Selck, H., V.E. Forbes, and T.L. Forbes. 1998. Toxicity and toxicokinetics of cadmium in *Capitella* sp. I: Relative importance of water and sediment as routes of cadmium uptake. *Marine Ecology Progress Series* 164: pp. 167–178.

Sverdrup, H.U., M.W. Johnson, and R.H. Fleming. 1942. *The Oceans: Their Physics, Chemistry, and General Biology.* Englewood Cliffs, N.J.: Prentice-Hall, Inc. 1087 pp.

Tetra Tech. 1986. *Recommended Protocols for Measuring Conventional Sediment Variables in Puget Sound.* Prepared for the U.S. Environmental Protection Agency, Region 10—Office of Puget Sound, Seattle. 46 pp.
http://www.wa.gov/puget_sound/Publications/protocols/protocol.html

U.S. Army Corps of Engineers (COE). 1995. Naturally occurring levels of ammonia and sulfide in pore water: An assessment of the literature. Environmental Effects of Dredging Technical Notes, EEDP-O 1-34. [Q: USCOE – what city? Vicksburg?]

U.S.EPA. 2000. *Bioaccumulation Testing and Interpretation for the Purpose of Sediment Quality Assessment; Status and Needs.* EPA-823-R-00-001. Washington, D.C.: U.S. EPA, Bioaccumulation Analysis Workgroup. http://www.epa.gov/ost/cs/biotesting/

U.S.EPA. 2001. *Methods for Collection, Storage and Manipulation of Sediments for Chemical and Toxicological Analyses: Technical Manual.* EPA 823-B-01-002. Washington, D.C.: U.S. EPA, Office of Water.

Wang, W.-X., I. Stupakoff, and N.S. Fisher. 1999. Bioavailability of dissolved and sediment-bound metals to a marine deposit-feeding polychaete. *Marine Ecology Progress Series* 178: pp. 281–293.

APPENDIX A

GLOSSARY

Definitions

accuracy	The closeness of a measured or computed value to its true value (e.g., sample mean to population mean).
acute toxicity test	A toxicity test in which the test organism is exposed to the contaminated material for only a small portion of its life cycle.
acid volatile sulfides	A class of operationally defined metal sulfides that react with aqueous acids at room temperature to liberate gaseous hydrogen sulfide and composed principally of amorphous ferrous and manganese sulfides.
bioaccumulation	The net result when the uptake of a chemical by a biological organism exceeds the depuration of the chemical from the organism.
bioassay	The determination of the presence or absence of a specific substance based on a biological response.
bioassessment	The assessment of environmental conditions by the use of biological organisms.
bioavailability	The availability of a substance to be taken up by biological organisms.
bioconcentration	The process by which a chemical is directly taken up (by absorption only) from water and is accumulated to levels greater than those found in the surrounding water.
bioeffect	A change in the condition or functioning of an organism, which reduces its potential viability, resulting from exposure to a toxic chemical(s).
biomagnification	The increase in tissue concentrations of a bioaccumulated chemical as the chemical passes up through two or more trophic levels.
biomarker	A biochemical, physiological, or histological indicator of either exposure to, or effects of, xenobiotic chemicals at the suborganismal or organismal level
bulk sediment	Whole or intact sediment, same as solid phase.

chronic toxicity test	In general, a toxicity test in which the test organism is exposed to the contaminated material for at least the duration of one life cycle (the definition breaks down for long-lived organisms).
correlation analysis	Analysis to determine the relationship between two independent variables.
endpoint (toxicity test)	The response of the organism that is used as a measure of toxicity.
exposure	The contact with and uptake of an abiotic substance.
flow-through	A continuous flow of water through the test container;
(toxicity test)	the water will be clean in sediment tests and will contain a fixed concentration of contaminants in a water test.
***in-situ* toxicity test**	A toxicity test performed in the environment of concern as opposed to a laboratory toxicity test.
liquid phase	The sediment elutriate water from which all particles have been removed either by filtration or centrifugation.
negative control	A sample known to be nontoxic to the test organisms and in which they can function normally.
nonparametric methods	Methods not dependent on the characteristics of population parameters for their validity.
null hypothesis	A hypothesis formulated with the hope of rejection in favor of an alternative hypothesis, with a known probability of being correct.
parameter (ecological)	A physical or chemical characteristic of an environment, e.g., sediment grain size or salinity.
parameter (statistical)	A characteristic of a statistical population, e.g., population mean.
parametric methods	Methods concerned with the characteristics of population parameters (e.g., population mean and variance).
population (biological)	A group of individuals of the same species between which genetic material freely flows.
population (statistical)	The totality of all possible observations of the variable of concern (takes into account method of observation).

pore water	The water present within the spaces between sediment particles; also called interstitial water.
positive control	One or a dilution series of water spiked with a toxic compound that produces a response (the endpoint) in the species used for the test.
power (statistical)	The probability of rejecting the null hypothesis and accepting the alternative hypothesis when the null hypothesis is in fact false.
precision (statistical)	A measure of the closeness of agreement of individual measures of the same quantity (closeness of replicate values).
pseudoreplication	The taking and analysis of nonindependent observations as if they were independent observations (i.e., replicates).
random sampling	Taking of independent observations; e.g., with regards to spatial observations, the location of one observation has no influence on the location of any other observation.
regression analysis	Analysis to relate the behavior of a dependent variable to that of an independent variable.
replicate	One of a number of independent observations the total of which make up a sample.
sample (statistical)	A subset of a population.
sampling (statistical)	The process of taking observations of a population.
sediment elutriate	The water phase produced by mixing clean water with the test material and either allowing the material to settle or centrifuging it out.
simultaneously extracted metals	
solid phase	Whole or intact sediment, same as bulk sediment.
static renewal	Water is periodically replaced in test container either to prevent build up of waste products or to maintain the concentration of the contaminants in a water toxicity test.
statistic	A characteristic of a sample, e.g., sample mean, sample standard deviation.
stratified random	The division of the area to be sampled into various strata

sampling	(subareas of similar characteristics) and then random sampling of each strata.
suspended phase	The sediment elutriate water still containing suspended particles.
systematic sampling	The first observation is selected randomly and each succeeding observation is taken at a fixed interval.
toxicity test	A test to determine the toxicity of an environmental sample utilizing the response of a biological organism.
type I error	Rejection of the null hypothesis when it is true.
type II error	Failure to reject null hypothesis when it is false.
uptake	the passing of a contaminant from the external environment surrounding an organism across a cell boundary layer and into the internal environment of the organism.
variability	The difference between replicate values (within sample variability) or between sample statistics (between sample variability).

Abbreviations and Acronyms

AET	apparent effects threshold
Ag	silver
AHH	aryl hydrocarbon hydroxylase
ALAD	delta-aminolevulinic acid dehydratase
ANOVA	analysis of variance
ASTM	American Society for Testing and Materials
AVS	acid volatile sulfides
BCF	bioconcentration factor
Cd	cadmium
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CI	confidence intervals
cm	centimeter
Cu	copper
DMSO	dimethyl sulfoxide
EC_{50}	effective concentration 50 percent
Eh	oxidation-reduction potential
EPA	United States Environmental Protection Agency

EPT	Ephemeroptera, Plecoptera, Trichoptera
EROD	ethoryresorufin-O-deethylase
FCA	foci of cellular alteration
GHS	glutathione
GST	glutathione transferases
H_2S	hydrogen sulfide
HBI	Hilsenhoff Biotic Index
Hg	mercury
KCI	potassium chloride
L	liter
LC_{50}	lethal concentration 50 percent
LOEC	lowest observed effective concentration
MESA	Marine Ecosystem Analysis (NOAA Program)
MFO	mixed function oxidase
ml	milliliter
ml/L	milliliter per liter

mm	millimeter
MN	micronuclei
MT	metallothioneins
NOAA	National Oceanic and Atmospheric Administration
NOEC	no observed effective concentration
NS&T	National Status and Trends (NOAA Program)
O&G	oil and grease
PAH	polycyclic aromatic hydrocarbons
Pb	lead
PCB	polychlorinated biphenyls
ppm	parts per million
ppt	parts per thousand
REMOTS™	remote ecological monitoring of the sea floor
RI	remedial investigation
RTG	rainbow trout gonad
SAIC	Scientific Applications International Corporation
SCE	sister chromatid exchange

SDN	specific or degenerative necrotic
SEM	simultaneously extracted metals
SFG	scope for growth
TIE	toxicity identification evaluation
TIC	total inorganic carbon
TOC	total organic carbon
TOM	total organic matter
TS	total sulfides
TVS	total volatile solids
Zn	zinc

APPENDIX B

PROTOCOLS

The paper version of this appendix included copies of the non-copyrighted protocols, principally the Puget Sound Protocols. The complete Puget Sound Protocols, are available on the Internet at **http://www.wa.gov/puget_sound/Publications/protocols/protocol.html,** while the copyrighted ASTM protocols can be purchased from the American Society for Testing and Materials via the Internet at **http://www.astm.org/**.

APPENDIX C

Toxicity Test
Sensitivity Comparison Table

Table C-1 Toxicity tests comparison summary.

Test Organisms	Test Type	Test	Treatments				Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
			Control	Exp.	Rep.	# per					
Hyalella azteca	S-R	10d		45	3/T	10	Survival		1-most mortality w/out food in sediment	•Test organisms-*H. azteca* 12-19d old; *C. tentans* 3rd instar; *L. varigatus* mixed adults •Tests were conducted with & without feeding in sediment •The nutritional quality of the test sediments was a major factor influencing *H. azteca* test results •Conclusions: •There is a potential for a significant percentage of false positive results in tests with clean sediments in which *H. azteca* & *C. tentans* is not fed •Addition of food to sediments reduced the variability of assay results w/respect to survival •Further attention should be given to the possible effects of sediment grain-size on growth of *C. tentans* •*L. variegatus* was the least sensitive organism	**1**
Chironomus tentans	S-R	10d		45	3/T	10	Survival & growth		2-<mortality w/out food in sediment; reduced growth		
Lumbriculus variegatus	S-R	10d		50	2/T	10	Surv./reprod./ growth		3-least sensitive-less mortality w/out food in sediment; # of worms increased and weight increased w/food addet to sediment		
Cerastoderm edule	F-T	17 wks	1	6	3	50	Mortality %	1	1-100% death @ 2 wks high conc. w/inc. to 80% @ low conc. 17 wks	•Data suggest: *C. edule* greater sensitivity TBT than *S. plana*; may be due to feeding regeme, tolerance, or ability to metabolize TBT •Few *N. diversicolor* were collected from any microcosms including the controls, and it is possible that some worms may have starved and others may have been ripe adults which spawned when temperature rose, and subsequently died •*C. edule* had a generally higher body burden of TBT than *S. plana* after 11 weeks, but *S plana* seems to support a much greater body burden before death occurred in the high level TBT microcosm •Mean high level TBT = 1.786 ±0.047 •Mean low level TBT = 0.106 ±0.997 •Mean Control level TBT = 0.007 ±0.008	**3**
Scrobicularia plana	F-T	11 wks	1	6	3	26	Mortality %	1	2-100% death @ 10 wks high; 78% @ low 11 wks		
Nereis diversicolor	F-T	17 wks	1	6	3	50	Mortality %	1	Not possible to rank		
Cirratulus cirratus	F-T	15 wks	1	6	3	50	Mortality %	1	3-14% death @ 11 wks high; 0%@ low		
Littorina litorea	F-T	12 wks	1	6	3	12	Mortality %	1	4-8% death @ 12 wks high; 0% @ low		
Rhepoxynius abronius	S	10d	1 pos.	1	5	20	Survival	1	2-> surv. than worms	•Both organisms were exposed to the same sediment contaminants & concentrations of metals, organics, phthalates, phenols. •One test measured PCBs •All three tests--storage of sediments can have a substantial effect on measurements of sediment toxicity; both for reference & contaminated sediments •Amphipod survival generally decreases w/storage time •Polychaete & Microtox mortality/bioluminescence varied unpredictably (responses either inc. or dec. w/storage time) •For both kinds of sediment, the test response declined dramatically after 12-14 weeks of storage	**4**
Neanthes (spp.)	S-R	20d		1	5		Growth (biomass)	1	1-< biomass, more sensitive		
Mytilus galloprovicialis		4wks			2	30	LT50	1	1--4.6 days @ 100 µg/l Ag	•Rank based on 100 µg/l Ag--the 1st lowest conc. for all the species •No direct relationship exists between bioaccumulation & toxicity, may be due to the physico-chemical forms of sequestration of Ag (e.g., Ag associated w/insol. forms which mainly consist of Ag-sulphide, a very stable compound whose tox. is therefore limited) •Sublethal effects: *C. gigas*--after 4wks, decrease of glycogen in intestinal tissue at 20 µg/l Ag *M. galloprovincialis*--dilation of ER in branchial cells, result death at 50-100 µg/l Ag *C. varia*--unable to attach, lost bysus & no thread produced at 20 µg/l Ag *S. plana*--after 2d exposure, pieces of siphons were lost at 50 µg/l Ag *D. polymorpha*--lost byssus, lesions GI walls, tubes lysed, degeneration of ovaries	**5**
Chlamys varia		4wks			2	30	LT50	1	2--4.8 days @ 100 µg/l Ag		
Crassostrea gigas		4wks			2	30	LT50	1	3--8.7 days @ 100 µg/l Ag		
Scrobicularia plana		2wks			2	30	LT50	1	4--10.4 days @ 100 µg/l Ag		
Dreissena polymorpha							LT50	1	5--28 days @ 400 µg/l Ag		

Test Organisms	Test Type	Test	Treatments				Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
			Control	Exp.	Rep.	# per					
Oncorhynchus mykiss		96h				10	LC50 (mg/l)	1	1--TNT @ 1.2 1--TNT/RDX @ 1.7 1--DNT +16 @ 7.2 1--Cd @ 00095 2--Zn @ 0.41 2--Pb @ 1.0 7--Cr @ 69.0 4--Malathion 0.17 3--Sevin @ 4.4 1--Chlordane @ 0.01 2--CH_4Oxchlor0.0072 1--Treflan @ 0.011 4--DDT @ 0.007	•*L. variegatus* appears much less sensitive to pesticides than the other invertebrates; in most cases, fish were also more sensitive than the worms •All of the compounds (except the nitroaromatics) were evaluated for sublethal effects on *L. variegatus*; in general, sublethal effects wee dose-related in terms of severity & frequency of occurrence &, except with Hg, Pb, & 2-4-D, occurred at levels below the LC50: •Fragmentation--affected worms broke up into large pieces (Cd, Hg, Pb, Sevin, chlordane, & malathion) •Clumping--formation of tight spherical bunches or clumps (Cr, Cu, Zn, malathion, & chlordane) •Localized swelling--small swellings or lumps along their bodies (Malathion, Sevin, & chlordane) •Mucus production--(Pb, Cd) •Overall swelling--during exposure, the diameter of the worms often doubled (Cd) •Color changes--worms exposed to Sevin turned dark brown, and those exposed to Zn were gray	**2**
Salvelinus fontinalis		96h				10	LC50 (mg/l)	1	2--Cu @ 0.017 5--Zn @ 0.96 6--Pb @ 4.0 6--Cr @ 50.0		
Ictalurus punctatus		96h				10	LC50 (mg/l)	1	2--TNT @ 2.4 3--TNT/RDX @ 5.3 2--DNT +16 @ 17.5 5--Malathion 9.0 5--2-4-D @ 500.0 8--Sevin @ 15.8 7--Chlordane @0.5 5--Treflan @ 0.42 6--DDT @ 0.016		
Lepomis macrochirus		96h				10	LC50 (mg/l)	1	3--TNT @ 3.0 2--TNT/RDX @ 4.2 1--DNT +16 @ 7.2 8--Cu @ 0.74 6--Zn @ 2.9 3--Malathion 0.103 4--2-4-D @ 375.0 4--Sevin @ 5.6 3--Chlordane @0.022 4--CH_4Oxchlor0.058 5--DDT @ 0.008		
		24h				10	LC50 (mg/l)	1	3--Treflan @ 0.1		
Pimephales promelas		96h				10	LC50 (mg/l)	1	4--TNT @ 3.1 4--TNT/RDX @ 6.0 3--DNT +16 @ 22.0 6--Cu @ 0.55 4--Zn @ 0.87 7--Pb @ 5.0 6--Malathion 12.5 5--Sevin @ 6.7 5--Chlordane @0.052 7--DDT @ 0.19		
Fundulus heteroclitus		96h					LC50 (mg/l)	1	4--Hg @ 2.0		

Test Organisms	Test Type	Test	Treatments				Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
			Control	Exp.	Rep.	# per					
Lumbriculus variegatus		48h			2	10	LC50 (mg/l)	1	5--TNT @ 4.9 5--TNT/RDX @ 8.8 5--Cu @ 0.23 6--DNT +16 @ 24.5 3--Cd @ 0.12 8--Zn @ 8.1 4--Pb @ 3.4 5--Cr @ 25.3 3--Hg @ 0.11 8--Malathion 30.9 3--2-4-D @ 122.2 7--Sevin @ 13.0 9--Chlordane @ 1.8 5--CH4Oxchlor 0.2 4--Treflan @ 0.3 9--DDT @ 0.13		**2** cont'd
		96h			2	10	LC50 (mg/l)	1	4--Cu @ 0.15 2--Cd @ 0.074 7--Zn @ 6.3 3--Pb @ 1.8 3--Cr @ 13.3 2--Hg @ 0.1 7--Malathion 20.5 3--2-4-D @ 122.2 6--Sevin @ 8.2 8--Chlordane @1.4 5--CH_4Oxchlor 0.2 4--Treflan @ 0.3 9--DDT @ 0.13		
Hyalella azteca		48h				5	LC50 (mg/l)	1	6--TNT @ 6.5 6--TNT/RDX @ 10.5 4--DNT +16 @ 22.7		
Daphnia magna		48h				5	LC50 (mg/l)	1	7--TNT @ 11.9 7--TNT/RDX @ 11.7 5--DNT +16 @ 23.5" 1--Cu @ 0.0098 1--Zn @ 0.1 1--Cr @ 0.05 2--Malathion 0.009 2--2-4-D @ 100.0		
Daphnia hyalina		64h				5	LC50 (mg/l)	1	1--DDT @ 0.001		
Daphnia pulex		48h					LC50 (mg/l)		1--Pb @ 0.6 1--Hg @ 0.006		
		48h					LC50 (mg/l)		1--Sevin @ 0.0064 4--Chlordane @0.029 3--CH_4Oxchlor0.0078 3--Treflan @ 0.24 2--DDT @ 0.0036		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Tanytarsus dissimilus		48h				5	LC50 (mg/l)	1	8--TNT @ 24.8 8--TNT/RDX @ 32.0 7--DNT +16 @ 47.0		**2** cont'd
Simulium		24h					LC50 (mg/l)		9--Cu @ 3.9		
Ephemerella		48h					LC50 (mg/l)		4--Cd @ 2.0 5--Pb @ 3.5 2--Cr @ 3.5 4--Hg @ 2.0		
Damsel fly nymph		48h					LC50 (mg/l)		10--Zn @ 69.5		
Aeroneuria		14d					LC50 (mg/l)		8--Pb @ 19.5 4--Hg @ 2.0		
Chironomus spp.		8h					LC50 (mg/l)		2--Chlordane @0.015 8--DDT @ 0.04		
Physa		24h					LC50 (mg/l)		3--Cu @ 0.05 3--Zn @ 0.8 4--Cr @ 17.3		
Gammarus lacustris		96h					LC50 (mg/l)		7--Cu @ 0.58 1--Malathion 0.0018 2--Sevin @ 0.04 6--Chlordane @0.16		
		24h					LC50 (mg/l)		1--CH_4Oxchlor0.0047 6--Treflan @ 8.8 3--DDT @ 0.0047		
Gammarus faciatus		48h					LC50 (mg/l)		1--2-4-D @ 3.2		
Asellus		48h					LC50 (mg/l)		9--Zn @ 56.0		
Chlorella spp.	Agar plates	96h		4 conc.		See notes	Growth inhibition	1	1--Cu @ 25 mg/l 2--Cr(VI) @ 100 mg/l no inhibition	•Metal toxicity to algal cultures, growth of streaks of metal treated cultures on 2% BG11 - agar plates •Of the 4 metals investigated, Cu, Cr, Pb, and Cr_2O_7 the former 2 were more toxic than the latter two	**6**
Scenedesmus spp.	Agar	96h		4 conc.		Notes	Growth inhibition	1	2--Cu @ 100 mg/l 1--Cr(VI) @ 100 mg/l	•4 concentrations--2.5, 1.0, 25, and 100 mg/l were the exposures •The algae were streaked out on 2% BG 11 - agar plates--(see number per column)	
Selenastrum spp.	Agar	96h		4 conc.		Notes	Grwth inhibit.	1	2--Cu @ 100 mg/l 1--Cr(VI) @ 100 mg/l	•The variable factor here may be the tolerance of the species & the amount of toxic metal accumulated	

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Ceriodaphnia dubia		48h	1 (E&S); 1 H_2O		3/T	20	Survival (N)	1	1-to elutriates 1-whole sediment	•Determinations of assay sensitivity based on comparisons between the elutriate phase on one toxicity assay and the solid phase of another toxicity assay should not be made •Cladocerans spent significant time feeding on the sediment surface, increasing the potential for toxicant uptake •Relative sensitivity will frequently vary w/the site and/or toxicants being evaluated •The current study confirms the premise that multiple test assays are necessary to both detect sediment toxicity and differentiate degrees of toxicity •Microbial activity in elutriates-ß-galactosidase & ß-glucosidase were the most sensitive	7
Daphnia magna		48h	1 (E&S); 1 H_2O		3/T	20	Survival (N)	1	2-to elutriates 2-whole sediment		
Hyalella azteca		48h	1 (E&S); 1 H_2O		3/T	10	Survival (J)	1	4-to elutriates 3-whole sediment		
Selenastrum capricornutum		48h	1 (E&S); 1 H_2O		3/T	$1x10^6$/ ml	Growth	1	3-to elutriates		
Daphnia magna	S-R	2d	1	5	3	20	LC50	1	1--Cu @ 299 µg/l aqueous Cu 2--Cu @ 428 µg/l aq Cu 3--Cu @ 719 µg/l aq Cu 4--Cu @ 1152 µg/l aq Cu	•*Daphnia* were added to the test beakers on day 8 of the 10 day exposure to *Chironomus* test beakers •Suvival data: Cu Dilutions: Control; 37.5; 75.0; 150; 300; 600 *D. magna* 15/20 14/20 7/20 0/20 0/20 0/20 *C. tentans* 11/15 13/15 8/15 1/15 0/15 0/15 *H. azteca* 15/15 13/15 13/15 0.15 0/15 0/15 *G. lacustris* 9/10 10/10 6/10 0/10 0/10 0/10	8
Chironomus tentans	S-R	10d	1	5	3	15	LC50	1			
Hyalella azteca	S-R	10d	1	5	3	15	LC50	1			
Gammarus lacustris	S-R	10d	1	5	3	10	LC50	1			

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Rhepoxynius abronius		10d	1 pos., 1 neg.	16 sites		20	Surv./reburial	1 USA	1-whole sediment	•Labs: *R. abronius* (USA-Swartz) *M. arenaria*, *C. gigas* (USA-Phelps) *R. abronius*, *N. arenaeodentata*, *C. gigas* (Canada-Chapman) *C. volutator*, *C. gigas*, *P. phosphoreum* (UK-Roddie/Butler) *C. gigas*, *B. sarsi* (The Netherlands-van den Hurk)	**9**
Rhepoxynius abronius		10d	1 pos., 1 neg.	16 sites		20	Surv./reburial/ avoidance	1 Canada	-whole sediment		
Neanthes arenaceodentata	S-R	20d	1 pos., 1 neg. & 1 ref	16 sites	2	5	Surv. & growth	1	-whole sediment	Limitations: •UK mean oyster larvae tests were below criteria for negative controls •*R. abronius* exposed to drilling sites more sensitive than expected •No control criteria for direct exposure in the Microtox assay •*M. arenaria* specimens were stressed prior to testing; sediments held too long •Differences between the German and drilling sites; sediment grain size, sediment sampling, sediment treatment	
Corophium vlutator		10d	1 pos., 1 neg.	16 sites		10	Survival/immobilization	1-UK	2-whole sediment		
Crassostrea gigas		2d	1 pos., 1 neg.	16 sites	2 lab reps for 7 sites	23×10^3	Surv./devel.	1 Canada	1-elutriate; whole sediment	Tests: •Originally for water column, adapted for sediment-oyster abnormality & Microtox •Acute lethal tests specifically for sediments-all others •Specific for sediment-sublethal tests (growth, reproduction, full life-cycle tests; e.g., *Neanthes*)	
Crassostrea gigas		24h	1 neg.		3 at 2 dilut	20×10^3	Survival/development	1-UK	-elutriate		
Crassostrea gigas		96h	1 neg.			30	Survival & Metamorphosis	1-USA	-whole sediment; inconsistant results		
Crassostrea gigas		24h	1 neg.			53×10^3	Survival	1-Netherlands	-elutriate		
Photobacterium phosphoreum		1h	1 clean, 1 ref. sed.	16 sites	20		Bioluminescence	1-UK	-whole sediment; turbidity problem		
Mya arenaria		3h	1 neg.				Burrowing	1-USA	-whole sediment; inconclusive results		
Bathyporeia sarsi		10d		16 sites	Few, not all	20	Surv./reburial	1-Netherlands	3-whole sediment		
Rhepoxynius abronius		10d	1	9	5	20	Survival & avoidance	1	2-survival 2/9 signif. diff. from control 2-avoidance 1/9 signif. diff. from control	•The responsiveness of the organisms-the number of statistical differences from reference; lethal (survival) endpoints were compared and sublethal endpoints were compared to rank the relative sensitivity •Exposure similarities: Few differences in concentration observed among stations for major metals (Al, Si, Fe, Mn, Mg, Ca, Na, Ti) 13 trace elements that were always detected (As, Cr, Cu, Ni, Zn, Pb, Hg, Sn, Ag) Organic compounds-station ranks--LPAHs, HPAHs & other OC--IS (Islais Waterway)>OA(Oakland)>SP(San Pablo Bay) Detected at all stations--DDT, DDE, DDD •Aggregate index--mean of the combined values for 5 chemical groups to achieve a single composite value that will identify contaminated sites; IS = 6.29, OA = 1.97, & SP = 1.00	**10**
Mytilus edulis		48h	1	9			Survival & abnormal developmt.	1	1-survival 8/9 signif. diff. from control 1-abn. dev. 5/9 signif. diff. from control		
Tigriopus californicus		4wks	1	9	8	1	Reprod.-# young/adult	1	1-# young 5/9 signif. diff. from control		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Photobacterium phosphoreum		5 min		68 chem-icals	Few-10%		Bioluminescence (EC50)	1	1-LC_{50} prediction from EC_{50} for alcohols (r^2=0.96), ketones (r^2=0.81), ethanes (r^2=0.72) 2-general groupings	•The reliability of the bacterial assay for predicting fish LC_{50} is greatest when the data are derived from a semiholologous series of chemicals rather than from general groups (e.g., inorganic vs organic) •Prediction of LC_{50} values from bacteria EC_{50} values not sufficient to justify except for order of magnitude estimates	**11**
Pimephales promelas	F-T	96h					Mortality (LC50)	1	1-general groups & not a semihomolo-gous series	•Semihomologous series-a plot of the endpoint against the number of carbon atoms for each chemical; it was proposed that the that if the EC_{50} had the same dependence on the carbon number as the fish LC_{50}, comparable mechanisms of toxicity wer operative & the two endpoints would be correlated	
Lubriculus variegatus	S-R	2 wks	1 site	6 sites	5	10	Survival	1	1--5/6 stations signif. diff. from controls	•Worm survival in the more toxic sediments from the harbor's center (sites 1, 2, 3) was greater for Lumbriculus than for the deep-water amphipod Pontoporeia *Diporeia hoyi* •The sites contained high levels of PCBs (2300 µg/g), PAHs, and metals associated w/the industrial area	**13**
Diporeia hoyi	S-R	1 wk	1 site	6 sites	5	5	Survival	1	2--2/6 stations signif. diff. from controls		
Leptocheirus plumulosus		10d 28d	1 ref. 1 ref.	5 dilut. 5 dilut.			Mortality Mort.gro.fert	4 3	3-10d mortality 2-10d/1-28d 1-Cd sensitivity 1-most conc./discrim.	•Protocols--DeWitt et al. '92; Schlekat et al. '92; Hall et al. '91; ASTM '90 E1367-90 Vol 11.04; McGee & Schlekat '93; Phelps & Warner '90; Phelps '90 •22 organic contaminants in sediment from Curtis Cr.; 5 metals (e.g., Cd, Cu, Ni, Pb, Zn); total SEM/AVS <1.0 •Conclusions-- •Mortality in most short-term (10d) tests were similar in their sensitivity to contaminated Chesapeake Bay sediment •Short-term sublethal responses were not more contaminant-sensitive than acute mortality for any species •Chronic L. plumulosus were more sensitive than the acute toxicity test responses by only 2-3x •Moderate among-lab variability in 10d L. plumulosus mortality, large between lab differences in 28d mortality, growth, and fertility due to methodologies used by each research group	**14**
Eohaustorius estuarius		10d	1 ref.	5 dilut.			Mort.rebury	2	4-10d mortality		
Lepidactylus dytiscus		10d	1 ref.	5 dilut.			Mort.rebury	1	2-10d mortality 2-least conc./discrim.		
Hyalella azteca		10d	1 ref.	5 dilut.			Mortality	1	5-10d mortality 1-Cd sensitivity 2-least conc./discrim.		
Streblospio bendicti		10d	1 ref.	5 dilut.			Mortality	1	1-10d mortality 2-least conc./discrim.		
Crassostrea gigas		96h	1 ref.	5 dilut.			Mort.meta-morphosis	1	1-Cd sensitivity		
Mya arenaria		43h	1 ref.	5 dilut.			Reburial	1	3-difficult to compare w/other species		

Test Organisms	Test Type	Test	Treatments Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Photobacterium phosphoreum *Culex pipiens* *Oryzias latipes* *Poecilia reticulata* *Uronema parduczi* *Aedes aegypti* *Hydra oligactis* *Xenopus laevis* *Pseudomonas putida* *Chlorella pyrenoidosa* *Selenastrum capriconutum* *Lymnea stagnalis* *Chilomonas paramecium* *Ambystoma mexicanum* *Scendedesmus pannonicus* *Entosyphon sulcatum* *Pimephales promelas* *Oncorhynchus mykiss* *Daphnia magna* *Daphnia pulex* *Microcystis aeruginosa*	5 min		1	4	2			1	20--least sensitive 1-6.93 > MAS 2--3.48 > MAS 3--2.54 > MAS 4--2.04 > MAS 5--1.99 > MAS 6--1.8 > MAS 7--1.74 > MAS 8--1.29 > MAS 9--1.19 > MAS 10--1.17 > MAS 11--1.09 > MAS 12--1.03 > MAS 13--0.96 > MAS 14--0.93 > MAS 15--0.75 > MAS 15--0.75 > MAS 16--0.7 > MAS 17--0.66 > MAS 18--0.65 > MAS 19--0.53 > MAS	•In comparing the sensitivity of the Microtox Assay System (MAS) w/that of a different test on aquatic toxicity, the geometric mean of the ratio for all of the tested chemicals of Microtox 5-min. EC10 & the concentration causing minimal effect in the other test expresses the number of times the specific test is more susceptible than the MAS •Test compounds: mercuric chloride, Cd nitrate, *n*-Propanol, *n*-Heptanol, ethylacetate, ethylpropionate, acetone, trichloroethylene, allylamine, aniline, benzene, pyridine, o-Cresol, Salicylaldehyde, pentachlorphenate •Chemicals tested: bacteria-14, green algae--12-14, protozoa-14, crustaceans-15, insects-15, coelenterates-15, molluscs-15, fishes-15, and amphibians-15	**12**
Daphnia magna		72h			6	70	Mortality	1	2--7% @ pH6	•Exposure concentration of TBTCl at 6.0 µg/l for both species •pH affects the toxicity of TBT; higher at higher pH values •TBT must be in soluable form, and not bound to organic materials or coloidal substances to be available to organisms •TBT is lipophilic	**17**
		72h			9	70	Mortality	1	1--22% @ pH8		
Thymallus thymallus	S-R	168h	1		2	300	Mortality	1	3--2.6% @ pH8.1		

Test Organisms	Test Type	Test	Treatments Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Paracorophium excavatum		10d	4	6 conc.	2	10	Survival	1	1-based on LC_{50} value of 18.3 mg/kg	•*P. hurleyi* sensitive to sand grain size-best survival in fine sand, and not in very fine, medium, or coarse sand; best in 16% salinity & not 8%; best in enriched sediment w/organic content 7% Protocol: Swartz et al. 1985	**18**
Proharpinia hurleyi		10d	4	6 conc.	2	10	Survival	1	2-based on LC_{50} value of 23 mg/kg	•Exposure was to Cd in sediment-range 2.8-70 mg/kg	
Crab zoea	S	96h					LC50	1	1--0.25 mg/l Cd	•This work primarily investigated the use, refinement & comparative sensitivity of a sea urchin sperm cell bioassay for marine pollution monitoring. This test exposed sea urchin or sand dollar sperm cells to test solutions for 60 min. prior to fertilization of the eggs. Elevation of the fertilization membrane was used as the test endpoint. An oyster sperm and a salmon sperm test (for brackish waters) were also investigated •The "validation" portion of this work compared the sensitivity of sperm cell tests to other common test organisms available in the Pacific Northwest	**15**
Sand dollar adult	S	72h					LC50	1	2--0.4 mg/l Cd		
Cabezon (sculpin)							LC50	1	3--Juv. 0.5 mg/l Cd		
Sea urchin emb.	S	120h					LC50	1	3--0.5 mg/l Cd		
Sea urchin adult	F-T	96					LC50	1	4--0.7 mg/l Cd		
Sea urchin sperm	S	1h					LC50	1	1--2 µg/l Cu		
Oyster embryo	S	48					LC50	1	2--6 µg/l Cu		
Sea urchin emb.	S	120h					LC50	1	2--6 µg/l Cu		
Oyster sperm							LC50	1	3--12 µg/l Cu		
Sand dollar sperm							LC50	1	4--26 µg/l Cu		
Crab zoea	S	96h					LC50	1	1--0.6 mg/l Pb		
Oyster embryo	S	48h					LC50	1	2--0.7 mg/l Pb		
Sea urchin sperm	S	1h					LC50	1	3--1.0 mg/l Pb		
Cabezon (sculpin)							LC50	1	4--juv. 1.5 mg/l Pb		
Mussel embryo	S	72h					LC50	1	1--4.4 µg/l Ag		
Salmon sperm							LC50	1	2--11 µg/l Ag		
Sea urchin emb.	S	120h					LC50	1	3--15 µg/l Ag		
Oyster embryo	S	48h					LC50	1	4--19 µg/l Ag		
Sea urchin emb.	S	120h					LC50	1	1--23 µg/l Zn		
Sand dollar sperm							LC50	1	2--28 µg/l Zn		
Sea urchin sperm	S	1h					LC50	1	3--148 µg/l Zn		
Cabezon (sculpin)							LC50	1	4--juv. 191 µg/l Zn		
Oyster sperm							LC50	1	1--0.4 µg/l DDT		
Sea urchin sperm	S	1h					LC50	1	2--1.0 µg/l DDT		
Crab zoea	S	96h					LC50	1	3--1.1 µg/l DDT		
Salmon sperm							LC50	1	4--2.4 µg/l DDT		
Sea urchin sperm	S	1h					LC50	1	1--1.4 µg/l Dieldrin		
Oyster embryo	S	48h					LC50	1	2-- 23 µg/l Dieldrin		
Crab zoea	S	96h					LC50	1	3--24 µg/l Dieldrin		
Mussel embryo	S	72h					LC50	1	4--48 µg/l Dieldrin		
Shiner perch	F-T	96h					LC50	1	1-Endosulfan1.1µg/l		
Salmon smolts	F-T	96h					LC50	1	2-Endosulfan1.7µg/l		
Sand shrimp	F-T	96h					LC50	1	3-Endosulfan2.3µg/l		
Crab zoea	S	96h					LC50	1	4-Endosulfan 15µg/l		
Sand shrimp	F-T	96h					LC50	1	1--Endrin 0.45 µg/l		
Shiner perch	F-T	96h					LC50	1	2--Endrin 0.5 µg/l		
Salmon smolts	F-T	96h					LC50	1	3--Endrin 1.2 µg/l		
Crab zoea	S	96h					LC50	1	4--Endrin 2.0 µg/l		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Hyalella azteca	F-T	10d	1	5 conc. spiked sed.	4/T	10	Survival-LC50 values	1	2-Dieldrin LC_{50} 7.6 µg/l (water only)	•Examination of the dieldrin 10d water only LC50 values for both species indicates that these organisms are two of the most sensitive freshwater invertebrate species tested with dieldrin •Growth was diminished (relative to control) in spiked sediment exposure •Some interaction may occur between sediment/dieldrin/H. azteca, limiting dieldrin exposure to this species	**19**
Chironomus tentans	F-T	10d	1	5 conc. spiked sed.	4/T	5	Survival-LC50 values; growth (wt)	1	1-Dieldrin LC_{50} 1.1 µg/l (water only)		
Photobacterium phosphoreum	S	15min			2		Bioluminescense	1	1-ave. rate = 1.53 for 5 metals & 8 compounds	•Protocols: Microtox-Beckman Instruments Interim Manual No. 11067B-9-80; *S. volutans*-Bowdre & Krieg 1974; P. fluorescens-ATCC-13525; *A. hydrophila*-ATCC-23213 •Metals tested: Hg, Zn, Cu, Pb, Ni •Other compounds tested: ∝Napthol, NaLauryl SO_4, NaCyanide, 3,3 dichlorobenzidine, Phenol, N-Nitrosodiethylamine, dichloromethane, Nitrotriacetic Acid	**16**
Spirillum volutans	S	18h	1 pos. & 1 neg.	log 4 conc.	2	25ml	Opticle density	1	2-ave. rate = 2.38 for 5 metals & 8 compounds		
Pseudomonas fluorescens	S	18h	1 pos. & 1 neg.	log 4 conc.	2	25ml	Opticle density	1	3-ave. rate = 2.61 for 5 metals & 8 compounds		
Aeromonas hydrophila	S	18h	1 pos. & 1 neg.	log 4 conc.	2	25ml	Opticle density	1	4-ave. rate = 3.76 for 5 metals & 8 compounds		
Hyalella azteca	S F-T	29d 29d	1 1	3 3	4 4	20 20	Mortality & growth (lgth)	1	1-Waukegan sediments 2-Homer sediments	•Survival of amphipods reduced in Waukegan treatments (10d & 29d), but not in control treatments; body length not affected; delay in sexual development observed in Waukegan-2 sediments after 29d F-T tests •Confounding factors--Sensitivity of *C. riparius*-1st instar 1-27x more sensitive than 4th instar to Cu and 127x more sensitive than 2nd instar to Cd exposure; indigenous animals may have affected emergence	**21**
Chironomus riparius	F-T	29d	1	3	4	50	Emergence	1	2-Waukegan sediments 1-Homer sediments		
Neanthes arenaceodentata		20d	1	12w/3 dilution	5		Mortality & biomass	1	2-mortality (25%) 2-biomass (41.6%)	•Statistical sensitivity-the number of statistical differences from the reference (control)--reported here as a percentage in the rank column •The authors classify *N. arenaceodentata* as moderately sensitive •Sediment contaminantes included 9 metals (Commencement Bay), 7 low molecular weight PAHs, 9 high molecular weight PAHs (Eagle Harbor), and polychlorinated biphenyls (Harbor Island)	**22**
Rhepoxynius abronius		20d	1	12w/3 dilution	5		Mortality		1-mortality (50%)		
Eohaustorius estuarius		20d	1	12w/3 dilution	5		Mortality & reburial		1-mortality (50%) 3-nonreburial (8.3%)		
Photobacterium phosphoreum							Bioluminescence		1-red. bioluminescence (100%)		
Photobacterium fischeri	F-T	5 min	1 blank	7		10 µl	Bioluminescence	1	2-mean = 0.75%	•Processed water from shale retorting & coal gasification used as toxicant substances •Rainbow trout were more sensitive to processed waters than luminescent bacteria (EC_{50} & LC_{50} correlation coeff. = 0.82, F-T) •Fathead minnow & bacteria were not significantly correlated (EC_{50} & LC_{50} correlation coefficient = 0.41, F-T) •Bacterial system more sensitive than fathead minnow in static tests (EC_{50} & LC_{50} correlation = 0.97) •Fish more sensitive to phenolic compounds than bacteria	**23**
Oncorhynchus mykiss	S	5min	1 blank	9		10 µl		1	1-mean = 1.20%		
Pimephales promelas	F-T			7			Mortality (LC50)	1	1-mean = 0.44%		
	F-T S	96h 24h		7 9			Mortality (LC50)	1	2-mean = 0.75%		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Photobacterium phosphoreum		5 min	1	35 com-pounds	3-5		Bioluminescence	1	2-Microtox generally more sensitive than the Tchan proceedure	•Compounds included metals (e.g., Hg, Na, Zn, Cd, Be, K, Ni), DMSO, hydrocarbons, herbicides Protocols: proceedure of Tchan utilizing a bacterium & an alga, & Microtox by Beckman Instruments, Inc.	**25**
Photobacterium phosphoreum & *Dunaliella tertiolecta*		1h & 24h	1	35 com-pounds	3-5	11.25 ml each spp.	Bioluminescence	1	3-least sensitive in general than Microtox or fish survey; exception-some herbicides were more sensitive	•Most of the test compounds were detected at lower concentrations by fish bioassays and implies that the microbial bioassays are less sensitive	
Fish species							LD50	Many (lit. search)	1-generally more sensitive than either microbial bio-assays		
Rhepoxynius abronius	S	10d	2	15 samples	5	20	Survival, avoidance	1	1-survival most sensitive; reburial & avoidance insensitive	•*M. edulis*-high concordance w/*R. abronius* tests; highest precision & discriminatory power •*R. abronius*-relatively high discriminatory power & intermediate precision •*A. abdita*-high analytical precision; not correlated w/sedimentological variables & relatively high correlation w/several toxicants •*S. purpuratus*-responsive to benzo(a)pyrene exposure resulting in cellular effects; intermediate sensitivity, precision and discriminatory power •*D. gyrociliatus*-resistant to endosulfan & pentachlorophenol; survival insensitive; intermediat sensitivity, precision & discriminatory power w/relatively high correlation w/PAHs	**24**
Ampelisca abdita	F-T	10d	2	15 smpls	5	20	Mortality, avoidance	1	2-surv. insensitive (< *R. abronius*); avoid. not sensitive		
Mytilus edulis	S	48h	2	15 smpls	5	100-900	Survival, norm development	1	1-embryo test high; abn. develpmt most sensitive		
Strongylocentrotus purpuratus	S	20min	4	15 smpls	5	20	Cell changes, fert., abnorm developmt	1	3-< *M. edulis* for development-inter-mediate sensitivity		
Dinophilus gyrociliatus	S	7d	5	15 smpls	5	20	Survival, eggs/female	1	4-rel. resistant to some chem.; eggs sensitive to complex effluents		
Leptocheirus plumulosus		10-20d	1	CC % 100,50, 25,12.2			Survival	1	1 (J) 2 (A)	•Temperature variations (2°C recommended) & feeding regime affect the growth and reproduction of *L. plumulosus* •Increase the number of tests/concentration, & number of replicates •Significant positive relationship between length & dry weight in *L. plumulosus* •Juveniles more sensitive to use for sublethal endpoints	**26**
		10-40d	1	12.5,6.3,3.1, 1.5			Growth (Lgth,Wt)				
Hyalella azteca		10-20d	1	100,50, 25,12.2	3	20	Survival	1	3 (J)		
Rhepoxynius abronius		4d 10d					Mortality LC50	1	2-30 to 100x more tolerant to TBT	•*R. abronius* 4d LC_{50} = 138 ppm, 10d LC_{50} = 29 ppm •*E. washingtonianus* LC_{50} = 4.7 ppm, 10d LC_{50} = 0.3 ppm •The decrease from 4 to 10 day was substantial, indicating that tributyltin is a slow-acting toxicant •Survivors exhibited differences in reburial behavior related to their tolerance of the toxicant •BCF for *R. abronius* (2000-3000) was an order of magnitude lower than for *E. washingtonianus*	**27**
Eohaustorius washingtonianus		4d 10d					Mortality LC50	1	1-non-tolerant to TBT		

Test Organisms	Test Type	Test	Treatments				Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
			Control	Exp.	Rep.	# per					
Hyalella azteca	F-T	10d	1	5 conc.		10	LC50	1	1-Cu, Zn, Cd, Ni, Pb, DDT, DDD, DDE 2- Dieldrin, Chlorpyrifos 3/4-Cd w/other spp. in WQCD & AQUIRE	•Water only tests in soft water (hardness 40 mg/l as $CaCO_3$) •Toxicity of the three species was compared to the genus mean acute & chronic data found in US EPA WQC documents & AQUIRE data base-they are reasonalbly represent the range of sensitivities of other aquatic species, & occasionally are amone the most sensitive species •No one species was consistently most sensitive to the 10 toxicants •It appears that *H. azteca* is more sensitive to broad spectum toxicants than the other two species •L. variegatus was never the most sensitive species tested; but more sensitive than midges for Cu & Pb •Sensitivity is a function of the test species &test regime: Static or flow-through, nominal or measured chemical concentrations? •Ranks are reported for metals from Water Quality Criteria Documents (WQCD) & AQUIRE data bases	**30**
Chironomus tentans	F-T	10d	1	5 conc.		10 or 5	LC50	1	3-Cu 2-Zn, DDT, DDD, DDE 1-Dieldrin, Chlorpyrifos 3/4-Zn, 5/7-Cu w/other spp. in WQCD & AQUIRE data base		
Lumbriculus variegatus	F-T	10d	1	5 conc.		10	LC50	1	1-Cu 3-Zn 2-Cd, Ni, Pb, DDE 3/4-Ni, 4/4-Zn, 4/5-Pb, 4/7-Cu w/other spp. in WQCD & AQUIRE		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Photobacterium phosphoreum							Bioluminescence		1-(15/27) organics 2-inorganic (pure) 1&2-oil refinery wastewater 1-fossil fuel water 4-oil sands tailing 1-sewage plant efflu. 1-industrial efflu. (4/5) 2-industrial efflu. (1/5) 1-sedimt. contam. (2/8) 2-sedimt. contam. (4/8) 3-sedimt. contam. (3/8) 3-sedimt. extracts	•The evaluations were made on chemical groupings. •Microtox was more sensitive than or as sensitive as the acute lethality tests for pure organics, but less sensitive to most inorganics. •Microtox not as sensitive to effluents or leachates w/high components of insecticides, herbicides, inorganics, pharmaceuticals or textiles, or highly lipophilic contaminants. •As complexity and toxicity of industrial effluents increased, the sensitivity of Microtox increased. •Limitations of the available data: variation of sensitivity with the compounds and organisms tested; lack of standardization in approach and failure to provide details; published reviews commonly failed to identify fish species and test details for cited data. •Data and studies available for comparison were separated by chemical groupings into organic chemicals, inorganic chemicals, municipal wastes, oil refinery wastewaters, industrial effluents/complex wastes, and sediment extracts or landfill leachates.	**28**
Daphnia							Acute lethal		1-(15/27) organics 1-inorganic (pure) 1-oil refinery water 2-oil sands tailing 1-industrial efflu. (1/5) 3-industrial efflu (4/5) 1-sedimt. contam. (5/6) 2-sedimt. contam. (1/6)	•Micotox has a lower CV than many other bioassays with decreasing variablility with increased toxicity. •Possible that lyophilized bacteria show an increase in sensitivity due to toxicants (multichlorinated organics) crossing cell membranes. •Variability of Microtox is higher w/metals than organics. •Microtox good for field screeningindustrialized effluents, greater sensitivity to effluents than pure compounds, more sensitive to commercial formulations of herbicides and pesticides than pure chemical forms. •Increased sensitivity to effluents may be due to pH, NH_3, and phenols. •Conclusions: Organics--as sensitive as acute lethal tests to most pure substances, more sensitive to complex compounds. Inorganics--not as sensitive as Daphnia and less sensitive than rainbow trout and fathead minnow, except Hg, As, and Co. Municipal wastes--favorable results w/Microtox. Petroleum wastes--useful test. Industrial wastes--good for screening for relative toxicity. Other wastes--less sensitive to insecticides, herbicides, pharmaceutical effluents, textile effluents, and highly lipophilic contaminants.	
Oncorhynchus mykiss							Acute lethal		1-(13/27) organics 4-inorganic (pure) 1-oil sands tailing 2-fossil fuel process 1-industrial efflu. (1/4) 2-industrial efflu. (3/4) 4-sedimt. contam. (1/1)		
Pimephales promelas							Acute lethal		1-(13/27) organics 3-inorganic (pure) 2-oil refinery wastewater 3-oil sands tailing, refinery, fossil fuel wastewater 2-industrial efflu. (1/1) 1-sedimt. contam. (1/2) 2-sedimt. contam. (1/2) 2-sedimt. extracts		
Amphipod									1-sedimt. extracts		
Oyster											

Test Organisms	Test Type	Test	Treatments Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Rhepoxynius abronius		10d	1	7/site; 3 stations	5	20	Mortality %	1	1--mortality (high)	•Bivalve test--estimated number of initial larvae not representa-tive of true number of initial larvae; high mortality due to unmeasured aspect; unusual sensitivity due to unevaluated factor	**31**
Dendraster excentricus		72h	2	7/site; 3 stations	5	$2x10^4$	Mortality % Abnorm. %	1	2--mortality (med.) 1--abnorm. (slightly)	•Amphipod test--bioassay appears to be the most sensitive •Echinoderm test--only positive numbers used to calculate the mean mortality (i.e., 3/7 stations)	
Neanthes spp.		20d	1	7/site; 3 stations	10		Mortality %	1	3--mortality (least)		
Crassostrea gigas		48h	1	7/site; 3 stations	5	$2x10^4$	Mortality % Abnorm. %	1	2--abnorm.		
Photobacterium phosphoreum		15min	1 NaCl	4 conc. 6-10 conc.			Bioluminesces		1-trichloroethaniol (ave. rank for organics (ARO)= 3.25)	•Protocols: Beckman Instruments, Inc. 1982. Microtox System Operating Manual. Fullerton CA. p. 8-11; sea urchin embryo-Jackim & Nacci (1984); Nacci & Jackim (1985) ASTM STP 891, pp. 382-394; sea urchin sperm-Dinnel, P., et al. (1983).	**29**
Arbacia punctulata		4h 1h					Growth (E) Mortality (S)		(ARO = 3.75) 1-pentanedione, methyl-propanol, methyl pentanedione, ethocxyethoxyethanol (ARO = 2.50)	•Conclusions: Some chemicals measures similarly by both rapid and acute tests Only parent compounds measured--not metabolic by-products Sperm cell tests overestimated the general toxicity of compounds In metal exposures, rapid tests measured toxicity similar to standard embryo tests, but may not have predictable relationships for acute toxicity in certain marine fish and macroinvertebrates	
Pimephales promelas	FT	96h					Mortality		1-hexachloroethane (ARO = 3.12)	Some rapid tests predicted general toxicity	
Daphnia magna	S	48h					Mortality		1-pentachlorophenol (ARO = 2.00)		
Menidia menidia	S	96h					Mortality		2-silver (ave. rank for metals (ARM)= 4.0)		
Mysidopsis bahia	S	96h					Mortality		1-lead (ARM = 3.6)		
Arbacia punctulata	S	48h					EC50		Ave. metals = 3.2 (E) Ave. metals = 3.2 (S)		
S. purpuratus	S	48h					LC50		Ave. metals = 1.2 (E) Ave. metals = 2.6 (S)		
S. droebachiensis	S	48h					LC50		Ave. metals = 2.6 (E) Ave. metals = 3.2 (S)		
C. gigas	S	48h					LC50		Ave. metals = 1.8 (E) Ave. metals = 1.8 (S)		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
Rhepoxynius abronius	F-T	7d		6		300	Higher conc. of toxic intermediate compound & covalent binding of xenobiotic to macromolecules	1	1	•Higher levels of covalent binding of B*a*P intermediates in *R. abronius* than in *E. washingtonianus* suggest a higher potentialfor disruption of cellular processes and subsequent impairment of the overall health of *R. abronius* in PAH-contaminated sediments •*R. abronius* was shown to be more susceptible than *E. washingtonianus* to contaminated sediments •There were notable differences in relative proportions of individual metabolites produced by *R. abronius* & *E. wasthintonianus* (e.g., higher amounts of B*a*P 7-8-diol to B*a*P-9-10-diol was > for *R. abronius*; B*a*P-7-8-diol is more cytotoxic, mutagenic and carcinogentic)	**33**
							same as above				
Eohaustorius washingtonianus	F-T	7d		6		300		1		•The concentration of [^{3}H]B*a*P in the sediment did not change significantly during the 7d exposure; average value was 5.1 pmol BaP/g sediment dry wt	
									2		
Photobacterium phosphoreum		5 min	1	4 dilution	2	0.5 ml	Bioluminescence EC50	1	2.75-metal average 4/4-NH4, KCN 2/4-Phenol, styrene, chloroform 1/4-dichloroethane 1.5-ave. complex effluents	•Ranking shown in decending order from greatest to least sensitive •The Microtox test generally demonstrated the greatest sensitivity to industrial effluents •In general, the rainbow trout was found to be the most sensitive for several toxicants •The daphnid assay was more sensitive than Microtox when the rainbow trout bioassay was not the most sensitive •Micotox may be a poor indicator of ammonia toxicity •Rainbow trout & Microtox were the most sensitive tests for the organic compounds •Daphnid & spirillum bioassays were generally the least sensitive tests for organic compounds •Rainbow trout were 22 & 30x more sensitive to Zn & Cu than the Microtox test; for these metals, the daphnid was the most sensitive with spirillum the least sensitive, especially for As & Hg •Regarding potassium cyanide toxicity, the rainbow trout was 11, 40, and 87x more sensitive than the spirillum, daphnid, & Microtox assays respectively •Limitations: Microtox proceedure requires the addition of 2% NaCl, & ideally the utilization of freshwater luminescent bacteria would be more desirable The pH of the diluent is distictly different from that of the aquatic environmentsl samples & may affect test relevance The revised Microtox proceedure was not available when this study was conducted Color of the effluents & turbidity may affect the instrument readings for bioluminescence in the Microtox tests	**32**
									2.25-metal average		
Oncorhynchus mykiss	S	96h		log series-dilution	1		Mortality LC50	1	1.6-NH4, KCN 1-Phenol, styrene, chloroform 2-dichloroethane 2.3-ave. complex effluents		
									1.25-metal average		
Daphnia magna	S	48h	1/series	log series-dilution	2	5	Mortality LC50	1	2.3-NH4, KCN 3-Phenol, styrene, chloroform 3-dichloroethane 1.7-complex effluents		
									3.5-metal average		
				log series-dilution					2-NH4, KCN 4-Phenol, styrene, chloroform 4-dichloroethane		
Spirillum volutans		5 min	1 neg. & 1 pos.		2 & 1	0.1 ml/l	Reduced reversal swimming	1			
Armandia brevis		20d	1 ref. 1 posi-tive	17 sites, 4 expo-sure		10	Growth (weight)	1	1-red. growth	•Reduced growth of polychaetes was significantly correlated with increasing proportions of contaminants in sediments, 14/17 •Growth of sand dollars was significantly correlated w/conc. of various contaminants in sediments; as contam. conc. of LPAHs, HPAHs, PCBs, & some individual selected elements increased, sand dollar growth decreased, 14/17 •Signif. amphipod mortality occurred in 2/9 sediments; signif. reduced reburial was not observed; neither mort. nor nonreburial correlated signif. w/sed. contam., org. carbon content, or grain size	**34**
Dendraster excentricus	F-T	28d	1 ref. 1 pos.	17 sites, 2 exp.		20	Growth (lngth/wt)	1	2-red. growth		
Rhepoxinius abronius		10d	1 ref. 1 pos.	9 sites, 1 exp.	5	10	Mortality Reburial	1	1-mortality		

Test Organisms	Test Type	Test	Treatments: Control	Exp.	Rep.	# per	Endpoint(s)	# Labs Testing	Relative Sensitivity Rank	Comments	Ref.
				BRH %						•Interlaboratory comparison study; protocol from EPA600/R-94/025	**35**
Ampelisca abdita	S	10d	1	7,20,33	5	20	Mortality	6	1	•Black Rock Harbor % contaminated sediment tested in experimental groups •All species identified moderate to highly contaminated sediments	
Eohaustorius estuarius	S	10d	1	9,25,42	5	20	Mortality	8	2	•More variability of data for low contam. sediments between labs •All species appropriate for measuring toxicity of sediments	
Leptocheirus plumulosus	S	10d	1	10,28,47	5	20	Mortality	7	3		
Photobacterium phosphoreus		5 min					Bioluminescense	1	1-Most sensitive test	•An effluent study comparing Microtox with fathead minnow and daphnia •No test was able to predict 100% presence of toxicant to another species	**37** as cited in **16**
Pimephales promelas		24h					LC50	1	2-81% agreement w/*P. promelas* & Micotox		
Daphnia magna		24h					LC50	1	3-62% agreement w/Microtox		
Photobacterium phosphoreum		5 & 15 min	KCr_2 ref Sea H_2O	6 sites			Bioluminescence EC50	1	1-solvent extract 2-solid phase 4-pore water	•Microtox tests: pore water, org. solvent extract, solid phase •Only the 3 Microtox and the *R. abronius* useful in assessing the degree of contamination at the regulatory levels of contaminants	**36**
Rhepoxynius abronius		10d	2 sites	5 sites	5	20	Survival LC50	1	3-	•Microtox test was the most sensitive bioassay for sediment toxicity assessment •*R. abronius* was relatively sensitive to the Halifax sediments; most suporting study for the Microtox bioassay	
		96h	$CdCl_2$		1	10					
Corophium volutator		10d	2 sites	5 sites	5	20	Survival LC50	1	5-	•Macoma exposure too short for significant uptake of contaminants •Sample size too small in winter flounder test, but lesions found	
		96h	$CdCl_2$		1	10				•Contaminants in Halifax Harbor: Metals (e.g., As, Cd, Cu, Hg, Pb, Zn); PAHs	
Neanthes sp.		20d	2 sites & $CdCl_2$	4 sites	5	5	Biomass	1	6-	•Protocols: Microtox--supplied by Microbics™ *Neanthes* sp.--Johns et al. 1990 *R. abronius* & *C. volutator*--Sed. Tox. Test. using Infaunal Amphipod *R. abronius* Environ. Canada, 1989 *Macoma*--method developed by Environ. Canada Atlantic Regional Laboratory *P. americanus*--Luna, 1968	
Macoma balthica		30d	1 ref site 1 cont.	1 sites	2	60	Bioaccumulation	1			
Pseudopleuronectes americanus	F-T	2wks	9	3			Histopathology	1			
Hyalella azteca									DDT, DDE, DDD		**38** cited in **20**
Chironomus tentans									DDT, DDE, DDD		
Hyalella azteca	S-R	10d	1 ref.	11	2	10	Survival	1	1-Cu-LC50 @ 31 µg/l	•LC50 values were for water only exposure to copper sulfate •Reduced survival from sediment exposure (Keweenaw Waterway sites: •*H. azteca*--8/11 were signif. toxic from control •*C. tentans*--7/11 were signif. toxic form control •*L. variegatus*--0/11 were signif. toxic from control; reproduction decreased 6/11 signif. diff. from control	**39**
Chironomus tentans	F-T	10d	1 ref.	11	2	10	Survival	1	3-Cu-LC50 @ 54 µg/l		
Lumbriculus variegatus	F-T	10d	1 ref.	11	3	10	Surv. & Reprod.	1	2-Cu-LC50 @ 35 µg/l		
Rhepoxynius abronius	S	10d	12 4 ref.	46	5	20	Survival	1	2-39% signif. mortality over ref. 3-35% signif. abnormal over ref.	•Microtox utilizing the saline extract procedure. •Agreement among the three bioassays was evaluated using correlation analysis. •Protocols--oyster-Chapman & Morgan (1983); amphipod-Swartz et al. (1985); Microtox-Schiewe et al. (1985). •Statistical comparisons showed that the results of the Microtox, oyster embryo and amphipod bioassays were significantly correlated and concurred in determination of the presence or absence of toxic effects in 41% of the sediment samples.	**40**
Crassostrea gigas		48h	5 H_2O 4 ref.	46			Survival Abnormal	1	1-63% signif. decreases over ref.		
Photobacterium phosphoreum		15min	1@0 2ref.tox. 4 ref.	46@100,50, 25,12.4 ref.			Bioluminescence	1			

Codes:

Test type S = static
F-T = flow-through
S-R = static-renewal

Duration d = days
h = hours

Sensitivity Rank 1 = highly sensitive
2 =moderately
3 = less than moderately sensitive
4 = least sensitive

Reps i.e., 3/T = three per treatment

Life Stage J = juvenile
A = adult
E = embryo
S = sperm
N = neonates

Table C-1 References:

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

2 Bailey, H. C., and D. H. W. Liu. 1980. *Lumbriculus variegatus*, a benthic oligochaete, as a bioassay organism. In: Eaton, J. C., P. R. Parrish, and A. C. Hendricks (eds.). Aquatic Toxicology. ASTM STP 707. American Society for Testing and Materials. pp. 205-215.

3 Beaumont, A. R., P. B. Newman, D. K. Mills, M. J. Waldock, D. Miller, and M. E. Waite. 1989. Sandy-substrate microcosm studies on tributyl tin (TBT) toxicity to marine organisms. In: Ros, J. D. (ed.). Topics in Marine Biology. *Scient. Mar*. 53:2-3 pp 737-743.

4 Becker, D. S., and T. C. Ginn. 1995. Effects of storage time on toxicity of sediments from Puget Sound, Washington. Environmental Toxicology and Chemistry 14:5 pp 829-835.

5 Berthet, B., J. C. Amiard, C. Amiard-Triquet, M. Martoja, and A. Y. Jeantet. 1992. Bioaccumulation, toxicity and physico-chemical speciation of silver in bivalve molluscs: Ecotoxicological and health consequences. The Science of the Total Environment 125: pp 97-122.

6 Brady, D., B. Letebele, J. R. Buncan, and P. D. Rose. 1994. Bioaccumulation of metals by *Scenedesmus*, *Selenastrum* and *Chlorella* algae. Water SA 20:3 pp 213-218.

7 Burton, G. A. Jr., B. L. Stemmer, and K. L. Winks. 1989. A Multitrophic level evaluation of sediment toxicity in Waukegan and Indiana Harbors. Environmental Toxicology and Chemistry 8: pp 1057-1066.

8 Cairns, M. A., A. V. Nebeker, J. H. Gakstatter and W. L. Griffis. 1984. Toxicity of copper-spiked sediments to freshwater invertebrates. Environmental Toxicology and Chemistry 3: pp 435-445.

9 Chapman, P. M., R. C. Swartz, B. Roddie, H. L. Phelps, P. van den Hurk, and R. Butler. 1992. An international comparison of sediment toxicity tests in the North Sea. Marine Ecology Progress Series. 91: pp 253-264.

10 Chapman, P. M., R. N. Dexter and E. R. Long. 1987. Synoptic measures of sediment contamination, toxicity and infaunal community composition (the Sediment Quality Triad) in San Francisco Bay. Marine Ecology Progress Series 37: pp 75-96.

11 Curtis, C., A. Lima, S. J. Lozano, and G. D. Veith. 1982. Evaluation of a bacterial bioluminescence bioassay as a method for predicting acute toxicity of organic chemicals to fish. In: Pearson, J. G., R. B. Foster, W. E. Bishop (Eds.). Aquatic Toxicology and Hazard Assessment: Fifth Conference. ASTM STP 766. American Society for Testing and Materials. pp 170-178.

12 De Zwart, D., and W. Slooff. 1983. The Microtox as an alternative assay in the acute toxicity assessment of water pollutants. Aquatic Toxicology 4: pp 129-138.

13 Dermott, R. and M. Munawar. 1992. A simple ans sensitive assay for evaluation of sediment toxicity using *Lumbriculus variegatus* (Muller). Hydrobiologia 235/236: pp 407-414.

14 DeWitt, T., B. McGee, R. Alden, H. Phelps, and F. Pinkney. 1992. A multi-species, multi-lab comparison of the toxicity of a Chesapeake Bay sediment. Poster presented at SETAC Annual Meeting, Cincinnati, OH. November 1992. 13 pp.

15 Dinnel, P. A. 1984. Metholodlogy and validation of a sperm cell toxicity test for testing toxic substances in marine waters. Ph.D. Dissertation, School of Fisheries, University of Washington, Seattle. In: Dinnel, P. A. 1990. Annotated Bibliography of Bioassays Related to Sediment Toxicity Testing in Washington State. Final Report. U.S. Army Corps of Engineers, Seattle District, Seattle, Washington. Contract No. E318900PD. pp. 122-124.

16 Dutka, B. J., and K. K. Kwan. 1981. Comparison of three microbial toxicity screening tests with the Micotox test. Bulletin Environmental Contamination and Toxicology 17: pp 753-757.

17 Fent, K., and P. W. Looser. 1995. Bioaccumulation and bioavailability of tributyltin chloride: Influence of pH and humic acids. Water Research 29:7 pp 1631-1637.

18 Hickey, C. W., and D. S. Roper. 1992. Acute toxicity of cadmium to two speces of infaunal marine amphipods (tube-dwelling and burrowing) from New Zealand. Bulletin Environmental Contamination and Toxicology 49: pp 165-170.

19 Hoke, R. A., P. A. Kosian, G. T. Ankley, A. M. Cotter, F. M. VanderMeiden, G. L. Phipps, and E. J. Curhan. 1995. Check studies with *Hyalella azteca* and *Chironomus tentans* in support of the development of a sediment quality criterion for dieldrin. Environmental Toxicology and Chemistry 14: pp 435-443.

20 Ingersoll, C. G., G. T. Ankley, D. A. Benoit, E. L. Brunson, G. A. Burton, F. J. Dwyer, R. A. Hoke, P. F. Landrum, T. J. Norberg-King, and P. V. Winger. 1995. Toxicity and bioaccumulation of sediment-associated contaminants using freshwater invertebrates: A review of methods and applications. Environmental Toxicology and Chemistry 14:11 pp 1885-1894.

21 Ingersoll, C. G., M. K. Nelson. 1990. Testing sediment toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In: Landis, W. G., W. H. van der Schalie (Eds.). Aquatic Toxicology and Risk Assessment: Thirteenth Volume, ASTM STP 1096. American Society for Testing and Materials, Philadelphia, PA. pp 93-109.

22 Johns, D. M., R. A. Pastorok, and T. C. Ginn. 1991. A sublethal sediment toxicity test using juvenile *Neanthes* sp. (polychaeta: Nereidae). In: Mayes, M. A. and M. G. Barron (Eds.). Aquatic Toxicology and Risk Assessment: Fourteenth Volume, ASTM STP 1124. American Society for Testing and Materials. Philadelphia, PA. pp 280-293.

23 Lebsack, M. E., A. D. Anderson, G. M. DeGraeve, and H. L. Bergman. 1981. Comparison of bacterial luminescence and fish bioassay results for fossil-fuel process waters and phenolic constituents. In: Branson, D. R. and K. L. Dickson (Eds.). Aquatic Toxicology and Hazard Assessment: Fourth Conference. ASTM STP 737. American Society for Testing and Materials. pp 348-356.

24 Long, E. R., M. F. Buchman, S. M. Bay, R. S. Carr, P. M. Chapman, J. E. Hose, A. L. Lissner, J. Scott, and D. A. Wolfe. 1990. Comparative evaluation of five toxicity tests with sediments from San Francisco Bay and Tomales Bay, California. Environmental Toxicology and Chemistry 9: pp 1193-1214.

25 McFeters, G. A., P. J. Bond, S. B. Olson, and Y. T. Tchan. 1983. A comparison of microbial bioassays for the detection of aquatic toxicants. Water Research 17:2 pp 1757-1762.

26 McGee, B. L., C. E. Schlekat, and E.Reinharz. 1993. Assessing sublethal levels of sediment contamination using the estuarine amphipod *Leptocheirus plumulosus*. Environmental Toxicology and Chemistry 12: pp 577-587.

27 Meador, J. P., U. Varanasi, and C. A. Krone. 1991. Comparative toxicity of tributyltin to two species of infaunal amphipods. In: Chapman, P., F. Bishay, E. Power, K. Hall, L. Harding, D. McLeay, M. Nassichuk and W. Knapp (Eds.). Proceedings of the Seventeenth Annual Aquatic

Toxicity Workshop: November 5-7, 1990, Vancouver, B.C. Vol. 2. Canadian Technical Report of Fisheries and Aquatic Sciences No. 1774. pp 894-895.

28 Munkittrick, K. R., E. A. Power, and G. A. Sergy. 1991. The relative sensitivity of Microtox, daphnid, rainbow trout, and fathead minnow acute lethality tests. Environmental Toxicology and Water Quality. 6: pp 35-62.

29 Nacci, D., E. Jackim, and R. Walsh. 1986. Comparative evaluation of three rapid marine toxicity tests: Sea urchin early embryo growth test, sea urchin sperm cell toxicity test and Microtox. Environmental Toxicology and Chemistry 5: pp 521-525.

30 Phipps, G. L., V. R. Mattson, and G. T. Ankley. 1995. The relative sensitivity of three freshwater benthic macroinvertebrates to ten contaminants. Archives Environmental Contamination and Toxicology 28: pp 281-286.

31 PTI Environmental Services. 1991. Reference area performance standards for Puget Sound. Final Report. Prepared for US EPA, Region 10, Seattle, WA and Washington Department of Ecology, Olympia, WA. EPA Contract 68-D8-0085, PTI Contract C744-28. September 1991. 76 pp. + appendices.

32 Qureshi, A. A., K. W. Flood, S. R. Thompson, and S. M. Janhurst. 1982. Comparison of a luminescent bacterial test with other bioassays for determining toxicity of pure compounds and complex effluents. In: Pearson, J. G., R. B. Foster, and W. E. Bishop (Eds.). Aquatic Toxicology and Hazard Assessment: Fifth Conference. ASTM STP 766. American Socienty for Testing and Materials. pp 179-195.

33 Reichert, W. L., B-T. Le Eberhart, U. Varanasi. 1985. Exposure of two species of deposit-feeding amphipods to sediment-associated [3H]Benzo[a]pyrene: Uptake, metabolism and covalent binding to tissue macromolecules. Aquatic Toxicology 6: pp 45-56.

34 Rice, C. A., P. D. Plesha, E. Casillas, D. A. Misitano, and J. P. Meador. 1995. Growth and survival of three marine invertebrate species in sediments from the Hudson-Raritan Estuary, New York. Environmental Toxicology and Chemistry 14:11 pp 1931-1940.

35 Schlekat, C. E., K. J. Scott, R. C. Swartz, B. Albrecht, L. Antrim, K.Doe, S. Douglas, J. A. Ferrette, D. J. Hansen, D. W. Moore, C. Mueller, and A. Tang. 1995. Interlaboratory comparison of a 10-day sediment toxicity test method using *Ampelisca abdita*, *Eohaustorius estuarius* and *Leptocheirus plumulosus*. Environmental Toxicology and Chemistry 14:12 pp 2163-2174.

36 Tay, K. L., K. G. Doe, S. J. Wade, J. D. A. Vaughn and R. E. Berrigan. 1991. Biological effects of contaminants in Halifax Harbour sediment. In: Chapman, P., F. Bishay, E. Power, K. Hall, L. Harding, D. McLeay, M. Nassichuk and W. Knapp (Eds.). Proceedings of the Seventeenth Annual Aquatic Toxicity Workshop: November 5-7, 1990, Vancouver, B.C. Vol. 1. Canadian Technical Report of Fisheries and Aquatic Sciences No. 1774. pp 383-426.

37 US EPA. 1981. Quality Assurance Newsletter. Vol. 4 No. 2.

38 West, C. W., G. L. Phipps, R. A. Hoke, T. A. Goldenstein, F. M. VanderMeiden, P. A. Kosian, and G. T. Ankley. 1994. Sediment core vs grab samples: Evaluation of contamination and toxicity at a DDT contaminated site. Ecotoxicology and Environmental Safety 28: pp 208-220.

39 West, C. W., V. R. Mattson, E. N. Leonard, G. L. Phipps, and G. T. Ankley. 1993. Comparison of the relative sensitivity of three benthic invertebrates to copper-contaminated sediments from Keweenaw Waterway. Hydrobiologia 262: pp 57-63.

40 Williams, L. G., P. M. Chapman, and T. C. Ginn. 1986. A comparative evaluation of marine sediment toxicity using bacterial luminescence, oyster embryo and amphipd sediment bioassays. *Marine Environmental Research* 19: pp 225-249.