Environmental Protection Series

Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant

Report EPS 1/RM/30
September, 1995

Canada

Environment Canada

Environnement Canada

Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant

Method Development and Applications Section
Environmental Technology Centre
Environment Canada

Report EPS 1/RM/ 30
September, 1995

Canadian Cataloguing in Publication Data

Ecological Services for Planning Ltd.

Guidance document on measurement of toxicity test
precision using control sediments spiked with a
reference toxicant

(Report : EPS 1/RM/30)
Issued also in French under title: Document d'orientation
sur la mesure de la précision d'essais de toxicité au
moyen de sédiments de contrôle additionnés d'un produit
toxique de référence.
ISBN 0-660-16426-4
Cat. no. En49-24/1-30E

1. Toxicity testing – Methodology
2. Marine sediments – Analysis.
3. Freshwater ecology
4. Water quality – Measurement – Methodology
I. Canada. Environmental Technology Advancement Directorate.
II. Canada. Environment Canada.
III. Title.
IV. Series Report (Canada. Environment Canada);
EPS 1/RM/30.

QD79.S4E36 1996 363.73'942'01 C96-980108-4

Catalogue No. En 49-24/1-30E
ISBN 0-660-16426-4

Readers' Comments

Comments regarding the content of this report should be addressed to:

Richard Scroggins
Method Development and Applications Section
Environmental Technology Centre
335 River Road
Environment Canada
Ottawa, Ontario
K1A 0H3

Cette publication est aussi disponible en français sous le titre : Document d'orientation sur la mesure de la précision d'essais de toxicité au moyen de sédiments de contrôle additionnés d'un produit toxique de référence. Pour l'obtenir, s'adresser à:

Publications de la Protection de l'environnement
Direction générale de l'avancement des technologies environnementales
Environnement Canada
Ottawa (Ontario)
K1A 0H3

Review Notice

This report has been reviewed by the staff of the Environmental Technology Advancement Directorate, Environment Canada, and approved for publication. Approval does not necessarily signify that the contents reflect the views and policies of Environment Canada. Mention of trade names of commercial products does not constitute recommendation or endorsement for use.

Abstract

This report provides guidance and recommendations for the selection of a reference toxicant, spiking procedures, and use of control sediments for assessing changes in sensitivity of the test organisms to toxicants and measuring precision of both intra- and inter-laboratory spiked-sediment toxicity tests. A reference test, in which a control sediment is spiked with a reference toxicant, may be used within a laboratory or among laboratories to measure the precision of spiked-sediment toxicity tests and to detect differences in the responses of test organisms exposed to sediment-associated contaminants over time. This constitutes one component of a continuous quality assurance/quality control program for sediment toxicity tests.

Four potential reference toxicants (cadmium, copper, pentachlorophenate, and fluoranthene) were considered in the selection process because these were the only chemicals with published information regarding sediment spiking and use in sediment toxicity tests. For tests using freshwater sediment, fluoranthene and copper (as either copper sulphate or copper chloride) are recommended as organic and inorganic reference toxicants, respectively. These recommendations reflect the information currently available and do not preclude the use of other chemicals as reference toxicants as more information becomes available.

Test organisms used in spiked-sediment reference toxicant test (SSRTT) should have well developed procedures for collection and maintenance or culture in the laboratory, as well as, standardized procedures for conducting toxicity tests with whole sediment.

Control sediments to be spiked should enable an acceptable level of survival of test organisms for the duration of the test. Due to lack of specific information, no one type of sediment (e.g., field-collected, formulated, or artificial) is recommended over another. Field-collected sediments are suggested for test organisms with a marrow tolerance range of sediment characteristics; however, for the most part, artificial or formulated sediments are more practical for use in a reference toxicant test. The control sediment should be consistent quality and should be prepared and spiked in exactly the same manner for each test. Wet-spiking methods are recommended for spiking sediments with reference toxicants.

General procedures are described for the wet-sediment rolling, slurry, and sediment suspension techniques. Regardless of the method used to spike the sediment, the homogeneity of the chemical-sediment mixture should be verified before conducting a toxicity test with a reference toxicant. Ideally, the chemical in the sediment should be in a "steady state" equilibrium with the pore water. In the event that a "steady state" equilibrium cannot be determined, the mixing time, contact time, settling time(s), and the time between the addition of the chemical to the sediment and the addition of the overlying water and/or test organisms must be reported. These times are currently being standardized for tests with formulated sediment.

Résumé

Le présent report est un document d'orientation qui contient des recommandations sur la sélection d'un produit toxique de référence et d'une méthode d'addition et sur l'ulilisation de sédiments de contrôle afin d'évaluer les changements de sensibilité des organismes à des produits toxiques et afin de mesurer la précision des essais de toxicité tant intralaboratoires qu'interlaboratoires menés sur des sédiments additionnés. Un essai de référence dans lequel un sédiment de contrôle est additionné d'un produit toxique de référence peut être utilisé dans un ou plusioers laboratoires afin de mesurer la précision des essais de toxicité sur des sédiments addtionnés et afin de mettre en évidence les différentes réactions, dans le temps, des organismes qui sont exposés à des contaminants liés aux sédiments. Il s'agit là d'un aspect d'un programme permanent de contrôle/assurance de la qualité pour les essais de toxicité sur des sédiments.

Quatre produits toxiques de référence (le cadmium, le cuivre, le pentachlorophénate et le fluoranthène) ont été examinés dans le processus de sélection parce que ce sont les seuls produits chimiques pour lesquels il existe des données publiées sur l'utilisation de sédiments addtionnés dans des essais de toxicités sur des sédiments. Pour ce qui est des essais menés avec des sédiments d'eau douce, le fluoranthène et le cuivre (sous forme de sulfate de cuivre ou de chlorure de cuivre) sont recommandés comme produits toxiques de référence organique et inorganique respectivement. Pour ce qui est des essais menés avec des sédiments marins, le fluoranthène et le chlorure de cuivre ouo le cadmium (sous le forme de chlorure de cadmium) sont recommandés comme produits toxiques de référence organique et inorganiques respectivement. Ces recommandations reflètent les données courantes et n'empêchent aucunement le recours à d'autres produits chimiques comme produits toxiques de référence si d'autres données deviennent disponibles.

Les organismes utilisés dans un essai sur des sédiments additionnés d'un produit toxique de référence (ESAPTR) devraient correspondre à des procédures bien élaborées de collecte et de conservation ou d'élevage en laboratorie, de même qu'à des procédures normalisées d'exécution des essais de toxicité sur des sédiments entiers.

Les sédiments de contrôle à additionner devraient donner lieu à un niveau acceptable de survie des organismes étudiés durant tout l'essai. À cause du manque d'information précise, un même type de sédiment (p. ex. recueilli sur le terrain, formulé ou artificiel) n'est pas recommandé plus qu'un autre. Les sédiments recueillis sur le terrain sont proposés lorsque les organismes étudiés manifestent une marge restreint de tolérance aux charactéristiques des sédiments; toutefois, dans la plupart des cas, les sédiments artificiels ou formulés sont plus commodes s'il s'agit d'un essai de produit toxique de référence. La qualitié du sédiment de contrôle devrait être uniforme et celui-ci devrait être préparé et additionné exactement de la même façon pour chaque essai. Les méthodes d'addtion par voie humide sont

recommandées lorsque les sédiments sont addtionnées de produits toxiques de référence. Des méthodes générales sont décrites pour les techniques de rotation des sédiments par voie humide, de boue liquide et de suspension des sédiments. Peu importe la méthode utilisées pour l'addition du sédiment, l'homogénéité du mélange produit chimique-sédiment devrait vérifiée avant l'exécution d'un essai de toxicité au moyen d'un produit toxique de référence. Idéalement, le produit chimique qui se trouve dans le sédiment devrait être en état d'équilibre "permanent" avec l'eau interstitielle. S'il n'est pas possible de déterminer la présence d'un équilibre "permanent", le temps de mélange, le temps de contact, le ou les temps de sédimentation et le temps écoulé entre l'addition du produit chimique au sédiment et l'addition de l'eau sus-jacente et (ou) des organismes étudiés doivent être signalés. Ces temps sont en voie de normalisation pour de qui des essais menés avec un sédiment formulé.

Foreword

Guidance on measurement of test precision using control sediments spiked with reference toxicant is part of a series of guidance manuals and recommended biological toxicity test methods that have been developed by Environment Canada for measuring and assessing the biological effects of toxic substances in freshwater, estuarine, and marine environments.

Recommended guidance and methods have been evaluated by Environment Canada and are favoured:

- *for use in Environment Canada and provincial aquatic and sediment toxicity testing laboratories;*
- *for testing that is contracted out by Environment Canada or requested from outside agencies or industry;*
- *in the absence of more specific instructions, such as are contained in regulations; and*
- *as a foundation for the provision of very explicit instructions as may be required in a legal protocol or standard reference method.*

These reports are intended to provide guidance and to facilitate the use of consistent, appropriate, and comprehensive procedures for obtaining data on toxic effects of samples of chemical, effluent, elutriate, leachate, receiving water, or sediment. The recommendations and methods described within this guidance manual for the selection of a reference toxicant and use of control sediment form the basis for the development of a test in which a control sediment is spiked with a reference toxicant. Recommendations are provided for the spiking of sediments; however, the actual test procedures are specified by the biological test method for a particular species of test organism. Although guidance is provided within this report, key original references should be consulted for details.

Table of Contents

List of Figures

List of Tables

Terminology

Note: All definitions are given in the context of the procedures in this report, and may not be appropriate in another context.

Grammatical Terms

Must is used to express an absolute requirement.

Should is used to state that the specified condition or procedure is recommended and ought to be met, if possible.

May is used to mean "is (are) allowed to".

Might is used to express the possibility that something could exist or happen.

Can is used to mean "is (are) able to".

General Technical Terms

Coefficient of variation (CV) is the standard deviation (SD) of a sample expressed as a percentage of the mean (100 SD/$\bar{x}$).

Control is a treatment in an investigation or study that duplicates the conditions and factors that might affect the results of the investigation, except the specific condition that is being studied. In an aquatic toxicity test, the control must duplicate the conditions of the exposure treatment(s), but must contain essentially no test substance (e.g., trace levels may be present). The control is used to determine the absence of measurable toxicity due to basic test conditions (e.g., grain size, temperature, health of organisms, or effects due to their handling). In a test with a control sediment spiked with a reference toxicant the unspiked control sediment becomes the actual experimental control treatment. In the event that a carrier other than water (e.g., solvent such as acetone) is used in spiking the reference toxicant, a spiked-solvent control is established in addition to the unspiked control treatment.

Homogenize means to make homogeneous by mixing to a uniform consistency and composition.

Quality Assurance (QA) refers to the management and technical practices (e.g., planning, control, assessment, reporting, remedial action) designed to ensure an end product of known or reliable quality.

Quality Control (QC) refers to the techniques and procedures used to measure and assess data quality and the remedial actions to be taken when data objectives are not realized.

Spiking refers to the addition of a known amount of chemical to a clean, control sediment. After the addition of the chemical, the sediment is mixed thoroughly to evenly distribute the test chemical throughout the sediment.

Warning chart is synonymous with control chart and refers to a chart of mean toxicity values prepared for reference toxicant tests by plotting the results of a successive series of tests on a chart where the x-axis represents the test date and the y-axis indicates the endpoint concentration (e.g., LC50, EC50). The chart also indicates a measure of the variability expected in the test results.

Terms for Test Substances

Artificial substrate is a substrate comprised of a synthetic substance (or substances) which is (are) relatively uniform in consistency and composition and which is (are) used in toxicity tests primarily to reduce stress effects associated with water-only toxicity tests with sediment-associated test organisms. It is often used synonymously with formulated sediment, but for the purpose of the guidance provided herein, it is not the same as a formulated sediment.

Chemical, in this report, is any element, compound, formulation, or mixture of chemical substances that might be mixed with, deposited, or found in association with sediment or water.

Clean water is seawater or fresh water that does not contain concentrations of toxicants which cause discernible distress to the test organisms or reduce their survival.

Clean sediment is sediment that does not contain concentrations of toxicants which cause discernable distress to the test organisms or reduce their survival.

Control sediment is a natural, artificial, or formulated, clean sediment or substrate of known physicochemical composition, and is of consistent quality. This sediment must not contain concentrations of contaminants which cause discernible stress to the test organisms or reduce their survival (e.g., trace levels may be present). The use of control sediment provides a basis for interpreting data derived form toxicity tests using test sediment(s) and also as a base sediment for spiking procedures.

Deionized water is fresh water that has been purified to remove ions from solution by passing it through resin columns and/or a reverse osmosis system.

Dilution/control water is the water used to prepare test solutions with specific concentrations of a reference toxicant or other test chemicals for waterborne exposures of test organisms or for spiking sediments. Dilution water is used as a control in waterborne toxicity tests, or as overlying water in a sediment toxicity test.

Distilled water is water that has been passed through a distillation apparatus of borosilicate or quartz glass, or other material, to remove non-volatile impurities.

Elutriate is an aqueous solution obtained after adding water to a solid waste (e.g., sediment, tailings, drilling mud, dredge material), shaking the mixture, then centrifuging or filtering, or decanting the supernatant.

Formulated sediment refers to a substrate that is produced by mixing together naturally occurring particulate, according to a recipe or formulation, to produce a sediment with specific characteristics in which organisms can burrow. It can be used as a control sediment in toxicity tests with reference toxicants.

Intersititial or pore water is the water occupying space between sediment particles. The amount of interstitial water in sediment is the ratio of the weight of water in the sediment to the weight of the whole sediment expressed as a percentage.

Reference sediment is a sample of whole sediment that is collected from a site within the general vicinity of a test sediment (i.e., same body of water). It can be used in toxicity tests as an indicator of localized conditions exclusive of the specific contaminant(s) of concern which may be present in the test sediment. It is frequently selected for biological testing because of tis physicochemical similarity (e.g., particle size, total organic content) to the test sediment(s).

Reference test refers to either a whole-sediment toxicity test in which a control sediment is spiked with a reference toxicant, or a waterborne toxicity test in which a reference toxicant is added to control/dilution water.

Reference toxicant is a standard chemical used to assess the sensitivity of organisms to establish confidence in the toxicity data obtained for a test material. In most instances, a toxicity test with a reference toxicant is performed to assess the sensitivity of the organisms at the time the test material is evaluated, and to assess the precision of results obtained by the laboratory over time. The toxicity test with the reference toxicant is performed in a manner consistent with that of the toxicity test for which test precision is of interest.

Sediment is natural particulate that has been transported and deposited in water. The term can also describe a substrate that has been artificially prepared or formulated from particulate and within which the test organisms can burrow.

Spiked control sediment is a control sediment that has been spiked with a specific amount of reference toxicant to achieve a specific concentration of reference toxicant in the sediment. This sediment serves as a positive control that can be used to determine whether the test organisms respond consistently over time to a specific concentration of a reference toxicant.

Spiked sediment is any sediment to which a test material such as a chemical, a mixture of chemicals, drilling mud, contaminated dredged material, or sludge has been added, and mixed thoroughly, for experimental purposes.

Stock solution is a concentrated aqueous solution of a reference toxicant. Measured volumes of a stock solution are added to a carrier (e.g., acetone) or dilution water to prepare the required concentrations of test solution for spiking sediments.

Substance is a particular kind of material having more or less uniform properties.

Wet-sieved sediment refers to a control sediment that has been sieved using dilution water previously aerated to achieve ≥90 % saturation and adjusted to the desired temperature and salinity (if marine or

estuarine test). The sieved sediment and dilution water mixture should be allowed to settle for a minimum of 12 h (i.e., overnight) before the overlying dilution water is decanted and discarded.

Test sediment is a field-collected sample of sediment, taken from a site thought to be contaminated with one or more chemicals, and intended for use in a bioassay. In some instances, the term may also apply to any sediment sample (including control and reference sediment) used in a test.

Test water is the water place over the layer of sediment in the test vessels. It also denotes the water used to manipulate the sediment, if necessary (e.g., for wet sieving), and as the control/dilution water for waterborne tests with reference toxicants.

Whole sediment or solid-phase sediment is an intact sediment with its associated pore water that has not been sieved. It is not a form or derivative of the sediment such as an elutriate or a resuspended sediment.

Toxicity Terms

Acute denotes events that occur within a short period (seconds, minutes, hours or a few days) in relation to the life span of the test organism.

Acute toxicity means a discernible adverse effect (lethal or sublethal) induced in the test organism(s) within a period of exposure to a test material. The period or exposure is short relative to the life span of the test organism.

Chronic denotes events that occur within a relatively long period of exposure, usually a significant portion of the like span of the organism (e.g., 10 % or more).

Chronic toxicity implies long-term effects that are related to changes in such things as metabolism, growth, reproduction, or ability to survive.

EC50 is the median effective concentration. That is, the concentration of material in sediment (mg/kg or percent by weight) or water (mg/L) that is estimated to cause a discernible sublethal effect to 50 % of the test organisms. In most instances the EC50 (together with its 95 % confidence limits) is statistically derived by analysis of an observed sublethal response (e.g., emergence, reburial in control sediment) for various test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 10 days).

Endpoint is (are) the variable(s) (i.e., time, reaction of the organisms, etc.) that indicate(s) the termination of a test. It can also refer to the measurement(s) or value(s) derived that characterizes the results of the test (e.g., EC50, LC50).

ICp is the inhibiting concentration for a (specified) percentage effect. It represents a point estimate of the concentration of test material that causes a designated percentage impairment in a quantitative biological function such as growth of a test organism. For example, an IC50 could be the concentration estimated to cause a 25 % reduction in growth of chironomid larvae, relative to the control. This term should be used for any toxicological test which measures a change in rate, such as

reproduction, growth, or respiration. (The term EC50, or median effective concentration, is limited to quantal measurements, i.e., number of individuals which show a particular effect.)

IC50 is the median inhibition concentration and represents the concentration of material in sediment (mg/kg) or water (mg/L), that is estimated to be lethal to 50 % of the test organisms. The LC50 and its 95 % confidence limits are usually derived by statistical analysis of mortalities in various test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 96-h LC50).

LC50 is the medial lethal concentration and represents the concentration of material in sediment (mg/kg) or water (mg/L), that is estimated to be lethal to 50 % of the test organisms. The LC50 and its 95 % confidence limits are usually derived by statistical analysis of mortalities in various test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 96-h LC50).

Lethal means causing death by direct or indirect action of the chemical used in the bioassay. For example, death of amphipods is defined as the cessation of movement after gentle prodding or other activity (e.g., a pleopod twitch).

Liquid-phase toxicity test means a toxicity test where organisms are exposed to sediment elutriate, or interstitial water in the absence of sediment.

LOEC is the lowest-observed-effect concentration. This is the lowest concentration of a test material to which organisms are exposed, that causes adverse effects on the organisms. Effects are detected by the observer and are statistically significant.

NOEC is the no-observed-effect concentration. This is the highest concentration of a test material to which organisms are exposed, that does not cause any observed and statistically significant adverse effect on the organism.

Spiked sediment reference toxicant test (SSRTT) is a whole-sediment toxicity test in which test organisms are exposed to a reference toxicant in control sediments spiked with a series of specific concentrations. For the purpose of this report, it can be used synonymously with reference test.

Static describes toxicity tests in which test waters or solutions are not renewed during the test.

Sublethal means detrimental to the test organism, but below the level that directly causes death within the test period.

Sublethal effect is an adverse effect on a test organism, below the level that directly causes death within the test period.

Sublethal concentration is a concentration of test material that does not cause death under defined test conditions.

Threshold effect concentration is calculated as the geometric mean of NOEC and LOEC. Chronic value or subchronic value are alternative terms that might be appropriate depending on the duration of exposure in the test.

Toxicity is the inherent potential or capacity of a material to cause adverse effects toward the exposed organism.

Toxicity test is a determination of the effect of a material on a group of selected organisms or a single species (e.g., *Rhepoxynius arbonius*), under defined conditions. An aquatic toxicity test usually measures either (a) the proportions of organisms affected (quantal) or (b) the degree of effect shown (graded or quantitative), after exposure to a specific test material (e.g., a sample of sediment).

Waterborne toxicity test is a toxicity test in which test organisms are exposed to specific concentrations of a toxicant in dilution water only, in the absence of sediment.

Water-renewal describes tests in which water in test vessels is renewed by frequent intermittent flow.

Acknowledgements

This document was written by Gladys Stephenson (Ecological Services for Planning Ltd., Guelph, Ontario). It is based on existing guidance documents and reports that describe methods and technologies for similar endeavours, as well as the most recent published scientific data and that provided by members of the scientific community who are actively involved in research that is relevant. L. Porebski and R. Scroggins (Environmental Protection Service, Environment Canada) acted as Scientific Authorities and provided technical advice and guidance throughout this endeavour. The contributions of P. Murdoch, A. Murdoch, K.E. Day, and J. Osborne (Environmental Protection Service, Environment Canada), as members of the ad-hoc Technical Sub-committee, are also greatly appreciated. The assistance of Roman Lanno in the preparation of an early draft of this guidance document is also acknowledged.

Members of the Inter-Governmental Aquatic Toxicity Group (IGATG), Regional Ocean Dumping Advisory Committee (RODAC), and the Laboratory Managers' Committee (LMC) (Environmental Protection Service, Environment Canada), who actively participated in the review of the initial draft, are gratefully acknowledged. Particularly useful guidance and technical expertise was provided by S. Yee (Environmental Protection Service, Pacific & Yukon Region, Environment Canada, West Vancouver, BC), and K. Doe and S. Wade (Environmental Conservation Service, Environment Canada, Dartmouth, NS).

Members of the Technical Sub-committee who actively participated in the development and review of this report are thanked accordingly. Special acknowledgment is made of the many technical contributions and information provided by the individuals whose names are listed in Appendix A, and by the following:

G. Ankley (USEPA, Duluth, MN), G. Atkinson (Science Policy Branch, Environment Canada, Ottawa), S. Bay (Southern California Coastal, Long Beach, CA), D. Bedard (Standards Development Branch, MOEE, Toronto, ON), R. Burgess and J. Hetsche (USEPA, Naragansett, RI), G.A. Burton (Wright State University, Dayton, OH), R. Eganhouse (U.S. Department of Interior, Geological Survey, Reston, VA), R. Gambrell (Louisiana State University, Baton Rouge, LA), D. Hansen (USEPA, Naragansett, RI), R. Hesselin (Department of Fisheries and Oceans, Winnipeg, MB), C. Ingersoll and M. Nelson (National Biological Survey, Columbia, MO), K. Keenleyside (Evaluation and Interpretation Branch, Environment Canada, Ottawa, ON), G. Krantzberg (Standards Development Branch, MOEE, Toronto, ON), P. Landrum (NOAA, Great Lakes Environmental Laboratory, Ann Arbor, MI), S. Lorenzato (State Water Resources Control Board, Sacramento, CA), D. McLeay (McLeay Associates Ltd., West Vancouver, BC), J. Percival (Geological Survey of Canada, Ottawa, ON), M. Redmond, J. Jones, D. Schultz, and R. Swartz (USEPA/Asci Corporation, Newport, OR), J. Rodgers (University of Mississippi, University, MI), B. Suedel (E.A. Engineering, Science & Technology, Inc., Hunter Valley, MD), and J. Skei (Norwegian Institute for Water Research, Oslo, Norway).

The comments and suggestions recommended by the reviewers of the final draft document were greatly appreciated:

G. Atkinson (Science Policy Branch, Environment Canada, Ottawa), D. Bedard (Standards Development Branch, MOEE, Toronto, ON), M. Buchman (NOAA, Office of Ocean Resources

Conservation and Assessment, Seattle, WA), G.A. Burton (Wright State University, Dayton, OH), J.W. Gorsuch (Eastman Kodak Company, Rochester, NY), C.W. Hickey (NIWA Ecosystems, Hamilton, New Zealand), M. Hinman (Biomedical Sciences, Inc., East Millstone, NJ), C. Ingersoll (National Biological Survey, Columbia, MO), K. Keenleyside (Evaluation and Interpretation Branch, Environment Canada, Ottawa, ON), K.J.M. Kramer (TNO Environmental Research and Energy, Den Helder, The Netherlands), P. Landrum (NOAA, Great Lakes Environmental Laboratory, Ann Arbor, MI), L. Maltby and C. Naylor (University of Sheffield, Sheffield, UK), G. Pagano (Instiuto Nazionale Per Lo Studio E La Curadei Tumori, Napoli), J. Percival (Geological Survey of Canada, Ottawa, ON), W. Windle (Commercial Chemicals Evaluation Branch, Environment Canada, Ottawa, ON), P.V. Winger (National Biological Survey, University of Georgia, Athens, GA).

Funding and technical guidance that enabled the development of this report were provided by the Marine Environment Division of the Office of Waste Management, Environment Canada, Ottawa through L. Porebski and J. Osborne, and the Environmental Technology Advancement Directorate, Environmental Protection Service, Environment Canada, Ottawa, through R. Scroggins.

Introduction

1.1 Background

Procedures for performing toxicity tests with samples of sediment are being developed by Environment Canada for use in Canadian laboratories for the assessment and management of toxic substances found in freshwater, estuarine, or marine sediments. Whole-sediment toxicity tests are one of a number of methods that can be used to assess sediment toxicity. Standard test methods or guidance are available for whole-sediment bioassays for a number of aquatic organisms (EC, 1992; 1994; ASTM, 1994a;b; 1995; USEPA, 1994 a;b).

Intra- and interlaboratory results from these tests might differ because of differences in sediment characteristics, methods of sediment manipulation, quality of dilution water, genetic history and sensitivity of test organisms, life stage of test organism, training and experience of technicians, etc. Waterborne toxicity tests with reference toxicants have been conducted by laboratories as part of quality assurance/quality control (QA/QC) programs for both waterborne and sediment toxicity tests to ensure that comparable results can be achieved within and between laboratories (EC, 1990). Because of the relative ease of conducting a waterborne toxicity test, the relative speed with which results are obtained, and the lack of guidance for conducting a test with a control sediment spiked with a reference toxicant, waterborne toxicity tests with reference toxicants have traditionally been used in conjunction with sediment toxicity tests. However, these water-only tests fail to assess the potential effects that the absence of sediment might have on tests and do not provide and adequate means to measure or assess organism sensitivity in tests where the endpoint is growth. These effects might introduce a source of unpredictable variability to the test. As a result, waterborne toxicity tests might be considered inadequate as reference tests for toxicity tests with whole-sediment bioassays, particularly for bioassays with growth of the test organism as the biological endpoint. Among other components of laboratory QA, laboratories will be required to conduct reference tests with a control sediment that has been spiked with a reference toxicant to monitor intra- and interlaboratory test precision of spiked-sediment tests.

This report specifically addresses the requirements of a standardized method for conducting a reference test using control sediments spiked using a reference toxicant for use in measuring intralaboratory whole-sediment precision, and monitoring test organism health, and the relative sensitivity of test organisms over time. The reference test may also be used to monitor interlaboratory test precision. Although procedures for selecting a reference toxicant and preparing the spiked control sediment are presented herein, the actual toxicity test procedures will be identical to those specified in the biological test method described for each species of test organism.

1.2 Use of Reference Toxicants in Sediment Toxicity Tests

The primary functions of a spiked-sediment reference toxicant test (SSRTT) are to monitor intralaboratory precision of spiked-sediment toxicity tests with a given species over time and to ensure that the tests organisms are of adequate health and sensitivity. To perform a SSRTT, a control sediment is spiked with a reference toxicant, as described herein, and the spiked-control sediment is then used to perform a toxicity test with a test organism for which

standard methods have been developed. The results from each toxicity test are compared with historical test performances to identify whether they fall within an acceptable range of variability. Data that fall outside established limits trigger a review of potential sources of variability. Variability in toxicity test results might be attributed to the health of the organisms, genetic differences in tolerance to toxicants between batches of test organisms, potential differences in sediment quality, and/or the operational consistency of technicians in both organism and sediment manipulation.

The interlaboratory precision of whole-sediment toxicity tests can be assessed through as series of round-robin analyses using a standard control sediment, an acceptable test species, a standard test method, and a reference toxicant. The test results from each laboratory are compared and a consensus among endpoints generally suggests that the performance of the laboratory is acceptable. The acceptability and precision of the results will be affected by differences in test organisms, interpretation of defined toxicity test procedures, nature of the dilution water, and technician consistency.

1.3 Rationale

At present, there is no standard method for the preparation of a spiked control sediment, nor is a recommended reference toxicant for conducting whole-sediment reference toxicant tests available. The purpose of this report is to recommend suitable reference toxicants and to provide guidance in the preparation of a spiked control sediment for conducting reference toxicity tests. There is a paucity of published data regarding the use of a spiked-control sediment test to monitor whole-sediment bioassay precision in a given laboratory over time. However, additional information or insights were provided by a number of scientists experienced in conducting spiked-sediment bioassays (Appendix A).

Copper (Cu) -spiked control sediment tests have been used to monitor whole-sediment bioassay precision during 10-d growth tests with *Chironomus tentans* (Geisy, 1992). However, the most common practice has been to use water-only reference toxicant bioassays to monitor variability in the response of organisms used in whole-sediment toxicity tests (Paine and MacPherson, 1991; MacPherson, 1992). Liquid-phase reference toxicity tests might be useful with epibenthic test organisms such as *Daphnia magna* or *Hyalella azteca*, but test organisms such as *Chironomus* spp. or marine amphipods might be stressed by the lack of suitable substrate, which in turn influences the response of the test organism to the waterborne toxicant (Pesch and Morgan, 1978; Burgess *et al.,* 1994). Artificial substrates such as inert glass beads or glass tubes have been used with infaunal test organisms in waterborne toxicity tests with reference toxicants (Day, 1993; Fremling and Mauck, 1980; Henry *et al.,* 1986). In comparative liquid-phase versus substrate tests with *Euhaustorius washingtonianus* and the reference toxicant cadmium chloride ($CdCl_2$), the presence of Cd-spiked substrate (fine-grained sand, resuspended in a range of concentrations of $CdCl_2$) reduced the variability in mortality data and resulted in consistently higher LC50s relative to the concurrent water-only toxicity tests (Yee *et al.,* 1992).

The SSRTT might also be useful for developing new methods and for defining the optimal toxicity test conditions for a particular species of test organism. The guidance provided herein will be updated as the science of spiking sediments progresses. This report is to be used by laboratories conducting spiked-sediment bioassays.

Reference Toxicant Selection

2.1 Recommended Reference Toxicant(s)

The use of the selection criteria presented in Table 1 resulted in the recommendation of copper (Cu) as an inorganic reference toxicant for whole-sediment reference tests. The available published data are for Cu as copper sulphate ($CuSO_4$) and copper chloride ($CuCl_2$). Copper chloride might be a more suitable compound for marine environments since the Cl^- anion is much less reactive with metals in sediments than SO_4^{-2} (Doe and Mudroch, 1994). The addition of sulphate could, under certain conditions, lead to the formation of AVS complexes and a subsequent reduction in bioavailability. Copper is used as a reference toxicant in spiked-sediment toxicity tests for the 10-d *Chironomus tentans* growth test (Geisy, 1992), and for 96-h tests with *Mulinia lateralis* in which mortality and growth are endpoints (Burgess *et al.,* 1994). Data are available to support the use of cadmium (Cd), as cadmium chloride ($CdCl_2$), as a reference toxicant. However, because of concerns regarding carcinogenic hazards to workers involved directly with the activities of spiking sediments, Cd is relegated to our second choice as a reference toxicant. These inorganic reference toxicants are spiked using water as a carrier which eliminates concerns regarding solvent effects on the partitioning of organic chemicals in spiked sediments (Nkedi-Kizza *et al.*, 1985).

There is a paucity of data available regarding the use of organic contaminants in spiked sediment toxicity tests. Fluoranthene has been added with some verification of homogeneity of mixing (Ditsworth *et al.,* 1990), and has successfully elicited responses from test organisms in toxicity tests (Swartz *et al.,* 1990; DeWitt *et al.,* 1989; 1992; Suedel *et al.,* 1993). When an organic reference toxicant is required, in addition to an inorganic, or when an inorganic reference toxicant is inappropriate, fluorathene is recommended.

2.2 Initial Selection of Potential Reference Toxicants

The initial screening of candidate reference toxicants for SSRTTs considered all potential toxicants suggested in Environment Canada (1990). This list included four organic (4-chlorophenol, dodecyl sodium sulphate, sodium pentachlorophenate, and phenol), and nine inorganic [cadmium chloride ($CdCl_2$), copper sulphate ($CuSO_4$), potassium dichromate (K_2CrO_4), potassium chromate ($K_2Cr_2O_7$), potassium chloride (KCl), sodium chloride (NaCl), silver nitrate ($AgNO_3$), zinc chloride ($ZnCl_2$), and zinc sulphate ($ZnSO_4$)] chemicals. However, in most cases, published information regarding the use of most of these chemicals as reference toxicants was available only for waterborne toxicity tests (Paine and MacPherson, 1991). Information on spiked-sediment toxicity tests was limited to Cd (Swartz *et al.,* 1985; Nebeker *et al.,* 1986; Birge *et al.,* 1987; Robinson *et al.,* 1988; DiToro *et al.,* 1990; Green *et al.,* 1993), Cu (Cairns *et al.,* 1984; Malueg *et al.,* 1986; Burgess *et al.,* 1994), pentachlorophenol (PCP) (Lydy *et al.,* 1990), and fluoranthene (DeWitt *et al.,* 1989; 1992; Swartz *et al.*, 1990; Suedel *et al.,* 1993). Therefore, only the latter four toxicants were considered for reference tests with spiked control sediments. This decision is simply a reflection of available information and does not preclude the use of other chemicals as reference toxicants as more information becomes available.

Each of the chemicals was evaluated according to the criteria outlined in Table 1. Other limiting

Table 1 Major Criteria for Selecting a Reference Toxicant for an Intralaboratory Spiked-sediment Reference Test

Criteria	Weighting 1 to 10[1]	Rank and Scoring for Each Potential Reference Toxicant			
		Cd	Cu	Fluoranthene	PCP
Be easily manipulated with the sediment and mix homogeneously	10	10 (2,7,9,18,19)	10 (4,10,16)	9 (2,12,14,17)	5 (13)
Have a reasonable equilibrium time with sediments (hours to days)	10	10 (3,7,9,11,18,19)	10 (4,5,10)	5 (1,12,14,17)	3 (13)
Have good dose-response curve for the desired endpoint with a given test organism	9	10 (3,6,7,11)	10 (4,5,10,16)	9 (1,12,17)	9 (13)
Be easily measured/analyzed accurately and precisely in water, sediments, or test organisms at the levels which elicit biological effects	8	10 (7,8,9,11,12, 18,19)	10 (10,16)	8 (1,12,14,17)	8 (13)
Be nontoxic to workers that are mixing or spiking sediments, or have standardized safety techniques for minimizing exposure	7	7 (15)	10 (15,16)	8	5
TOTAL[2]		**419**	**440**	**341**	**260**

[1] 10 = the highest level of importance

[2] Total, is the summation of the product of the weighting times the rank

(1) DeWitt *et al.* 1989;1992
(2) Ditsworth *et al.* 1990
(3) Nebeker *et al.* 1986
(4) Carins *et al.* 1984
(5) Geisy, 1992
(6) Birge *et al.* 1987
(7) Schuytema *et al.* 1984
(8) EVS Consultants 1991
(9) Ray *et al.* 1980
(10) Malueg *et al.* 1986
(11) Swartz *et al.* 1985
(12) DeWitt *et al.* 1992
(13) Lydy *et al.* 1990
(14) Swartz *et al.* 1990
(15) Environment Canada 1990
(16) Burgess *et al.* 1994
(17) Suedel *et al.* 1993
(18) Di Toro *et al.* 1990
(19) Green *et al.* 1993

criteria were considered, but not included in the evaluation when it became evident that supporting information was lacking. Although the rationale behind most of the selection criteria should be self-evident, some require discussion (see Section 2.3).

An important consideration which was not used as a selection criterion in the evaluation process is that a reference toxicant must elicit a different level of response form healthy and unhealthy test organisms (e.g., different LC50s). Assuming consistent experimental conditions among toxicity tests, this would allow the detection of organism responses that differ from established mean values. Such differences would indicate potential problems with the batch of test organisms used in the reference toxicant test. The detection of differences from established mean values has been assumed to be the primary function of the reference toxicant test. Little research has actually been conducted to verify this assumption in waterborne or liquid-phase toxicity tests (EC, 1990) and no information was available regarding this assumption for spiked control sediment tests. As a result, this criterion could not be used in the process of selecting a reference toxicant.

2.3 Discussion of Selection Criteria

2.3.1 Homogeneity of Sediment-chemical Mixture

One of the most important criteria for the use of a chemical in the preparation of a spiked sediment is that its chemical properties allow homogeneous mixing with the substrate. Although sediments are frequently spiked with metals (Ray *et al.,* 1980; Schuytema *et al.,* 1984), few studies have assessed the homogeneity of chemicals mixed within a sample. In one study a jar-rolling apparatus was constructed to prepare test sediments spiked with Cd or fluoranthene (Ditsworth *et al.,* 1990). In Cd-spiked sediment samples collected along a longitudinal axis of a horizontally lying mixing jar, coefficients of variation ranged from 2.2 to 10.9 % (mean 4.8 %) for Cd levels. The coefficient of variation did not increase with nominal Cd levels (as $CdCl_2$; range 3.5 to 14 mg Cd/kg) added to the sediment. In some cases, significant differences ($p < 0.05$) in Cd concentrations existed among sampling locations within jars. It should be noted that Cd concentrations were determined in the sediment matrix by acid extraction (1:3 HCl) and in no way reflects the bioavailable fraction of Cd. It was suggested by the authors that Cd can be mixed "reasonably well" using these techniques, but their criteria for mixing were never established. It should also be noted that this method of spiking has not been adequately tested with fine-grained sediment and the homogeneity of mixing might be markedly different in sediments with higher silt and clay content.

Burgess *et al.* (1994) measured copper concentrations at three depths (surficial, 4 cm, and 8 cm) in a muffled beach sand to which copper chloride had been added. Copper concentrations were 371.3, 364.3, and 376.5 mg Cu/kg dry sand, respectively, which suggests that the sediment-chemical substrate was relatively uniform. Stemmer *et al.* (1990) also showed low variance between subsample replicates of toxicity endpoints for two mixing methods which suggests that the bioavailability of the chemical has not been affected by the mixing methods and that mixing was effective (i.e., homogeneous).

Ditsworth *et al.* (1990) reported that mixing fluoranthene into one jar of sediment provided a coefficient of variation of 11.5 % across sample locations within the jar and no significant effect ($p > 0.05$) of sample location was found. An average coefficient of variation of 10 % was measured for sediment spiked with dieldrin using the rolling technique of Ditsworth *et al.* (1990) (Ankley, 1993).

The homogeneity of sediment-chemical mixtures produced using the sediment suspension method

(sediment:water ration of 1:3 to 5) was examined on a bulk sediment dry weight basis and the following coefficients of variation for sediment at various concentrations in different experiments were 4.6 ± 2.6 % and 4.7 ± 2.1 % for pyrene and phenanthrene, respectively (Landrum *et al.,* 1991), 6.9 ± 4.5 and 9.3 ± 3.4 % for pyrene and phenanthrene, respectively (Landrum *et al.,* 1992), and 6.6 ± 3.8, 5.8 ± 3.2, 7.8 ± 4.5, and 9.1 ± 5.0 % for pyrene, benzo-*a*-pyrene, hexachlorobiphenyl, and tetrachlorobiphenyl, respectively (Landrum, 1994). These relatively consistent coefficients of variation suggest that the suspension approach to spiking sediments is suitable for achieving homogeneous sediment-chemical mixtures.

It is recommended that the chemical be added to the sediment and mixed until the mixture is uniform in colour, texture, and degree of wetness. The homogeneity of the chemical distribution in the control sediment could be checked by the following nested sampling procedure. Select "*n*" samples from the spiked-control sediment and divide each of these samples after further mixing into "*m*" subsamples. Measure the concentration of the reference toxicant in each of the "*nm*" subsamples. Univariate procedures should be used to test the assumptions of normality of the data and homogeneity of variances. If it is necessary to transform the data, a log (x + 1) transformation is recommended. Apply analysis of variance procedures to the data, or transformed data, and the sources of variation will be **among** and **within** samples. If the material is homogeneous the F-test should not be significant. This exercise should be done to validate the spiking method for each combination of control sediment and reference toxicant. It is not necessary to routinely test the efficacy of the spiking method if it is consistent among batches.

2.3.2 Equilibration of the Reference Toxicant

Once a sediment has been spiked with the reference toxicant of choice, it is desirable to allow the mixture to reach equilibrium before commencing a whole-sediment toxicity test (ASTM, 1994c). The term "equilibration" is used in this report in the context of equilibrium partitioning and refers to the assumption that an equilibrium or "steady state" exists between the chemical sorbed to particulate sediment components and the pore water (Di Toro *et al.,* 1991). The time required to reach equilibrium is an important criterion for the preparation of a spiked sediment, particularly for nonionic organic reference toxicants. Equilibration times for spiked sediments vary widely among studies (Burton, 1991). The duration of contact between the toxicant (organic or inorganic) and sediment particles can affect both the partitioning and bioavailability of the toxicant (ASTM, 1994c; Landrum *et al.,* 1991; 1992; Landrum, 1994). This effect apparently occurs because toxicants initially sorb rapidly to labile sorption sites and subsequently onto more resistant sorption sites (Karickhoff, 1980; Di Toro *et al.,* 1982; Karickhoff and Morris, 1985; Landrum *et al.,* 1992). Kinetically controlled changes in toxicant partitioning between water and sediments might result in changes in bioavailability (Nkedi-Kizza *et al.,* 1985; Landrum, 1989; Landrum *et al.,* 1989), hence the contact time before testing will likely be potential source of variability of test results when conducting SSRTTs. Although, Suedel *et al.* (1993) demonstrated that there was little difference in the 10-d EC50 values for *Hyalella azteca* and *Chironomus tentans* when exposed to fluoranthene in tests with contact times of 5 min and 24 h, contact time before testing requires further investigation. It is also important to recognize that the quantity of toxicant used might exceed the complexation capacity of the test sediment system which is largely influenced by the redox conditions in the sediment, the distribution of particle size, and the organic matter content.

The partitioning dynamics and bioavailability of metals are affected by chemical and physical factors such as oxygen/redox gradients, pH,

temperature, adsorption and grain size (Burton, 1991). Jenne and Zachara (1987) show that a large portion of dissolved metals spiked into a sediment are adsorbed irreversibly to solids within several hours. Using a model anaerobic sediment, Oakley *et al.* (1980) also observed that the kinetics of metal partitioning were rapid, with equilibrium being reached within two to five days. For this reason, studies on the bioavailable (toxic) fraction should be conducted after this initial period. The length of time allowed for equilibration between spiking of a sediment with a metal and initiating the bioassay has not been standardized. Schuytema *et al.* (1984) spiked sediment slurries with Cd based upon conditional adsorption constants (Nelson *et al.,* 1981) to achieve desired final soluble Cd concentrations, but the actual time of equilibration between the spiking of the slurry and the commencement of lethality bioassays with *Daphnia magna* was not given. The 48-h LC50s for *D. magna* in these two studies ranged form 19 to 84 µg Cd/L (calculated free ion activity). Birge *et al.* (1987) mixed Cd-spiked sediment/water mixtures for 5 to 7 days and allowed them to settle overnight before the addition of test organisms. The 10-d LC50 for *Rhepoxynius abronius* exposed to Cd-spiked control sediments was 6.9 mg/kg. No justification for selected equilibration durations was presented and verification of equilibrium was not conducted in these studies. Sawyer and Burton (1994) also spiked a formulated sediment with Cd and found that sorption was stable within an hour of mixing and the 96-h LC50 values for *Hyalella azteca* were consistent with varying storage times.

After spiking sediment/water mixtures with Cu, Cairns *et al.* (1984) monitored equilibration by measuring waterborne Cu levels until a constant value was maintained. Mixture equilibrium was attained in 7 to 42 days, depending on the physicochemical composition of the sediment. It is more common in toxicity tests with sediment to monitor equilibration by measuring Cu concentrations in the pore water until a constant value is maintained (Ankley and Swartz, 1994). Although the equilibration times for metals appear to be in the order of hours to days, the equilibration of hydrophobic organics between water and sediment phases might take days to years (Karickhoff and Morris, 1985; Podoll and Mabey, 1987). Lydy *et al.* (1990) conducted 24-h lethality tests with *Chironomus riparius* 48 hours after test sediments were spiked with PCP. DeWitt *et al.* (1989) allowed sediment and water to equilibrate for 24 h after spiking with fluoranthene before addition of test organisms. More recently, DeWitt *et al.* (1992) allowed sediments spiked with fluoranthene to equilibrate for 5 weeks, at 4° C, before toxicity testing. Despite great differences of the sediments with respect to particulate organic carbon content, substrate grain size, and equilibration time before testing, the maximum difference in 10-d LC50 values was only threefold, ranging from 5.1 to 15 mg fluoranthene/kg (total measured; dry sediment basis). Similar observations were reported by Suedel *et al.* (1993). These results suggest that the equilibration time of fluoranthene can be quite rapid. None of these studies actually verified that equilibrium had been established. Equilibration time can be estimated by measuring reference toxicant levels in samples simultaneously collected from the overlying water, the pore water, and the sediment. When toxicant levels remain constant after sediment spiking, an approximation of equilibration has been reached. The accuracy of this estimation is unknown because of procedural limitations (i.e., the ability to isolate these various components and perform accurate and precise measurements). There are insufficient data available to suggest a single recommended pre-test equilibration duration for the reference toxicants at this point in time. The minimum time for porewater/sediment equilibration should be ascertained for each combination of control sediment and reference toxicant at each exposure concentration. Until these data are available, we recommend that the spiked sediments be allowed to equilibrate for four weeks which is consistent with recommendations from other organizations and agencies (ASTM, 1995; USEPA, 1994a).

2.3.3 Storage of Spiked-control Sediments

Storage conditions (e.g., temperature) and length of storage can also affect the toxicity of spiked sediment samples. Malueg *et al.* (1986) examined the effects of storing Cu-spiked sediments at 5° C or -20° C on toxicity to *Daphnia magna*. The soluble Cu fraction was measured at different times over the 25-week period that these sediments were stored and, in sediments not amended with peat, no consistent trends in soluble Cu were apparent. Levels decreased at one week, returned to initial levels by three weeks, increased two-fold by eight weeks, and returned to initial levels from 12 to 25 weeks. The addition of peat attenuated the release of total and soluble Cu into the water column so Cu concentrations were lower in the water of the test with peat amended sediments; however, no consistent trends in soluble Cu were apparent over time. Stemmer *et al.* (1990) examined the effect of time (post-spiking) and spiking sediment to *D. magna*. Toxicity decreased after 48 hours of storage, remained constant until two weeks post-spiking, and then increased again at three weeks for sediments stored at 4° C. For spiked sediments stored at 20° C, toxicity decreased markedly at 48 h and remained constant for the duration of the three-week study. The two spiking methods used (stirring and shaking), generally produced similar LC50 values (Stemmer *et al.*, 1990).

Field-collected control sediments which have had debris and interfering indigenous organisms removed should be stored at 4 ± 2° C until the control sediment is to be used in a test with a reference toxicant. A sample of the sediment large enough to perform an SSRTT is removed from cold storage and acclimated to the test temperature. All control sediments to be used in toxicity tests should be freshly spiked at the test temperature, allowed to equilibrate for four weeks at this temperature, and used in a test as soon as possible thereafter. Batches of sediments for use in a series of toxicity tests over time should not be spiked and stored for periods longer than six weeks.

2.3.4 Efficacy of Reference Toxicants

Potential reference toxicants for spiked control sediment tests should be characterized by a marked dose-response lethality curve for the desired endpoint with a given test organism. Lethality data are available for organisms exposed to Cd- (Birge *et al.,* 1987; Green *et al.,* 1993; Nebeker *et al.,* 1986; Robinson *et al.,* 1988; Swartz *et al.,* 1985), Cu- (Cairns *et al.,* 1984; Malueg *et al.,* 1986), and fluoranthene-spiked sediment (DeWitt *et al.,* 1989; 1992; Swartz *et al.,* 1990; Suedel *et al.,* 1993). Lethality data were also available for two other chemicals that are sometimes used as reference toxicants in tests with whole sediment, PCP and Se sediments (Lydy *et al.,* 1990). It is important to be aware of avoidance mechanisms by test organisms which can result in a non-linear dose-response relationship. Test organisms have been observed to leave the sediment at higher doses and as a result they are not exposed to a proportionally higher concentration (Kukkonen and Landrum, 1994; Landrum *et al.,* 1994). Organisms may also experience a depression in feeding activity which can affect exposure.

Although the most common endpoint of acute toxicity tests is lethality, sublethal responses during acute and chronic exposures can be used to assess toxicity in tests for some species (e.g., emergence and reburial of marine amphipods, Swartz *et al.,* 1985; growth of the marine polychaete, *Nereis arenaceodentata*, Dillon *et al.,* 1993). Growth has been commonly used as an endpoint in tests where sediment has been spiked with a toxicant (Ankley, 1994). The endpoint used in the SSRTT should be identified in the test protocol for that particular test species.

2.3.5 Analysis and Potential Hazard

Standard methods are available for determining Cu, Cd, and fluoranthene in both liquid and solid matrices (APHA *et al.,* 1989; Ozretich and Schroeder, 1986). The hazard of potential reference toxicants to technicians involved with spiking the sediment samples should also be considered. Of the three recommended

reference toxicants, Cu is the least toxic to humans at concentrations that would be used for spiking sediments. An example of the type of information contained in Material Safety Data Sheets (MSDS) is presented in Appendix C; however, the information on each MSDS is valid for only four years. Valid MSDS information must be readily available to all users of the chemicals.

Selection of Control Sediments for Spiking with a Reference Toxicant

3.1 Recommended Control Sediments

Sediments from two different sources may be used in SSRTTS: 1) natural sediments collected from freshwater, estuarine, or marine environments; and 2) prepared sediments (e.g., artificial substrates or formulated sediments). Due to lack of literature regarding the use of these different sediment types in spiked-sediment reference tests, there is no overwhelming evidence to suggest the use of one type over another at this time. Therefore, as long as the sediment of choice is easily spiked with the reference toxicant, has consistent and characterized physicochemical properties, and sustains an acceptable response criterion in test organisms (e.g., ≥90 % survival for marine amphipods) in the experimental controls, any of the three sediment types may be used.

The type of control sediment might vary with the objective of the test. To monitor the precision of spiked, whole-sediment, toxicity tests over time in a laboratory and enhance the ability to detect changes in test organism response, it is necessary to minimize the variation in physicochemical characteristics of the control sediment to be spiked. This is generally easier to accomplish with the use of a formulated sediment that is freshly spiked with a reference toxicant. An artificial or formulated sediment would be ideal for use in monitoring the health of different populations of test organisms and preparations. They impart a desired consistency in the substrate essential to the measure of the acceptability of the performance of technicians and facilities, as well as, procedures and methods, and eliminate problems with the interpretation of test results that are often attributable to interferences form indigenous organisms. These types of sediment can be standardized as to composition, spiking procedures, and storage, thereby reducing variability between tests, the use of artificial and formulated sediments also has the advantage of reducing both the variability between batches of sediment, which makes them ideal for interlaboratory performance evaluations, and the cost associated with the field collection of sufficient amounts of sediment for routine testing. The advantages and disadvantages of the use of formulated sediments are discussed elsewhere (Stephenson *et al.,* 1994; Suedel and Rodgers, 1994; Walsh *et al.,* 1991).

Field-collected control sediment has traditionally been used for monitoring and assessment purposes and in research. Field-collected sediments might be more suitable for these purposes as they are more likely to reflect actual environmental conditions.

Many organisms that are routinely used in sediment toxicity tests have a wide tolerance of sediment grain size and do not require a specific sediment for survival. These organisms (e.g., *Chironomus riparius, Hyalella azteca*) can often be cultured in the laboratory on artificial substrates (e.g., shredded paper) (Pascoe *et al.,* 1990). Organisms that are difficult to culture and must be field-collected prior to testing (e.g., marine amphipods) often have specific sediment characteristic requirements (e.g., grain size) for survival. For these organisms, either field-collected control sediments or sediments formulated to satisfy their specific requirements are recommended for a whole-sediment reference test.

Information regarding substrate composition requirements is lacking for most marine test species; however, there are considerable data available for freshwater test species (ASTM,

1995; USEPA, 1994a). The detailed substrate composition required for each test species should be outlined in test-method documents specific for that species. For example, Environment Canada (1992) provides known information (or reference to the lack of) regarding the influence of sediment grain size on laboratory survival of each of seven test species of marine amphipods. Future biological test methods and protocols should include, as a minimum, tolerance thresholds and optima for grain size and organic matter content.

Regardless of the source of the control sediment, it must be characterized before being used in a spiked-sediment toxicity test. At a minimum, all types of control sediment should be characterized for total organic carbon content, acid volatile sulphide, particle size distribution, pH, percent water, and the concentrations of the reference toxicant (Environment Canada, 1992). Additional analyses for field-collected sediments that might provide additional information useful for the interpretation of test results include biological and chemical oxygen demand, total ammonia, metals, petroleum hydrocarbons, synthetic organic compounds, and redox potential. Criteria for the selection of a control sediment are presented in Table 2.

Little information was available regarding the stability of field-collected control sediment used for toxicity testing over time. Few laboratories employ a spiked-sediment reference toxicant test on a routine basis. One of the most important criteria for the selection of a suitable control sediment for a reference toxicant test is adequate survival of the test organisms. Sediments for reference toxicant tests should have a grain-size composition and other physicochemical characteristics that allow acceptable survival or performance of the test organisms in the experimental controls (Paine and MacPherson, 1991). The use of artificial or formulated sediments in tests with reference toxicants has also been limited to only a few laboratories and there is little information available regarding their preparation or the survival of control organisms in these substrates. However, if the composition of the sediment is such that it both enables the control animals to survive and thrive and consistently provides a medium with stable physical and chemical characteristics, then these sediments (e.g., artificial or formulated) offer the greatest potential for use in SSRTTs (see Subsection 3.3). Recent research regarding the development of artificial and formulated sediments is summarized in Appendix D.

3.2 Field Collection of Control Sediment

Control sediment can be collected from any uncontaminated depositional area and may consist of fine-grained or coarse (sand) material, as long as the sediment enables the test organisms to survive and thrive in the experimental controls throughout the course of the toxicity test. A control sediment is sometimes a sample of sieved sediment that is collected concurrently with the test organisms. Guidance for the collection of control sediment from freshwater, estuarine, and marine environments is detailed in a companion document (EC, 1994).

Before a sediment can be used as a control sediment, it should be analyzed for its physicochemical properties to assess total organic carbon (TOC), acid volatile sulphide (AVS), and natural background levels of trace elements of interest and to ensure that the grain-size requirements of the test species are satisfied. Standard procedures for testing with the desired test organism should be consulted to determine potential grain-size effects on the survival of test organisms. The collection site is commonly the site where field-collected test organisms are found.

Sufficient control sediment should be collected to provide substrate for enough reference

Table 2 Criteria for Selecting a Control Sediment

1. Be reproducible between tests within a given lab (each new batch must be characterized before use).
2. Give consistent, acceptable control survival, as defined by the standardized test method, for the duration of the test.
3. Be consistently formulated to specifications (i.e., spiked homogeneously, moisture content, particle size, organic matter content).
4. Have a consistent and effective source of organic carbon.
5. Experience minimal changes in the physicochemical characteristics of the sediment with storage time.
6. Be non-hazardous to the health of workers manipulating the sediment.

toxicant tests to establish the limits of variability on a warning chart (i.e., 15 to 20 tests). When repeated testing over an extended period of time is planned, enough sediment should be stored in the dark at 4 ± 2° C to provide substrate for this extended period of time (e.g., up to a year). Because storage of whole-sediment samples over long periods of time might alter the physicochemical characteristics of the sediment and subsequently affect the partitioning dynamics of the reference toxicant, it is recommended that the characteristics be routinely monitored. The use of formulated or artificial sediment would circumvent the potential changes due to storage of natural sediment over time. This would also obviate having to frequently establish new warning charts since warning charts should be established with each new batch of control sediment. Alternatively, field-collected samples of sediment may be modified to mitigate this problem (Burgess *et al.,* 1994).

Since it must be demonstrated by physicochemical and biological characterization that the quality of sediment has not changed significantly during storage, we recommend that ammonia, pH, total organic carbon, AVS, and background levels of the reference toxicant be measured monthly or when a new subsample of the stored sediment is prepared for spiking.

3.3 Preparation of Artificial or Formulated Control Sediment

There are insufficient data on the preparation of either artificial or formulated sediments for use in whole-sediment toxicity tests to recommend one over the another, because they have not been used routinely in toxicity assessments. However, there is a growing awareness of the need (i.e., usefulness) for this type of test substance and, as a result, a number of on-going investigations have as their goal the development of a substrate that would be suitable for toxicity testing. These research programs have resulted in a number of preliminary “recipes” for formulated sediments (Appendix D). These formulations generally are comprised of sand, silt, and/or clay, with or without a source of organic carbon. They vary with respect to either the percentage composition of the constituents, or with respect to a formulation procedure.

Although one type of formulated sediment is not being recommended over another, the following attributes should be considered when selecting a formulation for a sediment. Ideally, a formulated sediment should:

- support the survival, growth, or reproduction of a variety of benthic invertebrates
- provide consistent acceptable biological endpoints for a variety of species
- be characterized by a consistency in terms of both the individual constituents in the formulation, as well as, its performance in a test
- be comprised of standard constituents that are readily available to all individuals and facilities
- be free from concentrations of contaminants that might cause adverse effects to test organisms (i.e., only trace levels permitted) (ASTM, 1995).

The general approach that is recommended when formulating a sediment is to

1. select the formulation that will be most appropriate for the test species of interest; this is usually based on texture (e.g., percent sand, silt, and/or clay);
2. select a source of organic carbon; and,
3. combining the constituents in a manner which will produce a homogeneous substrate in which test organisms will survive, grow, and reproduce.

The formulated sediment should have physicochemical characteristics within the tolerance limits of the test organisms (USEPA, 1994a). The consistent sources of the formulation constituents must be reliable and readily accessible to all (Appendix D).

General procedures recommended for generating a freshwater formulated sediment include washing sand with distilled water, drying the sand at 105° C, and sieving to retain the desired particle sizes such that a ratio of 2:1, fine silica sand (125 to 250 µm):medium sand (250 to 500 µm) is obtained. Crushed silt (1 to 2 µm) and/or clay (Allen R Clay or kaolin; median particle size 1.3 µm) can be obtained commercially. The clay should be ashed at 450° C for 1 h to remove the organic carbon often associated with the clay particle fractions (Suedel and Rodgers, 1994). The source of organic carbon used in the formulation of freshwater sediment has been variable; however, the majority of testing laboratories tend to use sphagnum peat, aged cerophyl (Hamr *et al.,* 1994) or conditioned maple leaves (Kemble *et al.,* 1994). Preliminary testing with alpha-cellulose (Ribeiro *et al.,* 1994) as the organic carbon source in formulated sediments indicates that it is an adequate, reliable, and consistent source of carbon that is free of the secondary effects that characterize the other carbon sources (e.g., lowered dissolved oxygen concentration, increased ammonium concentrations). Alternative types of organic matter for formulated sediment are presented in Appendix D. Aged cerophyl is currently recommended as the carbon source in formulated sediment. The mixture should be buffered with either $CaMg(CO_3)_2$ or $CaCO_3$. All constituents are mixed on a percent dry weight basis in the following ratios: sand (75 %), silt/clay (20 %), organic matter (4 %), buffer (1%), or in a ratio that is necessary to meet the requirements of a particular test species. These recommendations are based on the relatively recent research (Appendix D).

The constituents, excluding the organic matter, are mechanically mixed together while dry, then hydrated with an appropriate dilution water in a ratio 1:1 of 600 g of dry substrate to 600 mL of water. The mixture is agitated in a glass container (e.g., flask or jar) for 24 h then allowed to stand for 3 days at room temperature (23° C).

The overlying water should be aerated gently during this time. After three days, add the solution of "aged" cerophyl and the buffer. To age the cerophyl, hydrate with dilution water to which is added a 2-mL inoculum of field-collected control water. The inoculum serves as a source of bacteria. The cerophyl should be aged for a minimum of 7 d and not longer than 21 d, during which time the decomposition of the organic material by the mircoorganisms occurs. This overcomes problems associated with unacceptably high concentrations of ammonia and low dissolved oxygen concentrations which might develop during chronic tests when the cerophyl is not aged. The substrate should be mixed for an additional 1 h, after which the contents of the container are allowed to settle overnight. Excess overlying water is decanted and discarded. It is preferable to spike the formulated sediment with a reference toxicant as soon as possible. However, it can be stored at 4° C in the dark, for up to 7 days.

Only a few published studies regarding the preparation of artificial test sediments for used in reference toxicant tests were found; however, several research scientists provided additional information regarding the development of formulated sediments for toxicity testing. Most of this research addresses the development of formulated sediments and includes only a few estuarine/marine sediment formulations. Although some of these data are summarized (Appendix D), the reader is encouraged to consult the original papers.

Artificial sediment should satisfy specific requirements with respect to particle size, organic matter, pH, etc., dictated by the optimal conditions for survival of the test organisms. Chemical characterization of the artificial control sediment should be performed routinely on a monthly basis or more frequently as required.

Test Organism for Spiked-sediment Reference Toxicant Tests

All species of test organism used in SSRTTs should respond in a consistent, reproducible manner when exposed to a reference toxicant in the control sediment. Each test species should have standardized test methods for performing sediment toxicity tests (i.e., standards, guides, protocols, biological test methods).

Chironomus tentans, C. riparius, and *Hyalella azteca* are good examples of representative infaunal (chironomids) and epibenthic (amphipod) test organisms for freshwater SSRTTs. These organisms have well-defined test methods, are relatively easily cultured, and exhibit acceptable control survival in a wide range of sediments. They have also proven to be more sensitive to fluoranthene than those species traditionally used in water-only toxicity tests (Phipps *et al.,* 1994; Suedel and Rodgers, 1994). For saltwater or estuarine tests, any of seven species of marine amphipods recommended by Environment Canada (1992) should prove suitable for SSRTTs.

Standard methods for use of *H. azteca, C. riparius, C. tentans,* and *Hexagenia* spp. are being developed by both the ASTM (as annexes to ASTM E 1383, ASTM, 1994b), the USEPA (1994a), and Environment Canada (1997a; 1997b). As laboratory culture and test and maintenance procedures are improved, other test organisms, such as *Hexagenia limbata, Diporeia* spp., *Lumbriculus* sp., *Tubifex tubifex,* and *Neanthes* spp. may be used for SSRTTs. Procedures for conducting new sediment assays using other test organisms need to be fully standardized to reduce the variability of results obtained in toxicity tests with a chemical (e.g., reference toxicant) and control, reference, or test sediments.

For each test organism used for toxicity testing by a laboratory, a separate set of SSRTTs must be conducted. As an example, if a marine amphipod (e.g., *Eohaustorius washingtonianus)* were used in testing for an Ocean Dumping Permit, SSRTTs must be conducted with that specific amphipod, not a related species. In addition, the field-collected organisms used in the SSRTTs should be collected from the same source as the organisms being used in the whole-sediment toxicity tests. The organisms should be acclimated to the test conditions before conducting a test as described in the toxicity test method for that particular species. Test organisms should be of a similar age or size class. If animals are collected from the field for laboratory tests, they ideally should be available year-round and exhibit a consistent response to spiked-control sediments.

Acquisition and Handling of Chemicals

All reference toxicants considered here (Section 2) are inexpensive and readily available in high purity. Recommended procedures for handling and storage may vary from chemical to chemical. In general, protective clothing is always advisable (e.g., gloves, safety goggles), contact with skin should be avoided, and inhalation or harmful vapours should be prevented by use of a respirator or fumehood. All chemicals should be stored in well-labelled containers, in a cool, dry, ventilated area, away from reactive materials or flame. Valid Material Safety Data Sheets (MSDS) should be readily available to all personnel using the chemicals. They are available from the supplier and contain useful information on the safe handling and storage of chemicals. They also provide information on the proper methods for disposal of the chemicals. An example of the type of information contained in a MSDS is provided in Appendix C.

Universal Procedures for Conducting Reference Toxicant Tests with a Control Sediment

Guidance for the control of toxicity test precision for exposure of organisms to reference toxicants in water is provided by Environment Canada (1990). Some of the information provided (EC, 1990) on the use of reference toxicants is adapted here for continuity. Methods for accurate preparation of test solutions are outlined in Appendix E of Environment Canada (1990).

Once a sediment is spiked with either copper, cadmium, or fluoranthene, the toxicity test must be conducted according to methods and procedures specified in a standard toxicity test for the test species of interest. Therefore, the test conditions, duration, and measurable endpoints will comply with those recommended in the test method for a particular test organism.

The results of a SSRTT can be used to prepare warning charts (EC, 1990). Warning charts should be developed and maintained for each species used in sediment toxicity tests by a particular laboratory (see Section 7). The following universal procedures can be used in conducting SSRTT with Cu, Cd, or fluoranthene.

6.1 Chemical Properties, Labelling, and Storage

Information should be obtained on the properties of the chemical to be tested, including water solubility, vapour pressure, chemical stability, dissociation constants, and biodegradability. Where aqueous solubility is in doubt or problematic, acceptable procedures used previously for preparing aqueous solutions of the chemical should be obtained and reported. Other available information such as structural formula, degree of purity, nature and percentage of significant impurities, presence and amounts of additives, *n*-octanol–water coefficient (K_{ow} or P_{ow}), and the equilibrium constant K_p, should be obtained (where appropriate) and recorded.

Chemical containers must be sealed, and labelled or coded (chemical name, lot number, supplier, date received) upon receipt. Storage conditions (e.g., temperature, protection from light) are frequently dictated by the nature of the chemical. Standard operating procedures for chemical handling and storage in the laboratory must be followed and guidance is provided on Material Safety Data Sheets (MSDS) available from the chemical supplier.

6.2 Preparation of Spiked Sediment

6.2.1 Handling of Field-collected Sediment Before Spiking

Field-collected sediment samples for use in SSRTTs must be collected, transported, and stored following the procedures outlined in a companion document (EC, 1994) and, if given, in the specified biological test method to be followed (e.g., EC, 1992; 1997a; 1997b). Collected control sediment should be placed in a container made of nontoxic material which can be sealed and transported to the laboratory.

In the laboratory, the sediment should be placed into a sorting tray and hand-picked with tweezers to remove large particles (e.g., rocks and debris) and indigenous organisms which might interfere directly or indirectly with the test organisms. If the microscopic examination for the presence of interfering endemic species reveals that it might not be possible to pick out the organisms, the sediment should be pressure sieved through a mesh with a pore size chosen in consideration of the toxicity test and test organism, predators and/or competitors that might be present, and the

nature of the sediment (e.g., particle size distribution, quantity, and type of debris). The debris and organisms should be removed from the sediments as soon as possible before storage to reduce deterioration of sediment quality form decomposition of dying infauna or organic debris. The sediments should then be stored in sealed containers a 4 ± 2° C until they are to be spiked with a reference toxicant. The field-collected control sediments should be stored for no longer than 12 months.

In the event that the control sediment cannot be sieved with pressure, it can be sieved with water (i.e., wet-sieved). A portion of the dilution water, previously adjusted to the desired test temperature and the required salinity (if marine or estuarine test), and aerated to ensure a dissolved oxygen value ≥90 % saturation, should be used for this sieving. Water overlying the sieved control sediment should be removed carefully and discarded. The objective is to remove the overlying water and discard it without discarding resuspended sediment fines (i.e., silt/clay fraction, < 0.063 mm). Wet-sieved control sediment must be allowed to settle for at least 12 hours (i.e., overnight) and should be kept under previously described storage conditions during the settling period. Storage containers should be stored on an angle to facilitate decanting or siphoning overlying water.

Before spiking, the major physical and chemical characteristics of the sediment should be measured. A subsample of the control sediment should be analyzed using standard methods for a least the following: moisture content, pH, ammonia, total organic carbon, acid volatile sulphide, particle size distribution (percentage gravel, coarse and fine sand, silt, and clay), wet density, and background levels of the reference toxicant (e.g., Cu, Cd, or fluoranthene). Further characterization may include analyses for total volatile residue, porewater salinity (before sieving in the laboratory), biochemical and/or chemical oxygen demand, oxidation-reduction potential (Eh), metals, acid volatile sulphide, total chlorinated organic content, chlorinated organic compounds, and polycyclic aromatic hydrocarbons (ASTM, 1994b;c). It is particularly important to know the total organic carbon (TOC) concentration in control sediments if the reference toxicant is a non-ionic organic compound (e.g., fluoranthene) and the acid volatile sulphide (AVS) concentration in control sediments if the reference toxicant is an inorganic compound (e.g., $CdCl_2$ or $CuCl_2$) (Carlson *et al.,* 1991; Ankley *et al.*, 1993).

The moisture content of the control sediment must be determined before spiking to standardize spiking on a dry weight basis. Precisely weigh (to the nearest 0.1 g) triplicate 2.0 g samples of wet sediment and place them into weighed, aluminum pans that have been previously oven-dried at 105° C for 24 h, allowed to cool to room temperature, and stored in a desiccator. Sediment and samples should be dried overnight at 105° C (Yee *et al.,* 1992), allowed to cool to room temperature is a desiccator, reweighed, and percent moisture determined by difference.

The mean wet density (e.g., mg/cm^3) also should be calculated by determining the wet weight of measured volumes of sediment, in triplicate, and averaging the values.

Depending on the test design and intent, test water (i.e., that used as water overlying sediment in the test) and control/dilution water (i.e., that used to prepare dilutions of test chemicals) may be artificial or reconstituted seawater or fresh water, or an uncontaminated supply of natural seawater or fresh water. If tap water is used, it should be dechlorinated and carbon filtered. Natural or reconstituted seawater can be adjusted to the required salinity, as in biological test methods for sediment, by the addition of dry ocean salts or brine (if too brackish), or distilled water (if too saline) (EC, 1992). Regardless of the choice of test water, it must be demonstrated that the test organisms can survive and thrive (i.e., exhibit normal behaviour). For consideration of the fate and behaviour of the

inorganic reference toxicants when added to seawater consult the summary in Appendix E and the review by McLusky *et al.* (1986).

Control/dilution water must be adjusted to the required test temperature before use. The water should have a dissolved oxygen content ≥90 % of the air-saturation value. As necessary, the required volume of water should be aerated vigorously (oil-free compressed air passed through air stones) immediately before use, and its dissolved oxygen content checked to confirm that ≥90 % saturation has been achieved.

6.2.2 Methods for Spiking Sediment

Primary methods used to spiked sediments with contaminants include wet-spiking and dry-spiking. Air-dried sediments have been successfully spiked in laboratory studies examining dose-response relationships between organic sediment contaminants and bioassay endpoints (Clark *et al.,* 1987; Foster *et al.,* 1987; Keilty *et al.,* 1988a; b). Air drying, however, can result in the loss of volatile compounds and changes in the sediment characteristics, especially particle size. The presence of air and air drying have been shown to change metal availability and complexation (Kersten and Förstner, 1987). Air drying also adds another step to the spiking procedure, increasing the time requirements of an already complex procedure, and risking contamination during handling. The drying of field-collected sediments for spiking is, therefore, not recommended.

Wet-spiking techniques are currently the most acceptable for the preparation of a spiked sediment, and three basic mechanical techniques are available. These methods differ in the amount of water present in the mixture during spiking and are best described as:

a. wet sediment rolling (Ditsworth *et al.,* 1990);

b. slurry spiking (Birge *et al.,* 1987); and

c. sediment suspension spiking (Cairns *et al.,* 1984; Schuytema *et al.,* 1984; Stemmer *et al.,* 1990).

The advantages and disadvantages of these methods have been summarized by Stephenson *et al.* (1994). In addition to these techniques, sediments may be spiked by stirring with a scoop or spatula, as long as the homogeneity of the mixture is verified. In all cases, accurately prepared, measured volumes of a stock solution should be prepared and mixed with the control sediment so that it is distributed throughout the sediment.

a. Wet Sediment Rolling Technique.

The wet sediment rolling technique for conducting SSRTTs required a specific jar-rolling apparatus described by Ditsworth *et al.* (1990). The method has been used to consistently produce relatively large volumes of metal-spiked sediments of a homogeneous nature. It is particularly effective for spiking sediments with nonionic organic compounds. The primary disadvantage is that the mixing apparatus must be constructed or purchased.

The jar-rolling apparatus used by Ditsworth *et al.* (1990) consists of eight parallel, horizontal rollers powered by an electric motor through a reduction gear, belts and pulleys, which rotate cylindrical vessels containing the substrate mixtures. This or a similar apparatus should be used if the wet sediment rolling technique is employed. Mixing is accomplished gravimetrically by slowly rolling the jars at less than the critical speed at which the contents would centrifuge. Gallon-size jars are rolled at approximately 15 rpm, about 14 % of critical speed. Substrates are saturated with dilution water before rolling. Optimally wetted, individual substrate particles adhere to each other and to the wall of the revolving jar until they cascade or tumble down the surface of the substrate mass.

Each jar should first be loaded with the required amount of wet base sediment (calculated mass of

dry sediment required for the test), before introduction of the toxicant. Several 1-cm diameter holes of different depths should be punched into the sediment to provide more surface area for the initial distribution of the toxicant. Each jar load of sediment should be spiked with a pre-determined volume of the stock solution or an equal volume of a dilution thereof. A volumetric pipette should be used to distribute each aliquot onto the top surface and into the holes of the sediment in each jar. Substrates must be spiked sequentially, proceeding from low to high concentrations of toxicant, to minimize the cross-contamination potential. Control substrates must be prepared by adding an equivalent volume of dilution water to a jar loaded with unspiked sediment. After spiking, all jars and their contents should be processed identically.

The rims of the mouth or each jar should be wiped free of particulate, the lids securely seated, and the jars place horizontally on the rollers of the mixing apparatus. Rolling should be done at the temperature stipulated in the toxicity test method. Jars must be closely monitored during the first hour to ensure proper mixing of substrates. After rolling for approximately 15 min, mixing efficiencies of the substrates can be judged visually. If any substrate displays cohesiveness, indicated by agglomerating or balling, all the jars must be opened and an aliquot of dilution water (50 mL) added to each substrate to increase the fluidity. Jars must be recapped and returned to the roller. This procedure should be repeated as necessary until the operator visually observes that all substrates are tumbling without forming balls, Addition of water in small aliquots minimizes the possibility of over-saturating the substrates, preventing them from tumbling, and requiring that excess water be decanted before substrates are loaded into test chambers. Normally, jars can be rolled until the end of the workday (minimally 2 h), removed from the rollers, gently shaken to settle substrate that adhered to the walls, set upright and stored overnight in the dark. Potentially toxic material should not be left rolling unattended. Prolonged rolling (e.g., > 1 week) should be avoided to minimize changes that might occur in the distribution of particle size. Oxidation of the sediments might also occur during rolling. The following morning, jars can be rolled for two hours to remix into the substrate any interstitial water that exuded overnight. Immediately after mixing, appropriate amounts of substrate for the desired test method should be randomly distributed to test chambers. The dilution water can be carefully added to minimize disturbance of the sediment and the test containers should remain undisturbed during the four-week equilibration period.

Although it has been verified that the wet sediment rolling technique can produce a homogeneous mixture of sediment and reference toxicant, data are limited to sediments with low silt and clay content. The efficacy of the method has yet to be demonstrated with sediments with moderate to high silt or clay content.

b. Slurry Technique.

The slurry technique (Birge *et al.,* 1987; Francis *et al.*, 1984; Landrum *et al.,* 1991;1992) requires a minimum of equipment. It differs from the sediment suspension technique with respect to either the ratio of water to sediment or the manner in which the toxicant is added. A sample of control sediment equivalent to 250-g dry weight sediment is placed in a 500-mL Erlenmeyer flask. Using a 25-mL aliquot of distilled, de-ionized water, add a sufficient concentration of the reference toxicant to obtain the desired sediment-enrichment level (mg/kg dry weight basis). Control (unspiked) sediment should receive a 25-mL aliquot of distilled, de-ionized water. Seal the flasks with tin foil and tape and maintain in a shaker for five days (Birge *et al.,* 1987) or shake vigorously for 60 seconds, twice daily for seven days (Francis *et al.,* 1984) to facilitate homogeneous distribution of the added toxicant. After mixing, the sediment suspensions should be centrifuged to remove water. The moisture of the sediment should be

approximately 15 to 20 % after centrifugation. After removal of excess water, the prepared sediment can be placed in the bioassay exposure chambers and covered with dilution water according to the specific test methods. Sediment exposure chambers should be allowed to stand for at least 12 hours before introducing test organisms (Birge *et al.,* 1987), to allow fine sediment particles to settle.

Before this technique is used for SSRTTs, homogeneity of the spiked sediment must be verified. As well, the shaking times required to obtain a homogeneous mixture must be standardized for the reference toxicants recommended herein, and may well be considerably different than the five days suggested. Although the method has been used successfully with organic compounds its usefulness for evaluating toxicity of control sediments spiked with inorganic compounds has yet to be demonstrated.

c. Sediment Suspension Technique.
The sediment suspension technique (Cairns *et al.,* 1984; Schuytema *et al.,* 1984; Stemmer *et al.,* 1990; Landrum and Faust, 1991) is the simplest, requires the least equipment, and homogeneity of toxicant distribution has been verified (Landrum *et al.,* 1991; 1992; Landrum, 1994). Place dilution water (700 mL) and sediment (200 g, dry weight) in a 1-L beaker. Add the desired amount of toxicant dissolved in 100 mL of dilution water to the beaker to achieve a 4:1 (v:w) water-to-sediment ratio. These absolute measures are not mandatory as long as the water-to-sediment ratio is maintained. The mixture should be stirred at a moderate speed with a stir bar, or mechanical stirrer, for a minimum of 4 h. Allow the sediment in the beakers to settle overnight at the appropriate test temperature as specified in the test method. Decent and discard the excess water overlying the sediment and distribute the sediment-chemical mixture to the test containers.

This approach was used to spike sediments with raidolabelled compounds (^{3}H-TCDD and ^{14}C-OCDD) for bioavailability studies with oligochaetes (Muir, 1993). The compounds were added in acetone (1 mL) to water (500 mL) which was subsequently added to a slowly stirring wet sediment(5 kg; wet density 1.6 kg/L). Homogeneity was assessed by measuring subsamples of sediment and pore water. After a 10-d equilibration period, the sediment-chemical mixture was distributed to the test containers and overlying water was added carefully to minimize disturbance of the spiked-control sediment. The concentrations in the sediment in the test containers varied among replicates by 10 % for TCDD and 26 % for OCDD. Although, homogeneity of the test mixture had been demonstrated using the sediment suspension approach, none of the sediment-chemical mixtures assessed included the reference toxicants recommended for the purposes of a SSRTT.

6.2.3 Mixing a Spiked Sediment

The efficacy of the mixing method in all wet-spiking techniques must be verified (i.e., homogeneity of the test mixture) before they may be used in an SSRTT. Three or more subsamples of the spiked sediment should be randomly sampled to determine the content of the substance being tested. A coefficient of variation of ≤ 5 % is desirable, and has been achieved with Cd-spiked substrates prepared using the rolling technique (Ditsworth *et al.,* 1990). The slurry and suspension techniques provide an opportunity for the spiking of the proper amount of sediment for use as individual replicate of a test, while the jar-rolling method is more suitable for spiking larger batches of sediment, which may be divided into pseudoreplicates.

No published studies were found that report the mixing time required to achieve a homogeneous mixture. It has been suggested that the mixing time should be limited to a few hours. It has also been recommended that the spiking temperature

be kept to a minimum (e.g., to the storage temperature of 4° C), to minimize the rapid alteration of the sediment's physicochemical and microbiological characteristics which might occur at higher temperatures, and which could subsequently alter bioavailability and toxicity (ASTM, 1994a). It is recommended that spiking be performed at the test temperature to avert changes in the toxicity and/or bioavailability of the toxicant that might result from changes in chemical equilibria at test temperatures higher than 4° C.

Organic compounds are generally added by means of a carrier solvent such as acetone or methanol to ensure that they are soluble and that they remain in solution during mixing. Metals are generally added in aqueous solutions (e.g., distilled water or saltwater for freshwater or marine sediments, respectively). When organic solvents are sued as carriers, they are often added directly to the sediment (Adams *et al.,* 1985; Muir *et al.,* 1982; McLeese *et al.,* 1980) and the carrier evaporated before addition of water. This approach seems to result in compounds being sorbed to sediment at different sites than when water is used as a carrier (Håkansen, 1984). Word *et al.* (1987) compared several sediment-labelling techniques using methylene chloride, ethanol, and glycine as carriers. They found glycine was the most effective carrier after seven days of mixing. In most cases, the chemical (e.g., fluoranthene) is either coated onto the walls of the flask (i.e., "shell coating", see the following) and an aqueous slurry (sediment in water in various proportions) is added (Schults, 1992), or the carrier containing the chemical is added directly to the slurry. When the sediment-to-water ratio is adjusted for optimal mixing, sediments that are too dense to mix by slurrying the water have been successfully mixed using a rolling mill (Ditsworth *et al.,* 1990; Swartz *et al.,* 1985). It has been suggested that a more homogeneous distribution of the reference toxicant among particles of different sizes within the sediment might be possible when the chemical is added drop-by-drop to the sediment suspension or slurry while it is being mixed (Gambrell, 1993). Regardless of the method of mixing, care should be taken to ensure that a homogeneous mixture is achieved. The use of a polar, water soluble carrier such as methanol has little effect on the partitioning of nonpolar compounds to dissolved organic matter at concentrations up top 15 % carrier by volume (Webster *et al.,* 1990). Another study shows that changes in partitioning of a factor of approximately two might occur with 10 % methanol as a cosolvent for anthracene sorption (Nkedi-Kizza *et al.,* 1985). The effect of carrier volume on partitioning of organic chemicals in sediments is equivocal and the used of solvents might be either directly or indirectly (i.e., increase sediment oxygen demand due to degradation of the solvent by microorganisms) toxic to the organisms, so caution should be taken to minimize the amount of carrier used. In addition, the use of a carrier such as acetone might result in faster equilibration of spiked organics (Schults, 1992). The term "shell-coating" has been used to describe a method of coating the sides of the mixing jars with a toxicant-carrier mixture. The solvent is allowed to evaporate before addition of the sediment. The time between the spiking of the compound and the use of the test sediment is variable (ASTM, 1994c) and seems to affect the biological availability of compounds (Malueg *et al.,* 1986; Landrum and Poore, 1988; Landrum, 1989; Landrum *et al.,* 1992).

If a solvent other than water is used, both a sediment solvent control and a sediment control with water only, must be included in the test. The solvent control must contain the highest concentration of solvent present in the test sediment and must use the solvent from the same batch used to make the stock solution. If a solvent is to be used as a carrier for a reference toxicant, a toxicity test using the same type of sediment and batch of test organisms should be performed to determine whether the growth, survival, or reproduction of the test organisms are related to the concentration of the solvent

above the range that will be used in the reference test. If there is a solvent effect in the range of concentrations that will be used in the reference test, the solvent is unacceptable and an alternative must be found.

6.3 Chemical Confirmation of Sediment Spiking and Determination of Exposure Concentrations

The periodic chemical confirmation of actual toxicant levels spiked into the sediment is necessary to verify that the nominal exposure concentrations represent the actual exposure concentrations in the sediment and pore water and to demonstrate that the toxicant concentrations changed minimally over the duration of the test. Therefore, it is recommended that chemical stock solutions, pore water (isolation of sediment pore water can be accomplished by the methods described in EC, 1994), chemical-sediment mixtures (bulk dry-weight analyses), and test solutions (if studied) be analyzed to determine exact chemical concentrations to which test organisms are exposed. Measuring tissue residues or whole-body residues in both the organisms that die and survive the exposure will also provide useful information for estimating the effective-dose values of critical body residues.

Toxicant levels should be determined in both the pore water and sediment. In instances where chemical concentrations are to be measured, sample aliquots should be taken from the high, medium, and low test concentrations, preferably all of test concentrations, at the beginning and end of the test, as a minimum. These should be preserved, stored, and analyzed according to best proven methodologies available for determining the concentration of the particular chemical in aqueous solution or adsorbed to sediment.

It may be difficult to collect a volume of pore water sufficient for the determination of fluoranthene at the lower concentrations. If only a small volume of pore water is required for chemical analyses, a syringe especially equipped with a filter (0.45 µm) may be carefully inserted into the middle of the sediment column, after the overlying water has been carefully decanted with minimal disturbance to the surface fines, and a volume of pore water slowly extracted. Where larger volumes of pore water are required, the pore water extracted from the spiked sediment in extra test containers by centrifugation at 10 000 G and 4° C, for 30 min with a large capacity centrifuge. Explicit detail regarding the various methods of extraction of pore water are provided in a companion document (EC, 1994; Subsection 2.9.3).

In some circumstances it may not be possible to analyze samples of both the pore water and sediment. In the event that concentrations of the reference toxicant are determined only in samples of pore water, the following guidance is provided. If a **constant fraction** of the reference toxicant is adsorbed to the sediment, the pore water concentrations will remain in a logarithmic series and porewater samples from sediments spiked with low, medium, and high concentrations of the reference toxicants can be collected and the concentrations measured. If however, a constant amount of the reference toxicant is adsorbed to the control sediment (i.e., adsorption is not proportional), then the porewater concentrations will no longer be in a logarithmic series and concentrations of the reference toxicant should be determined in the porewater samples collected from each treatment (i.e., each exposure concentration of the reference toxicant).

The exposure concentrations should be determined at least four times per year or approximately once every six tests, depending on the type and frequency of test being conducted, or when a new batch of animals is used for tests, to confirm that nominal concentrations are acceptable representations of actual concentrations (i.e., within 10 %).

6.4 Preparing Test Mixtures

The number of replicates required in a test is test-specific and depends primarily on the species of test organism and the endpoint of interest. Therefore, the number of replicates should be stipulated in the toxicity test method or the protocol being followed. For example, a minimum number of six replicates (five with 20 amphipods per replicate, and one for monitoring sediment and water quality) must be prepared for each chemical concentration in the marine amphipod sediment bioassay (EC, 1992). If no guidance on the number of replicates is provided in the toxicity test method, it is recommended that no reference test should be conducted with less than three replicates per test concentration.

For reference toxicant tests, at least five concentrations plus a control are normally prepared. An appropriate geometric dilution series may be used, in which each successive concentration of chemical in sediment is at least 50 % of the previous (e.g., 10, 5, 2.5, 1.25, and 0.63 mg/kg); however, all exposure concentrations must be made by directly spiking the sediment. The exposure concentrations should not be achieved by diluting a sediment-chemical mixture with a "clean" sediment. Chemical concentrations in sediment should be calculated and expressed as µg/g or mg/kg dry weight, and those measured in the pore water should be expressed as mg/L (Swartz *et al.,* 1985; 1988). Test concentrations may also be selected form other appropriate logarithmic dilution series (see Appendix B). In order to select a suitable range of effective concentrations, a preliminary or range-finding test which covers a broader range of concentration may be conducted. Concentrations should be selected to elicit responses of 0 and 100 %, as well as, partial responses above and below the 50 % level.

6.5 Testing Frequency

Ideally, SSRTTs should be conducted continuously for each type of acute toxicity test that is being performed by the laboratory to minimize the time lag before detection of an abnormal condition in test organism stocks. This frequency is often impractical with tests involving laboratory-cultured test organisms. The appropriate testing interval should be determined by experience gained in developing a base of reference toxicant data (e.g., after 15 to 20 tests). Testing every three, preferably two, months is recommended as a minimum once warning charts (see Section 7) have been established. For organisms that are not cultured in the laboratory, an additional stipulation is that all stocks be tested upon arrival and just before the stock is exhausted to determine whether:

a. the sensitivity of the stock to the reference toxicant is similar to that of previous stocks; and

b. the sensitivity of the stock to the reference toxicant changed significantly during holding in the laboratory.

Spiked-sediment reference toxicant tests should be conducted more frequently (e.g., when organisms are introduced into the laboratory or protocols are changed) to establish warning limits early in the program. It is recommended that a spiked-sediment reference test be conducted concurrent with each toxicity test during the initial few months, to establish a coefficient of variation (see Section 7) for the test. Once approximately ten tests have been completed without greatly modifying the coefficient of variation, the frequency can be decreased. As a guideline for waterborne reference toxicant tests, a coefficient of variation for a biological endpoint of ≤30 % is generally considered acceptable (EC, 1990). Until more data become available for SSRTTs, it is recommended that a coefficient of variation of ≤30 % be used as the maximum acceptable variation both within and between tests. All laboratories would be well advised not to report the findings of new tests until consistent test results with reference toxicants can be demonstrated.

6.6 Test Observations and Measurements

A qualitative description of each chemical-sediment mixture and of the overlying water should be made when the SSRTT is being established. This might include observations of the colour, texture, and homogeneity of each chemical-sediment mixture, and observations of the colour and opacity of the overlying water. Any change in appearance of the test mixture or overlying water noted during the test, or upon its termination, should be recorded. Daily measurements (e.g., temperature) of the quality of each chemical-sediment mixture being tested (including the control sediment) and of the overlying water should be made and recorded as described in the specific test method being followed.

The concentrations of the reference toxicant should be measured in the sediment and/or pore water (see Subsection 6.3). Unless there is good reason to believe that the chemical measurements are not accurate, toxicity results for any tests in which concentrations are measured should be calculated and expressed in terms of those average measured concentrations determined for the whole sediment and the pore water.

6.7 Test Endpoints and Calculations

In most instances, the primary endpoint for SSRTTs will be an LC50 value (based upon percent mortality). If a suitable range of spiked-sediment concentrations is studied, the data derived for each test concentration can be used to calculate the median lethal concentration (LC50) together with its 95 % confidence limits. To estimate an LC50, mortality data at the termination of the study are corrected for control mortality and combined for all replicates at each concentration. If mortality is not $\geq$50 % in at least one concentration, the LC50 cannot be estimated. If there is no mortality at a certain concentration, that information is used, being an effect of 0 % mortality. However, if successive concentrations yield a series of 0 % mortalities, only one such value should be used in estimating the LC50, and that should be the highest concentration of the series, i.e., the zero-effect that is "closest to the middle" of the distribution of data. Similarly, if there were a series of successive complete mortalities at the high concentrations in the test, only one value of 100 % effect would be used, again the one "closest to the middle", i.e., the 100 % effect at the lowest of these concentrations. Use of only one 0 % and one 100 % effect applies to analyzing the data by computer program or by hand plotting on a graph (see the following). Using additional values of 0 % and/or 100 % might distort the estimate of LC50.

Various computer programs for calculating the test endpoint may be used. Stephan (1977) developed an LC50 program that uses probit, moving average, and binomial methods, and adapted it for the IBM-compatible personal computer. It also calculates LC50s using logit. This program in the BASIC language is recommended, and is available on diskette[1] from Environment Canada (Pacific and Yukon Region, 224 Esplanade St., North Vancouver, BC, V7M 3S7). An efficient micro-computer program for probit analysis is also available from Hubert (1987), and other satisfactory computer and manual methods (APHA *et al.,* 1989; USEPA, 1991) may be used.

The recommended program of Stephan (1977) provides estimates of LC50 and confidence limits by each of its three methods, if there are at least two partial mortalities in the set of data. For smooth or regular data, the three results will likely be similar, and values from the probit analysis should be taken as the preferred ones and reported. The binomial estimate might differ somewhat from the others. If the results do not

[1] Through the courtesy of Dr. Charles E. Stephan (USEPA, Duluth, Minnesota).

include two partial mortalities, only the binomial method functions, and it can be used to provide a best estimate of the LC50 with conservative (wide) confidence limits.

Any computer-derived LC50 should be checked by examining a plot, on logarithmic-probability scales, of percent mortality at the end of the study for the various test concentrations (APHA *et al.,* 1989). Any major disparity between the estimated LC50 derived from this plot and the computer-derived LC50 must be resolved.

A manual plot of mortality-concentration data and derivation of the estimated LC50 are illustrated in Figure 1. In this hypothetical example, there were 100 amphipods (5 replicates of 20 organisms per concentration) tested at each of five concentrations).

This figure was based on concentrations of 1.8, 3.2, 5.6, 10, and 18 mg chemical/kg sediment (dry weight) causing mortalities of 0, 20,40, 90, and 100 % of test amphipods exposed to the respective concentrations for 10 days. The concentrations expected to be lethal to 50 % of the amphipods can be read by following across from 50 % (broken line) to the intersection with the best-fit line, then down to the horizontal axis for an estimated LC50 (5.6 mg/kg). In fitting a line such as that in Figure 1, more emphasis should be assigned to points that are near 50 % mortality. Logarithmic-probability paper ("log-probit", as in Figure 1) can be purchased in technical bookstores, or ordered through them.

Computer programs gave very similar estimates to the example shown in Figure 1. The LC50s (and 95 % confidence limits) were as follows:

Probit analysis of Hubert (1987):		5.56 (4.28 to 7.21)
Stephan (1977):	probit	5.58 (4.24 to 7.37)
	moving average	5.58 (4.24 to 7.33)
	binomial	6.22 (1.8 to 10)

Confidence limits were not estimated in the binomial method, but two concentrations were selected from the test as outer limits of a range, within which the true confidence limits would lie.

Sublethal-effect data derived from multiple-concentration tests can be analyzed to calculate median effective concentrations (EC50s) and their 95 % confidence limits. Separate EC50s should be determined for each of the sublethal responses quantified (e.g., percentage of amphipods emerged from sediment at Day 10; percentage not showing reburial in control sediment at the termination of the test). Statistical procedures for the calculation of these endpoints are according to those described earlier for determining LC50s.

If the reference toxicant were spiked into the control sediment using a solvent carrier, one of the experimental controls would consist of the control sediment spiked with the solvent only, in addition to the experimental control which would consist of the control sediment only (i.e., no solvent, no reference toxicant). The test results would be considered unacceptable if more than 20 % of the test organisms in either treatment were to die during the period of exposure. If both a clean sediment control and a solvent control were used in a test, endpoints determined for each control should be compared statistically. The chi-square test (Steel and Torrie, 1960) may be applied for this comparison. If a statistically significant difference is found between the two controls, only the solvent control may be used for the calculation of the EC50 and/or LC50. If no statistically significant difference is found, the data from both controls may be pooled for meeting the acceptability of the test and as the basis of calculating the LC50 and, as appropriate, EC50 (ASTM, 1994c). If there is a statistically significant difference between the solvent control and the non-solvent control, then the test should be repeated using a reference toxicant that either requires no solvent or a solvent that does not cause a response that differs from that of the non-solvent control.

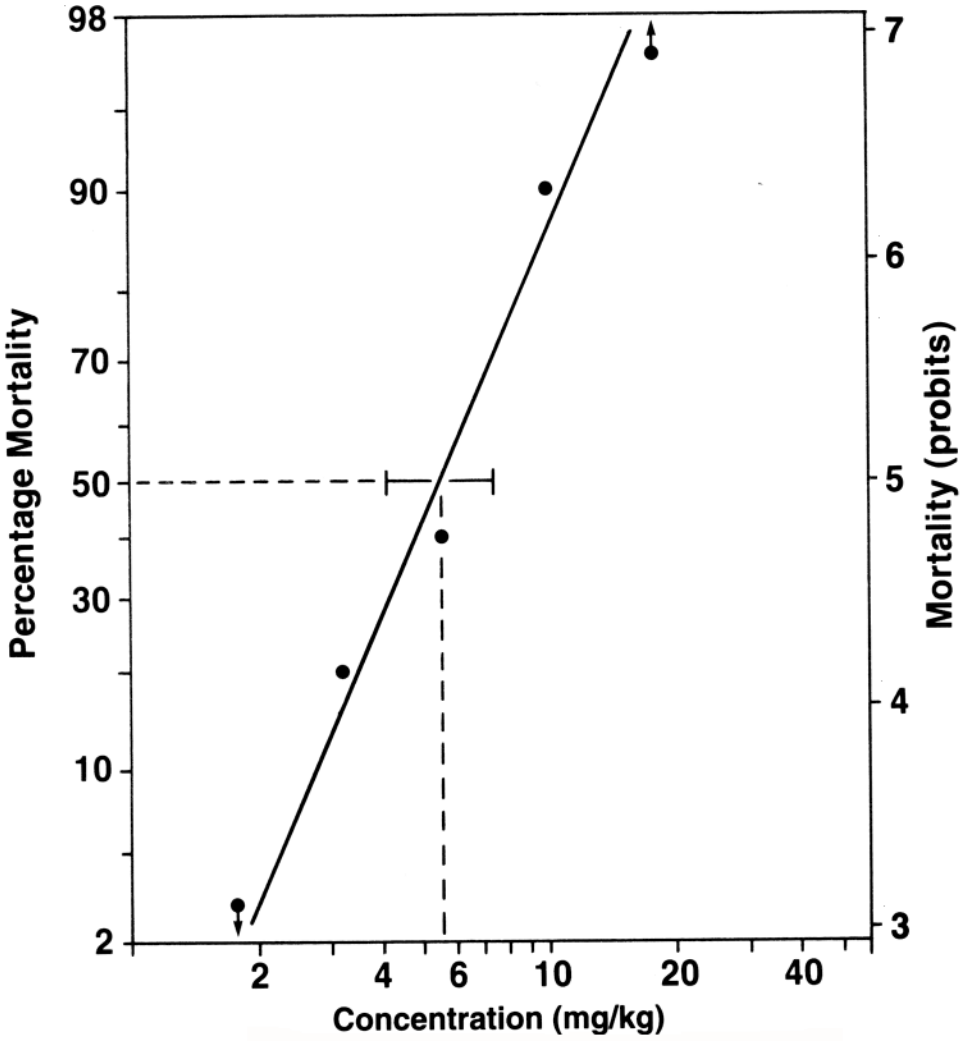


Figure 1 Estimating a Median Lethal Concentration by Plotting Mortalities on Logarithmic-probability Paper

Warning Charts

The discussions in this section assume a working knowledge of basic statistical functions and concepts (e.g., mea, standard deviation, confidence intervals, linear regression, significant difference), and that the reader is familiar with the development and role of warning charts in QA/QC programs (EC, 1990). The necessary familiarity with statistical methods can be acquired by reading some basic texts on statistical concepts (Sokal and Rohlf, 1981; Zar, 1984).

7.1 Establishing and Updating Warning Charts

The results from SSRTTs are used to produce warning charts that demonstrate the ability of laboratory personnel to obtain consistent, precise results with a given test organism and test method. The warning chart is also referred to as a control or mean chart.

The warning chart is prepared by plotting the results of a successive series of SSRTTs on a chart where the x-axis represents the test date or test number and the y-axis indicates the effect concentration. In acute toxicity tests, the y-axis values are usually LC50 and EC50 estimates which are continuous variables that estimate the concentration, or log concentration, that results in a 50 % effect and are usually reported with calculated confidence intervals. In chronic toxicity tests, a common method of analysis involves hypothesis testing, which involves variables that can only be one of the tested concentrations (i.e., the NOEC or LOEC) or the geometric mean between the two (i.e., chronic value or threshold-effect concentrations, TEC). These discrete variables are not appropriate for charting as described herein. An alternative effect variable in chronic tests is a change in rate or graded variable termed the ICp (e.g., IC50, IC25, etc.) for which confidence limits can also be calculated (Norberg-King, 1988).

The mean and standard deviation of a set of SSRTT data can be used to define a range of "normal" or "acceptable" variability in the test. For example, the mean LC50 (arithmetic or transformed) and standard deviation can be calculated for a series of toxicity tests within a single laboratory over a period of time. Given a sufficiently large sample size (e.g., 15 to 20 data points), the concentrations that equal two times the standard deviation above and below the mean ($\bar{x} \pm 2$ SD) represent the upper and lower 95 % confidence limits, respectively, for that data set. These lines ("warning limits") are then plotted on the warning chart (Figure 2). At the 95 % confidence level, 1 in 20 analyses (5 %) would be expected to fall outside of the limits by chance alone. Interpretation of outlying data is discussed in Section 7.2.

The concentrations that equal the mean plus or minus three times the standard deviation ($\bar{x} \pm 3$ SD) represent the 99.7 % confidence limits (which are referred to hereafter as 99 %). At this confidence level, the probability of data falling outside of the limits by chance alone is only 0.3 % (one out of every 333 tests). Inclusion of the 99 % limits on the warning chart is useful in interpreting the severity of outlying data. Severe outliers (outside the 99 % limit) should not be used in any subsequent recalculation of limits.

One of the assumptions underlying the statistics previously described is that a sufficient number of tests has been conducted to give a representative range if variability. To be certain that this is the case, 15 to 20 tests might be necessary (Dux, 1986). This might required

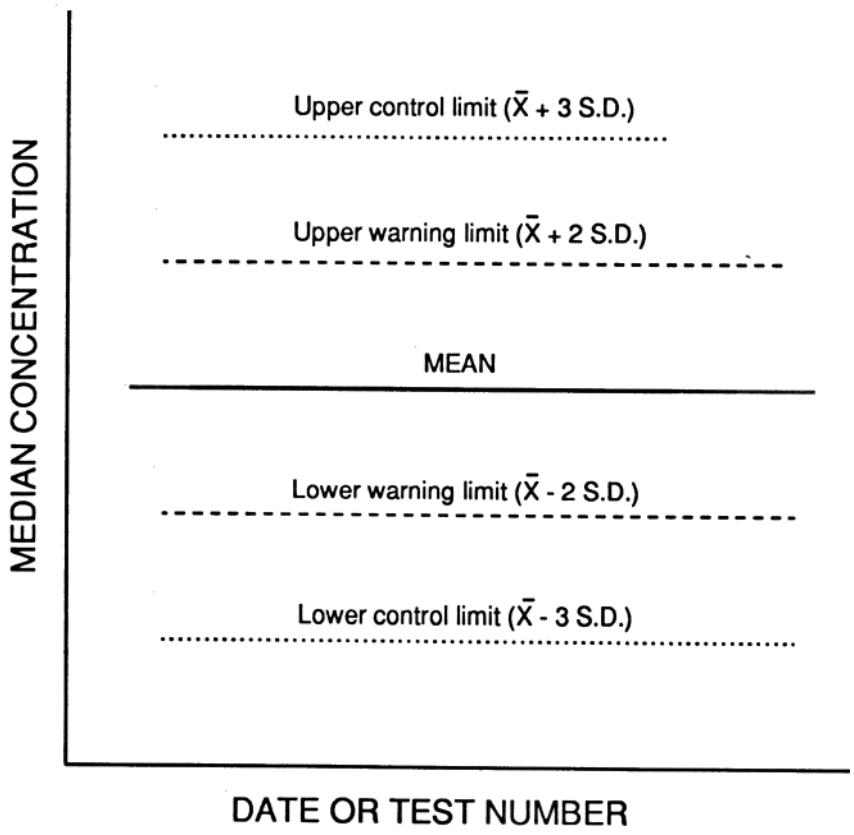


Figure 2 Example of a Warning Chart

considerable time (particularly for chronic tests), and the laboratory will probably want to estimate the toxicity test precision before that time. The USEPA requires that a minimum of five tests be conducted before 95 % confidence limits are established (Weber *et al.,* 1989). The laboratory should be aware that until a large number of tests has been completed, the limits are likely to change with the addition of each new data point to the data set. The limits will generally stabilize over time provided the laboratory cultures are healthy and the sensitivity of the individuals in the culture to the toxicant does not change. Field-collected organisms might exhibit seasonal differences in sensitivity.

Another statistical assumption is that the data are normally distributed. The Shaprio-Wilks test of normality (Sokal and Rohlf, 1981) should be preformed before establishing the control chart. If the raw data are normally distributed, the arithmetic mean and standard deviation are used. Otherwise, a transformation must be performed to normalize the data. Experience has shown that logarithmic transformation will usually result in normality of non-normal LC50 data, but this should be confirmed for each data set. Such data may be charted on the transformed or original scale. If an arithmetic scale is used, the control chart will show the logarithmic (geometric) mean of the data, and the associated 95 % and 99 % confidence limits will not be equidistant about the mean.

If a logarithmic transformation does not result in data normality, a suitable transformation must be found. An Arcsin transformation is often used to convert percentage data (binomial) to normally distributed data (Dowdy and Wearden, 1983). Other possibilities include square-root, cube-root, quadratic-root, and natural logarithmic transformations of the data. The laboratory might want to use the maximum likelihood method of Box and Cox (1964) to choose an optimum transformation.

Separate control charts should be prepared for each reference toxicant-test, species-test method combination. Each new test result (LC50, EC50, or IC50) should be compared against established warning limits, included in the data set. Interpretation of unusual or outlying data is discussed in Section 7.2.

Another consideration in the establishment of warning data or control limits is that data which show a high degree of variability will result in a large standard deviation about the mean, causing warning limits to be wide. Therefore, although a laboratory might not be generating consistent results, it might be able to demonstrate that the data are within the warning limits. No accepted standards regarding the width of the 95 % confidence limits have been found among regulatory authorities that have implemented reference toxicant testing requirements (e.g., USEPA). An objective coefficient of variation (%CV = 100 SD/$\bar{x}$) of 20 % for each water-only toxicity test is suggested (EC, 1990). It is recognized, however, that such factors as the degree of standardization of each test method will also affect test reproducibility.

A higher CV (e.g., 30 %) might be more realistic for some tests. It will not be possible to set specific limits on the width of control limits until sufficient data have been collected from laboratories across Canada demonstrating the degree of reproducibility that can be achieved.

Data should be stored electronically using spreadsheet software to facilitate recalculation of the mean and standard deviation for each data set. The charts can be plotted manually, but they can be more conveniently plotted and undated using the spreadsheet software packages that are commercially available.

7.2 Data Interpretation

7.2.1 Warning Limits

As discussed previously, at the 95 % confidence level, 5 % of the test results would be expected to fall outside of the warning limits due to chance. An outlier should prompt a review of the test system. A mistake in stock solution preparation, a dilution calculation error, or stressed or undernourished test organisms are only some of the possible factors. It is particularly important to examine other QA/QC measures in the laboratory. Examination of control survival during the test, reproductive success of cultured organisms, time to first brood and size of first brood in invertebrate cultures (and in tests when appropriate), dissolved oxygen levels, test temperature, etc., will provide important clues as to whether the outlier occurred by chance or, more likely, was due to a change or problem in the test system.

If an outlier (of warning limits) can be attributed to a specific problem in the test system (e.g., dilution error, miscalculation of data, poor organism health), the value should not be included in recalculation of the limits. If outlier appears to represent normal variability, it should be included in the data set (see Section 7.2.2).

Data from other toxicity tests of the same type (i.e., tests with the organisms from the same batch and same test method) conducted during the period of time (i.e., corresponding to that of the reference toxicant test) might need to be flagged as suspect if the reason for the outlier is not identified, or if it is traced to a factor common to the other test. For example, if an

outlying reference toxicant result for *Chironomus tentans* was attributable to poor culture conditions at that time (e.g., crowding), then other *Chironomus tentans* tests might be suspect. Alternatively, if the outlier were traced to a mistake in the preparation of the stock solution for sediment spiking, the results of concurrent sediment tests may be quite acceptable, since sediments for the two tests are prepared by different procedures. In either case, the test data in question (e.g., test results for field-collected sediment) should be reported with a note detailing interpretation of test results for the reference toxicant and other relevant QA/QC data. The reference toxicant datum (e.g., LC50) that is to be reported with each set of routine test data should be that generated by the most recent reference toxicant test. The period of time for which each reference toxicant datum applies, therefore, depends on the chosen test frequency. Over time, the frequency of data falling outside the 95 % limits should be close to 5 %. If the frequency exceeds 5 %, miscalculation of the limits or a deterioration in precision is indicated. A frequency of less than 5 % might also indicate miscalculation of the warning limits or demonstrate improved test precision. In the latter case, the laboratory might wish to re-establish the warning limits based on more recent data to more closely monitor and maintain the enhanced precision.

7.2.2 Confidence Limits

An outlier from the 99 % confidence limits (e.g., control limit, Figure 2) is unlikely to occur by chance alone. This datum is considered "out-of-control". The test system should be reviewed as outlined in Section 7.2.1. Even if a specific cause cannot be found to account for the outlier, it should not be attributed to chance. Concurrent data for that test system should always be flagged as suspect, The outlier should not be used in recalculation of 95 % and 99 % confidence limits.

7.2.3 Data Trends

It is not only important to monitor whether or not each of the data points falls inside or outside established warning limits, but also to monitor trends or patterns that develop in the data. Out-of-control data can be prevented by early detection of a trend. Probability theory dictated that the probability of any single value falling above or below the mean line is 50 % or ½ (assuming random sources of variation). The probability of two consecutive points being on the same side of the lines is 25 % or 1/4.

The probability of "n" points being on the same side of the line is therefore $1/(2^n)$. If n = 5, the probability is only about 3 % that this occurred through chance alone. Therefore, if five or more consecutive points are on the same side of the mean line, some action should be taken to detect a source of bias (Dux, 1986).

7.2.4 Training New Technicians

Results from tests where a control sediment is spiked with a reference toxicant can be used to judge the progress of new personnel. New technicians should be required to conduct a series of reference toxicant tests until they are able to demonstrate the ability to consistently generate results within established warning limits.

Record Keeping and Data Reporting

The raw data sheets (bench sheets) for each test type should be filed together and kept in a central, easily accessible location. It is extremely important that all relevant records be included on the bench sheets (paper, electronic, or both) such as test data, stock solution preparation date, unusual conditions, test technician, test data, etc. Thorough documentation can reduce the time and expense associated with tracing the source of out-of-control data.

There should be a QA officer in charge of monitoring and updating laboratory QA/QC procedures. The QA officer should also be responsible for scheduling the reference toxicant tests. The schedule should comply with the requirements discussed previously.

Once a reference toxicant test is complete and has been analyzed, the data should be checked by the laboratory supervisor. Bench sheets and results should then be passed on to the QA officer for comparison with existing control limits. Warning charts should then be updated to include the new data, provided that the data are within control limits. Outlier data should trigger an immediate investigation. As discussed in Section 7.2, the investigation might indicate that other data obtained using the same test method are suspect. In this event, it would be desirable for the testing laboratory to repeat the suspect tests after corrective action is taken. This is usually impossible, however, due to limited sample volumes and/or sample aging. Alternatively, the laboratory should report the results of the reference toxicant test, with all suspect data, including an interpretation of the results as to data quality.

The record of the results of an acceptable spiked, control sediment, reference toxicant test should include the following:

- name of the investigator, name of the facility and location of test, dates of the beginning and end of the test;
- source of control sediment and/or constituents including the procedures or conditions of collection, handling, transportation, storage, formulation, and disposal of sediment;
- source of reference toxicant, lot number, and purity;
- source of solvent (if applicable), lot number, purity, concentration(s) used in the solvent control;
- source and chemical composition of overlying water;
- source, history, and age of test organisms, and brood stock (if applicable), taxonomic verification of test species, including name of the person who identified the organisms, and the taxonomic key(s) used, life stage, means and ranges of the body weight and lengths of the test organisms, holding time, acclimation time, culture methods and/or conditions, source and composition of food, and feeding frequency;
- methods for preparing the exposure concentrations;
- experimental design including the number of treatments [e.g., test concentrations and control(s)], number of replicates per treatment, the number of organisms per replicate, and test measurements and frequency;
- test conditions including a description of test chambers (e.g., lighting, temperature,

photoperiod), the depth and volume or weight of sediment and overlying water in each type of replicate, a description of the type of test containers (e.g., size, dimensions, and nature of the material), and description of aeration before or during a test;

- description of the method(s) for spiking the control sediment with the reference toxicant including the frequency and length of mixing times, settling times, and time from addition of the chemical to the sediment to the addition of the test organism to the test container (i.e., time for equilibration);

- composition of sediment, pore water, and overlying water including reference to or description of the analytical methods; and

- biological endpoints for tests, summary of observed effects, tabular summary of data description of methods used for the statistical analyses of data and the results.

Future Research

Routine use of solid-phase sediment toxicity tests for scientific and regulatory purposes necessitates the establishment of tests where a control sediment is spiked with a reference toxicant. However, the science of spiking sediments with chemicals is relatively new and there is a need for information in several areas critical to the development of appropriate test methods. The key areas that required further research are:

1) developing and evaluating techniques for mixing chemicals homogeneously in sediments;

2) studying the effects of storage and equilibration time on the toxicity and bioavailability of the toxicant in the spiked sediments;

3) developing artificial or formulated sediments for use in whole-sediment reference tests;

4) demonstrating whether the SSRTTs are superior to water-only tests;

5) assessing whether one sediment-spiking approach is superior to another for the candidate reference toxicants;

6) assessing the relationship of the effect concentration with the various sediment characteristics to determine the influence of each on toxicity;

7) developing, improving, and standardizing the methods for isolating pore water and determining the amount of freely dissolved toxicant in the isolated pore water.

To date, there are insufficient data to demonstrate the homogeneity of sediment-chemical mixtures. The homogeneity of mixing techniques should be verified for different types of sediment and both the organic and inorganic reference toxicants. Tests should also be conducted to determine whether, or to what extent, mixing techniques alter particle size of control sediments. Another area of potential concern is whether homogenization of the control sediment before spiking and oxidation of the sediment during spiking will affect its binding properties and toxicological characteristics.

Additional research is required to answer the question of how long a spiked sediment should be allowed to equilibrate before commencing a toxicity test. Standard methods are also necessary for the actual process of measuring equilibration. Equilibration from a toxicological perspective might be quite different than chemical equilibrium. Research on the relationship of porewater concentrations to the spiked concentrations, and the sediment concentrations, should be investigated in light of the mortality rates observed.

To eliminate the variation associated with the field collection of control sediments, it is highly recommended that a number of suitable "standard" artificial or formulated sediments be developed for marine, estuarine, and freshwater toxicity tests with whole sediment. Two attributes essential to these sediments are a consistency in the physicochemical characteristics between batches (e.g., particle-size distribution, organic carbon content), and an acceptable and consistent performance of test organisms over the duration of a test. There is a need for a round-robin analyses to evaluate the merits of the SSRTT for assessing interlaboratory test precision with spiked sediments.

Research is also necessary to further investigate the efficacy of the candidate chemicals and other

potential reference toxicants for use in reference tests with spiked control sediments, especially other potential organic compounds.

Fluoranthene may not be the best organic reference toxicant for SSRTTs, but it is the only one for which there is suitable information to assess its worth as a reference toxicant.

References

Adams, W.U., R.A. Kimerle, and R.G. Mosher, "Aquatic Safety Assessment of Chemicals Sorbed to Sediments", pp. 429–453, In: *Aquatic Toxicology and Hazard Assessment,* R.D. Cardwell, R. Purdy, and R.C. Bahner (eds.), Amercian Society for Testing and Materials, Philadelphia, PA (1985).

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 sulphide. *Environ. Toxicol. Chem., 12:*315–320.

Ankley, G. and R. Swartz, United States Environmental Protection Agency (USEPA), personal communication (1994).

APHA, AWWA, and WPCF (American Public Health Association, American Water Works Association, and Water Pollution Control Federation), *Standard Methods for the Examination of Water and Wastewater*, 17th edition, Washington, DC (1989).

ASTM (American Society for Testing and Materials), "Standard Guide for Conducting 10-day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods", ASTM E 1367–92, In: *The Annual Book of ASTM Standards,* Vol. 11.04 (1994), Philadelphia, PA (1994a).

ASTM (American Society for Testing and Materials), "Standard Guide for Conducting Sediment Toxicity Tests with Freshwater Invertebrates", ASTM E 1383–94, In: *The Annual Book of ASTM Standards,* Vol. 11.04 (1994), Philadelphia, PA (1994b).

ASTM (American Society for Testing and Materials), "Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing, ASTM E 1391–90, In: *The Annual Book of ASTM Standards,* Vol. 11.04 (1994), Philadelphia, PA (1994c).

ASTM (American Society for Testing and Materials), "Standard test methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates", ASTM E XXXXX. In: *The Annual Book of ASTM Standards,* Vol. 11.04, Philadelphia, PA (1995).

Birge W.J., J. Black, S. Westerman, and P. Francis. 1987. "Toxicity of sediment-associated metals to freshwater organisms: biomonitoring procedures", pp. 199–218, In: *Fate and Effects of Sediment-bound Chemicals in Aquatic Systems,* D. Dickson, A. Maki, and W. Brungs (eds.), Pergamon Press, New York (1987).

Box, G.E.P. and D.R. Cox. 1964. "Analysis of transformations", *J. Royal Statistical Society, Series B, 26:*211–252.

Burgess, R.M., B.A. Rogers, S.A. Rego, J.M. Corbin, and G.E. Morrison. 1994. Sand spiked with copper as a reference toxicant material for seidment toxicity testing: a preliminary evolution. *Arch. Envrion. Contam. Toxicol. 26:*163–168.

Burton, Jr., G.A. 1991. Assessing the toxicity of freshwater sediments. *Environ. Toxicol. Chem. 10:*1585–1627.

Cairns, M.A., A.V. Nebeker, J.H. Gakstatter, and W.L. Griffis, 1984. Toxicity of copper-spiked sediments to freshwater invertebrates. *Environ. Toxicol. Chem., 3:*435–445.

Carlson, A.R., G.L. Phipps, V.R. Mattson, P.A. Kosian, and A.M. Cotter. 1991. The role acid-volatile sulphide in determining cadmium bioavailability and toxicity in

freshwater sediments. *Environ. Toxicol. Chem., 10:*1309–1319.

Clark, J.R., J.M. Patrick, Jr., J.C. Moore, and E.M. Lores. 1987. Waterborne and sediment-source toxicities of six organic chemicals to grass shrimp *(Paleomonetes pugio)* and amphioxus *(Branchiostoma caribaeum). Arch. Environ. Contam. Toxicol., 16:*401–407.

Clements, W.H., Colorado State University, personal communication (1995).

Day, K., Environment Canada, personal communication (1993).

DeWitt, T.H., R.C. Swartz, and J.O. Lamberson. 1989. Measuring the acute toxicity of estuarine sediments, *Environ. Toxicol. Chem., 8:*1035–1048.

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. Smith. 1992. The influence of organic matter quality on the toxicity and partitioning of sediment-associated fluoranthene. *Environ. Toxicol. Chem., 11:*197–208.

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

Ditsworth, G.R., D.W. Schults, and J.K.P. Jones. 1990. Preparation of benthic substrates for sediment toxicity testing. *Environ. Toxicol. Chem., 9:*1523–1529.

Di Toro, D.M., L.M. Horzempa, M.M. Casey, and W. Richardson. 1982. Reversible and resistant components of PCB-absorption-desorption. Adsorbent concentration effects. *J. Great Lakes Res., 8:*336–349.

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 sulphide. *Environ. Toxicol. Chem., 9:*1487–1502.

Di Toro, D.M., C.J. Zarba, D.J. Hansen, W.J. Berry, R.W. Swartz, C.E. Cowan, S.P. Pavlou, H.E. Allen, N.A. Thomas, and P.R. Paquin. 1991. Technical basis for establishing sedimnet quality criteria for non-ionic organic chemicals using equilibrium partitioning. *Environ. Toxicol. Chem. 10:*1541–1583.

Doe, K. and A. Murdoch, Environment Canada, personal communication (1994).

Dowdy, S. and S. Weardon. 1983. *Statistics for Research,* pp. 299–301. John Wiley and Sons, New York, NY.

Dux, J.P. 1986. *Handbook of Quality Assurance for the Analytical Chemistry Laboratory.* Van Nostrand Reinhold Company, New York, NY.

Environment Canada, Guidance Document for Control of Toxicity Test Precision Using Reference Toxicants, EPS 1/RM/12, Conservation and Protection, Ottawa, ON.

Environment Canada. 1992. Acute test for sediment toxicity using marine of estuarine amphipods. Environmental Protection Series, EPS 1/RM/26. Conservation and Protection, Ottawa, ON.

Environment Canada. 1997a. Test for Growth and survival in sediment using larvae of freshwater midges *(Chironomus tentans* or *Chironomus riparius)*. Environmental Protection Series, EPS 1/RM/32. Conservation and Protection, Ottawa, ON.

Environment Canada. 1997b. Test for Growth and survival using the freshwater amphipod

Hyalella azteca. Environmental Protection Series, EPS 1/RM/33. Conservation and Protection, Ottawa, ON.

Environment Canada. 1994. Guidance Document for the Collection and Preparation of Sediments for Physicochemical Characterization and Biological Testing. Environmental Protection Series, EPS 1/RM/29. Conservation and Protection, Ottawa, ON.

EVS Consultants. 1991. Phase V studies of 10-day tests for sediment toxicity using marine or estuarine infaunal amphipods. EVS Consultants, North Vancouver, BC. Project No. 3/047-24.2, 37 p. + app.

Foster, G.D., S.M. Baksi, and J.C. Means. 1987. Bioaccumulation of trace organic contaminants from sediment by Baltic clams *(Macoma balthica)* and soft-shelled clams *(Mya arenaria). Environ. Toxicol. Chem., 6:*969–976.

Francis., P.C., W.J. Birge, and J.A. Black. 1984. Effects of cadmium-enriched sediment on fish and amphibian embryo-larval stages. *Ecotoxicol. Environ. Safety, 8:*378-387.

Fremling, C.R. and W.L. Mauck. 1980. Methods for using nymphs of burrowing mayflies (Ephremeroptera; *Hexagenia*) as toxicity test organisms, pp. 81–97. In: *Aquatic Invertebrate Bioassays,* A.L. Buikema, Jr. and J. Cairns, Jr., (eds.), ASTM STP 715, American Society for Testing and Materials. Philadelphia, PA.

Gambrell, R., Louisiana State University, personal communication (1993).

Geisy, J., University of Michigan, personal communication (1992).

Green, A.S., G.T. Chandler, and E.R. Blood. 1993. Aqueous-, porewater-, and sediment-phase cadmium: toxicity relationships for a meiobenthic copepod. *Environ. Toxicol. and Chem., 12:*1497–1506.

Håkanson, L. 1984. Sediment sampling in different aquatic environments: statistical aspects. *Water Resour. Res, 20:*41–46.

Hamr, P., R.S. Kirby, P. Gillis, K.E. Day. 1994. Development of methodologies for whole-sediment toxicity tests with benthic invertebrates: a) optimum age for initiation and duration of the *Hyalella azteca* survival and growth test, b) Formulation of artificial sediment(s) for use in reference toxicity tests with benthic invertebrates. Prepared for Technology Development Directorate, Environment Canada, Ottawa, ON, 28 p.

Hanes, E.C., J.J.H. Ciborowski, and L.D. Corkum. 1991. Standardized rearing materials and procedures for *Hexagenia*. A benthic aquatic bioassay organism. Report prepared for the Research Advisory Committee, Ontario Ministry of Enviornment and Energy, Toronto, ON.

Henry, M.G., D.N. Chester, and W.L. Mauck. 1986. Role of artificial burrows in *Hexagenia* toxicity tests: recommendations for protocol development. *Environ. Toxicol. Chem., 5:*553–559.

Hickey, C. and P. Roper, NIWA Ecosystems (unpublished data) (1995).

Hubert, J.J., "PROBIT2: A Microcomputer Program for Probit Analysis", Dept. of Mathematics and Statistics, University of Guelph, Guelph, Ontario (1987).

Jenne, E.A. and J.M. Zachara. 1987. Factors influencing the sorption of metals. pp. 83–98, In: *Fate and Effects of Sediment-bound Chemicals in Aquatic Systems.* K.L. Dickson, A.W. Maki, and W.A. Brungs (eds.), Pergamon Press, Elmsford, NY.

Karickhoff, S.W. 1980. Sorption of hydrophobic pollutants in natural sediments. pp. 193–205, In: *Contaminants and Sediments.* R.A. Baker (ed.). Ann Arbor Science Publishers, Ann Arbor, MI.

Karickhoff, S.W. and K.R. Morris. 1985. Sorption dynamics of hydrophobic pollutants in sediment suspensions. *Environ. Toxicol. Chem., 4:*469–479.

Keilty, T.J., D.S. White, and P.F. Landrum. 1988a. Sublethal repsonses to endrin in sediment by *Limnofrilius hoffmeisteri* (Tuificidae) and in mixed-culture with *Stylodrilius heringianus* (Lumbriculidae). *Aquat. Toxicol., 13:*251–270.

Keilty, T.J., D.S. White, and P.F. Landrum. 1988b. Sublethal responses to endrin in sediment by *Stylodrilius heringianus* (Lumbriculidae) as measured by a ^{137}Cesium marker layer technique. *Aquat. Toxicol., 13:*227–250.

Kemble, N.E., F.J. Dwyer, and C.G. Ingersoll. 1994. Development of a formulated sediment for use in whole-sediment toxicity testing. SETAC Abstract, Annual Meeting of the Society of Environmental Toxicology and Chemistry, Denver, CO.

Kersten, M. and U. Förstner. 1987. Cadmium associations in freshwater and marine sediment, pp. 51–88. In: *Cadmium in the Aquatic Environment,* J.O. Nriagu and J.B. Sprague (eds.), John Wiley & sons, NewYork, NY.

Kukkonen, J. and P.F. Landrum. 1994. Toxicokinetics and toxicity of sediment-associated pyrene to *Lumbriculus variegatus* (Oligochaeta). *Environ. Toxicol. Chem., 13:*1457–1468.

Landrum, P.F. 1989. Bioavailability and toxicokinetics of polycyclic aromatic hydrocarbons sorbed to sediments for the amphipod, *Pontoporeia hoyi. Environ. Sci. Tech., 23:*588–585.

Landrum, P.F. and R. Poore. 1988. Toxicokinetics of selected xenobiotics in *Hexagenia limbata. J. Great Lakes Res., 14:*427–437.

Landrum, P.F. and W.R. Faust. 1991. Effect of variation in sediment composition on the uptake rate coefficient for selected PCB and PAH congeners by the amphipod, *Diporeia* sp., pp. 263–279. In: Aquatic Toxicology and Risk Assessment. MA. Mayes and M.G. Barron (eds.), ASTM STP 1124, American Associaton for Testing and Materials. Philadelphia, PA.

Landrum, P.F., W.R. Faust, and B.J. Eadie. 1989. Bioavailability and toxicity of a mixture of sediment-associated chlorinated hydrocarbons to the amphipod *Pontoporeia hoyi,* pp. 315–329, In: Aquatic Toxicity and Hazard Assessment: Volume 12, ASTM STP 1027. U.M. Cowgill and L.R. William (eds.), American Society for Testing and Materials. Philadelphia, PA.

Landrum P.F., B.J. Eadie, and W.R. Faust. 1991. Toxicokinetics and toxicity of a mixture of sediment-associated polycyclic aromatic hydrocarbons to the amphipod *Diporeia* sp. *Environ. Toxicol. Chem., 10:*35–46.

Landrum P.F., B.J. Eadie, and W.R. Faust. 1992. Variation in the bioavailability of polycyclic aromatic hydrocarbons to the amphipod *Diporeia* spp. with sediment aging. *Environ. Toxicol. Chem., 11:*1197–1208.

Landrum, P.F., W.S. Dupuis, J. Kukkonen. 1994. Toxicokinetics and toxicity of sediment-associated pyrene and phenanthrene in *Diporeia* spp.: Examination of equilibrium-partitioning theory and residue-based effects for assessing hazard. *Environ. Toxicol. Chem., 13:*1769–1780.

Landrum, P.F., NOAA, personal communication (1994).

Lydy, M.J., K.A. Bruner, K.A. Fry, and S.W. Fisher. 1990. Effects of sediment and the route of exposure on the toxicity and accumulation of neutral lipophilic and moderately water soluble metabolizable compounds in the midge, *Chironomus riparius*, pp. 140–164. In: *Aquatic Toxicology and Risk Assessment: Volume 13,* W.G. Landis and W.H. van der Schalie (eds.). ASTM STP 1906, American Society for Testing and Materials. Phildelphia, PA.

MacPherson, C. EVS Consultants, Ltd., personal communication (1992).

Malueg, K.W., G.S. Schuytema, and D.F. Krawczyk. 1986. Effects of sample storage on a copper-spiked freshwater sediment. *Environ. Toxicol. Chem., 5:*245–253.

McLeese, D.W., C.D. Metcalfe, and D.S. Pezzack. 1980. Uptake of PCBs form sediments by *Nereis virens* and *Crangon septemspinosa. Arch. Environ. Contam. Toxicol., 9:*507–518.

McLusky, D.S., V. Bryant, and R. Campbell. 1986. The effects of temperature and salinity on the toxicity of heavy metals to marine and estuarine invertebrates. *Oceangr. Mar. Biol. A Rev., 24:*481–520.

Murdoch, A., Environment Canada, personal communication (1994).

Muir, D.C.G., N.P. Grift, B.E. Townsend, D.A. Metner, and W.L. Lockhart. 1982. Comparison of the uptake and bioconcentration of fluoridone and terbutryn by rainbow trout and *Chironomus tentans* in sediment and water systems. *Arch. Envrion. Contam. Toxicol., 11:*595–602.

Muir, D., Fisheries and Oceans Canada, personal communication (1993).

Naylor, C. and Q.C.J. Rodrigues. 1994. The use of the OECD earthworm soil (a formulated sediment) in tests with *Chironomus riparius.* SETAC abstract, *15th Annual Meeting of the Society for Environmental Toxicology and Chemistry,* Denver, CO.

Nebeker, A.V., S.T. Onjukka, M.A. Cairns, and D.F. Krawczyk. 1986. Survival of *Daphnia magna* and *Hyalella azteca* in cadmium-spiked water. *Environ. Toxicol. Chem., 5:*933–938.

Nelson, P.O., A.K. Chung, and M.C. Hudson. 1981. Factors affecting the fate of heavy metals in the activated sludge process. *J. Water Pollut. Control Fed., 53:*1323–1333.

Nkedi-Kizza, P., P.S.C. Rao, and A.G. Hornsby. 1985. Influence of organic cosolvents on sorption of hydrophobic organic chemicals by soils. *Environ. Sci. Technol., 19:*975–979.

Norberg-King, T. 1988. An interpolation estimate for chronic toxicity: The IC50 approach. Natural Effluent Toxicity Assessment Centre Technical Report 05–88, Sept. 1988, Duluth, MN.

Oakley, S.M., C.E. Delphey, K.J. Williamson, and P.O. Nelson. 1980. Kinetics of trace metal partitioning in model anoxic marine sediments. *Water Res., 14:*1067–1072.

Ozretich, R.J. and W.P. Schroeder. 1986. Determination of selected neutral priority organic pollutants in marine sediment, tissue and reference materials utilizing bond-phase sorbents. *Anal. Chem., 58:*2041–2048.

Paine, M.D. and C.A MacPherosn. 1991. Phase IV studies of the 10-day tests for sediment toxicity using marine or estuarine infaunal

amphipods. Project No. 3/047-33. EVS Consultants, Vancouver, BC, 71 p.

Pascoe, D., A.F. Brown, B.M.J. Evans, and C. McKavanagh. 1990. Effects and fate of cadmium during toxicity tests with *Chironomus riparius* - the influence of food and artificial sediment. *Arch. Environ. Contam. Toxicol., 19:*872–877.

Pesch, C.E. and D. Morgan. 1978. Influence of three sediment types on copper toxicity to the polychaete *Neanthes arenaceodentata. Water Res., 12:*747–751.

Phipps, G.L., V.R. Mattson, and G.T. Ankley. 1994. The relative sensitivity of three benthic test species to 10 chemicals. *Arch. Environ. Toxicol. Chem*. (in review).

Podoll, R.T. and W.R. Mabey. 1987. Factors to consider in conducting laboratory sorption/desorption tests, pp. 83–98. In: *Fate and Effects of Sediment-bound Chemicals in Aquatic Systems.* K.L Dickson, A.W. Maki, and W.A. Brungs (eds.) Pergamon Press, Elmsford, NY.

Ray, S., D. McLeese, and D. Pezzack. 1980. Accumulation of cadmium by *Nereis virens*, *Arch. Environ. Contam. Toxicol., 9:*1–8.

Ribeiro, R., R. Margalho, F. Goncalves, and A.M.V.M. Soares. 1994. A novel formulated sediment for toxicity testing with a highly controlled composition sediment. SETAC Abstract, Annual Meeting of the Society of Environmental Toxicology and Chemistry, Denver, CO.

Robinson, A.M., J.O. Lamberson, F.A. Cole, and R.C. Swartz. 1988. Effects of culture conditions on the sensitivity of a Phoxocephalid amphipod, *Rhepoxynius abronius,* to cadmium in sediment. *Environ. Toxicol. Chem., 7:*953–959.

Sawyer, L.N. and G.A. Burton, Jr. 1994. Validation of various formulated sediment recipes for use in toxicity assessments. SETAC Abstract, Annual Meeting of the Society of Environmental Toxicology and Chemistry, Denver, CO.

Schults, D., USEPA, personal communication (1992).

Schuytema, G.S., P.O. Nelson, K.W. Malueg, A.V. Nebeker, D.F. Krawczyk, A.K. Ratcliff, and J.H. Gakstatter. 1984. Toxicity of cadmium in water and sediment to *Daphnia magna*. *Environ. Toxicol. Chem., 3:*293–308.

Sokal, R.R. and F.J. Rohlf. 1981. *Biometry* 2nd edition, W.H. Freeman and Company, New York, NY, 859 p.

Steel, R.G.D. and J.H. Torrie. 1960. *Principles and Procedures of Statistics*. McGraw-Hill Book Co., New York, NY. 633 p.

Stemmer, B.L., G.A. Burton, Jr., and S. Leibfritz-Frederick. 1990. Effect of sediment test variables on selenium toxicity to *Daphnia magna. Environ. Toxicol. Chem., 9:*381–389.

Stephan, C.E., "Methods for Calculating an LC50", p. 65–84, In: *Aquatic Toxicology and Hazard Evaluation,* F.L. Mayer and J.L. Hamelink (eds.), ASTM STP 634, American Society for Testing and Materials, Philadelphia, PA (1977).

Stephenson, G.L., K.E. Day, R. Scroggins, R. Scott Kirby, and P. Hamr. 1994. Current status of Environment Canada's guidance on control of test precision using a spiked sediment toxicity test: what to spike and how?, In: *Environmental Toxicology and Risk Assessment:* ASTM STP__. T.W. La Point and C.G. Ingersoll (eds.), American Society for Testing and Materials, Philadelphia, PA (in press).

Suedel, B.C. and J.H. Rodgers. 1994. Development of formulated reference sediments for freshwater and estuarine sediment testing. *Environ. Toxicol. Chem., 13:*1163–1175.

Suedel, B.C., J.H. Rodgers, and P.A. Clifford. 1993. Bioavailability of fluoranthene in freshwater sediment toxicity tests. *Environ. Toxicol. Chem., 12:*155–165.

Swartz, R.C., G.R. Ditsworth, D.W. Schults, and J.O. Lamberson. 1985. Sediment toxicity to a marine infaunal amphipod: Cadmium and its interactrion with sewage sludge. *Mar. Environ. Res., 18:*133–153.

Swartz, R.C., P.F. Kemp. D.W. Schults, and J.O. Lamberson. 1988. Effects of mixtures of sediment contaminants on the marine infaunal amphipod, *Rhepoxynius abronius. Environ. Toxicol. Chem., 7:*1013–1020.

Swartz, R.S., D.W. Schults, T.H. DeWitt, G.R. Ditsworth, and J.O. Lamberson. 1990. Toxicity of fluoranthene in sediment to marine amphipods: A test of the equilibrium partitioning approach to sediment quality criteria. *Environ. Tox. Chem., 9:*1071–1080.

USEPA. 1991. "Methods for measuring the acute toxicity of effluents to freshwater and marine organisms", 4th edition, United States Environmental Protection Agency, Cincinnati, Ohio. Report EPA/600/4-90-027.

USEPA. 1994a. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. United States Environmental Protection Agency, Duluth, MN. EPA 600/R-94/024.

USEPA. 1994b. Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods. Office of Research and Development, United States Environmental Protection Agency. Washington, DC. EPA/600/R-94/025.

Walsh, G.E., D.E. Weber, L.K. Esry, M.T. Nguyen, J. Noles, and B. Albrecht. 1991. Synthetic substrata for propagation and testing of soil and sediment organisms. *Pedobiologia, 26:*1–10.

Weber, C.I., W.H. Peltier, T.J. Norberg-King, W.B. Horning II, F.A. Kessler, J.R. Menkedick, T.W. Neiheisel, P.A. Lewis, D.J. Klemm, Q.H. Pickering, E.L. Robinson, J.M. Lazorchak, L.J. Wymer and R.W. Freyberg. 1989. Short-term methods for estimating the chronic methods of effluents and receiving waters to freshwater organisms. 2nd edition, USEPA. EPA/600/4-89/001.

Webster, G.R.G., M.R. Servos, G.G. Choudhry, L.P. Sarna, and G.C.G. Muir. 1990. Methods for dissolving hydrophobics in water for studies of their interactions with dissolved organic matter, pp. 191–192, In: Abstracts. Vol. 27, Advances in Chemistry Series, presented at *193rd National Meeting of the American Chemical Society, Division of Environmental Chemistry.*

Word, J.Q., J.A. Ward, L.M. Franklin, V.I. Cullinan, and S.L. Kiesser. 1987. Evaluation of the equilibrium partitioning theory for estimating the toxicity of the nonpolar organic compound DDT to the sediment dwelling amphipod *Rhepoxynius abronius.* Battelle/Marine Research Laboratory Report (Task 1, WA56). Sequim, WA. 60 p.

Yee, S., M. Van Rikxoort, and D. McLeay. 1992. The effect of holding time on *Eohaustorius washingtonianus* during ten-day sediment bioassays and reference toxicant tests. Report prepared for Environment Canada and the Inter-Governmental Aquatic Toxicity Group, North Vancouver, BC, 53 p.

Zar, J.H. 1984. *Biostatistical Analysis.* 2nd edition, Prentice-Hall, Inc., N.J. 718 p.

Individuals who Provided Information on Reference Toxicant Tests using a Spiked Control Sediment

Richard Swartz
USEPA
Hatfield Marine Science Center
Newport, OR 97365

Kristin Day
National Water Research Institute
Environment Canada
867 Lakeshore Blvd., P.O. Box 5050
Burlington, Ontario L7R 4A6

John Geisy
Michigan State University
Department of Fisheries and Wildlife
#13 Natural Resources Building
East Lansing, MI 48824

Don Schults
USEPA
Hatfield Marine Science Center
Newport, OR 97365

Ken Doe
Aquatic Toxicology
EP Laboratory Division
Environment Canada, C&P
Dartmouth, Nova Scotia

Robert Gambrell
Louisiana State University
Baton Rouge, LA 70810

Derek Muir
Freshwater Institute
Department of Fisheries and Oceans
501 University Crescent
Winnipeg, MB R3T 2N6

Cathy MacPherson
EVS Consultants
195 Pemberton Ave.
North Vancouver, BC V7P 2R4

Alena Murdoch
Lakes Research Branch
National Water Research Institute
867 Lakeshore Road
Burlington, Ontario L7R 4A6

Jill Jones
AScI Corporation
2111 S.E. Marine Science Dr.
Newport, OR

Chris Ingersoll
National Biological Survey
4200 New Haven Road
Columbia, MO 65201

Gary Ankley
Environmental Research Lab
6201 Congdon Boulevard
Duluth, MN 55804-1136

Peter Landrum
Great lakes Environmental Research Lab
2205 Commonwealth Boulevard
Ann Arbor, MI 48105-1593

Chris Hickey
NIWA Ecosystems
P.O. Box 11-115
Hamilton, New Zealand

Logarithmic Series of Concentrations Suitable for Toxicity Tests (adapted from Environment Canada, 1992)

Column (number of concentrations between 10.0 and 1.00, or between 1.0 and 0.10)*

1	2	3	4	5	6	7
10.0	10.0	10.0	10.0	10.0	10.0	10.0
3.2	4.6	5.6	6.3	6.8	7.2	7.5
1.00	2.2	3.2	4.0	4.6	5.2	5.6
0.32	1.00	1.8	2.5	3.2	3.7	4.2
0.10	0.46	1.00	1.6	3.3	2.7	3.2
	0.22	0.56	1.00	1.5	1.9	2.4
	0.10	0.32	0.63	1.00	1.4	1.8
		0.18	0.40	0.68	1.00	1.3
		0.10	0.25	0.46	0.72	1.00
			0.16	0.32	0.52	0.75
			0.10	0.22	0.37	0.56
				0.15	0.27	0.42
				0.10	0.19	0.32
					0.14	0.24
					0.10	0.18
						0.13
						0.10

* A series of five (or more) successive concentrations may be chosen from a column. Mid-points between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed as percentage by weight (e.g., mg/kg) or weight-to-volume (e.g., mg/L). As necessary, values may be multiplied or divided by any power of 10. Column 1 might be used if there was considerable uncertainty about the degree of toxicity. More widely spaced concentrations (differing by a factor <0.3) should not be used. Columns 4 to 7 might be useful for toxicants that have an abrupt threshold of effect.

Information Typically Provided in a Material and Safety Data Sheet

MSDS Identification

Record Number	435002
Language	English
Product Name(s)	Cadmium chloride
Product Id.	MSDS 2350-1; PIN2570; PCN 2350-1
Date of MSDS	1991-05-22

Manufacturer Information:
Address
Emergency Telephone Numbers
Disclaimers

Product Identification:

Name/Synonyms	Cadmium chloride
Chemical Family	Inorganic salt
Chemical Formula	$CdCl_2$
Product Use	Laboratory solvent

Hazardous Ingredients:
100 % cadmium chloride

TLV Units	0.05 mg/m^3 (as Cd)

Physical Data:

Physical state	solid
Odour/appearance	white, odourless, crystals
Vapour pressure	10 mm Hg at 656° C
Boiling point	960° C
Melting point	568° C
pH	4.0 to 6.5 (5 % sol'n)
Specific gravity	4.0

Shipping Information:

PIN	2570
T.D.G.	Class 6.1
Pkg. Group	II

Reactivity Data:

Chemical stability	normally stable
Incompatible with	potassium, strong oxidizers, acids
Hazardous decomp. compds.	cadmium oxides, chlorine compds.

Fire and Explosion Data:
Flammability
Extinguishing media
Flash point
Autoignition Temperature
Hazardous combustion products
Sensitivity to impact
Sensitivity to static discharge

Toxicological Properties and Health Data:
LD50 (oral, rat)
Effects of acute exposure to product
Inhaled
In contact with skin
In contact with eyes
Ingested
Effects of chronic exposure to product
Carcinogenicity
Teratogenicity
Reproductive effects
Mutagenicity
Synergistic products

Preventive Measures:
Engineering controls
Respiratory protection
Eye protection
Skin protection
Other personal protection
Leak and spill procedures
Waste disposal
Handling procedures and equipment
Storage requirements

First Aid Measures:
Eyes
Skin
Inhalation
Ingestion

Formulated Sediments: A Summary of Current Reserach

There are a number of ways to formulate freshwater sediments. These procedures usually involve similar constituents, only the relative composition of the constituents, or the procedures for combining the constituents differ (Table 3). Generally the source of the constituents is comparable.

Table 3 Relative Composition of Constituents (ASTM, 1994; USEPA, 1994)

Reference	Sources of Material	Formulation	Procedures
Walsh *et al.* 1991	Mystic White Sand No. 85, 45, 18 from New England Silica Clay/silt from Engelhard Corp.	**% Dry Weight** Coarse sand (500 to 1500 µm) 0.6 Medium sand (250 to 499 µm) 8.7 Fine sand (63 to 249 µm) 69.2 Silt 10.2 Clay 6.4 Organic matter (peat moss) 4.9	Grind peat moss and sieve with a screen having a pore size of 840 µm. Mix constituents dry and hydrate.
Clements, W.H. 1995	Mystic White Sand No. 45 Clay/silt ASP® 400 from Engelhard Corp.	**Dry Weight (g)** Sand 1242 Silt/Clay 219 Dolomite 7.5 Peat Moss 31.5 Humic acid 0.15	Rinse peat moss then soak for 5-d in de-ionized water with daily renewal of water. Remove from water and air dry, grind, and sieve through the following pore sizes 1.18 mm (discard the material retained on the sieve), 1.00 mm; 0.85 mm; 0.60 mm, and 0.425 mm. Combine the above material such that the peat moss has an average particle size of 840 µm. Wash sand and dry at 105° C. Combine the constituents in their dry form and mix on a rolling mill for 1 h and store (dry) until ready to use.
Hanes *et al.* 1991	Silica sand 180 to 500 µm Lewiscraft® sculpting clay No Name Potting Soil®	**% Dry Weight** Sand mixture 42 Clay 42 Soil 16	Sieve sand and retain two particle sizes 90 to 180 µm and 180 to 250 µm which are then combined in a ratio of 2:1. Air dry the potting soil and sieve with a 1-mm screen. Determine percent moisture in clay and soil after drying for 24 h at 60 to 100° C. Correct for the percent moisture of constituents when combining on the basis of dry weight. The mixture is autoclaved for 20 minutes and then stored until required for use.

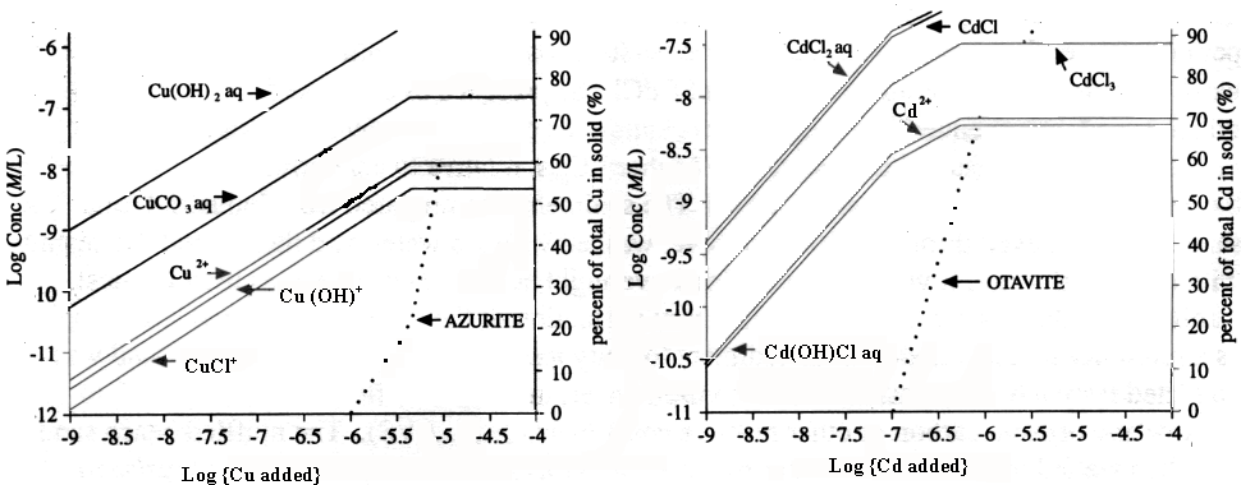
Table 3 Relative Composition of Constituents (cont'd)(ASTM, 1994; USEPA, 1994)

Reference	Sources of Material	Formulation	Procedures
Suedel and Rodgers, 1994	Mystic White Sand No. 18, 90 from New England Silica Inc. Silt-ASP® 400 Clay-ASP® 600 and 900 Clay as montmorillonite/kaolin from Engelhard Corp. Dolomite from Ward's Natural Science Establishment Inc. Humus from Sims Bark Co. Inc.	Formulated to match natural sediments, therefore variable composition: **% Dry Weight** Sand 7.1 to 92.0 Silt 8.0 to 92.9 Clay 0 to 3.5 Organic matter 0.17 to 8.4	Sand is dry sieved to provide three different size classes 500-2000, 250-500, and 50-200 µm. Silt, clay, and dolomite are ashed at 550° C for 1 h to remove organic matter. Humus is dried at 70° C and milled to 2.0 mm. Dolomite is added as 1 % of the silt requirement. Constituents are combined dry and hydrated with dilution water before use. A conditioning period of 7 d is required.
Hamr, P. *et al.* 1994	Fine silica and from American Colloid Company Allen R Clay from Stochem Inc. Dolomite from Redland Quarries Cerophyl from Sigma Chemical	**% Dry Weight** Sand 25 to 75 Silt 33 Clay 25 to 75 Cerophyl 1 to 6 Dolomite 0.5	Sand is washed and dried then combined with the dry silt and clay. The mixture is hydrated with dilution water (600–800 g) solid phase with 600 mL water) and the slurry swirled for 24 h. The slurry was allowed to settle for 3 d (conditioning) with gentle aeration of the overlying water. Dolomite in solution and aged cerophyl were added and mixed for 1 h. The slurry was allowed to stand overnight and the excess overlying water decanted and discarded before the enriched sediment is used in a test.
Naylor and Rodrigues, 1994		**% Dry Weight** Sand 69 Kaolin 20 Peat moss 10 $CaCO_3$ 1	Acid wash the sand and sieve to obtain particle sizes from 40 100 mm. Grind peat moss and sieve through a 2-mm screen (do not dry completely because it will float when substrate mixture is hydrated). Constituents are mixed by dry weight (i.e., adjusted for percent moisture). Mix for 2 h in a soil shaker.

The number of research projects addressing the development of artificial or formulated marine sediments is smaller than that for freshwater sediments (Walsh *et al.,* 1991). There appears to be a tendency to using collected sand from beaches as the base sediment for toxicity assessments. The sand is manipulated and prepared for testing by either washing, wet sieving, and oven drying the sand, or modifying the sand by placing it in a muffle oven at high temperature for a fixed period of time. In comparative liquid-phase versus substrate tests with *Eohaustorius washingtonianus* and the reference toxicant $CdCl_2$ the presence of Cd-spiked substrate (fine-grained sand, resuspended in a range of concentrations of $CdCl_2$ reduced the variability in mortality data and resulted in consistently higher LC50s, relative to concurrent tests performed in the absence of substrate (Yee *et al.,* 1992). The sediment was prepared from industrial sand. The sand was wet-sieved through a 600 µm sieve, washed with tap water, and then rinsed thoroughly with de-ionized water. It was then oven-dried overnight at 105° C before it was used in a test. Similar procedures were used in a preliminary evaluation of the use of sand spiked with copper as a reference toxicant material for sediment toxicity testing (Burgess *et al.*, 1994). Sand was collected from a beach, washed with de-ionized water, and then muffled at 450° C for 6 h to remove and organic carbon of other reactive constituents (e.g., AVS). The muffled beach sand was then wetted with de-ionized water by manually swirling the mixture, and the water/sand slurry was amended with crystalline $CuCl_2$ to produce 1.5 kg of a Cu-sand mixture with a concentration of 614 mg Cu/kg sand. This spiked substrate was allowed to equilibrate for five months before it was used in toxicity tests. Hickey and Roper (1995) observed a noticeable difference in the rates of drift and movement of shellfish exposed to muffled and unmuffled fresh sand. This difference was mitigated by 5 days of preconditioning with overlying water. This preconditioning period would be concurrent with the equilibration period for spiking.

Fate of Copper and Cadmium Added to Natural Seawater

The following graphs outline changes in the concentrations of various dissolved Cu and Cd species after the addition of the metals to the seawater (A. Murdoch, 1994). The concentrations of Cu and Cd in natural seawater (assuming salinity of 35 ‰ and an ionic strength of 0.714) are plotted on the y axis, and increasing concentrations of added Cu or Cd on the x axis. Both concentrations are given in *M*/L. The dotted line indicates the percentage of precipitated Cu [as azurite, $Cu_3(CO_3)_2$] or Cd (as otavite, $CdCO_3$) after the addition of about 0.63 mg Cu/L, and 98 % of Cu precipitates as azurite by the addition of about 6.3 mg/L. For Cd, otavite starts to precipitate after the addition of 0.01 mg Cd/L, and almost 80 % of Cd becomes precipitated as otavite after the addition of 0.11 Cd/L.



Solubility of Copper and Cadmium in Seawater at 25° C with pH 8.21